

**WHO GETS THE FLOOR?
GENDER DIFFERENCES IN SPEAKING, INFLUENCE, AND EGALITARIAN
PREFERENCES IN GROUPS**

EMILY A. DRISDELLE

A THESIS SUBMITTED TO THE FACULTY
OF GRADUATE STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF ARTS

GRADUATE PROGRAM IN PSYCHOLOGY
YORK UNIVERSITY
TORONTO, ONTARIO

JUNE 2026

© EMILY DRISDELLE, 2026

Abstract

A growing body of research suggests that volubility—how much a person speaks during a social interaction—is linked to leadership attainment, such that those who speak more are more likely to emerge as leaders. Some evidence also indicates that women speak less than men in group settings, raising the possibility that gender differences in volubility contribute to women’s lower representation in leadership roles. Across five studies ($N_{\text{total}} = 1,931$), I examined gender differences in volubility, informal leadership, and preferences for how voice is allocated within groups. Study 1 found that women spoke less than men and that this disparity mediated the relationship between gender and leadership attainment. Studies 2a-2b showed that women were more likely to prefer egalitarian voice allocation and viewed ability-aligned arrangements less favorably than men. Studies 3a-3b found that a hybrid speaking structure balancing equality and performance received the strongest support across gender.

Keywords: Volubility, Speaking time, Gender differences, Group dynamics, Leadership, Egalitarianism, Meritocracy, Rank psychology

Acknowledgements

To my supervisor, Dr. Joey Cheng, I extend my sincere gratitude for your support, mentorship, and the knowledge you have imparted throughout my Master's degree. Your commitment to excellence has inspired me throughout my graduate training, and I will carry it with me beyond my time at York University. Thank you for your investment in my growth as a student and researcher. It has been a privilege to build this project under your guidance.

Dr. Jennifer Steele, Dr. C. Ward Struthers, and Dr. Alexander Coutts, thank you for serving on my committee. I appreciate you making time to attend my defense and engage with my work.

Thank you to the *Leadership, Competition, and Collaboration Lab* for their continued support and for collecting and sharing part of the data used in this research.

To my former undergraduate supervisor and beloved friend, Dr. Maryanne Fisher, thank you for sparking my interest in psychology and research. You have played a meaningful role in shaping my trajectory as both a student and a person, for which I am grateful.

Lastly, I would like to express my appreciation for my incredible parents, grandparents, and friends for their continued support and encouragement. To my dad, thank you for pushing me to try new things—and for so graciously helping me move around the country (more than once) in pursuit of them. I love you more than anything!

Table of Contents

Abstract	ii
Acknowledgements	iii
Table of Contents	iv
List of Tables	viii
List of Figures	ix
Who Gets the Floor? Gender Differences in Speaking, Influence, and Egalitarian Preferences in Groups.....	1
The Current Research	5
Study 1	9
Method	10
Participants.....	10
Procedure	10
Measures	12
Analytic Approach	13
Results.....	14
Do Systematic Gender Differences in Volubility Emerge During Group Interactions, Such That Men Speak More Than Women (RQ1)?	14
Does the Gender Difference in Volubility Help Explain Women’s Lower Likelihood of Attaining Informal Leadership Within Groups (RQ2)?	16
Robustness Checks.....	17

Discussion	18
Study 2	19
Study 2a	20
Method	20
Participants.....	20
Procedure	21
Measures	21
Analytical Approach	25
Results.....	26
Do Women Hold Stronger Preferences for Egalitarian Distributions of Decision-Making Power, Whereas Men Prefer Ability-Aligned Structures in Which More Capable Individuals Receive Greater Voice (RQ3)?.....	26
Exploratory Analysis: Examining Speaking Arrangement Choice by Gender and Self- Perceived Performance	27
Exploratory Analysis: The Association Between Self-Reported Egalitarian Attitudes and Choice of Preferred Decision-Making Arrangement	30
Discussion	30
Study 2b	32
Method	33
Participants.....	33
Measures	33
Analytical Approach	34

Results	35
Do Women Hold Stronger Preferences for Egalitarian Distributions of Decision-Making Power, Whereas Men Prefer Ability-Aligned Structures in Which More Capable Individuals Receive Greater Voice (RQ3)?	35
What Psychological Factors Contribute to Gender Differences in Support for Ability- Aligned Speaking Arrangements (RQ4)?	36
Exploratory Analysis: Examining Speaking Arrangement Choice by Gender and Self- Perceived Performance	40
Discussion	41
Study 3	42
Study 3a	43
Method	43
Participants	43
Procedure	44
Measures	44
Analytical Approach	45
Results	46
Is There a Speaking Rule Group Structure That Can Bridge the Preference Divide Between Genders (RQ5)?	46
Exploratory Analysis: Examining Speaking Arrangement Preference by Gender and Self- Perceived Performance	48
Discussion	51

Study 3b	53
Method	54
Participants.....	54
Measures	55
Analytical Approach	55
Results.....	56
Is There a Speaking Rule Group Structure That Can Bridge the Preference Divide Between Genders (RQ5)?	56
What Psychological Factors Contribute to Gender Differences in Support for Ability- Aligned Speaking Arrangements (RQ4)?	58
Exploratory Analysis: Examining Speaking Arrangement Preferences by Gender and Self- Perceived Performance	63
Discussion.....	64
General Discussion	66
Theoretical Contributions	69
Practical Implications.....	71
Limitations and Future Directions	71
Conclusion	73
References.....	75

List of Tables

Table 1. Guiding Research Questions	9
Table 2. Studies 2a and 2b: Speaking Rule Options for a Hypothetical Group Problem-Solving Activity	23
Table 3. Studies 3a and 3b: Revised Speaking Arrangement Options for a Hypothetical Group Problem-Solving Activity	45

List of Figures

Figure 1. Study 1: Men Are More Voluble Than Women	15
Figure 2. Study 1: Volubility Mediates the Effect of Gender on Leadership Attainment	17
Figure 3. Study 2a: More Women Prefer Egalitarian Speaking Arrangements Than Men	27
Figure 4. Study 2a: Gender Differences in Speaking Arrangement Style Preferences Across Self-Perceived Performance	29
Figure 5. Study 2b: Overall, Both Men and Women Prefer the Ability-Aligned Speaking Arrangement	36
Figure 6. Study 2b: Gender Differences in Subjective Evaluations of the Merits of Each Speaking Arrangement.....	39
Figure 7. Study 2b: Ability-Aligned Speaking Arrangement Most Preferred Across Gender and Self-Perceived Performance.....	41
Figure 8. Study 3a: Hybrid Speaking Arrangement Most Preferred Overall but Gender Differences in Second Most Preferred Arrangement.....	48
Figure 9. Study 3a: Gender Differences in Speaking Arrangement Style Preferences Across Self-Perceived Performance	50
Figure 10. Study 3b: Hybrid Speaking Arrangement Most Preferred Across Gender, No Gender Difference in Second Most Preferred Arrangement	58
Figure 11. Study 3b: Gender Differences in Subjective Evaluations of the Merits of Each Speaking Arrangement.....	62
Figure 12. Study 3b: Hybrid Speaking Arrangement Most Preferred Across Gender and Self-Perceived Performance	64

Who Gets the Floor? Gender Differences in Speaking, Influence, and Egalitarian Preferences in Groups

Across a wide range of group settings, men consistently speak more than women, raising the question of whether a systematic gender difference in volubility—defined as how much individuals speak during group interaction—reliably emerges. This pattern has been observed in academic and professional contexts, where men are approximately two and a half times more likely to ask questions during seminars (Carter et al., 2018) and pose about 75% more questions during conference question-and-answer sessions (Jarvis et al., 2022). Similar gender differences extend beyond modern organizational environments: in contemporary hunter-gatherer societies, men participate more than women during community meetings and rank up to 67% higher in verbal influence (von Ruden et al., 2018). Together, these findings suggest that gender differences in speaking behavior are robust and appear across diverse social and cultural contexts.

Despite this evidence, much of the existing research comes from naturally occurring field settings that are difficult to control and may reflect contextual constraints rather than intrinsic interaction patterns. In many real-world environments, group members do not have equal opportunities to speak, and concerns about social evaluation or backlash may shape who chooses to participate (Williams & Tiedens, 2016). Indeed, women who speak more are viewed less favorably than their less voluble women peers and *equally* voluble men counterparts (Brescoll, 2011). Critically, the lack of anonymity in prior research makes it difficult to determine whether observed gender differences in volubility reflect genuine preferences, subconscious conformity to interaction norms, or strategic self-restraint driven by fear of social punishment for speaking

up. Little is also understood about the reasons why women speak less, including potential gender differences in how men and women may choose to allocate voice within groups.

To date, relatively little research has examined gender differences in volubility in controlled, laboratory-like group contexts that both afford all members equal opportunity to contribute and conceal social identities. By minimizing reputational risk and social evaluation, such contexts allow for a clearer test of whether gender differences in speaking persist even when external constraints and anticipated backlash are mitigated. Building on prior work, the present research has two broad goals. First, I examine gender differences in volubility under controlled group conditions that allow equal opportunities to speak and minimize reputational concerns. Moreover, beyond documenting the existence of this gap, the current research advances the literature by systematically investigating novel questions regarding the consequences of the volubility gap for influence and leadership outcomes. Second, I begin to explore the psychological factors that contribute to gender differences in speaking behavior. Specifically, I investigate gender-based preferences for how voice and influence ought to be distributed within groups, as well as the design of group structures that may be favored by both genders by balancing egalitarian and performance-based principles.

From Speaking to Leading: Consequences of Gender Differences in Volubility

One potentially critical consequence of the gender gap in volubility is its impact on influence and informal leadership within groups. Individuals who speak more tend to exert greater social influence and are more likely to be perceived as leaders (Mullen et al., 1989). Indeed, conversational features such as speaking time and interruption frequency can predict leadership perceptions with remarkable accuracy (Sanchez-Cortes et al., 2010). This association likely reflects a tendency to interpret frequent speaking as a signal of competence and

confidence, even when actual task ability is held constant (Anderson et al., 2012). Consistent with this interpretation, volubility has been shown to outweigh objective performance when people form judgments about who is fit to lead (Nikoleizig et al., 2021). Although prior research documents both a gender gap in volubility and a gender gap in leadership emergence, these two patterns have largely been examined in isolation (cf. Loignon et al., 2025). As a result, the extent to which gender differences in speaking directly contribute to gender disparities in informal leadership remains in need of systematic empirical testing. In the current work, I posit that gender differences in volubility may help explain why women are less likely than men to emerge as informal leaders in group settings (Badura et al., 2018; Eagly & Karau, 1991).

Explaining the Volubility Gender Gap: Psychological and Normative Factors

If gender differences in volubility are consequential for influence and leadership attainment, an equally essential question concerns what gives rise to this speaking disparity in the first place. Women's lower volubility may not reflect lower competence or leadership ambition (Netchaeva et al., 2022), but instead differences in how men and women believe decision-making power should be distributed within groups. One possibility is that women are more likely to favor egalitarian interaction norms that emphasize equal participation and shared voice, whereas men are more likely to favor ability-aligned norms in which more capable individuals are granted greater opportunity to contribute. Preliminary evidence for this divergence in preferences appears early in development, with girls as young as three years old showing a stronger tendency than boys to allocate resources equally (Benenson et al., 2019). In adulthood, women are more supportive of income redistribution (Alesina & La Ferrara, 2005) and are more willing to share earnings with lower-performing competitors (Cassar & Rigdon, 2021a)—a pattern that is particularly pronounced among high-performing women (Cassar &

Rigdon, 2021b). These findings suggest that women may view speaking time as a shared group resource and, as a result, intentionally limit their own participation to create space for others and promote collective decision-making.

Closing the Gap: Designing Group Interaction Rules that Work for Everyone

If gender differences in volubility arise from competing fairness ideals rather than fixed differences in ability or motivation, then modifying group structures to accommodate these differing preferences may help reduce gender gaps in participation and influence. Rather than privileging one ideal over the other, a group structure that ensures equal opportunity to contribute while also allowing expertise to shape decisions may be more broadly appealing and more effective. This idea aligns with a growing body of research showing that structural features of work environments can systematically deter women's participation and advancement. In particular, prior work on labor markets demonstrates that women are disproportionately less likely than men to enter highly competitive work settings, even when they are equally qualified (Niederle & Vesterlund, 2007). Field experiments on job-entry decisions show that competitive compensation regimes—such as those that emphasize individual competition for pay—can discourage women from applying, whereas features such as teamwork, reduced uncertainty, and less overt competition attenuate these gender differences (Flory et al., 2015). Importantly, these effects do not appear to reflect lower ability or ambition among women, but rather differences in preferences regarding competition, uncertainty, and the social structure of work environments.

By analogy, group interaction structures that emphasize ability-aligned speaking—where a small number of high performers dominate discussion—may similarly deter women who hold stronger egalitarian preferences, not only from speaking but also from wanting to participate fully or emerge as leaders. In contrast, arrangements that signal inclusion, shared influence, and

fairness may make leadership roles more attractive and attainable. Thus, identifying group speaking arrangements that balance egalitarian and ability-aligned principles may be critical for creating environments in which both men and women feel motivated to contribute and lead. Accordingly, the present research explores speaking arrangements that reconcile these competing fairness ideals in ways that appeal broadly across individuals, rather than catering to the preferences of any single gender, and that promote participation, influence, and leadership emergence for all group members.

The Current Research

The present research consists of a series of studies designed to address several interconnected goals. First, I sought to document whether a gender disparity in volubility reliably emerged in group interactions. Second, I examined the consequences of this disparity by testing whether speaking less was associated with lower informal leadership emergence. Third, I explored precursors of the disparity by examining factors that may underlie women's tendency to speak less and to allow others to take the floor—women may prefer more egalitarian distributions of voice and decision-making power, whereas men may prefer more ability-aligned arrangements in which more capable individuals speak more. Fourth, I further probed the psychological basis for women's greater apparent preference for egalitarianism, by comparing men and women's subjective evaluation of the different speaking arrangements across several domains, including perceived fairness, anticipated performance, anticipated sense of belongingness, and leadership aspiration. Fifth, I began to explore ideas for potential structural interventions—specifically, alternative speaking arrangements that could be implemented in real groups—that may be favored by most individuals and across genders, and thus, allow both men and women to thrive while reducing barriers to women's leadership emergence. Together, these

studies move from establishing the phenomenon of gender gap in volubility, to identifying its consequences, to tentatively explaining why it occurs, and ultimately to testing promising solutions. These broader goals guide the development of five general research questions that I address across studies (for an overview of my research questions, see Table 1).

Study 1 examined whether gender differences in volubility emerge during collaborative group problem-solving and whether these differences contribute to leadership emergence in virtual settings. Previously unacquainted individuals were placed into small groups and collaborated in real time to solve a NASA decision-making task through text-based group communication. This task is widely used across the social sciences to study teamwork, influence, and collaboration because it requires groups to share information, negotiate priorities, and reach consensus (Bottger, 1984; Hall & Watson, 1970). Individuals completed the task independently, before the group interaction, to capture their objective influence on subsequent group decisions. The use of an anonymous text-based format allowed “speaking” behavior to be quantified precisely while minimizing visual and status cues that might otherwise shape interaction. Chat transcripts were analyzed to assess gender differences in speaking frequency (Research Question 1), and to test whether individuals who speak less are in turn less likely to wield influence and emerge as informal leaders (Research Question 2).

Studies 2a and 2b shifted focus from consequences to potential precursors of women’s lower volubility by examining gender differences in preferred decision-making structures (Research Question 3). Specifically, I surmised that women’s lower volubility could reflect a preference for more egalitarian approaches to group decision-making, whereas men may prefer structures in which speaking opportunities are allocated in proportion to one’s demonstrated ability. To examine this, participants were presented with a hypothetical group task and asked to

choose between two discussion formats: one in which speaking time was distributed equally among all group members regardless of differences in ability, and another in which individuals who were more skilled were given greater opportunity to contribute. If women's lower volubility reflects a desire to promote shared influence and inclusion, they should be more likely than men to prefer the egalitarian speaking rule, whereas men should be more likely to prefer the ability-aligned rule.

In Study 2b (and later in Study 3b), I began to probe the psychological basis that underlies women's greater support for egalitarian speaking arrangements and men's greater support for ability-aligned speaking arrangements (Research Question 4). I tested how men and women subjectively evaluate different group discussion structures across several theoretically relevant dimensions (i.e., their expressed preference), to complement my assessment of their preferred speaking rule choice (i.e., their stated preference) in other studies. Specifically, participants evaluated each speaking rule in terms of its perceived fairness, anticipated impact on team performance, expected effects on teammates' sense of belongingness and inclusion, and their own desire to assume a leadership role under each group structure. This approach allowed me to examine whether men's stronger support for ability-aligned arrangements may stem from perceiving these structures as more fair, more effective, or more personally advantageous or enjoyable for leadership, whereas women may perceive fewer interpersonal or motivational benefits under such an arrangement. Together, these evaluations help clarify the psychological factors that may contribute to gender differences in preferences for group interaction norms.

Finally, in Studies 3a and 3b, I extended this work to a more practical domain, in which I sought to design a group structure that accommodated *both* egalitarian and ability-aligned preferences, thereby appealing to all genders and, by extension, reducing barriers to women's

leadership. Extending the design of Studies 2a and 2b, which offered only two speaking choice options, I assessed participants' preference for three speaking rules: a fully egalitarian rule, a fully ability-aligned rule, and a newly introduced hybrid rule (Research Question 5). Under the hybrid rule, all group members could contribute equally during an initial round of discussion regardless of demonstrated skill, after which more skilled members were given additional opportunities to share their input. This structure preserves equal participation and inclusion early in the discussion, while still allowing expertise to subsequently define decision-making. By examining how men and women chose among these three options, these studies allowed me to identify speaking arrangements—and more generally, associated social norms—that do not cater to the preferences of any one gender but instead appeal broadly across genders.

The data and analysis scripts for all studies are accessible on the Open Science Framework (OSF; see <https://osf.io/kvews>).

Table 1*Guiding Research Questions Across Studies*

Research Question (RQ)	Study Method and Variables Measured	Study
RQ1. Do systematic gender differences in volubility emerge during group interactions, such that men speak more than women?	Volubility (i.e., speaking frequency) in ad hoc, real-time, task-based group interactions	Study 1
RQ2. Does this gender difference in volubility help explain women's lower likelihood of attaining informal leadership within groups?	Volubility and informal leadership outcomes in ad hoc, real-time, task-based group interactions	Study 1
RQ3. What underlies women's lower levels of speaking in group settings? Specifically, do women hold stronger preferences for egalitarian distributions of decision-making power, whereas men prefer ability-aligned structures in which more capable individuals receive greater voice?	Speaking rule preference: egalitarian vs. ability-aligned (revealed preference)	Study 2a Study 2b
RQ4. What psychological factors contribute to gender differences in support for ability-aligned speaking arrangements?	Subjective evaluation of each speaking rule in: perceived fairness, anticipated team performance, anticipated teammate sense of belongingness, and desire to lead (stated preference)	Study 2b Study 3b
RQ5. Is there a speaking rule group structure that can bridge the preference divide between genders—such as one that simultaneously upholds egalitarian and ability-aligned principles—and thereby presumably minimize gender differences in participation and influence?	Speaking rule preference: egalitarian vs. ability-aligned vs. hybrid (revealed preference)	Study 3a Study 3b

Study 1

Study 1 examined whether gender differences in volubility emerge even in anonymous group settings in which participants' identities were concealed, thereby reducing pressures associated with gender norms and social backlash. Participants engaged in a real-time collaborative decision-making task through anonymous text-based discussion, allowing speaking

behavior to be quantified while minimizing visible identity cues. Beyond testing whether men speak more than women in this context, this study also examined whether volubility mediated the association between gender and informal leadership attainment within groups.

Method

Participants

In total, 331 participants (198 women; $M_{\text{age}} = 33.8$ years, $SD = 12.3$ years) recruited from Prolific partook in an online live group interaction study for approximately 30 minutes.¹ Participants were fluent in English, residents of Canada or the United States (U.S.), and at least 18 years old. They were informed that they would interact with three others via a text-based group chat in real-time to solve a puzzle. To make the task incentive compatible, teams competed to earn a performance-dependent monetary bonus prize of £8, in addition to the £3 each participant received as compensation for their individual participation.

The majority of participants (79.2%) identified as White/Caucasian or Asian ($n = 181$ and $n = 81$, respectively), 6.6% ($n = 22$) were Latin or South American, 6% Black ($n = 20$), 1.5% multiracial ($n = 5$), and 6.6% ($n = 22$) selected the option “other.” All study procedures were approved by the Human Participants Review Sub-Committee, York University’s Ethics Review Board.

Procedure

Participants provided informed consent and responded to a brief set of demographic questions. Next, they completed the Lost on the Moon task individually, where they ranked 15 items for their survival in a hypothetical moon crash-landing scenario from most to least

¹ Twenty-six individuals were excluded from analyses based on not completing the study in its entirety or failure to be matched with three other participants to form a group of four. An additional five participants who identified as neither a man nor woman were excluded from the analyses, given the explicit interest in gender differences for this research.

important (Hall & Watson, 1970). Following this, individuals were notified that they would be matched with three others in real time to complete the same survival task as a team. Before starting the group interaction, participants reported their goals and expectations for leadership in the upcoming collaboration portion of the study.

The majority of groups were mixed gender. Eleven teams (13% of all teams) were comprised entirely of men or women. Deliberation took place virtually via an online group chat. Teams had 12 minutes to reach a consensus before submitting their final answers. All members submitted their group's rankings separately.² Once finished, participants responded to questions probing their experience and perceptions of their teammates.

The Lost on the Moon exercise is commonly used to assess team problem-solving in a contextually novel problem space where deciphering correct ideas from wrong ones is difficult (Cheng et al., 2013; Yetton & Bolger, 1982). These circumstances incentivize communication, deliberation, and idea sharing, which stimulate discussions and enable observation about gendered patterns in discussion participation and subsequent leadership emergence.

Studying discussion dynamics where identities are concealed can be beneficial for multiple reasons. Participants are unable to use physical cues to infer competence and confer social influence accordingly. Instead, they must form their judgements based solely on the content and rate of others' contributions during discussion. These conditions also offer a way to examine speaking behavior in a somewhat "socially sanitized" environment by eliminating social pressures to act in stereotypically conforming ways (i.e., women may speak more freely,

² Sixteen groups did not reach a consensus—these teams were excluded from models designed to assess leadership, as no clear leader could be identified ($n = 63$ participants). Consensus within a team was operationalized as occurrences where at least two members submitted identical rankings for all 15 items. In cases where two or fewer members submitted rankings that deviated from the consensus, those responses were corrected to align with the rest of the group.

uninhibited by fear of presenting as too dominant or overbearing). This allows for an explicit test of true speaking preference.

Measures

Volubility

Volubility was measured as the total number of words each individual typed into the group text-based discussion ($M = 103.82$, $SD = 68.51$). The number of phrases ($M = 18.65$, $SD = 10.56$) and characters ($M = 516.46$, $SD = 342.18$) typed were collected as alternate measures of volubility for robustness check purposes.

Objective Informal Leadership

Informal leadership was operationalized as the objective influence each person actually exercised over their team's final ranking for all 15 survival items. The degree to which each group member shaped final decisions was calculated as the sum of the absolute value of the deviation between their personal ranking at the beginning of the study and their team's ranking for each item, multiplied by minus one, such that smaller negative values (closer to zero) reflect higher informal objective leadership ($M = -46.02$, $SD = 15.79$).

Objective Task Performance

Task performance was measured as the correctness of participants' individual rankings based on the official NASA Lost on the Moon task answer key. Similar to objective influence, objective task performance was operationalized as the sum of the absolute difference between one's ranking and the correct NASA ranking across all 15 survival items, multiplied by minus one such that scores closer to zero indicate greater task performance (i.e., greater similarity between an individual's ranking and the answer key; $M = -51.65$, $SD = 13.94$).

Desire to Lead

Participants indicated their desire to influence their team's final decision on a Likert scale ranging from 1 ("no desire to lead at all") to 7 ("very much desire to lead"; $M = 4.1$, $SD = 1.5$).

Analytic Approach

My primary analysis is a mediation model that examines the effect of gender (predictor) on leadership emergence (outcome) through volubility (mediator). All regressions were modelled using cluster-robust standard errors (CRSEs), clustering by group, as a practical solution to the interdependence of observations and non-constant error variance that typically characterizes hierarchical data (Cameron & Miller, 2015; McNeish et al., 2017).³ Unless otherwise specified, all reported coefficients (β) are standardized beta coefficients. Confidence intervals for the associated parameter estimates were derived using the nonparametric bootstrapping method, which resampled groups 1,000 times with replacement.

The data were uploaded to R (version 4.4.0) using the *readxl* package (Wickholm, 2019a) and cleaned with *dplyr* (Wickholm, 2019b; Wickholm et al., 2019c). The *psych* package facilitated the calculation of means and standard deviations (Revelle, 2025). Levene's test statistics were obtained using the *car* package (Fox & Weisberg, 2019). Mediation models with cluster-robust standard errors were built using the *mediation* (Tingley et al., 2014), *lmtest* (Zeileis & Hothorn, 2002) and *sandwich* packages (Zeileis et al., 2020). A and b path coefficients with robust standard errors were modelled using the *estimatr* package (Blair et al.,

³ Independence of observations and homoscedasticity are two key assumptions in ordinary least squares (OLS) estimation. Failure to account for violations of these assumptions can produce biased estimates of the standard error and invalidate associated confidence interval estimates, given that confidence intervals are derived from the standard error. The cluster-robust method of standard error estimation is a common approach when dealing with the inconveniences of correlated data and non-constant error variance when fixed effects, but not random effects, are of interest (as is the case in the present research; McNeish et al., 2017). This method reduces the type 1 error rate by adjusting standard error estimates, often making them more conservative, to account for within-group variability when examining between-group differences.

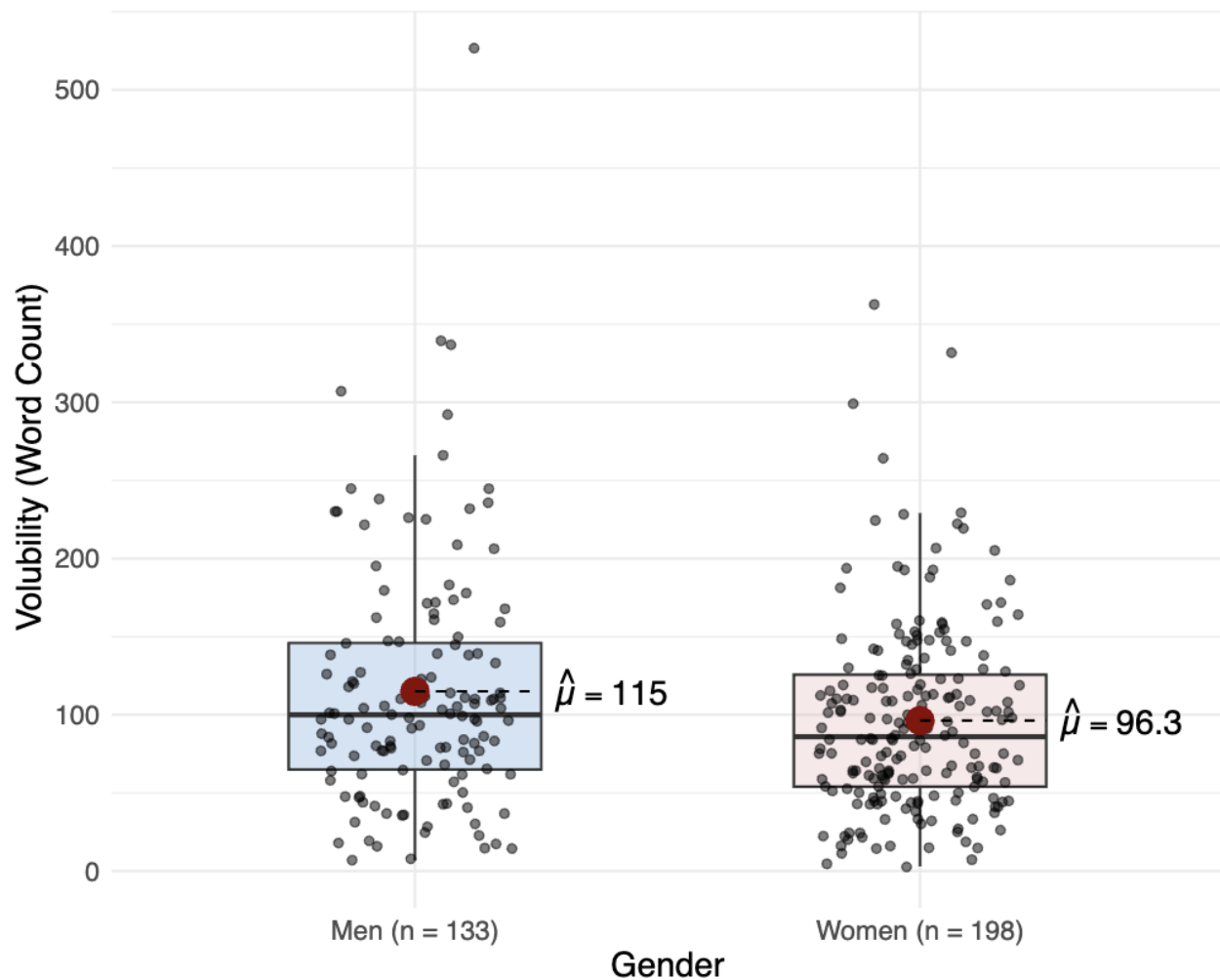
2025). Cohen's d s were calculated using *effectsize* (Ben-Shachar et al., 2020) and *ggplot2* was used to create the boxplot (Wickham, 2016).

Results

Do Systematic Gender Differences in Volubility Emerge During Group Interactions, Such That Men Speak More Than Women (RQ1)?

As I hypothesized and consistent with previous research, men were more voluble (operationalized as typing more words) during the group discussion, $t(232.28) = 2.33, p = .02, d = 0.27, 95\% \text{ CI } [0.05, 0.50]$ (see Figure 1).⁴ Importantly, however, this gender disparity in volubility cannot be attributed to differences in leadership aspiration or objective task performance. Men and women expressed similar interest in influencing their team's decisions during the collaboration portion of the study, $t(329) = 1.32, p = .19, d = .15, 95\% \text{ CI } [-0.07, 0.37]$ (see Figure 2), and did not differ in their objective performance on the survival task, $t(329) = 1.45, p = .15, d = 0.16, 95\% \text{ CI } [-0.06, 0.38]$.

⁴ The same qualitative results emerged when dropping four outliers who were more than three standard deviations above the mean volubility from analyses. Among the trimmed subset ($N_{\text{total}} = 264$), men typed more words than women, $t(195.42) = 2.36, p = 0.02, d = 0.31, 95\% \text{ CI } [0.06, 0.56]$.

Figure 1*Study 1: Men Are More Voluble Than Women*

Note. The red dots represent the mean number of words typed by men and women in the virtual group chat, respectively. Mean values are listed to the immediate right of each box. The enlarged horizontal line immediately below the red dot represents the median, with whiskers (i.e., vertical lines spanning outward of each box) signifying the 95% confidence interval. The same qualitative results emerge when excluding outliers ($n = 4$) who were more than three standard deviations above the mean volubility from analysis—men were more voluble than women.

Does the Gender Difference in Volubility Help Explain Women's Lower Likelihood of Attaining Informal Leadership Within Groups (RQ2)?

A mediation model tested whether volubility (mediator) causally influenced the effect of gender (independent variable) on leadership attainment (outcome variable). One thousand bootstrapped samples were drawn with replacement to estimate the 95% confidence intervals for the model effects.⁵

Volubility (total word count) mediated the effect of gender on leadership attainment, and this effect was statistically significant (indirect effect = -0.07, bootstrapped 95% CI [-0.15, -0.01], $p = .02$). Men spoke more than women, and their higher volubility led to more leadership (see Figure 2).⁶

Notably, the total effect of gender on leadership did not reach statistical significance ($\beta = -0.22$, 95% CI [-0.15, -0.01], $p = .07$). Nor did the direct effect ($\beta = -0.15$, 95% CI [-0.39, 0.08], $p = .20$). Although Baron and Kenny's (1986) popular approach to mediation stipulates a statistically significant direct effect and total effect should be present (in addition to non-zero a path and b path coefficients), modern approaches to mediation have demonstrated ways in which statistically significant indirect effects may present in the absence of these effects (Hayes, 2009; Kenny & Judd, 2014; Shrout & Bolger, 2002; O'Rourke & MacKinnon, 2018).⁷ Therefore, the

⁵ The confidence interval for the indirect effect can be derived by treating the sample as a population. This is achieved by resampling the sample (in this case, 1,000 times) with replacement to generate the sampling distribution of the indirect effect from which a 95% confidence interval is calculated (Bollen & Stine, 2009; Hayes, 2009). This method, otherwise known as bootstrapping, is particularly useful because it is not founded on the assumption that the sampling distribution of the indirect effect is normal (unlike the parametric Sobel test). Although the a path ($X \rightarrow M$) and b path ($M \rightarrow Y$) are typically normally distributed, their product (the indirect effect) is often not (MacKinnon et al., 2004).

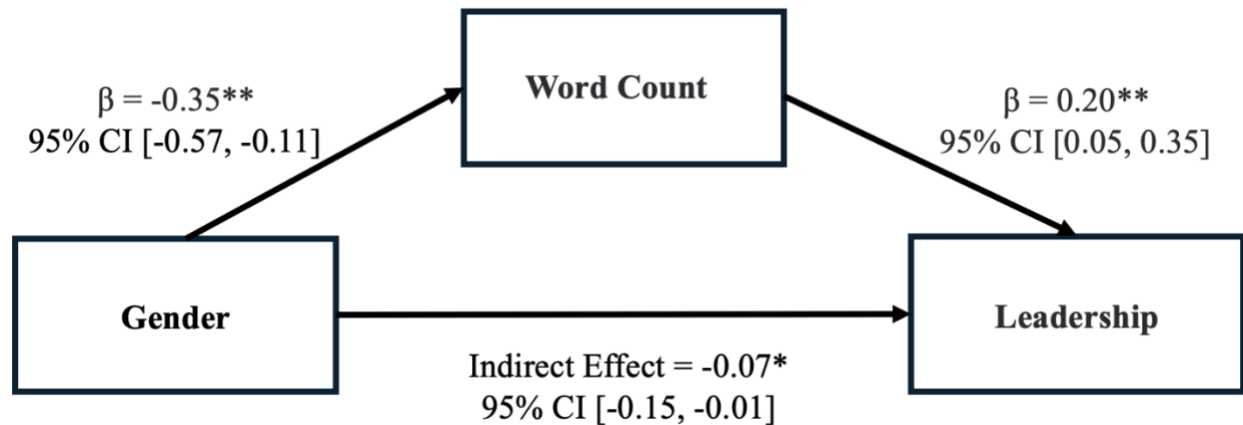
⁶ The same qualitative results were obtained when I removed the four participants whose word counts were above three standard deviations from the mean ($N_{\text{total}} = 264$). Volubility continued to significantly mediate the effect of gender on leadership attainment (indirect effect = -0.07, bootstrapped 95% CI [-0.160, -0.002], $p = 0.04$).

⁷ In the present case, despite the large sample size, the power to detect an indirect effect is much larger than the power when testing the total effect. This is likely why the total effect did not reach statistical significance (O'Rourke & MacKinnon, 2018).

identified causal influence of gender on leadership through volubility should be considered meaningful.

Figure 2

Study 1: Volubility Mediates the Effect of Gender on Leadership Attainment



Note. Mediation model showing the hypothesized mediating role of volubility in explaining the association between gender (coded as “0” for men and “1” for women) and leadership emergence ($N = 268$ in 67 groups). All model coefficients are standardized. Word count represents the total number of words an individual typed during the group discussion. The indirect effect was calculated using 1,000 bootstrapped samples. Significance levels $p < .01$ and $p < .05$ are denoted by “**” and “*”, respectively.

Robustness Checks

I ran two additional mediation models to ascertain the robustness of the mediating effect of volubility found above when volubility was operationalized as word count. Both sets of results—based on phrase count and character count—produced the same qualitative results. A significant, non-zero mediating effect was found for both phrase count (indirect effect = -0.05,

bootstrapped 95% CI [-0.12, 0.00], $p < .05$) and character count (indirect effect = -0.07, bootstrapped 95% CI [-0.15, -0.01], $p = .01$), when they were entered as mediators of the association between gender and leadership emergence.

Discussion

Study 1 tested whether gender differences in volubility persist in an environment in which people's identities were concealed. Despite prior evidence that men are more voluble than women (e.g., Carter et al., 2018; Jarvis et al., 2022; von Ruden et al., 2018), prior research has been limited in its ability to disentangle speaking behavior from pressures to conform to gender norms, which are more salient in in-person group situations where identities are more readily observable. The present study addressed this limitation by examining participation in an anonymous group context and found that women continued to speak less than men in this setting. Moreover, I found that not only did men systematically speak more, but this behavioral difference predicted an advantage in leadership attainment compared to women. These insights parallel research exploring the effect of volubility on social influence in face-to-face situations, which link higher volubility to increased leadership (Anderson et al., 2012; Loignon et al., 2025; MacLaren et al., 2020; Sanchez-Cortes et al., 2010; Sanchez Cortes et al., 2012). Importantly, these effects emerged despite lack of a gender difference in domain expertise or desire to lead one's team. Women's skill level was statistically indistinguishable from that of men, indicating objective task performance was not a prerequisite for leadership in these groups. Women also expressed equivalent aspirations to influence decisions as men; this finding is particularly relevant as it counters meta-analytic work suggesting women's lesser prominence as leaders could be tied to their lower desire to lead (Netchaeva et al., 2022). Given women's divulged

eagerness to lead, it is imperative to identify why their (speaking) behavior operated in ways that were non-conducive to leadership acquisition.

Arguably, the anonymous design element of the group interaction in this research reduced (if not eliminated entirely) the pressure to conform to gender roles as gender identities were concealed from other participants (Brescoll, 2011; Williams & Tiedens, 2016); this strengthens my confidence that women participated in the group discussions at the rate they desired, and at a lower degree of hindrance by social pressures characteristic of in-person collaboration. Yet, why women may deliberately speak less than men remains largely unanswered. Identifying the primary motivation for women's lower volubility, relative to men, is essential to advancing the current understanding, which largely attributes women's lower participation rate to gendered social norms. On a more applied note, understanding the underpinnings of women's lower volubility is essential to developing interventions targeted at reducing the gender gap in discussion participation in professional settings and, in turn, increasing women's entry to leadership. I pivot to this question next and in the following studies begin to explore the role of preferences in shaping men's and women's speaking behavior in an effort to understand why women often speak less than men in group situations.

Study 2

While a number of factors may be at play, one reason why women often speak less than men in group situations is women's greater preference for egalitarian social norms, leading them to regulate their own participation in ways that create greater space for others to contribute. I explored this possibility in the present research by testing whether women show greater preference for discussion structures that distribute speaking opportunities equally across group members, regardless of relative differences in ability or performance. This prediction is tested

across two studies—Study 2a and Study 2b—by surveying men’s and women’s endorsement of two distinct discussion arrangements: one that allocates speaking opportunity equally across members, and an ability-aligned format that affords highly skilled members more opportunity to speak.

In addition, I examined two questions for exploratory purposes. First, I examined whether gender differences in speaking-rule preference varied as a function of self-perceived performance or ability, given the possibility that high-performing men and women may differ especially strongly in whether they prefer to translate their ability into greater voice, whereas the gender gap may be minimal among lower performers. Second, I administered a battery of questionnaires designed to assess individual differences in self-reported egalitarian attitudes (e.g., social dominance orientation, support for gender equality) in order to test whether these beliefs may be associated with participants’ preferred discussion arrangement choice.

Study 2a

Method

Participants

A sample of 201 participants (50.7% women; $n = 102$), averaging 43 years of age ($SD = 12.2$) took part in the current study which was advertised as an inquiry concerning preferences about group work.⁸ Participants were recruited through Prolific. The ethnic breakdown of the sample was 65.7% ($n = 132$) White/Caucasian, 11.4% ($n = 23$) Black, 6% ($n = 12$) Chinese, 5% ($n = 10$) Hispanic/Latine, 5% ($n = 7$) South Asian, 2% ($n = 4$) mixed origin/multiracial, 1.5% ($n = 3$) Southeast Asian, 1%, ($n = 2$) Filipino, 1% ($n = 2$) West Asian, 0.5% ($n = 1$) Arab, and 2.5% ($n = 5$) selected the option “other.”

⁸ Three and five individuals incorrectly responded to the first and second attention checks, respectively. These responses ($n = 8$) were excluded from analyses.

Participants were informed they would be asked to envision hypothetical scenarios and respond to questions about their preferences for leadership under the imagined circumstances. £0.5 was provided as compensation. English fluency, Canadian or American residence, and an age of 18 years or older were mandatory prerequisites to participate; respondents were pre-screened through Prolific based on these criteria. The Human Participants Review Sub-Committee, York University's Ethics Review Board, formally approved all study methods before data collection commenced.

Procedure

Participants were first presented with instructions for the same Lost on the Moon task used in Study 1. They were asked to imagine working with three hypothetical group members to solve the task via a synchronous online text-based chat.⁹ Then, participants were asked to anticipate their performance in terms of relative rank within the group and indicate their desire and expectations for leadership in the hypothetical scenario. Next, they selected their preferred speaking arrangement in the context of the group interaction and concluded the survey by completing a battery of egalitarianism scales.

Measures

Speaking Arrangement Style Preference

Decision-making style preference was measured by surveying the endorsement of two speaking rules presented to participants in the form of a table (see Table 2).¹⁰ Hypothetical task performance was denoted by a column that listed performance scores for four group members

⁹ The hypothetical conditions of the group exercise were designed to mirror the method of Study 1. The imagined nature of the group task makes the true rank of the responder ambiguous and allows for a more accurate test of true decision-making preference, given that individuals do not know where they rank in their respective groups.

that ranged from 45% (least skilled) to 90% (most skilled).¹¹ The two right-most columns outlined different ways in which the group could collaborate to solve the survival task. In the first arrangement (ability-aligned rule), highly skilled members had more opportunities to contribute to the discussion (indexed by a higher allotted word count). In the second arrangement (egalitarian rule), word count was distributed equally across members, allowing for equal contribution to the group discussion regardless of relative performance score. Note that the word count permitted for both arrangement options across all group members always totaled 500 words. Only the distribution of words allocated across members varied with respect to performance score between arrangements.¹²

¹¹ Task performance varied in the scenario to ensure the egalitarian arrangement was costly and truly elicited egalitarian preference. Equalizing the number of words that high and low performers can contribute during the deliberation process suppresses the decision-making influence of more knowledgeable members which creates a trade-off: equalized opportunity to contribute at the expense of taking decision-making opportunity away from more skilled contributors.

¹² To control for order effects, one of two versions of the arrangement table were randomly displayed to participants. The content across both tables was identical. Only the order of the columns that detailed the speaking arrangement options varied. Randomly counterbalancing the order that the egalitarian and less egalitarian designs were presented helped control for response bias favoring the first listed choice (Miller et al., 1998).

Table 2*Studies 2a and 2b: Speaking Rule Options for a Hypothetical Group Problem-Solving Activity*

	Performance score (0-100%)	Words I want to allocate to each member	
		Option A	Option B
Member C	90%	250 words	125 words
Member B	75%	125 words	125 words
Member D	60%	75 words	125 words
Member A	45%	50 words	125 words

Note: Participants viewed one of two speaking arrangement options for a hypothetical group problem-solving task. The performance score column depicts each group member's hypothetical performance score on a survival task, wherein higher values reflect greater task performance. In Option A, speaking time was allocated based on performance scores. In Option B, speaking time was allocated equally, regardless of skill.

Self-Reported Egalitarian Preference

Social Dominance Orientation (SSDO). A shortened version of the SSDO (Pratto et al., 2013), which is comprised of four items, was adapted to assess attitudes towards social dominance (e.g., “in setting priorities, we must consider everyone in a group”).^{13 14} This scale, despite shortened, demonstrates strong validity across nations and cultures (Pratto et al., 2013). Greater endorsement of social dominance is associated with lower willingness to protect minorities, support the poor, and promote women leaders. The four-item composite demonstrated good internal consistency ($\alpha = .85$, average inter-item $r = .60$; $M = 5.5$, $SD = 1.4$).

¹³ The SSDO item verbiage was modified to better align with Study 2's focus on within-group dynamics. The SSDO targets group-level beliefs. For example, instead of “in setting priorities, we must consider all groups”, the item was rephrased to “in setting priorities, we must consider *everyone within a group*.” The original 7-point Likert scaling for SSDO items was preserved.

¹⁴ The SSDO uses items from an earlier iteration of the SDO—SDO₆. Each item maps onto one of the four dimensions included in the most recent version of the scale (SDO₇; Ho et al., 2015). These dimensions include: pro-trait dominance, con-trait dominance, pro-trait antiegalitarianism, and con-trait antiegalitarianism.

Societal Egalitarianism. A subset of five items from the American National Election Studies Survey (ANES; 2020) measured beliefs about equality on a 5-point Likert scale where 1 indicated “strongly agree” and 5 indicated “strongly disagree” (e.g., “if people were treated more equally in this country, we would have fewer problems”). The subsequent societal egalitarianism composite demonstrated a strong internal consistency ($\alpha = .92$) and average inter-item correlation ($r = .69$; $M = 4$, $SD = 1.1$).¹⁵

Economic Egalitarianism. Five items were adapted from prior work (Wiwad et al., 2019) to measure attitudes about economic inequality ($M = 5.3$, $SD = 1.7$).¹⁶ Each item was rated on a 7-point Likert scale where higher scores reflect reducing economic inequality and endorsing egalitarian economic policies (e.g., “we need to do everything possible to reduce economic inequality in the world today”). The five items were averaged to create an economic egalitarianism composite with high internal consistency ($\alpha = .95$; average inter-item $r = .80$).

Gender Equality. Gender equality was captured with three items originating from the Taylor Nelson Sofres Gallup poll (Jakobsson et al., 2010; e.g., “do you think that gender equality is important?”). Higher values reflect greater support for gender equality ($M = 7.6$, $SD = 1.8$). The composite showed relatively weak internal consistency ($\alpha = .60$; average inter-item $r = .34$).

¹⁵ The ANES matrix response options are scaled in the direction opposite of the SSDO (presented immediately before) and the Wiwad (2019) items (presented directly after) to promote attentive responding. Since the ANES and SSDO measure similar constructs at different levels (group vs. societal), responses should be positively correlated. Negative correlations at the participant level could reflect straight lining or inattentive responding. In addition, an instructed response item (IRI) was embedded in the ANES question matrix to assess attentiveness toward the end of the survey (Hauser et al., 2018; Gummer et al., 2021; Oppenheimer et al., 2009; Shamon & Berning, 2020). Five individuals failed to respond correctly—their data were excluded from all analyses.

¹⁶ In addition, two items from the International Social Survey Program (ISSP; 2018) were administered to assess economic egalitarianism. Given the two-item composite demonstrated lower internal consistency ($\alpha = 0.61$; average inter-item $r = .47$) than the primary measure adapted from Wiwad and colleagues (2019), it was dropped from subsequent analyses.

Self-Perceived Performance

Participants ranked their self-perceived performance on the Lost on the Moon survival task relative to three random others. Respondents selected a number from 1 to 4 to indicate their rank, where response options ranged from “1—I would be the top performer” to “4—I would be the bottom performer” ($M = 2.8$, $SD = 0.7$). In the analyses reported below, individuals who placed themselves in rank 1 or 2 were considered ‘top performers’, whereas those who placed themselves in ranks 3 or 4 were considered ‘bottom performers.’

Analytical Approach

Speaking rule preference was first assessed descriptively by looking at the spread in endorsement percentage across genders and arrangement options. To formally analyze these data, I next used binomial logistic regression models to test whether gender predicted speaking arrangement choice.¹⁷ As in previous studies, R version 4.4.0 was used to analyze the data. The data were cleaned first in Excel, then loaded into R with the *readxl* package (Wickholm 2019a).¹⁸ Scale items were transformed into composites using *dplyr* (Wickholm 2019b, Wickholm et al., 2019c). The *psych* package facilitated the calculation of means, standard deviations, and the assessment of internal consistency for the battery of egalitarianism measures (Revelle, 2025). Cohen’s *d* was calculated using *effectsize* (Ben-Shachar et al., 2020). I used the *glm()* function from base R to conduct logistic regressions. Base R was also used to perform the chi-square analyses alongside the *lsr* package which calculated Cramér’s Vs (Navarro, 2015). To create the bar graphs that illustrate speaking arrangement preferences across gender and perceived

¹⁷ Generalized linear modelling was conducted due to the binary categorical dependent variable and predictor.

¹⁸ Eight inattentive responders who failed to correctly answer at least one (of two) IRIs were excluded from analyses.

performance, I used the packages *scales* (Wickham et al., 2025), *ggplot2* (Wickham, 2016), and *RColorBrewer* (Neuwirth, 2022).

Results

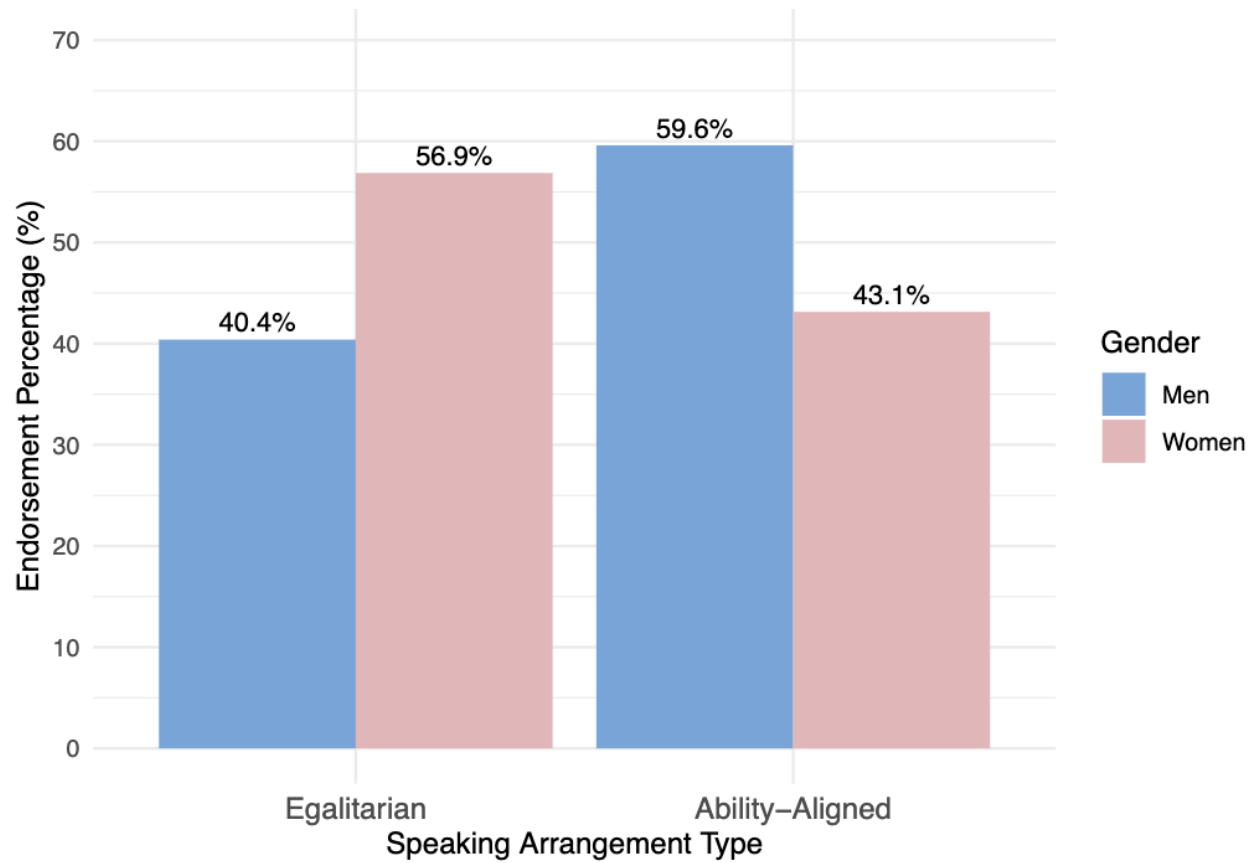
Do Women Hold Stronger Preferences for Egalitarian Distributions of Decision-Making Power, Whereas Men Prefer Ability-Aligned Structures in Which More Capable Individuals Receive Greater Voice (RQ3)?

The gender breakdown of preferred speaking arrangement choice is shown in Figure 3. As expected, descriptively, slightly more than half of the women in the sample—56.9% ($n = 58$)—preferred the egalitarian speaking arrangement, whereas 43.1% ($n = 44$) of women chose the ability-aligned arrangement option. Also as expected, men demonstrated the opposite pattern: nearly three fifths (59.6%; $n = 59$) opted for the discussion structure that prioritized top performers by giving them more voice during the group discussion, and 40.4% ($n = 40$) chose the egalitarian distribution of decision-making power.

To formally test for a gender difference in choice, I used a binomial logistic regression model to regress speaking arrangement (ability-aligned vs. egalitarian; the binary outcome) on gender (women coded ‘1’ and men coded ‘0’; the dummy-coded predictor). Consistent with the descriptive results, women had 94% higher odds of choosing the egalitarian speaking arrangement compared to men, $OR = 1.94$, $SE = 0.29$, 95% CI [1.11, 3.43], $p = .02$.

Figure 3

Study 2a: More Women Prefer Egalitarian Speaking Arrangements Than Men



Note. Bars represent the within-gender percentage of men ($n = 99$) and women ($n = 102$) in their choice of egalitarian or ability-aligned speaking arrangement.

Exploratory Analysis: Examining Speaking Arrangement Choice by Gender and Self-Perceived Performance

Next, I explored whether the observed gender difference in arrangement preference found above varies for those with high versus low self-perceived performance. Figure 4 illustrates the gender breakdown in speaking arrangement choice by self-perceived performance.

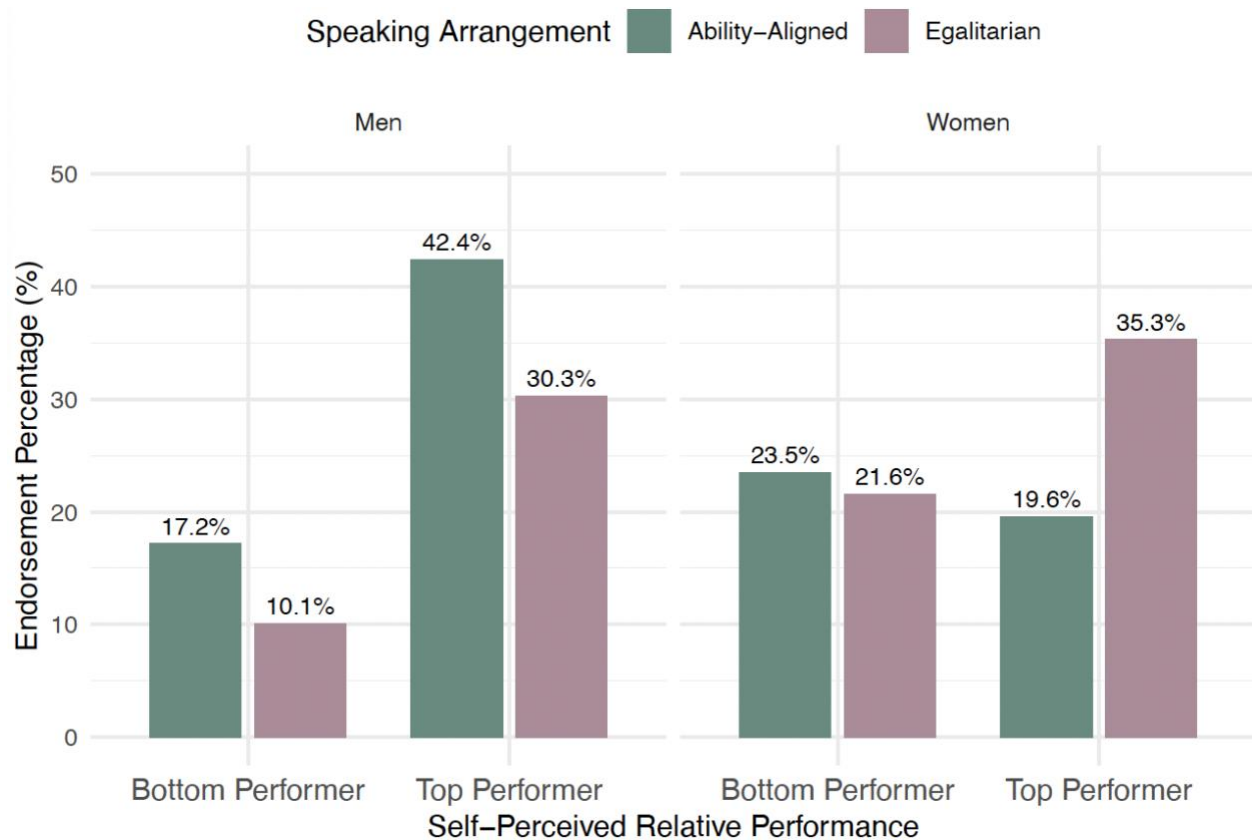
Among top performing women ($n = 56$), 63.5% ($n = 36$) preferred the egalitarian arrangement over the ability-aligned one (35.7%; $n = 20$). By contrast, top performing men ($n = 72$) demonstrated the opposite trend, as more than half (58.3%; $n = 42$) chose the ability-aligned design, whereas a minority of 41.7% ($n = 30$) chose the egalitarian option. Indeed, a chi-square test of independence confirmed that the association between gender and arrangement preference among this group of high performers was statistically significant, $\chi^2(1, N = 128) = 6.45, p = .01$, Cramér's $V = .21$.

Examining gender differences in choice among low performers, interestingly, the association between gender and arrangement choice did not extend to this group, $\chi^2(1, N = 73) = 0.08, p = .37$, Cramér's $V = .08$. Over half (52%; $n = 24$) of women bottom-performers ($n = 46$), and over half (63%; $n = 17$) of men bottom-performers ($n = 27$), chose the ability-aligned arrangement, relative to the 47.8% ($n = 22$) and 37% ($n = 10$) who chose the egalitarian arrangement, respectively.

These data suggest that gender differences in arrangement preference may be most pronounced among individuals who perceive themselves to be high performers. However, this conclusion should be interpreted cautiously. Relatively few men classified themselves as lower performers compared to women, $t(199) = 2.53, p = .01, d = 0.36, 95\% \text{ CI } [0.08, 0.64]$, resulting in substantially lower statistical power to detect gender differences within this subgroup. In fact, 72.7% ($n = 72$) of all men reported that they would rank as one of the top two members in the group. In contrast, women's relative performance projections were more equally split with 54.9% ($n = 56$) anticipating ranking among the top two performers. Consequently, the absence of a significant gender effect among lower performers should be viewed as tentative until replicated with larger samples.

Figure 4

Study 2a: Gender Differences in Speaking Arrangement Style Preferences Across Self-Perceived Performance



Note. The bar graph illustrates the proportion of men ($n = 99$) and women ($n = 102$) within each gender who preferred each type of speaking arrangement across levels of self-perceived performance in a hypothetical group of four participants. “Bottom Performer” refers to participants who estimated their own performance to fall in ranks 3 or 4, and “Top Performer” refers to participants who estimated their own performance as rank 1 or 2.

Exploratory Analysis: The Association Between Self-Reported Egalitarian Attitudes and Choice of Preferred Decision-Making Arrangement

To examine whether the association between choice of speaking arrangement—which may be considered a more behavioral expression of true preference that is less biased than self-reports prone to biases and impression management—is associated with self-reported egalitarian attitudes, I examined the point-biserial correlations between each egalitarianism measure and binary speaking arrangement choice (ability-aligned = 0, egalitarian = 1). None of the self-reported egalitarianism measures were significantly associated with speaking arrangement choice (ps ranged from .61 to 1.00).^{19 20}

Discussion

Study 2a explored whether men and women have different preferences about how voice should be allocated within groups during the decision-making process. The present work extends Study 1 in search of a possible explanation for women's lower volubility beyond what is often assumed in the literature—that women's speaking behavior is largely governed by a social pressure to conform to gendered norms (Brescoll, 2011)—as the fear of backlash effect is less compatible with Study 1 findings when participants' identities are concealed. As predicted, this study revealed that women preferred the conversation structure that distributed speaking opportunities equally within groups, despite variance in within-group task performance. By contrast, men preferred conversation structures that prioritized more knowledgeable members by allocating them more time to speak during the decision-making process. Interestingly, the

¹⁹ Economic egalitarianism was not associated with speaking rule choice, $r_{pb}(199) = .02, p = 1.00$. Social dominance orientation did not predict speaking arrangement preference, either, $r_{pb}(199) = .06, p = 1.00$. Social equality, $r_{pb}(199) = .00, p = 1.00$, and self-reported gender equality beliefs, $r_{pb}(199) = .04, p = .61$, likewise showed no statistically significant association with speaking arrangement choice.

²⁰ Independent-samples t -tests revealed that women scored significantly higher than men on attitudes toward gender equality, $t(184.67) = -2.25, p = .03, d = -0.32, 95\% \text{ CI } [-0.60, -0.04]$. There were no gender differences in attitudes toward economic egalitarianism, social dominance orientation, or societal egalitarianism.

gendered effect on arrangement preference appears to be performance-dependent. Notably, these gendered patterns are most pronounced among high-performers: whereas top performing women preferred to share the floor with others—despite the egalitarian speaking arrangement offering them less power to influence group decisions—top performing men instead preferred the ability-aligned speaking arrangement. There was no statistically significant association between gender and choice among self-perceived bottom performers.

From a bird's-eye view, Study 2a provides preliminary evidence that men and women feel differently about how decision-making power should be shared in small groups. Tying these insights to Study 1, the attitudinal divergence towards asymmetrical participation rate during group discussions in Study 2a mirrors the directionality of the gender difference in participation rates and consequent leadership disparity illustrated during the previous group interaction. Here, men are more welcoming of conversation dynamics that allocate more speaking time to some members over others. As revealed in Study 1, they are also more likely to be the members who dominate discussions. Importantly, women's preference to share the speaking floor coincides with their lesser likelihood to dominate conversations, and ultimately, influence group decisions. This raises the possibility that women may deliberately choose to be less voluble with the intention of leaving space for others to participate during the decision-making process. Consequently, their desire to share the conversation space with others may undermine their ability to shape the course of discussions and, ultimately, attain social influence.

Although the present findings reveal a gender difference in decision-making power allocation preference, given the moderate sample size, it remains uncertain whether these patterns represent both populations. Further research is needed to verify whether these results generalize to a larger sample—an issue I turn to in the next study.

Study 2b

Study 2b tested whether the observed gender difference in speaking arrangement preference from Study 2a would generalize in a sample nearly three times as large ($N = \sim 600$). The sampling technique, study procedure, and data analysis process were identical to that of Study 2a, with one exception.

In Study 2b, we measured *both* revealed preferences—participants' behavioral choice between two available speaking rules, as was done in Study 2a—and stated preferences, which address their subjective evaluations of each speaking rule in terms of fairness, anticipated performance, belongingness, and leadership aspiration. In economics literature, revealed preferences reflect what people actively choose, while stated preferences involve survey-based questions to capture people's attitudes or beliefs (Samuelson, 1948). Measuring both is important because stated judgments may not always align with revealed choices, and exploring both allows me to examine whether gender differences in revealed choices are consistent (or discrepant) with their subjective stated evaluations. It is possible, for example, that I may find a mismatch across the types of preferences, such as a consistent gender difference in stated preference, but not in revealed preferences. Assessing both types of preferences therefore allows for a more comprehensive look into not only the decision-making power structure preferred by men and women (Research Question 3), but also the psychological factors that underpin men's and women's presumed differential preferences for group interaction norms (Research Question 4).

Method

Participants

In total, 592 participants (297 women; $M_{\text{age}} = 37.5$ years, $SD = 12.8$ years) from Prolific were invited to take part in this replication study.²¹ Over half (52.5%; $n = 311$) of the sample identified as White/Caucasian. The racial breakdown of the other half of the sample is as follows: 6.1% Chinese ($n = 36$), 20.1% Black ($n = 119$), 5.9% South Asian ($n = 35$), 2.9% Hispanic/Latine ($n = 17$), 1.9% Southeast Asian ($n = 15$), 1.5% Filipino ($n = 9$), 2% mixed origin/multiracial ($n = 13$), 2.2% selected option “other” ($n = 10$), 0.8% Indigenous ($n = 5$), 1.2% Japanese ($n = 7$), 1.5% West Asian ($n = 9$), and 1% Arab ($n = 6$).

The study eligibility criteria were consistent with those of Study 2. Each participant received £1.5. In accordance with ethical guidelines, all study procedures were approved by the Human Participants Review Sub-Committee, York University’s Ethics Review Board.

Measures

Speaking Arrangement Preference

Speaking arrangement preference was captured using the same item used in Study 2a.

Self-Perceived Performance

Participants indicated their self-perceived performance using the same item as in Study 2a. Below, I again report findings by grouping individuals who self-placed in ranks 1 or 2 as “top performers”, whereas those who self-placed in ranks 3 or 4 as “bottom performers.”

²¹ A total of 599 participants completed the questionnaire. However, five inattentive responders who failed to pass an instructed response item embedded in the survey, and two participants who did not identify as “man” or “woman”, were excluded from analyses to ensure data quality and fulfill sampling requirements.

Subjective Evaluations of Each Speaking Arrangement

Participants rated each speaking arrangement on four dimensions detailed below, on 10-point scales ranging from “1—not at all” to “10—very much.” Each dimension was rated using a single item.

Fairness. The perceived fairness of each speaking arrangement was captured using the following item: “to what extent is this speaking arrangement fair?” ($M = 6.5$, $SD = 1.4$).

Team Performance. Participants indicated how conducive each speaking arrangement was to high team performance through the following question: “to what extent will this speaking arrangement lead to better team performance?” ($M = 6.7$, $SD = 1.3$).

Sense of Belongingness. Participants indicated how conducive each speaking arrangement was to group members’ sense of belongingness by responding to the following item: “to what extent will this speaking arrangement lead to higher feelings of belongingness?” ($M = 6.3$, $SD = 1.4$).

Desire to Lead. The leadership appeal of both speaking arrangements was captured by asking participants, “how much would you want to lead a group that follows this speaking arrangement?” ($M = 6.3$, $SD = 1.6$).

Analytical Approach

All descriptive and inferential analyses involving speaking rule preference and self-perceived performance were conducted identically to those in Study 2a (see Study 2a Analytical Approach section for more information). To assess how ratings for each construct varied across the two speaking arrangement options, I used *tidyr* (Wickham, 2026) to convert the data from wide to long format, ran *t*-tests to examine gender differences in ratings with base R, and calculated corresponding Cohen’s *ds* with *effectsize* (Ben-Shachar et al., 2020). The *ggplot2*

package (Wickham, 2016), *scales* (Wickham et al., 2025), and *RColorBrewer* (Neuwirth, 2022) were used to graph speaking arrangement preferences across groups and scenario-based ratings across gender.

Results

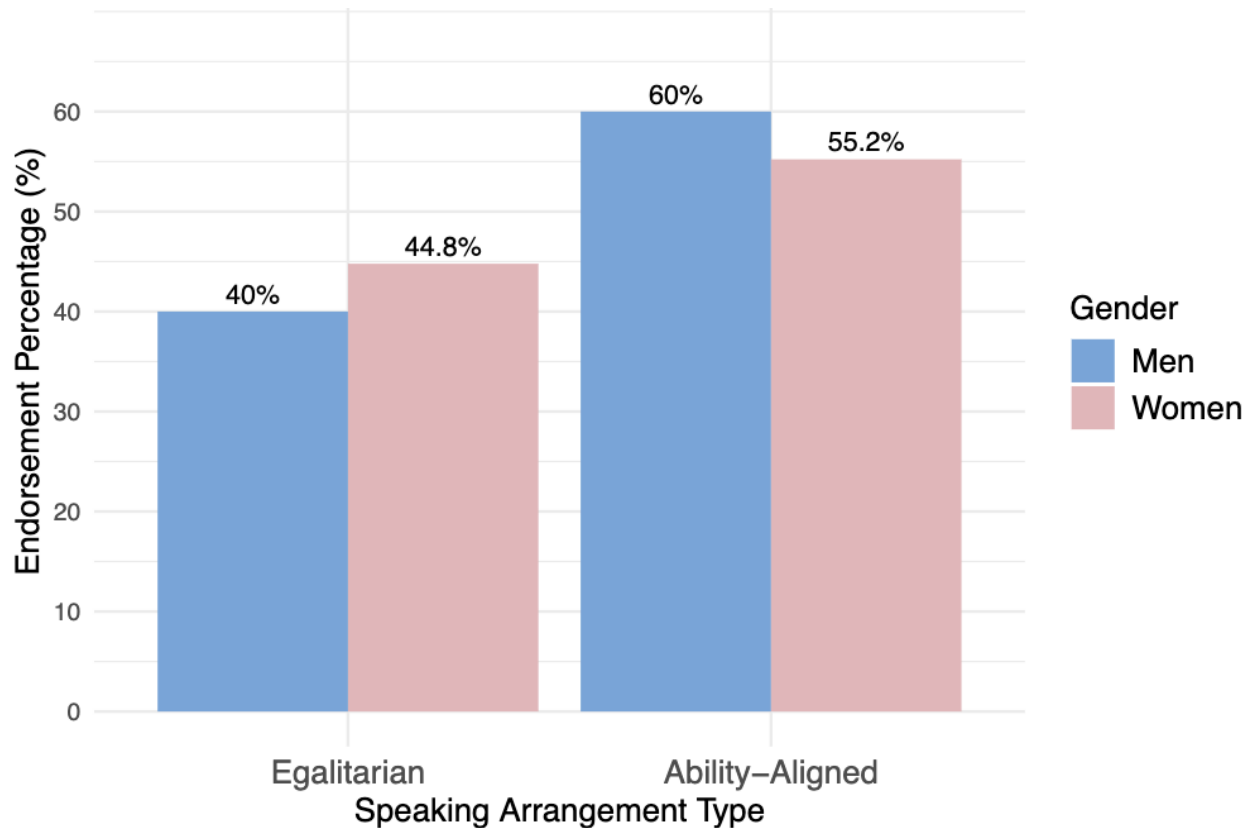
Do Women Hold Stronger Preferences for Egalitarian Distributions of Decision-Making Power, Whereas Men Prefer Ability-Aligned Structures in Which More Capable Individuals Receive Greater Voice (RQ3)?

The gender breakdown of speaking arrangement endorsement is illustrated in Figure 5. Descriptively, the majority of men (60%; $n = 177$) and women (55.2%; $n = 164$) preferred the ability-aligned distribution of speaking time. Notably, the egalitarian design was most popular with women (44.8%; $n = 133$) compared to men (40%; $n = 118$). This pattern diverges from that found in Study 2a.

To formally test the association between gender and speaking arrangement preference, a binomial logistic regression was used to regress speaking arrangement choice (egalitarian coded ‘1’ and ability-aligned coded ‘0’) on gender (women coded ‘1’ and men coded ‘0’). Unlike in Study 2a, although trending in the same direction, gender did not predict the likelihood of choosing the egalitarian speaking arrangement at a level of statistical significance, $OR = 1.22$, $SE = .17$, 95% CI [0.88, 1.68], $p = .24$.

Figure 5

Study 2b: Overall, Both Men and Women Prefer the Ability-Aligned Speaking Arrangement



Note. Bars represent the within-gender percentage of men ($n = 295$) and women ($n = 297$) in their choice of egalitarian or ability-aligned speaking arrangement.

What Psychological Factors Contribute to Gender Differences in Support for Ability-Aligned Speaking Arrangements (RQ4)?

Although the preceding analyses reveal gender differences in participants' stated preference for speaking arrangements, they do not clarify why men and women differ in these choices. To better understand the psychological basis of these preferences, I next examined whether men and women differed in how they subjectively evaluated egalitarian and ability-

aligned arrangements across several theoretically relevant dimensions, including perceived fairness, anticipated team performance, belongingness, and leadership aspiration. I report these results in turn. All between-gender comparisons were conducted using independent samples *t*-tests, whereas all within gender comparisons were conducted using paired samples *t*-tests.

As can be seen from Figure 6, between-gender comparisons revealed that men evaluated the ability-aligned arrangement as significantly more fair than women did, $t(589.67) = 3.19, p = .001, d = 0.26, 95\% \text{ CI } [0.10, 0.42]$. Yet, there was no gender difference in fairness ratings of the egalitarian arrangement, $t(588.93) = -0.04, p = .97, d = -0.003, 95\% \text{ CI } [-0.16, 0.16]$.

Furthermore, examining within genders, the egalitarian arrangement was regarded as more fair than the ability-aligned arrangement by both women, $t(296) = 13.26, p < .001, d = 0.77, 95\% \text{ CI } [0.64, 0.90]$, and men, $t(294) = 11.14, p < .001, d = 0.65, 95\% \text{ CI } [2.10, 3.00]$.

Similarly, men evaluated the ability-aligned arrangement as more likely to lead to higher team performance than women did, $t(584.07) = 2.44, p = .02, d = 0.20, 95\% \text{ CI } [0.04, 0.36]$. However, no gender difference in anticipated team performance emerged for the egalitarian arrangement, $t(587.16) = -0.23, p = .82, d = -0.02, 95\% \text{ CI } [-0.18, 0.14]$. Within gender comparisons revealed that both women, $t(296) = -3.03, p = .003, d = -0.18, 95\% \text{ CI } [-0.29, -0.06]$, and men, $t(294) = -5.33, p < .001, d = -0.31, 95\% \text{ CI } [-0.43, -0.19]$, perceived the egalitarian arrangement to be less conducive to high team performance than the ability-aligned arrangement.

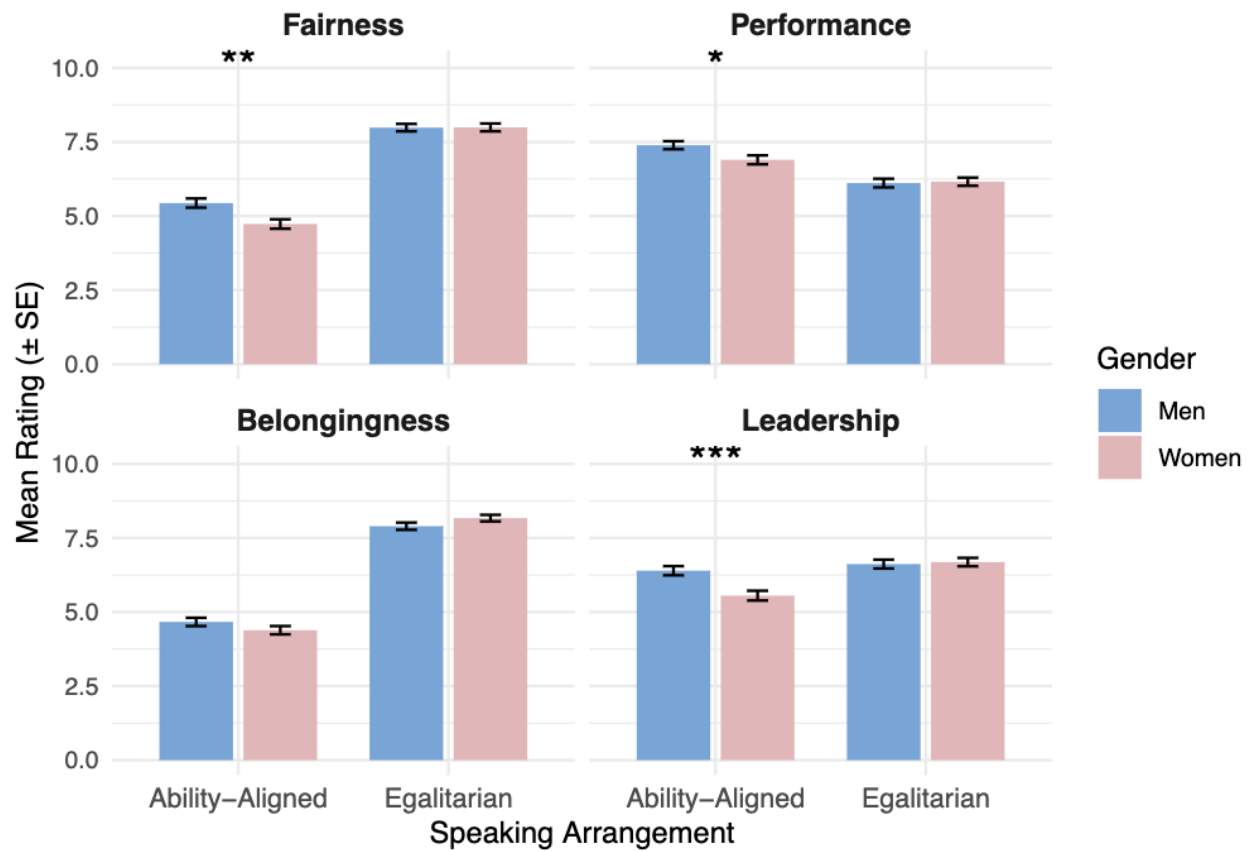
Not surprisingly, men expressed a stronger desire to lead under the ability-aligned conditions than women did, $t(587.87) = 3.71, p < .001, d = 0.31, 95\% \text{ CI } [0.14, 0.47]$. However, they did not significantly differ in their desire to lead under the egalitarian conditions, $t(588.48) = -0.32, p = 0.75, d = -0.03, 95\% \text{ CI } [-0.19, 0.13]$. Notably, within-gender comparisons show that

while women wished to lead in the egalitarian context more than the ability-aligned one, $t(296) = 4.48, p < .001, d = 0.26, 95\% \text{ CI } [0.14, 0.38]$, the same was not true for men, $t(294) = 0.91, p = .36, d = 0.05, 95\% \text{ CI } [-0.06, 0.17]$.

Lastly, men and women did not differ in their perceptions of belongingness under the egalitarian arrangement, $t(583.63) = -1.67, p = .10, d = -0.14, 95\% \text{ CI } [-0.30, 0.02]$, or the ability-aligned arrangement, $t(590) = 1.42, p = .16, d = 0.12, 95\% \text{ CI } [-0.04, 0.28]$. However, within-gender comparisons indicated that both women, $t(296) = 18.56, p < .001, d = 1.08, 95\% \text{ CI } [0.93, 1.22]$, and men, $t(294) = 16.48, p < .001, d = 0.96, 95\% \text{ CI } [0.82, 1.10]$, rated the egalitarian speaking arrangement as more likely to foster a higher sense of belongingness than the ability-aligned arrangement.

Figure 6

Study 2b: Gender Differences in Subjective Evaluations of the Merits of Each Speaking Arrangement



Note. Mean ratings for fairness, anticipated team performance, anticipated sense of belongingness, and leadership aspiration are shown for each speaking arrangement as evaluated by each gender. Standard errors are represented by the error bars. Asterisks indicate statistically significant gender differences in ratings within each speaking arrangement based on independent-samples *t*-tests. $p < .05^*$, $p < .01^{**}$, $p < .001^{***}$.

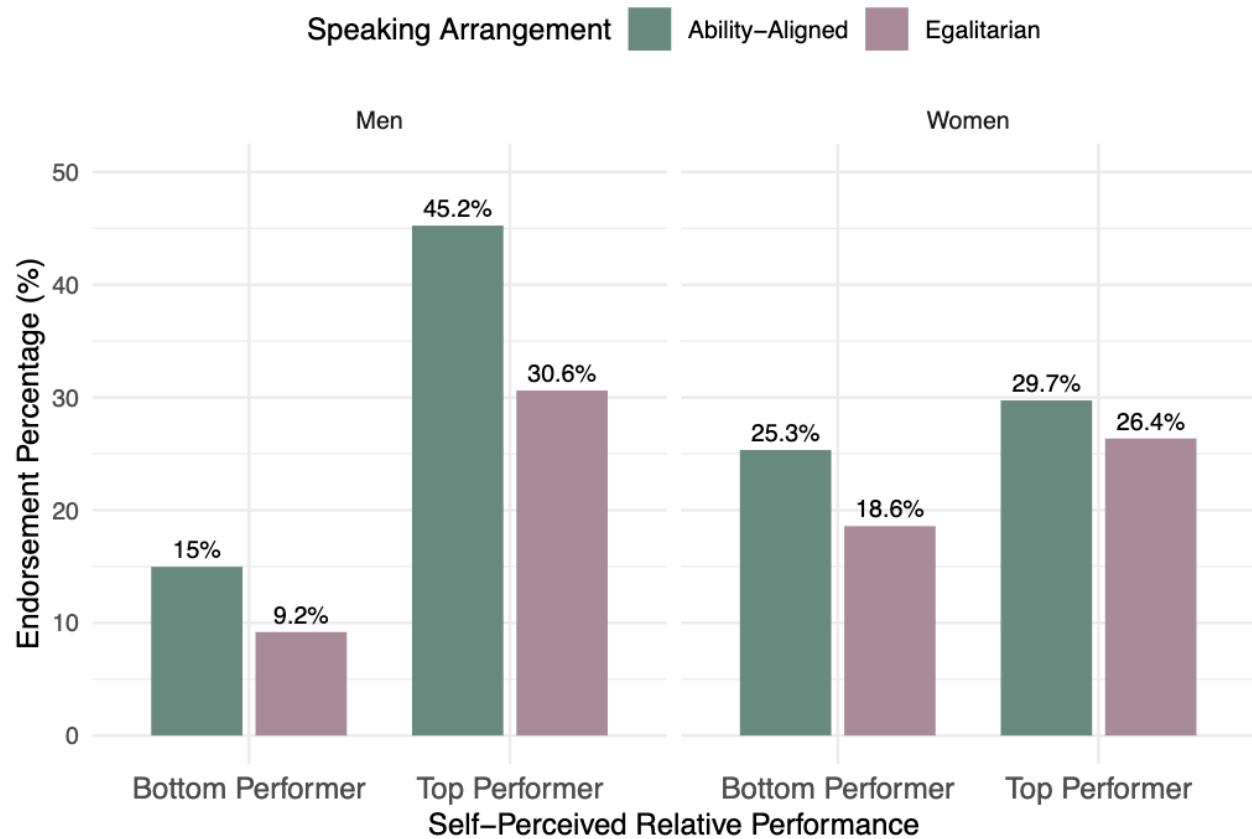
Exploratory Analysis: Examining Speaking Arrangement Choice by Gender and Self-Perceived Performance

Figure 7 illustrates the gender breakdown in speaking arrangement choice across levels of self-perceived performance. As in Study 2a, I explored whether the gendered pattern of speaking arrangement preference emerged among top performers in this sample. Descriptively, the ability-aligned design was the most endorsed decision-making style across gender and performance level. However, men who considered themselves to be top performers opted for the ability-aligned arrangement at a higher rate than others—including women top performers and low performers of either gender. Among top performing men (75.9% of all men; $n = 223$), 59.6% ($n = 133$) opted for the ability-aligned speaking arrangement, relative to 40.4% ($n = 90$) who chose the egalitarian design. By contrast, top performing women's (56.1% of all women; $n = 166$) preferences were more evenly split, with 53% ($n = 88$) and 47% ($n = 78$) choosing these arrangements, respectively. However, unlike Study 2a, a chi-square test showed no association between gender and speaking rule preference among top performers, $\chi^2(1, N = 389) = 1.70, p = .19$, Cramér's $V = .06$. Despite the lack of a significant relationship found here, these patterns are nevertheless consistent with those found in Study 2a.

Similarly, as was found in Study 2a, gender was not associated with speaking arrangement choice among bottom performers, $\chi^2(1, N = 201) = 0.35, p = .56$, Cramér's $V = .03$. Among women who self-placed as bottom performers, ($n = 130$), less than half (42.3%; $n = 55$) preferred the egalitarian speaking arrangement, relative to 57.7% ($n = 75$) who preferred the ability-aligned arrangement. Self-perceived bottom performing men ($n = 71$) showed the same pattern, less than half (38%; $n = 27$) chose the egalitarian arrangement, whereas 62% ($n = 44$) chose the ability-aligned arrangement.

Figure 7

Study 2b: Ability-Aligned Speaking Arrangement Most Preferred Across Gender and Self-Perceived Performance



Note. The bar graph depicts the proportion of men ($n = 295$) and women ($n = 297$) within gender who preferred each speaking arrangement across two levels of self-perceived performance in a hypothetical group of four participants. “Bottom Performer” refers to participants who estimated their own performance as rank 3 or 4, whereas “Top Performer” refers to participants who estimated their performance as rank 1 or 2.

Discussion

The gender difference in decision-making power allocation preference found in Study 2a was not fully replicated in this larger sample. In contrast to Study 2a, most women in this

study—similar to most men—preferred the ability-aligned speaking arrangement over the egalitarian arrangement. Despite this finding, however, several results are nevertheless consistent with a gender difference in preference. First, women’s preferences were more evenly split across the two discussion options relative to men, who endorsed the ability-aligned design at a higher rate over the egalitarian one. Second, women consistently perceived the ability-aligned arrangement as less desirable. When evaluating this distribution rule, they rated it as less fair, associated with lower team performance, and were less interested in leading in that context, compared to men. Although it should also be noted that, interestingly, no gender differences were found in evaluations of the egalitarian structure. Together, these findings indicate that while there are some mixed findings when it comes to gender differences in revealed preference (speaking rule choice) when considered in conjunction with Study 2a, this study revealed further support for gender differences in stated preferences (subjective evaluations), particularly when it comes to women’s tendency to disfavor the ability-aligned arrangement.

Study 3

Emerging evidence from Studies 2a and 2b suggests that women view the ability-aligned speaking arrangement less favourably on multiple dimensions and may even endorse it at lower levels than men. Women’s less favorable appraisal of this discussion format is consistent with evidence from prior work demonstrating that women not only anticipate backlash for speaking more but are asymmetrically penalized for doing so (Brescoll, 2011). Study 2b findings suggest that women may anticipate such a personal disadvantage of higher speaking time, as reflected in their lower ratings of fairness and expressed desire to lead under the conditions that disproportionately allocate voice across group members. Instead, these women may be more satisfied by approaches that allocate power more fairly while still optimizing team performance.

In other words, attempts to blend hierarchy in voice allocation with some degree of egalitarianism could appeal more broadly to women and possibly men as well (RQ5). Studies 3a and 3b test this by presenting participants with a choice of three arrangements: an egalitarian format, an ability-aligned format, and a new hybrid format that blends features of the two aforementioned structures.

Study 3a

Method

Participants

The sample included 205 participants (101 women; $M_{\text{age}} = 41.2$ years, $SD = 12.5$ years) who were recruited through Prolific.²² In terms of ethnicity, 64.9% of the sample identified as White/Caucasian ($n = 133$), 11.2% Chinese ($n = 23$), 8.8% Black ($n = 18$), 3.9% South Asian ($n = 8$), 2.4% Hispanic/Latine ($n = 5$), 2.4% Southeast Asian ($n = 5$), 2% Filipino ($n = 4$), 1.5% mixed origin/multiracial ($n = 3$), 1.5% selected option “other” ($n = 3$), 0.5% Indigenous ($n = 1$), 0.5% Japanese ($n = 1$), and 0.5% West Asian ($n = 1$). A pre-screen filtered respondents based on the eligibility criteria of a minimum age of 18 years old, resident status in Canada or the U.S, and English fluency.

Participants were informed they would be asked to share their preferences for leadership in hypothetical scenarios. They received £0.5 as compensation for their completion of the five-minute questionnaire. All aspects of the present research were approved by the Human Participants Review Sub-Committee, York University’s Ethics Review Board.

²² 16 participants were excluded from analyses because they dropped out of the study before responding to the speaking rule preference item.

Procedure

The study procedure was similar to Study 2a and 2b, wherein participants were asked to imagine completing a hypothetical Lost on the Moon task with three other randomly selected participants, before proceeding to indicate their preferred speaking rule preference.²³

Measures

Speaking Arrangement Preference

Decision-making style preference was again measured by surveying the endorsement of three speaking rules presented to participants in the form of a table (see Table 3). Several modifications were made compared to Studies 2a and 2b. Speaking *time* was the currency for social influence (instead of the number of typed words as used in Studies 2a and 2b). In the group scenario, two rounds of deliberation comprised group discussions (as opposed to one round in Studies 2a and 2b). Two of the three arrangement options presented were conceptually similar to the speaking rules presented in Study 2a and 2b—all members spoke an equal amount during both rounds of discussion, regardless of performance (the egalitarian arrangement), or more skilled members dominated both conversations, respectively (the ability-aligned arrangement). The new third speaking rule option introduced—the hybrid option—contained elements of both speaking rules. Specifically, during the first round of discussion, all members spoke for an equal amount of time. Then, only top performers shared their input during the second round of discussion.

²³ One of three tables outlining the speaking rule options was randomly presented to participants. Across tables, the order in which speaking rule options were presented varied. The information was identical in each table. This method was utilized to curtail primacy bias (Miller et al., 1998).

Table 3

Studies 3a and b: Revised Speaking Arrangement Options for a Hypothetical Group Problem-Solving Activity

Performance Score (0-100%)		Speaking time I want to allocate to each member (mins)					
		Option A		Option B		Option C	
		Round 1	Round 2	Round 1	Round 2	Round 1	Round 2
Member C	90%	2	1	4	3	2	3
Member B	75%	2	1	2	1	2	1
Member D	60%	2	1	1	0	2	0
Member A	45%	2	1	1	0	2	0

Note. Performance scores represent (hypothetical) task performance on the Lost on the Moon exercise. Each speaking arrangement illustrates a different way that speaking time (in minutes) can be distributed within a group. Option A distributes speaking time equally during both discussion rounds. In Option B, more speaking time is allocated to high-performing members in both rounds. Option C allows each member to speak an equal amount during the first round and prioritizes speaking time for top performers in the second discussion round.

Self-Perceived Performance

Self-perceived performance was measured in a similar way to the preceding studies by asking participants to rank themselves in the group relative to three other randomly selected participants ($M = 2.3$, $SD = 0.7$). As in Studies 2a and 2b, in the results reported below, individuals who self-placed in ranks 1 or 2 were considered “top performer”, whereas those who self-placed in ranks 3 or 4 were considered “bottom performer.”

Analytical Approach

The speaking arrangement preference was analyzed by calculating the percentage of men and women who endorsed each speaking arrangement option. Multinomial logistic regressions

were used to model how gender related to speaking arrangement preference, given that the outcome variable is a choice between three available options. Chi-square tests of independence tested the association between gender and speaking arrangement preference among top performers.

The data were analyzed using R version 4.4.0. The *readxl* package (Wickholm, 2019a) facilitated the loading of the data into the environment. Performance level composites were created using the *dplyr* package (Wickholm, 2019b; Wickholm et al., 2019c). The *psych* package was used to calculate means and standard deviations (Revelle, 2025) and Cohen's *ds* were obtained using *effectsize* (Ben-Shachar et al., 2020). Multinomial logistic regressions were modelled with *nnet* (Venables & Ripley, 2002). Chi-square analyses were run using base R and Cramér's *Vs* were calculated with *lsr* (Navarro, 2015). Three packages were used to create bar graphs to examine speaking arrangement preferences across groups, visually (*scales*, *ggplot2*, and *RColorBrewer*; Neuwirth, 2022; Wickham, 2016; Wickham et al., 2025).

Results

Is There a Speaking Rule Group Structure That Can Bridge the Preference Divide Between Genders (RQ5)?

As predicted, the hybrid speaking arrangement was the most frequently preferred choice in *both* men and women (see Figure 8). More than two fifths (42.4%; $n = 87$) of the entire sample opted for this decision-making style. When broken down by gender, 39.4% ($n = 41$) of men and 45.5% ($n = 46$) of women chose the hybrid option. The second most preferred speaking rule was the ability-aligned arrangement; 29.3% ($n = 60$) of the entire sample chose this arrangement. Notably, this figure was largely driven by men—36.5% ($n = 38$) of whom voted for this discussion design. Only 21.8% ($n = 22$) of women voted to give more voice to higher performers

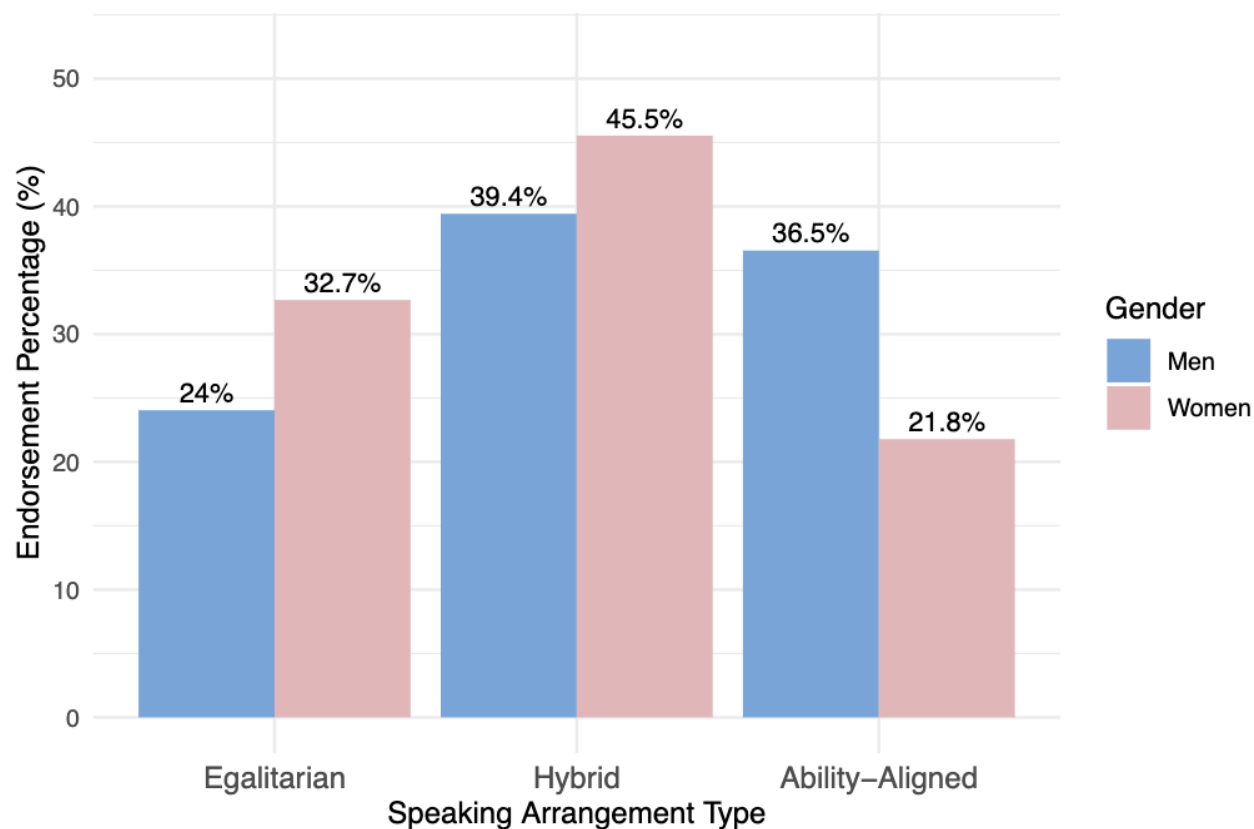
during both rounds of deliberation throughout the group exercise. Finally, less than one third (28.3%; $n = 58$) of the entire sample chose the egalitarian speaking arrangement, making it the least preferred discussion style by less than 2% compared to the ability-aligned arrangement. Endorsement rates varied between gender as 32.7% ($n = 33$) of women and 24% ($n = 25$) of men chose this speaking arrangement.²⁴ Figure 10 illustrates the spread in speaking arrangement choice across gender.

To formally test whether gender predicted speaking arrangement preference, I ran a multinomial logistic regression, regressing speaking arrangement choice (ability-aligned, hybrid, and egalitarian; the outcome) on gender (women coded '1' and men coded '0'). The ability-aligned arrangement was designated as the reference category. As predicted, results showed that gender was a significant predictor of speaking arrangement preference. Relative to men, women had marginally higher odds of preferring the hybrid arrangement over the ability-aligned one, $OR = 1.94$, $SE = 0.34$, 95% CI [0.99, 3.80], $p = .054$. In other words, women had approximately 94% higher odds of choosing the hybrid design relative to the ability-aligned design than men did. Moreover, women's odds of choosing the egalitarian arrangement versus the ability-aligned structure were 128% higher than men's, $OR = 2.28$, $SE = .38$, 95% CI [1.09, 4.77], $p = .03$.

²⁴ To ensure participants' expressed preferences for the egalitarian speaking arrangement were not motivated by an aversion to the aspect of the hybrid design which "silences" lower performing group members during the second phase of discussion by allocating them no time to speak, I piloted a modified version of this speaking arrangement which gave the bottom two performers one minute each to speak during the second round (compared to zero minutes) to assess whether the spread in speaking arrangement choice would differ ($N = 101$; $n = 50$ men). There was virtually no difference in the spread of speaking arrangement preference within or between gender across the original hybrid design presented in Study 3a and the modified pilot design. These findings provide support against the possibility that individuals may have chosen the egalitarian speaking arrangement because the hybrid option "silences" members during the second round of discussion.

Figure 8

Study 3a: Hybrid Speaking Arrangement Most Preferred Overall but Gender Differences in Second Most Preferred Arrangement



Note. The bar graph illustrates the proportion of men ($n = 104$) and women ($n = 101$) who chose the egalitarian, hybrid, and ability-aligned speaking arrangements, respectively.

Exploratory Analysis: Examining Speaking Arrangement Preference by Gender and Self-Perceived Performance

To assess whether self-perceived performance mattered for men's and women's power allocation preferences, speaking arrangement choice by performance level was assessed descriptively. Figure 9 displays the arrangement preferences by gender and self-perceived

performance. Endorsement rates for each decision-making structure varied across gender among top performers ($n = 51$ women; $n = 77$ men). The hybrid design was most strongly preferred among top-performing women (47.1%; $n = 24$), followed by the egalitarian (35.3%; $n = 18$), and then ability-aligned arrangement 17.6% ($n = 9$). Men's preferences were a sharp contrast to that of women's. Their most endorsed format was the ability-aligned design (40.3%; $n = 31$), followed by the hybrid (37.7%; $n = 29$) and then egalitarian (22.1%; $n = 17$).

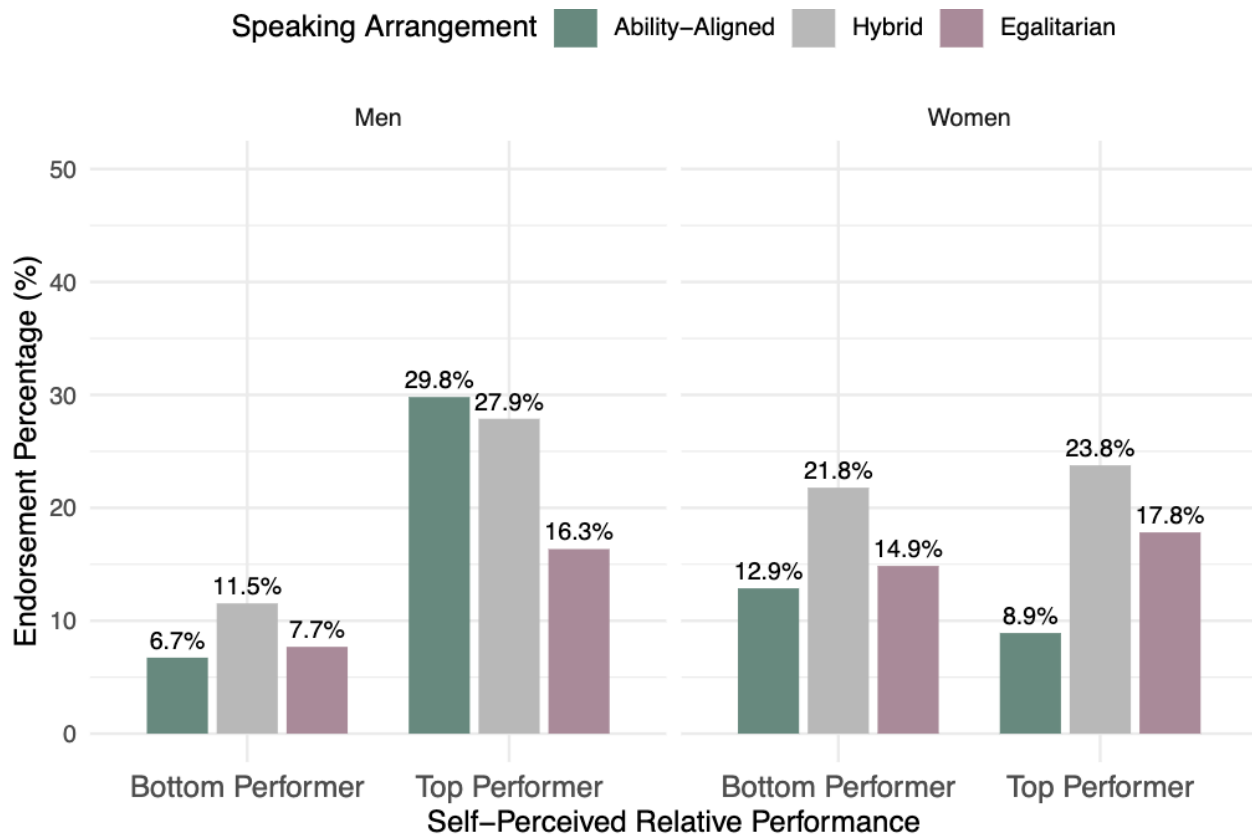
Given the overrepresentation of men who estimated their performance would fall in the top two members of a four-person group (74%; $n = 77$), potentially obscuring rank effects, a chi-square test of independence examined the association between gender and speaking arrangement preference (ability-aligned, hybrid, and egalitarian) within top performers and confirmed this relationship was statistically meaningful, $\chi^2(2, N = 128) = 7.63, p = .02$, Cramér's $V = .24$. Not surprisingly, the association between gender and speaking arrangement preference among top performers persisted when considering ability-align and egalitarian preferences, exclusively, $\chi^2(1, N = 75) = 6.78, p = .01$, Cramér's $V = .24$.

Interestingly, among bottom-performers, speaking arrangement preferences were nearly identical across gender. Almost half of women (44%; $n = 22$) and men (44.4%; $n = 12$) preferred the hybrid speaking arrangement, making it the most endorsed across gender at this performance level. Bottom-performers' second most-supported discussion format was the egalitarian arrangement which was chosen by 30% of women ($n = 15$) and 29.6% of men ($n = 8$), with the ability-aligned speaking arrangement as the least preferred by women (26%; $n = 13$) and men (25.9%; $n = 7$). A chi-square test of independence revealed no significant association between gender and arrangement preference (considering all three arrangement types), $\chi^2(2, N = 77) = 0.00, p = 1.00$, Cramér's $V = .00$. Examining the relationship between gender and speaking

arrangement preference among the egalitarian and ability-aligned choices, only, still no association emerged, $\chi^2(1, N = 43) = 0.00, p = .99$, Cramér's $V = .00$.

Figure 9

Study 3a: Gender Differences in Speaking Arrangement Style Preferences Across Self-Perceived Performance



Note. The bar graph illustrates the proportion of men ($n = 104$) and women ($n = 101$) within gender who preferred each type of speaking arrangement, by self-perceived performance in a hypothetical group of four random participants. “Bottom Performer” refers to participants who estimated their performance as rank 3 or 4. “Top Performer” refers to participants who estimated a personal rank of 1 or 2.

Discussion

Consistent with predictions, overall, *both* men and women preferred the hybrid speaking rule which allocated speaking time equally during the first round of the group discussion and dedicated more speaking time to more skilled members during the second round. This provides novel evidence that men's and women's preferences for decision-making power allocation converge when components of egalitarian and performance-based principles are integrated into the decision-making process. On the whole, this suggests that the judgements men and women make about how voice is best allocated in a group are both rooted in fairness *and* performance-oriented motivations.

However, the weight people place on fairness and performance appears to vary depending on their sense of relative performance within a team and their gender. Study 3a results revealed a distinct pattern of self-perceived performance when it comes to speaking arrangement choice between men and women such that high performing women were more egalitarian in their preferences for speaking dispersion compared to men who chose discussion structures that were more self-serving. Descriptively speaking, the most pro-egalitarian group was high performing women. Seeing as this discussion dynamic would directly limit their ability to shape group decisions relative to other members, despite being the most qualified to make the correct choices, this highlights their true egalitarian nature. Alternatively, men who anticipated high performance tended to select the decision-making structure that rewarded more knowledgeable members with more voice during both rounds of discussion. Importantly, consistent with Study 2a findings, these gendered patterns in preference were performance-dependent as self-perceived bottom performers showed no such association between gender and speaking arrangement preference.

Men's substantially higher self-perceived performance compared to women sparks questions about the role of perceived performance in decision-making power allocation preference. Although survival task performance was not explicitly tested, one can infer that men were overconfident in their performance estimates based on the Study 1 findings that revealed no gender difference in survival task performance. This trend aligns with previous work that finds women are more accurate in their assessments of their abilities, unlike men, who often demonstrate high levels of confidence even when they lack competence (Lundeberg et al., 1994). One possibility is that men's overconfidence may have led them to erroneously underestimate the probability that they would perform poorly, unlike women, who may have chosen the arrangement that distributed power outcomes more equally as insurance in case they were one of the bottom-performing group members in the scenario. Ultimately, this begs the question of whether men's choices would have been different if they knew their true performance score (especially in instances where they were bottom-performers), considering how only roughly one quarter of bottom-performing men chose the ability-aligned speaking rule (which would undermine their ability to contribute to the group discussion) relative to the two-fifths of self-perceived high-performing men.

Importantly, when considering the two speaking arrangement options presented in Studies 2a, exclusively, the effect of gender on decision-making power allocation preference replicates in Study 3a. After accounting for individuals who chose the hybrid arrangement, women were more inclined to select the egalitarian distribution of speaking time, relative to men, who preferred the ability-aligned structure. Given the strong endorsement of the hybrid design in Study 3a, this suggests that when forced to decide between group conversation dynamics that favour fairness *or* objective performance, women gravitate toward the former and men, the latter.

This aligns with the notion that women tend to be more communal and democratic in their approach to leadership, whereas men are more task-oriented (Eagly & Johnson, 1990; Eagly, 2009).

In summary, Study 3a highlights a potential common ground between men and women on decision-making style dynamics. Overall, both genders prefer discussion styles that balance fairness and knowledge prioritization during the consensus-building process. In this way, enforcing discussion designs that allow members to contribute equally during the initial stages of consensus building and awarding more influence to more skilled members during the latter stages can curtail the extent to which members feel disregarded, while still positioning teams to perform well. These findings provide encouragement that balancing competing power allocation preferences in mixed-gender groups is more feasible than initially thought.

Study 3b

To verify the robustness of the gender convergence in decision-making power allocation preferences identified in Study 3a, I again tested whether the hybrid option would be the most preferred speaking arrangement across gender (RQ5) in a larger sample ($N = \sim 600$). Increasing the number of observations in this replication study by nearly three times that of the previous study will increase confidence in the generalizability of this finding.

As in Study 2b, I also examined the psychological factors underlying gender differences in speaking rule preferences by assessing how men and women evaluated each speaking rule structure in terms of fairness, anticipated team performance, belongingness, and leadership motivation (RQ4). These evaluations help clarify why ability-aligned, egalitarian, and hybrid arrangements may differentially appeal to men and women.

Method

Participants were recruited using Prolific. To determine the target sample size, an *a priori* power analysis—informed by the effect sizes obtained from the Study 3a multinomial logistic regression that showed an effect of gender on speaking arrangement preference—was conducted. Specifically, I used the estimates derived from the contrasts modelling gender predicting hybrid and egalitarian speaking arrangement preference relative to the ability-aligned design (reference group) for the power calculation.²⁵ A sample size of $N = 299$ was suggested to detect an effect with 80% power. To be conservative, approximately twice the minimum number of observations were collected (target $N = 600$). All study procedures, predictions, and the analysis plan were pre-registered on AsPredicted (<https://aspredicted.org/it7fv7.pdf>).

Participants

A total of 602 (303 women, $M_{age} = 38$ years, $SD = 13.6$ years) individuals from Prolific participated in this research.²⁶ Inclusion criteria matched that of Studies 1, 2a, 2b, and 3a: using a pre-screener, individuals who were fluent in English, 18 years old or older, and residents of Canada or the U.S were invited to take part in the study. Participants earned £1.50 as compensation for their completion of the 15-minute survey. Before beginning, respondents were informed that they would be asked to complete a task independently, envision hypothetical group work scenarios, and report their preferences for leadership under those conditions. More than

²⁵ At the time of the analysis, the Study 3a multinomial model was approximated with women as the reference group for the gender predictor. Consequently, odds ratios of approximately $OR = 0.52$ (hybrid versus ability-aligned contrast) and $OR = 0.44$ (egalitarian versus ability-aligned contrast) were used to determine the required sample size—these differ from the odds ratio values cited in the Study 3a results section above. Later on, the reference category for the model was subsequently reversed for more intuitive interpretability, which led to this change. Note that the coefficients referenced in the main text are the reciprocal odds ratios of those used in the power analysis here. To be clear, the reparameterization does not change the underlying effect sizes used to estimate the Study 3b sample size, as multinomial logit coefficients are not dependent on the choice of reference category; it only changes how the coefficients are expressed depending on the coding scheme.

²⁶ In total, 636 participants completed the survey. Thirty-three participants failed to pass the attention check, and one identified with a gender other than “man” or “woman.” These respondents ($n = 34$) were excluded from analyses.

half of the sample identified as White/Caucasian (53.5%; $n = 322$), 6.5% were Chinese ($n = 39$), 16.1% were Black ($n = 97$), 4.5% were South Asian ($n = 27$), 4.3% were Hispanic/Latine ($n = 26$), 3.5% were Southeast Asian ($n = 21$), 1.5% were Filipino ($n = 9$), 5% were mixed origin/multiracial ($n = 30$), 2% selected option “other” ($n = 12$), 1% were Indigenous ($n = 6$), 1.2% were Japanese ($n = 7$), 0.5% were Arab ($n = 3$), and 0.5% were West Asian ($n = 3$). All study procedures were approved by the Human Participants Review Sub-Committee, York University’s Ethics Review Board.

Measures

All measures were nearly identical to those used in Study 3a.²⁷ The only exception is the addition of the subjective evaluations from Study 2a. For this component, participants rated each speaking arrangement (i.e., hybrid speaking arrangement included, unlike Study 2b) along the same four dimensions using the same items as in Study 2a: fairness ($M = 6.6$, $SD = 1.5$), team performance ($M = 6.9$, $SD = 1.3$), sense of belongingness ($M = 6.4$, $SD = 1.4$), and leadership motivation ($M = 6.4$, $SD = 1.7$).²⁸

Analytical Approach

The analytical procedure was identical to that in Study 3a. In addition, chi-square goodness of fit tests were conducted with the *lsr* package (Navarro, 2015) to determine whether the distribution of speaking rule preferences deviated from chance (i.e., equal proportions across speaking arrangement options). I also examined and graphed how evaluations of each speaking arrangement varied across gender using *t*-tests. To do this, I used *tidyr* (Wickham, 2026) to

²⁷ Self-perceived performance was scaled differently than in previous studies. In Study 3b, participants were asked to indicate where they anticipated their performance on the survival task would rank among a group of 99 randomly selected others, wherein smaller scores reflected higher performance, and larger scores reflected poorer performance.

²⁸ Congruent with the Study 3a method, one of three speaking arrangement tables were randomly presented to participants. Each table contained the same three speaking arrangements, but the order in which they were presented across the table columns differed to reduce order effects (Miller, 1998).

convert the data from wide to long format, base R to run the *t*-tests, and *ggplot2* (Wickham, 2016), *scales* (Wickham et al., 2025), and *RColorBrewer* (Neuwirth, 2022) to create the associated visualizations.

Results

Is There a Speaking Rule Group Structure That Can Bridge the Preference Divide Between Genders (RQ5)?

To examine whether the hybrid speaking arrangement appeals more broadly across genders, I next compare the rates at which participants selected the egalitarian, ability-aligned, and hybrid discussion structure as their preferred arrangement. In descriptive terms, the hybrid speaking arrangement was preferred by the majority of men and women (see Figure 10). Overall, across gender, 51.5% ($n = 310$) of the entire sample chose this structure. Within gender, 52.2% ($n = 158$) of all women and 50.8% ($n = 152$) of all men, chose the hybrid option. Interestingly, the ability-aligned and egalitarian options were equally endorsed across gender with 24.3% of the entire sample ($n = 146$) choosing these designs, respectively. Men (24.4%; $n = 73$) and women (24.1%; $n = 73$) chose the ability-aligned rule at similar relative frequencies. This trend also emerged with the egalitarian choice: roughly a quarter of all men (25.8%; $n = 74$) and women (23.8%; $n = 72$) chose this format.

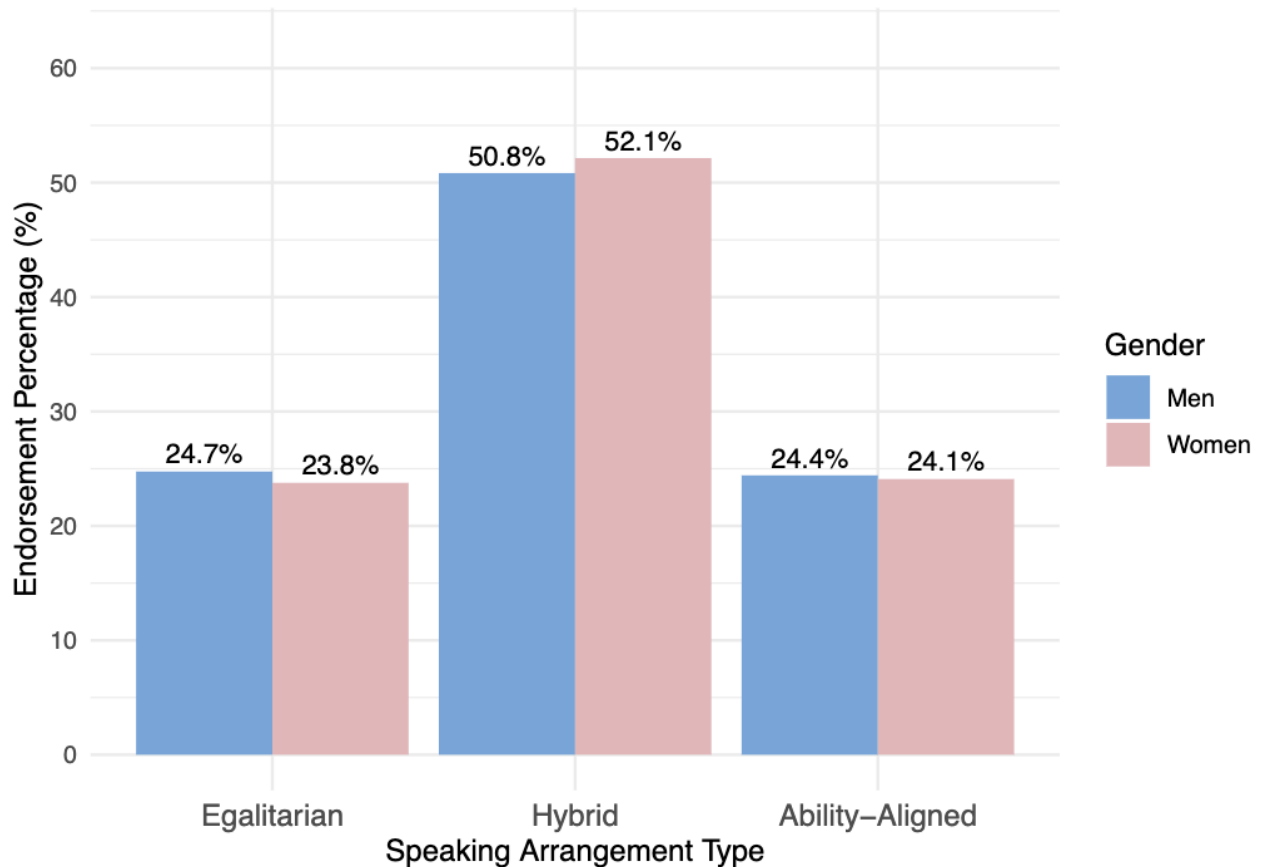
As foreshadowed by the raw choice counts between gender above, chi-square goodness of fit test confirmed that the three speaking arrangements were not equally preferred when collapsing across gender, $\chi^2(2, N = 602) = 89.36, p < .001$. Bonferroni-adjusted pairwise comparisons confirmed that the hybrid speaking rule was preferred more than both the ability-aligned and egalitarian arrangements ($ps < .001$), whereas preferences for the ability-aligned and egalitarian arrangements did not differ ($p = 1.00$). Unlike in Study 3a, however, when I tested for

gender difference in choice, a chi-square test of independence revealed no association between gender and speaking arrangement choice, $\chi^2(2, N = 602) = 0.12, p = .94$, Cramér's $V = .01$.

As an alternative test of gender difference in choice, I conducted a multinomial logistic regression with arrangement choice as the outcome (with the ability-aligned speaking rule as the reference category) and gender (men = 0, women = 1) as the predictor. The model output indicated that men and women did not differ in their odds of choosing the hybrid arrangement (OR = 1.04, $SE = 0.20$, 95% CI [0.70, 1.54], $p = .85$) or the egalitarian arrangement (OR = 0.97, $SE = 0.23$, 95% CI [0.61, 1.54], $p = .91$), relative to the ability-aligned arrangement.

Figure 10

Study 3b: Hybrid Speaking Arrangement Most Preferred Across Gender, No Differences in Second Most Preferred Arrangement



Note. The bar graph illustrates the percentage of men ($n = 299$) and women ($n = 303$) who chose each speaking arrangement.

What Psychological Factors Contribute to Gender Differences in Support for Ability-Aligned Speaking Arrangements (RQ4)?

As in Study 2b, I examined whether men and women differed in their subjective perceptions of all three speaking arrangements on the dimensions of fairness, anticipated team performance, belongingness, and leadership aspiration. The results, visualized in Figure 11, are

reported below. Gender comparisons were conducted using independent samples *t*-tests. Within-gender comparisons were conducted using paired samples *t*-tests.

There was a marginal gender difference in perceived fairness of the ability-aligned arrangement such that men rated it as more fair than women did, $t(599.85) = 1.86, p = .06, d = 0.15$, 95% CI [-0.01, 0.31]. Consistent with Study 2b, no gender differences in perceived fairness emerged for the egalitarian arrangement, $t(599.99) = -0.76, p = .45, d = -0.06$, 95% CI [-0.22, 0.10]. Novel to this study, I also found no gender difference in the perceived fairness of the hybrid arrangement, $t(590.85) = 1.26, p = .21, d = -0.10$, 95% CI [-0.06, 0.26]. Within-gender comparisons further revealed that women rated the egalitarian arrangement as more fair than the ability-aligned arrangement, $t(302) = 14.66, p < .001, d = 0.84$, 95% CI [0.71, 0.97], but less fair than the hybrid arrangement, $t(302) = 8.93, p < .001, d = -0.51$, 95% CI [0.39, 0.63]. Women also rated the hybrid arrangement as more fair than the ability-aligned arrangement, $t(302) = 10.48, p < .001, d = 0.60$, 95% CI [0.48, 0.72]. Among men, similar to women, they rated the egalitarian arrangement as more fair than the ability-aligned arrangement, $t(298) = 11.78, p < .001, d = 0.68$, 95% CI [0.55, 0.81], and the hybrid arrangement, $t(298) = 7.42, p < .001, d = 0.43$, 95% CI [0.31, 0.55]. They also rated the hybrid arrangement as more fair than the ability-aligned one, $t(298) = 8.34, p < .001, d = 0.48$, 95% CI [0.36, 0.60].

As was found in Study 2b, although men anticipated higher team performance than women did under the ability-aligned arrangement, $t(594.23) = 2.44, p = .02, d = 0.20$, 95% CI [0.04, 0.36], there was no gender difference in performance evaluations under the egalitarian arrangement, $t(599.82) = -1.75, p = .08, d = -0.14$, 95% CI [-0.30, 0.02]. Novel to this study, I found that men also anticipated higher team performance than women under a hybrid arrangement, $t(587.96) = 2.50, p = .01, d = 0.20$, 95% CI [0.04, 0.36]. Among women, there was

no significant difference in perceptions of team performance between the egalitarian arrangement and the ability-aligned arrangement, $t(302) = -0.44, p = .66, d = -0.03, 95\% \text{ CI } [-0.14, 0.09]$. However, the hybrid arrangement was rated more favorably in this respect compared to the egalitarian arrangement, $t(302) = -3.51, p < .001, d = -0.20, 95\% \text{ CI } [-0.32, -0.09]$, and the ability-aligned arrangement, $t(302) = 3.73, p < .001, d = 0.21, 95\% \text{ CI } [0.10, 0.33]$. Yet, men rated the egalitarian arrangement as less conducive to team performance relative to the ability-aligned format, $t(298) = -4.58, p < .001, d = -0.26, 95\% \text{ CI } [-0.38, -0.15]$, and the hybrid one, $t(298) = -8.45, p < .001, d = 0.49, 95\% \text{ CI } [-0.61, -0.37]$. Interestingly, as women, they also perceived the hybrid arrangement more favourably on this dimension compared to the ability-aligned arrangement, $t(298) = 3.50, p < .001, d = 0.20, 95\% \text{ CI } [0.09, 0.32]$

Also consistent with Study 2b, men reported a higher desire to lead than women under the ability-aligned arrangement, $t(599.98) = 3.07, p = .002, d = 0.25, 95\% \text{ CI } [0.09, 0.41]$. The same gender pattern was also found for the hybrid arrangement, $t(598.93) = 2.81, p = 0.005, d = 0.23, 95\% \text{ CI } [0.07, 0.39]$. Again, as in Study 2b, there was no gender difference in desire to lead under the egalitarian arrangement, $t(599.37) = -1.51, p = 0.13, d = -0.12, 95\% \text{ CI } [-0.28, 0.04]$. Within women, desire to lead was highest within the egalitarian arrangement compared to the ability-aligned arrangement, $t(302) = 6.39, p < .001, d = 0.37, 95\% \text{ CI } [0.25, 0.48]$. The hybrid arrangement was also preferred over the ability-aligned arrangement at a level that reached statistical significance, $t(302) = 8.33, p < .001, d = 0.48, 95\% \text{ CI } [0.36, 0.60]$. Interestingly, women showed no difference in leadership aspirations between the egalitarian arrangement and the hybrid arrangement, $t(302) = 0.68, p = .50, d = 0.04, 95\% \text{ CI } [-0.07, 0.15]$. Among men, there was no difference in leadership aspirations between the egalitarian and ability-aligned arrangements, $t(298) = 1.57, p = .12, d = 0.09, 95\% \text{ CI } [-0.02, 0.20]$. However, men did express

higher levels of leadership aspiration under the hybrid conditions than the ability-aligned ones, $t(298) = 7.19$ $p < .001$, $d = 0.42$, 95% CI [0.30, 0.53] and the egalitarian arrangement, $t(298) = -3.79$ $p < .001$, $d = -0.33$, 95% CI [-0.33, -0.10].

Finally, as in Study 2b, there was no gender difference in belongingness ratings for the ability-aligned arrangement, $t(600) = 1.38$, $p = .17$, $d = 0.11$, 95% CI [-0.05, 0.27], the egalitarian arrangement, $t(599.10) = -0.18$, $p = .86$, $d = -0.01$, 95% CI [-0.17, 0.15], or the hybrid arrangement, $t(587.25) = 1.78$, $p = .08$, $d = 0.15$, 95% CI [-0.01, 0.31]. Examining within gender, women perceived the egalitarian arrangement to yield a higher sense of belongingness than the hybrid arrangement, $t(302) = 12.28$, $p < .001$, $d = 0.71$, 95% CI [0.58, 0.83], and the ability-aligned arrangement, $t(302) = 18.11$ $p < .001$, $d = 1.04$, 95% CI [0.90, 1.18]. They also rated the hybrid arrangement more favourably on this dimension relative to the ability-aligned arrangement, $t(302) = 11.12$ $p < .001$, $d = 0.64$, 95% CI [0.51, 0.76]. These same patterns emerged among men: the egalitarian arrangement was associated with higher sense of belonging than the hybrid arrangement, $t(298) = 11.00$ $p < .001$, $d = 0.64$, 95% CI [0.51, 0.76], and the ability-aligned arrangement, $t(298) = 16.70$ $p < .001$, $d = 0.97$, 95% CI [0.83, 1.10]. Men also assessed the hybrid arrangement to foster a higher sense of belongingness than the ability-aligned arrangement, $t(298) = 11.53$ $p < .001$, $d = 0.67$, 95% CI [0.54, 0.79].

Figure 11

Study 3b: Gender Differences in Subjective Evaluations of the Merits of Each Speaking Arrangement

Arrangement



Note. The figure illustrates the mean ratings for fairness, anticipated team performance, sense of belonging, and leadership motivations across speaking arrangements and gender. Error bars represent standard errors and asterisks show significant gender differences in ratings within a speaking arrangement based on independent-samples *t*-tests. $p < .05^*$, $p < .01^{**}$, $p < .001^{***}$.

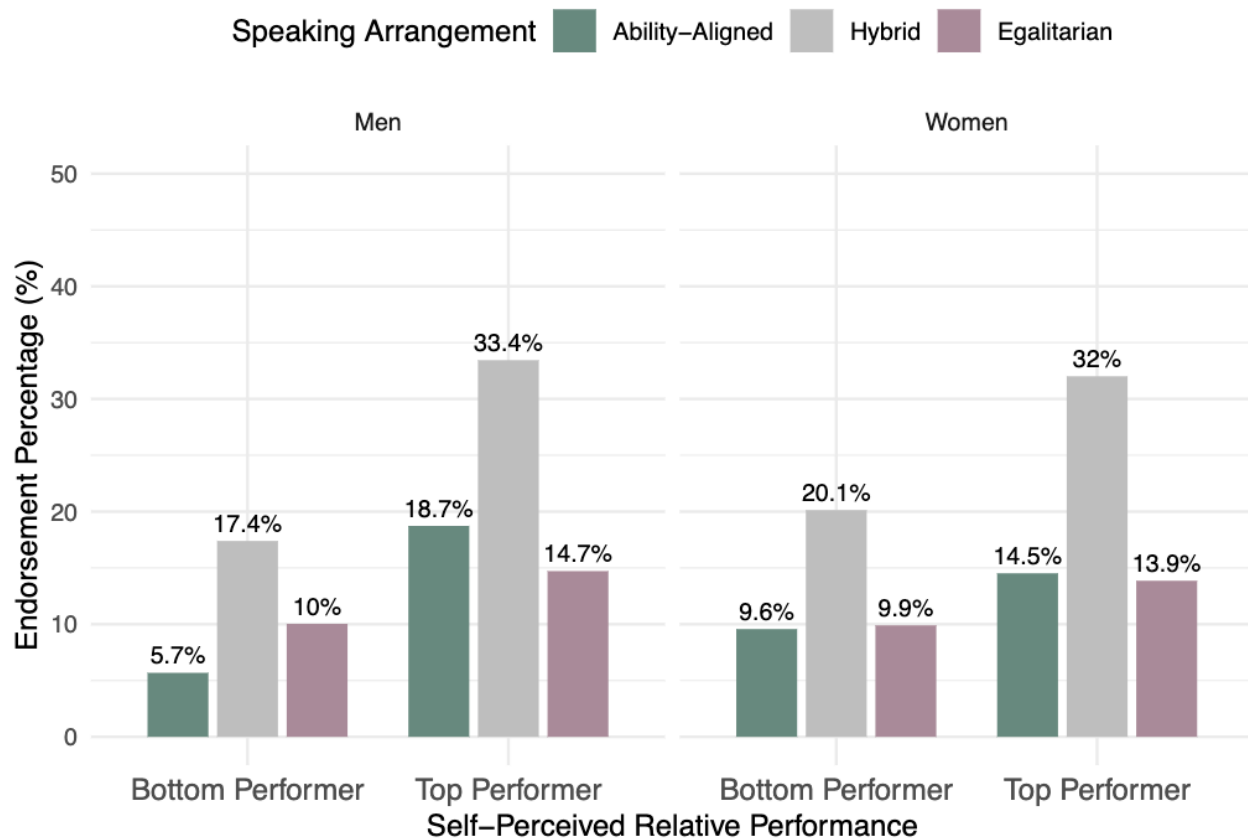
Exploratory Analysis: Examining Speaking Arrangement Preferences by Gender and Self-Perceived Performance

Figure 12 displays the breakdown of speaking rule preference by gender and self-perceived performance. Among top performing women ($n = 183$), 53% ($n = 97$) chose the hybrid design, 24% ($n = 44$) chose the ability-aligned design, and 23% ($n = 42$) chose the egalitarian design. Top performing men's preferences ($n = 200$) mirrored this pattern with 50% ($n = 100$) choosing hybrid, 28% ($n = 56$) ability-aligned, and 22% ($n = 44$) choosing egalitarian. A chi-square test of independence revealed gender was not associated with decision-making style preference among top performers when considering all three arrangement options, $\chi^2(2, N = 383) = 0.78, p = .68$, Cramér's $V = .05$, or comparing ability-aligned and egalitarian choice, exclusively, $\chi^2(1, N = 186) = 0.44, p = .51$, Cramér's $V = .04$.

In line with Studies 2a and 2b, and 3a and 3b, there was no association between gender and speaking arrangement preference among bottom performers when considering all three arrangement options, $\chi^2(2, N = 219) = 1.85, p = .40$, Cramér's $V = .09$, or considering the ability-aligned and egalitarian arrangement choices, only, $\chi^2(1, N = 106) = 1.79, p = .18$, Cramér's $V = .11$. Among bottom performing women ($n = 120$), 50.8% ($n = 61$) chose the hybrid arrangement, 25% ($n = 30$) chose the egalitarian arrangement, and 24.2% ($n = 29$) chose the ability-aligned arrangement. Bottom performing men's choices ($n = 99$) followed a similar pattern: 52.5% ($n = 52$) chose the hybrid arrangement, 30.3% ($n = 30$) chose the egalitarian arrangement, and 17.2% ($n = 17$) chose the ability-aligned arrangement.

Figure 12

Study 3b: Hybrid Speaking Arrangement Most Preferred Across Gender and Self-Perceived Performance



Note. The figure displays speaking arrangement preference across gender ($N = 602$; 303 women) and levels of self-perceived performance in a group of four participants. “Top Performer” represents those who anticipated a rank among the top half of 100 random others. “Bottom Performer” is those who anticipated a rank among the bottom half of 100 random others.

Discussion

In this study, we again found evidence for the popularity of the hybrid model, converging with Study 3a. Over half of men and half of the women sampled expressed a preference for the hybrid speaking arrangement in which all members speak an equal amount during the first half

of discussion, with more knowledgeable members contributing more during the latter half. These findings align with Study 3a which identified that the hybrid arrangement was men's and women's most preferred conversation design overall. Study 3b therefore provides strong evidence that when presented with an option that merges fairness and performance-based components of decision-making, previously observed gender differences in voice allocation preferences subside. The hybrid option is preferred by the largest proportion of participants, across genders.

Interestingly, when focusing on the gender breakdown of egalitarian and ability-aligned speaking preferences, men's and women's second choices did not differ in this larger sample. I put forth two possible explanations for this null finding. First, and perhaps the most likely, is that the appeal of the hybrid arrangement was so overwhelmingly high for both genders that the number of observations for the remaining arrangements was comparably low and do not paint an accurate picture of men's and women's preferences for these decision-making styles. In other words, there is limited statistical power to detect potential gender differences in secondary preferences, and this raises concerns about the representativeness of the data for these two speaking rules. A second possibility is that there is no gender difference in endorsement for the egalitarian and ability-aligned arrangements when presented with the hybrid option, and this was not captured in Study 3a due to the considerably smaller sample size. There were 118 observations total for the egalitarian and ability-aligned rules combined across both genders after excluding hybrid votes in Study 3a. This limits the statistical power to detect reliable gender differences which could have produced the observed results here, misleading one to infer a gender difference does exist in decision-making allocation preferences, even in the presence of the hybrid rule.

As in Study 2b, I found evidence that women viewed the ability-aligned speaking rule as less fair, less effective for team performance, and expressed a lower interest in leadership under these conditions relative to men. These evaluations suggest that ability-aligned decision-making approaches appeal differently to men and women, raising concerns about their use in mixed-gender groups. By contrast, arrangements that balance fairness and performance-based principles appear broadly appealing across genders and may represent a more effective approach.

General Discussion

I investigated gender differences in volubility, influence, and preference for how these commodities should be allocated within groups, across a series of five studies and nearly 2000 participants ($N_{\text{total}} = 1,931$). My research was guided by five questions, including: whether men and women differ in their volubility (RQ1) and if so, how it contributes to women's lower leader emergence (RQ2)? After this, I explored whether women's preferences for decision-making power allocation differ from men's (RQ3), and the psychological factors that may contribute to this difference (RQ4). Then, I tested whether men's and women's decision-making style preferences converge under conditions that integrate egalitarian and ability-aligned principles (RQ5).

Study 1 results show that women speak less than men during informal group interactions, even when their identities are concealed, and this has negative consequences for their leadership attainment. Volubility mediated the effect of gender on leadership such that men spoke more than women, and this caused their higher obtained social influence during the group discussion. These findings were consistent when using alternative measures of volubility—characters and phrases typed in the group chat. Importantly, these effects emerged despite no meaningful gender differences in objective task performance, supporting the notion that speaking more may inflate

perceptions of competence (Badura et al., 2018; MacLaren et al., 2020). Despite gender differences in obtained influence, women did not report a lower desire to lead during the group exercise relative to men which challenges explanations that chalk up gender differences in volubility to women's aversion or disinterest in being leaders (Netchaeva et al., 2022).

Studies 2a and 2b extended Study 1 by surveying gender differences in voice allocation preferences during a hypothetical group problem-solving context. Although there was a demonstrable gender difference in preference, with the majority of women preferring more egalitarian distributions of power in Study 2a, this finding did not replicate in the larger Study 2b sample which found that while the egalitarian speaking rule was most popular amongst women, overall, the majority of them preferred the ability-aligned design. However, despite the apparent gender alignment in preference, evidence suggests this design appeals to men more than women as reflected in their more favourable assessments of fairness, performance, and motivation to lead groups that follow ability-aligned rules.

To explore if there are power allocation dynamics that appeal more broadly to men *and* women, Studies 3a and 3b surveyed interest in a third speaking rule that combined elements of the egalitarian and ability-aligned rules introduced in Studies 2a and 2b. Indeed, there was substantial interest in the hybrid design across gender. Notably, hybrid rule preference dominated in the Study 3b sample with more than half of men and women choosing this arrangement over the other two options, providing compelling evidence that decision-making processes that balance fairness and performance may be an effective approach for mixed gender groups. Further support for this idea is found in men's and women's statistically non-significant difference in ratings of perceived fairness and anticipated belongingness under the hybrid structure. On the other hand, women judged the ability-aligned design less favourably, as observed in Study 2b.

Specifically, they perceived this format as less fair, less beneficial for team performance, and were less motivated to lead under these conditions, indicating gender differences in evaluative valence toward ability-aligned decision-making.

A common thread across the current research is the role of self-perceived performance on decision-making power allocation preference. Anticipated performance level appears to exhibit differential effects on preferences between men and women, especially among those who anticipate being a top performer. In Studies 2a, 2b, and 3a, within top-performing men, the ability-aligned arrangement was preferred the most. In contrast, only one sample of top-performing women (Study 2b) preferred the ability-aligned arrangement over the egalitarian one, and this difference in preference was modest.

Across Studies 2a to 3b, the findings revealed a somewhat mixed pattern for revealed preferences, but a notably consistent pattern for expressed preferences. Specifically, Studies 2a and 3a found significant gender differences in speaking arrangement preference, with women showing greater preference for egalitarian rules and men showing greater preference for ability-aligned rules, whereas Studies 2b and 3b showed similar directional trends that did not reach statistical significance. In contrast, gender differences in expressed preferences were highly consistent across Studies 2b and 3b: women evaluated ability-aligned speaking arrangements as less fair, less likely to produce strong team performance, and less desirable as a context for leadership compared to men. Importantly, men and women did not differ reliably in their evaluations of egalitarian speaking arrangements, suggesting that the primary gender difference centered on the perceived desirability of ability-aligned structures (rather than egalitarian structures).

Theoretical Contributions

My work contributes to prior leadership research in three meaningful ways. First, by confirming that gender disparities in volubility not only exist in face-to-face group settings as typically examined in past work (e.g., Carter et al., 2018; Jarvis et al., 2022) but persist in socially sanitized environments where burdens to conform to gendered speaking norms are reduced. Moreover, not only are women less voluble when their gender identities are concealed, but also when they are equally as knowledgeable and motivated to lead as men. This counters the notion that women may speak less on account of lower interest in leadership (Elprana et al., 2015; Netchaeva et al., 2022; Schuh et al., 2015) and suggests that gender disparities in volubility may reflect personal preferences for different conversational structures, and more broadly, power allocation.

Second, my findings align with social role theory (Eagly, 1987; for a brief review see Eagly & Wood, 2012) by demonstrating that women resonate with more communal approaches to group decision-making, while men often gravitate to more agentic and hierarchical styles. They also extend this body of work by clarifying *why* these gendered patterns may emerge: not only do more women choose egalitarian speaking arrangements, but they also perceive and evaluate ability-aligned formats differently, associating more agentic structures with lower fairness, inclusivity, and team performance. This suggests that the speaking rules surveyed in this research are psychologically distinct and valued differently across gender, which is consistent with past work identifying gender differences in fairness beliefs (Jaffee & Shibley, 2000; Vanutelli et al., 2020), communion (Cassar & Rigdon, 2021a, 2021b; Eagly, 2009), and competitiveness (Conson & Gneezy, 2009; Saccardo et al., 2017).

Third, the current findings, which demonstrate a consistent gender difference in stated preferences—and a somewhat less consistent but still notable gender difference in revealed preferences—converge with broader research on gendered social preferences. For example, Benenson’s work (2014) suggests that girls and women tend to value equality and avoid hierarchical ranking more than boys and men. Cassar and Rigdon (2021) similarly find that women are more inclined toward fairness and equal redistribution, particularly when others may benefit. Extending this literature, the present research is among the first to systematically examine how men and women differ in their preferences for the distribution of power and influence within group decision-making structures. By connecting previously disparate literatures on volubility and gendered egalitarian preferences, these findings open new avenues for understanding how egalitarian preferences may contribute to gender differences in voice allocation, speaking behavior, and leadership emergence within groups.

More generally, my work contributes to leadership and organizational research by identifying the conditions under which women express interest in leadership, as well as contexts in which men’s and women’s leadership motivations align. While prior research has focused on gender *differences* in leadership, characterizing women’s leadership as more democratic, collaborative, and interpersonally oriented, relative to men’s more performance-focused and autocratic style (Badura et al., 2018; Eagly & Johnson, 1990; Eagly & Johannesson-Schmidt, 2001), my results show that these themes surface when women and men evaluate ability-aligned decision-making structures. Yet, at the same time, I contribute new knowledge about the conditions under which men’s and women’s power allocation preferences converge.

Practical Implications

The present findings provide practical implications for improving women's visibility in leadership roles. As evidenced here, and across multiple studies (i.e., Carter et al., 2018; Jarvis et al., 2022, von Ruden et al., 2018), women tend to speak less in group settings. This gender difference matters because speaking more is associated with higher leadership, as marked by higher objective influence on group decisions in Study 1. It follows that attempts to reduce gender differences in volubility could improve women's representation in leadership. Indeed, recent evidence has shown that once speaking time is accounted for, gender does not reliably predict leader emergence, suggesting that reducing gender disparities in participation may also reduce gender differences in leader emergence (Loignon et al., 2025).

In addition, this research gives insight into the leadership conditions that women perceive as fair and position them to contribute without incurring the social penalties associated with speaking more. Simply enforcing protocols that disproportionately award women more voice over other group members is likely insufficient and could result in social backlash for women who act in gender non-conforming ways by speaking more (Williams & Tiedens, 2016). Interventions should first prioritize equal participation and belonging within groups before shifting toward ability-aligned contribution structures where more skilled members can impart their knowledge after group needs are met.

Limitations and Future Directions

The novelty of this research provides new insights into pathways to improving women's visibility as leaders and sheds light on the psychological factors that influence gender differences in support for ability-aligned decision-making structures. Yet, future inquiries can improve the present work in two important ways.

For consistency's sake, the Lost on the Moon survival task was the sole decision-making context assessed in this research. Despite the uniquely abstract nature of space, and the associated methodological benefits for making observations about volubility, decision-making, and leadership (Yetton & Bolger, 1982), one may argue this problem space could be perceived as biased toward men. A large body of research has identified a common tendency for individuals to associate science, technology, engineering, and mathematics-related (STEM) capabilities with men (Charlesworth & Banaji, 2019; Cvencek et al., 2011; McMichael et al., 2026). This sentiment is echoed in men's chronic overrepresentation as self-perceived top performers on the survival task across studies despite no observed gender differences in task performance in Study 1. Given this, I offer two lines of inquiry for future research. First, could women have deferred to more equalized power allocation structures on account of lower confidence in the survival task subject matter to preserve their decision-making power in the event they themselves perform poorly (which is not rewarded in ability-aligned formats; Charness & Gneezy, 2012; Cronson & Gneezy, 2009; Dohmen et al., 2011)? And second, could the opposite process affect men? That is, is men's partiality to ability-aligned designs a product of their (over)confidence/anticipation of being the top performers (Buser et al., 2020; Soll & Klayman, 2004) who unequivocally benefit from those group structures by receiving more airtime during the decision-making process? One way to test these questions is to substitute the survival task used in this research for Raven's Advanced Progressive Matrices, a pattern recognition exercise that tests general cognitive abilities (Raven et al., 1998), or another alternative (perhaps even more women-centric) exercise.

Additionally, given the observed self-perceived performance-dependent effects on decision-making preferences, future research should examine whether these patterns hold when

participants know their true relative performance. One limitation to the present research is that participants did not know their actual performance score. Instead, respondents were asked to state their *anticipated* relative rank among an unknown group of hypothetical participants. Consequently, it is unclear whether voice allocation preferences would differ if respondents knew exactly where they fell in the group order, as is characteristic of many real-world settings. Consider how, in high-status hierarchy settings (e.g., armed forces, fire departments, sports teams), knowing one's rank may amplify or constrain voice allocation preferences depending on where an individual falls in the chain of command (Pfrombeck et al., 2022).

Lastly, although the current research has high statistical power with nearly 2000 individuals sampled across studies, all participants resided in Canada or the U.S. Follow-up studies should explore whether the observed gender differences replicate in non-Western, Educated, Industrialized, Rich and Democratic (non-WEIRD) societies, given previous research demonstrating WEIRD samples are not always representative of humans more generally (Henrich et al., 2010). For example, subsequent research could explore whether men consistently prefer ability-aligned decision-making formats over egalitarian ones in more collectivist societies (e.g., countries located in Asia, Africa, Latin America, etc.) which are characterized by greater prioritization of relational harmony (Kim et al., 2006; Leung & Bond, 1984) and more equalized distributions of resource allocation amongst ingroup members (Jiao & Zhao, 2023).

Conclusion

Gender asymmetries in leadership attainment are commonly thought to be driven by women's lower interest in leadership and fear of violating social norms. While a growing body of research has established that gender differences in volubility—a leadership catalyst—do exist, to my knowledge, alternative explanations for women's lower volubility have yet to be

examined. In this research, I explored whether gender disparities in speaking time reflected different beliefs about how power should be allocated within groups and found that women appear to view speaking time as a resource and exhibit more egalitarian preferences toward its distribution within groups compared to men. I also identified a group structure that appealed more broadly, across gender. Overall, this research offers a novel explanation for why women speak systematically less than men, identifies performance-preserving methods to minimize group speaking disparities, and ultimately, the gender gap in leadership emergence.

References

- Alesina, A., & La Ferrara, E. (2005). Preferences for redistribution in the land of opportunities. *Journal of Public Economics*, 89(5), 897–931.
<https://doi.org/10.1016/j.jpubeco.2004.05.009>
- American National Election Studies. (2021). The ANES 2020 Time Series. University of Michigan and Stanford University.
<https://electionstudies.org/data-center/2020-time-series-study/>
- Anderson, C., Brion, S., Moore, D. A., & Kennedy, J. A. (2012). A Status-Enhancement Account of Overconfidence. *Journal of Personality and Social Psychology*, 103(4), 718-735.
<https://doi.org/10.1037/a0029395>
- Badura, K. L., Grijalva, E., Newman, D. A., Yan, T. T., & Jeon, G. (2018). Gender and leadership emergence: A meta-analysis and explanatory model. *Personnel Psychology*, 71(3), 335–367. <https://doi.org/10.1111/peps.12266>
- Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). Regression diagnostics: *Identifying influential data and sources of collinearity* (1st ed.). Wiley. <https://doi.org/10.1002/0471725153>
- Benenson, J. F., Durosky, A., Nguyen, J., Crawford, A., Gauthier, E., & Dubé, É. (2019). Gender differences in egalitarian behavior and attitudes in early childhood. *Developmental Science*, 22(2), Article e12750. <https://doi.org/10.1111/desc.12750>
- Ben-Shachar, M. S., Lüdtke, D., & Makowski, D. (2020). “effectsize: Estimation of Effect Size Indices and Standardized Parameters.” *Journal of Open Source Software*, 5(56), Article 2815. <https://doi.org/10.21105/joss.02815>
- Blair, G., Cooper, J., Coppock, A., Humphreys, M., & Sonnet, L. (2025) *estimatr: Fast Estimators for Design-Based Inference*. R package version 1.0.6,

<https://CRAN.R-project.org/package=estimatr>

- Bollen, K. A., & Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological Methodology*, 20, 115. <https://doi.org/10.2307/271084>
- Bottger, P. C. (1984). Expertise and air time as basis of actual and perceived influence in problem solving groups. *Journal of Applied Psychology*, 69, 214–221.
<https://doi.org/10.1037/0021-9010.69.2.214>
- Brescoll, V. L. (2011). Who takes the floor and why: Gender, power, and volubility in organizations. *Administrative Science Quarterly*, 56(4), 622–641.
<https://doi.org/10.1177/0001839212439994>
- Buser, T., Grimalda, G., Putterman, L., & Van Der Weele, J. (2020). Overconfidence and gender gaps in redistributive preferences: Cross-Country experimental evidence. *Journal of Economic Behavior & Organization*, 178, 267–286.
<https://doi.org/10.1016/j.jebo.2020.07.005>
- Cameron, C. A., & Miller, D. L. (2015). A practitioner’s guide to cluster-robust inference. *Journal of Human Resources*, 50(2), 317–372. <https://doi.org/10.3368/jhr.50.2.317>
- Carter, A. J., Croft, A., Lukas, D., & Sandstrom, G. M. (2019). Correction: Women’s visibility in academic seminars: Women ask fewer questions than men. *PLOS ONE*, 14(2), Article e0212146. <https://doi.org/10.1371/journal.pone.0212146>
- Cassar, A., & Rigdon, M. L. (2021a). Prosocial option increases women’s entry into competition. *Proceedings of the National Academy of Sciences*, 118(45), Article e2111943118.
<https://doi.org/10.1073/pnas.2111943118>

- Cassar, A., & Rigdon, M. L. (2021b). Option to cooperate increases women's competitiveness and closes the gender gap. *Evolution and Human Behavior*, 42(6), 556–572.
<https://doi.org/10.1016/j.evolhumbehav.2021.06.001>
- Charlesworth, T. E. S., & Banaji, M. R. (2019). Gender in science, technology, engineering, and mathematics: Issues, causes, solutions. *The Journal of Neuroscience*, 39(37), 7228–7243.
<https://doi.org/10.1523/JNEUROSCI.0475-18.2019>
- Charness, G., & Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization*, 83(1), 50–58.
<https://doi.org/10.1016/j.jebo.2011.06.007>
- Cheng, J. T., Tracy, J. L., Foulsham, T., Kingstone, A., & Henrich, J. (2013). Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology*, 104(1), 103–125.
<https://doi.org/10.1037/a0030398>
- Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2), 448–474. <https://doi.org/10.1257/jel.47.2.448>
- Cvencek, D., Meltzoff, A. N., & Greenwald, A. G. (2011). Math–gender stereotypes in elementary school children. *Child Development*, 82(3), 766–779.
<https://doi.org/10.1111/j.1467-8624.2010.01529.x>
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3), 522–550.
<https://doi.org/10.1111/j.1542-4774.2011.01015.x>

- Eagly, A. H. (1987). *Sex Differences in Social Behavior: A social-role interpretation*. Lawrence Erlbaum
- Eagly, A. H., & Johnson, B. T. (1990). Gender and leadership style: A meta-analysis. *Psychological Bulletin*, 108(2), 233–256. <https://doi.org/10.1037/0033-2909.108.2.233>
- Eagly, A. H., & Karau, S. J. (1991). Gender and the emergence of leaders: A meta-analysis. *Journal of Personality and Social Psychology*, 60(5), 685–710. <https://doi.org/10.1037/0022-3514.60.5.685>
- Eagly, A. H., & Johannesen-Schmidt, M. C. (2001). The leadership styles of women and men. *Journal of Social Issues*, 57(4), 781–797. <https://doi.org/10.1111/0022-4537.00241>
- Eagly, A. H. (2009). The his and hers of prosocial behavior: An examination of the social psychology of gender. *American Psychologist*, 64(8), 644–658. <https://doi.org/10.1037/0003-066X.64.8.644>
- Eagly, A. H., & Wood, W. (2012). Social role theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 458–476). Sage Publications Ltd. <https://doi.org/10.4135/9781446249222.n49>
- Elprana, G., Felfe, J., Stiehl, S., & Gatzka, M. (2015). Exploring the sex difference in affective motivation to lead: Furthering the understanding of women's underrepresentation in leadership positions. *Journal of Personnel Psychology*, 14(3), 142–152. <https://doi.org/10.1027/1866-5888/a000137>
- Flory, J. A., Leibbrandt, A., & List, J. A. (2015). Do competitive workplaces deter female workers? A large-scale natural field experiment on job entry decisions. *The Review of Economic Studies*, 82(1), 122–155. <https://doi.org/10.1093/restud/rdu030>

- Fluit, S., Kunst, J. R., Bierwiazzonek, K., & Von Soest, T. (2023). Self-esteem trajectories over three decades predict opposition to social equality in midlife. *Proceedings of the National Academy of Sciences*, 120(1), Article e2212906120.
<https://doi.org/10.1073/pnas.2212906120>
- Fox, J., & Weisberg, S. (2019). *An R companion to applied regression*. (3rd ed). SAGE Publications. <https://www.john-fox.ca/Companion/>
- Garfield, Z. H., & Hagen, E. H. (2020). Investigating evolutionary models of leadership among recently settled Ethiopian hunter-gatherers. *The Leadership Quarterly*, 31(2), 101290.
<https://doi.org/10.1016/j.leaqua.2019.03.005>
- Gummer, T., Roßmann, J., & Silber, H. (2021). Using instructed response items as attention checks in web surveys: Properties and implementation. *Sociological Methods & Research*, 50(1), 238–264. <https://doi.org/10.1177/0049124118769083>
- Hall, J., & Watson, W. H. (1970). The effects of a normative intervention on group decision-making performance. *Human Relations*, 23, 299–317.
<https://doi.org/10.1177/001872677002300404>
- Hauser, D. J., Ellsworth, P. C., & Gonzalez, R. (2018). Are manipulation checks necessary? *Frontiers in Psychology*, 9(998). <https://doi.org/10.3389/fpsyg.2018.00998>
- Hayes, A. F. (2009). Beyond baron and kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, 76(4), 408–420.
<https://doi.org/10.1080/03637750903310360>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), 61–83.
<https://doi.org/10.1017/S0140525X0999152X>

Ho, A. K., Sidanius, J., Kteily, N., Sheehy-Skeffington, J., Pratto, F., Henkel, K. E., Foels, R., & Stewart, A. L. (2015). The nature of social dominance orientation: Theorizing and measuring preferences for intergroup inequality using the new SDO7 scale. *Journal of Personality and Social Psychology*, 109(6), 1003-1028.

<http://dx.doi.org/10.1037/pspi0000033>

International Social Survey Programme. (2018). ISSP 2019 – Social Inequality V: Source questionnaire.

https://issp.org/wpcontent/uploads/2021/04/ISSP_Social_Inequality_V_source_questionnaire.pdf

Jaffee, S., & Hyde, J. S. (2000). Gender differences in moral orientation: A meta-analysis.

Psychological Bulletin, 126(5), 703–726. <https://doi.org/10.1037/0033-2909.126.5.703>

Jakobsson, N., & Kotsadam, A. (2010). Do attitudes toward gender equality really differ between Norway and Sweden? *Journal of European Social Policy*, 20(2), 142–159.

<https://doi.org/10.1177/0958928709358790>

Jarvis, S. N., Ebersole, C. R., Nguyen, C. Q., Zhu, M., & Kray, L. J. (2022). Stepping up to the mic: Gender gaps in participation in live question-and-answer sessions at academic conferences. *Psychological Science*, 33(11), 1882–1893.

<https://doi.org/10.1177/09567976221094036>

Jiao, J., & Zhao, J. (2023). Individualism, collectivism, and allocation behavior: Evidence from the ultimatum game and dictator game. *Behavioral Sciences*, 13(2), 169.

<https://doi.org/10.3390/bs13020169>

Kassambara, A. & Patil, I. (2023). *ggcorrplot: Visualization of a correlation matrix using 'ggplot2.'* R package version 01. 4. 1. [10.32614/CRAN.package.ggcorrplot](https://doi.org/10.32614/CRAN.package.ggcorrplot)

- Kenny, D. A., & Judd, C. M. (2014). Power anomalies in testing mediation. *Psychological Science*, 25(2), 334–339. <https://doi.org/10.1177/0956797613502676>
- Kim, H. S., Sherman, D. K., Ko, D., & Taylor, S. E. (2006). Pursuit of comfort and pursuit of harmony: Culture, relationships, and social support seeking. *Personality and Social Psychology Bulletin*, 32(12), 1595–1607. <https://doi.org/10.1177/0146167206291991>
- Leung, K., & Bond, M. H. (1984). The impact of cultural collectivism on reward allocation. *Journal of Personality and Social Psychology*, 47(4), 793–804. <https://doi.org/10.1037/0022-3514.47.4.793>
- Loignon, A. C., Bergeron, D. M., Johnson, M. A., Dunn, A. M., Li, C. S., Martinez, H., ... & Woznyj, H. M. (2025). Signaling with babble? Exploring the effects of gender and speaking time on leader emergence. *The Leadership Quarterly*, 36(4), Article 101879. <https://doi.org/10.1016/j.leaqua.2025.101879>
- MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, 39(1), 99–128. https://doi.org/10.1207/s15327906mbr3901_4
- MacLaren, N. G., Yammarino, F. J., Dionne, S. D., Sayama, H., Mumford, M. D., Connelly, S., Martin, R. W., Mulhearn, T. J., Todd, E. M., Kulkarni, A., Cao, Y., & Ruark, G. A. (2020). Testing the babble hypothesis: Speaking time predicts leader emergence in small groups. *The Leadership Quarterly*, 31(5), Article 101409. <https://doi.org/10.1016/j.leaqua.2020.101409>
- McMichael, S. L., West, S. G., & Kwan, V. S. Y. (2026). The enduring gender gap in stem: A meta-analysis of gender differences in self-efficacy in stem fields. *Behavioral Sciences*, 16(1), 141. <https://doi.org/10.3390/bs16010141>

- McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods*, 22(1), 114–140.
<https://doi.org/10.1037/met0000078>
- Miller, J. M., & Krosnick, J. A. (1998). The impact of candidate name order on election outcomes. *The Public Opinion Quarterly*, 62(3), 291–330.
<https://www.jstor.org/stable/2749662>
- Mullen, B., Salas, E., & Driskell, J. E. (1989). Salience, motivation, and artifact as contributions to the relation between participation rate and leadership. *Journal of Experimental Social Psychology*, 25(6), 545–559. [https://doi.org/10.1016/0022-1031\(89\)90005-X](https://doi.org/10.1016/0022-1031(89)90005-X)
- Navarro, D. (2015). Learning statistics with R: A tutorial for psychology students and other beginners [R package version 0.5.1]. <https://learningstatisticswithr.com>
- Netchaeva, E., Sheppard, L. D., & Balushkina, T. (2022). A meta-analytic review of the gender difference in leadership aspirations. *Journal of Vocational Behavior*, 137, Article 103744.
<https://doi.org/10.1016/j.jvb.2022.103744>
- Neuwirth, E. (2022). *RColorBrewer: ColorBrewer palettes* [R package version 1.1-3].
<https://CRAN.R-project.org/package=RColorBrewer>
- Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics*, 122(3), 1067-1101.
<https://doi.org/10.1162/qjec.122.3.1067>
- Nikoleizig, L., Schmukle, S. C., Griebenow, M., & Krause, S. (2021). Investigating contributors to performance evaluations in small groups: Task competence, speaking time, physical expressiveness, and likability. *PLoS ONE*, 16(6), Article 0252980.
<https://doi.org/10.1371/journal.pone.0252980>

- Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, 45(4), 867–872. <https://doi.org/10.1016/j.jesp.2009.03.009>
- O'Rourke, H. P., & MacKinnon, D. P. (2018). Reasons for testing mediation in the absence of an intervention effect: A research imperative in prevention and intervention research. *Journal of Studies on Alcohol and Drugs*, 79(2), 171–181. <https://doi.org/10.15288/jsad.2018.79.171>
- Pfrombeck, J., Levin, C., Rucker, D. D., & Galinsky, A. D. (2022). The hierarchy of voice framework: The dynamic relationship between employee voice and social hierarchy. *Research in Organizational Behavior*, 42, 100179. <https://doi.org/10.1016/j.riob.2022.100179>
- Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, 67(4), 741–763. <https://doi.org/10.1037/0022-3514.67.4.741>
- Pratto, F., Çidam, A., Stewart, A. L., Zeineddine, F. B., Aranda, M., Aiello, A., Chryssochoou, X., Cichocka, A., Cohrs, J. C., Durrheim, K., Eicher, V., Foels, R., Górska, P., Lee, I.-C., Licata, L., Liu, J. H., Li, L., Meyer, I., Morselli, D., ... Henkel, K. E. (2013). Social dominance in context and in individuals: Contextual moderation of robust effects of social dominance orientation in 15 languages and 20 countries. *Social Psychological and Personality Science*, 4(5), 587–599. <https://doi.org/10.1177/1948550612473663>
- Raven, J., Raven, J. C., & John Hugh Court. (1998). *Manual for Raven's progressive matrices and vocabulary scales*. Oxford Psychologists Press.

- Revelle, W. (2025). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University, Evanston, Illinois. R package version 2. 5. 6.
<https://CRAN.R-project.org/package=psych>
- Samuelson, P. A. (1948). Consumption theory in terms of revealed preference. *Economica*, 15(60), 243-253.
- Sanchez-Cortes, D., Aran, O., Mast, M. S., & Gatica-Perez, D. (2010). Identifying emergent leadership in small groups using nonverbal communicative cues. Proceedings of the 12th ACM International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction, 1–4. <https://doi.org/10.1145/1891903.1891953>
- Saccardo, S., Pietrasz, A., & Gneezy, U. (2018). On the size of the gender difference in competitiveness. *Management Science*, 64(4), 1541–1554.
<https://doi.org/10.1287/mnsc.2016.2673>
- Shamon, H., & Berning, C. C. (2020). Attention check items and instructions in online surveys: Boon or bane for data quality? *Survey Research Methods*, 14(1), 55-77.
<https://doi.org/10.18148/SRM/2020.V14I1.7374>
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, 7(4), 422–445.
<https://doi.org/10.1037/1082-989X.7.4.422>
- Schuh, S. C., Hernandez Bark, A. S., Van Quaquebeke, N., Hossiep, R., Frieg, P., & Van Dick, R. (2014). Gender differences in leadership role occupancy: The mediating role of power motivation. *Journal of Business Ethics*, 120(3), 363–379.
<https://doi.org/10.1007/s10551-013-1663-9>

- Soll, J. B., & Klayman, J. (2004). Overconfidence in Interval Estimates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(2), 299–314.
<https://doi.org/10.1037/0278-7393.30.2.299>
- Tingley D, Yamamoto T, Hirose K, Keele L, Imai K (2014). “mediation: R Package for Causal Mediation Analysis.” *Journal of Statistical Software*, 59(5), 1–38.
<https://www.jstatsoft.org/v59/i05/>
- Vanutelli, M., Meroni, F., Fronda, G., Balconi, M., & Lucchiari, C. (2020). Gender differences and unfairness processing during economic and moral decision-making: A fNIRS study. *Brain Sciences*, 10(9), 647. <https://doi.org/10.3390/brainsci10090647>
- Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (4th ed), Springer.
<https://www.stats.ox.ac.uk/pub/MASS4/>
- von Rueden, C., Alami, S., Kaplan, H., & Gurven, M. (2018). Sex differences in political leadership in an egalitarian society. *Evolution and Human Behavior*, 39(4), 402–411.
<https://doi.org/10.1016/j.evolhumbehav.2018.03.005>
- Welch, B. L. (1947) The generalizations of “Student’s” problem when several different population variances are involved. *Biometrika*, 34(1-2), 28-35.
<https://doi.org/10.1093/biomet/34.1-2.28>
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
<https://ggplot2.tidyverse.org>
- Wickham, H., François, R., Henry, L., & Müller, K. (2019a). *dplyr: A grammar of data manipulation*. <https://CRAN.R-project.org/package=dplyr>
- Wickham, H., & Bryan, J. (2019b). *readxl: Read excel files*.
<https://CRAN.R-project.org/package=readxl>

- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., ... Yutani, H. (2019c). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686.
<https://doi.org/10.21105/joss.01686>
- Wickham, H., Pedersen, T., & Seidel, D. (2025). *scales: Scale functions for visualization* [R package version 1.4.0]. <https://scales.r-lib.org>
- Wickham, H., Vaughan, D., Girlich, M. (2026). *tidyr: Tidy messy data* [R package version 1.3.2]. <https://tidyr.tidyverse.org>
- Williams, M. J., & Tiedens, L. Z. (2016). The subtle suspension of backlash: A meta-analysis of penalties for women's implicit and explicit dominance behavior. *Psychological Bulletin*, 142(2), 165–197. <https://doi.org/10.1037/bul0000039>
- Wiwad, D., Mercier B., Maraun, M. D., Robinson A. R., Piff, P.K., Aknin, L. B., & Shariff, S. P. (2019) The support for economic inequality scale: Development and adjudication. *PLoS ONE*, 14(6), Article e0218685. <https://doi.org/10.1371/journal.pone.0218685>
- World Values Survey Association. (2022). *World Values Survey Wave 7 (2017-2022)*.
<https://www.worldvaluessurvey.org/WVSDocumentationWV7.jsp>
- Yetton, P. W., & Bottger, P. C. (1982). Individual versus group problem solving: An empirical test of a best-member strategy. *Organizational Behavior and Human Performance*, 29(3), 307–321. [https://doi.org/10.1016/0030-5073\(82\)90248-3](https://doi.org/10.1016/0030-5073(82)90248-3)
- Zeileis, A. (2014). *ineq: Measuring inequality, concentration, and poverty*. <https://CRAN.R-project.org/package=ineq>

Zeileis, A., & Hothorn T (2002). Diagnostic checking in regression relationships. *R News*, 2(3), 7–10. <https://CRAN.R-project.org/doc/Rnews/>

Zeileis, A., Köll, S., & Graham, N. (2020). Various versatile variances: An object-oriented implementation of clustered covariances in R. *Journal of Statistical Software*, 95(1). <https://doi.org/10.18637/jss.v095.i01>