Making Contribution the Default: Gender, Merit, and Idea Sharing in Groups

Jingnan (Cecilia) Chen* Shan Gui† Daniel Houser‡ Erte Xiao§

This version: 31 July 2026

Latest version

Abstract

In many workplaces, qualified women are less willing than their male counterparts to contribute ideas to the group. As a result, women remain under-represented in idea contribution, particularly in male-stereotyped fields. Using an online experiment with tasks that vary in gender stereotype, we test whether an easily implementable default can reduce this gender gap. We consider two types of defaults: (i) random defaults, where an individual is randomly defaulted to the leading position to contribute, from which they can opt out; and (ii) merit-based assignment, where an individual's default position is determined by relative performance on a previous task. We find that both random and merit-based defaults substantially increase people's willingness to contribute ideas. While women subjected to defaults are more likely to contribute ideas across all areas, they remain less likely to contribute in male-typed than female-typed environments. Finally, we find both merit-based and random defaults have a similar positive impact on group performance. The implication is that organizations using random defaults can avoid the complexity of identifying merit ex ante while still enjoying the benefits of increased participation.

Keywords: Defaults, Gender, Merit, Stereotypes, Idea contribution, Groups.

*Jingnan Chen: j.chen2@exeter.ac.uk, University of Exeter.

†Shan Gui: guishan@shufe.edu.cn, School of Economics; Key Laboratory of Mathematical Economics; Shanghai University of Finance and Economics.

‡Daniel Houser: dhouser@gmu.edu, George Mason University.

§Erte Xiao: Erte.Xiao@monash.edu, Monash University. We thank Cesar Martinelli, Kevin McCabe, Johanna Mollerstrom, Thomas Stratmann, and participants at Central South University Seminar 2023, APEE International Conference 2024, and Shanghai University of Finance and Economics Brownbag 2024, for helpful comments. The authors acknowledge the funding support of Australian Research Council Discovery Project (DP240101021), the National Natural Science Foundation of China (No. 72503130), and the Interdisciplinary Center for Economic Science (ICES) at George Mason University. Declarations of interest: none.

1 Introduction

Women are generally less willing than men to share their ideas in group settings, particularly in male-stereotyped fields and tasks (see, e.g., Babcock and Laschever, 2007; Cooper and Kagel, 2016; Coffman, 2014; Chen and Houser, 2019). Under-contributing imposes substantial costs at both individual and organizational levels. It limits group performance by preventing the full utilization of available human capital. It can also constrain women's career advancement by reducing their visibility and perceived competence, and thereby contributes to persistent gender gaps in pay and leadership representation. Despite growing recognition of this problem, existing interventions have shown limited effectiveness in encouraging women's participation in contributing ideas (Coffman, 2014). The persistence of gender participation gaps highlights the continuing need for more effective and feasible to implement policy solutions.

This paper investigates whether changing default contribution mechanisms can address women's under-participation in idea sharing. We build on emerging evidence that switching from opt-in to opt-out defaults can successfully encourage women's participation in leadership and competitive settings. This approach leverages behavioral nudges to achieve participation goals that have proven difficult through traditional interventions like debiasing or institutional reform. The underlying insight is that current defaults in our institutions often implicitly discourage women from participating by requiring active opt-in decisions. Erkal et al. (2022) demonstrate that switching from opt-in to opt-out selection processes significantly reduces gender leadership gaps. He et al. (2021) show similar effects for competitive participation. These findings suggest that defaults may represent a low-cost, scalable intervention for addressing gender participation gaps across various organizational contexts.

We extend this literature by exploring whether default mechanisms can also promote gender equity in idea contribution. In particular, we examine two central questions: (i) Does a default increase women's likelihood of contributing ideas in a group? and (ii) How does merit-based assignment interact with default settings? While encouraging broader participation is desirable, doing so indiscriminately could dilute the quality of contributions if less qualified individuals are nudged to speak up. One possible solution is to implement a merit-based default, where only individuals with high demonstrated competence are defaulted to contribute. This approach could enhance both the effectiveness and perceived legitimacy of group decision-making. However, merit-based assignment comes with practical challenges: organizations must define and identify "merit" ex ante, a process that can be costly, subjective, and exclusionary. Moreover, pre-selecting qualified individuals may not be efficient if many decline to participate. This paper aims to provide descriptive evidence on the impact of merit by comparing merit-based and random default mechanisms, and thus offers practical insights into whether merit-based assignment delivers sufficient benefits to

justify its implementation costs.

By extending the default mechanism to the context of idea contribution, our study examines whether making idea contribution the default can reduce the influence of gender stereotypes on participation. Prior research suggests that the decision to contribute is shaped not only by competence, but also by whether the domain is perceived as gender-congruent (Coffman, 2014). In particular, individuals may be less willing to speak up in settings stereotypically associated with the opposite gender, with women often showing greater sensitivity to this effect. This gender stereotype effect may partly reflect social norms regarding when contribution is considered appropriate. If defaults affect behavior by altering perceptions of what is expected or socially endorsed (McKenzie et al., 2006; Davidai et al., 2012; Everett et al., 2015), then requiring individuals to contribute by default may encourage participation even in domains where gender stereotypes would otherwise discourage it.

We test this possibility in a controlled experiment and find results consistent with prior evidence that women are less willing to contribute ideas in the absence of a default. Introducing defaults significantly increases participants' overall willingness to contribute and eliminates the overall gender gap in choosing the leading position. However, we find little evidence that defaults reduce the influence of gender stereotypes on women's contribution decisions. Women become more likely to contribute across all areas, but they remain less likely to contribute in male-typed than in female-typed contexts. We also find no evidence that merit-based assignment significantly strengthens the default effect relative to random assignment. These findings suggest that, if the primary goal is to promote participation, random defaults can be as effective as merit-based assignment, while avoiding the complexity of identifying merit ex ante.

The remainder of this paper is organized as follows: Section 2 reviews the related literature; Section 3 outlines the experimental design and procedure; Section 4 discusses the experimental results; and Section 5 concludes the paper.

2 Literature Review

Many initiatives have been carried out to promote gender diversity in the workplace. Trainings are designed to help women improve skills and build confidence. Women are advised to act more like a man (e.g., lean in) to make gains. The success of these programs, however, is limited and can sometimes backfire (Bohnet, 2016). In addition, interventions such as affirmative action and quotas have been used to force gender equity. Such mechanisms face criticisms and concerns regarding potential counterproductive consequences of their use (Matsa and Miller, 2013; Gangadharan et al., 2016; Afridi et al., 2017).

Our study makes a distinctive contribution to three strands of literature. First, we add to the

literature on gender stereotypes and the willingness to contribute ideas (Coffman, 2014; Chen and Houser, 2019; Bordalo et al., 2016). This literature observes that women (man) are underconfident about their performance in male-typed (female-typed) domains and consequently less willing to contribute their ideas to the group in gender incongruent domains. These findings are also consistent with the prediction from Bordalo et al. (2019) on women's shying away from male-type domains, and the findings from Benjamin et al. (2010) that argue norms related to social identity affect economic decisions. Little success has been documented in the area of mitigating the gender gap in the willingness to contribute ideas. We contribute to this literature by showing that default can potentially alleviate the gender gap in the willingness to contribute the ideas (although it does not change the gender stereotype effect).

Second, we extend research on the power of default to a new context - group decision-making. Previous studies have investigated the power of default in many domains, such as organ donation (Johnson and Goldstein, 2003), retirement saving programs (Madrian and Shea, 2001), donation (Urminsky, 2016), reducing gender gap in voluntary leadership, and willingness to compete (Erkal et al., 2022; He et al., 2021). Jachimowicz et al. (2019) conduct a meta-analysis summarizing 55 studies on default effects from 35 articles. They show that on average, the default option leads to 27.24 % more people staying with the pre-selected default. They highlight, however, that how effective defaults are and why defaults can take effects vary in environments. Dinner et al. (2011) propose three psychological channels underlying the default effect. One is called the endorsement channel. That is, individuals are more likely to stay with defaults if they believe the default option is motivate by good intentions (Tannenbaum et al., 2017). Another channel through endowment effect (Kahneman and Tversky, 1979). The third channel is through costly deviation from default options (Johnson et al., 2012). By investigating the relevance of merit in setting up the default, our study sheds light on whether the legitimacy of the default option matters.

Third, we contribute to the literature on perceptions toward merit and luck. Research has shown that people react differently to merit versus luck-based outcomes. For example, people are more likely to cheat when the payoffs are luck-based versus performance-based (Kajackaite, 2018; Gravert, 2013). Meanwhile, the willingness to redistribute is stronger when inequalities are not merit-based (Konow, 2002; Cappelen et al., 2007). In addition, when bargaining over legitimized wealth (effort-based), dictators act more selfishly than they do with unearned wealth (Cherry et al., 2002). Reactions toward merit-based mechanisms and luck-based mechanisms are also rooted in culture: Almås et al. (2020) find that Americans are more meritocratic, while Scandinavians favor egalitarianism. We extend this line of study and investigate the potential reactions to different default mechanisms. In contrast to previous findings, our results indicate that the power of default seems to be overwhelming in that individuals do not react differently to default positions that are gained by luck, as compared to those that are founded in merit. The findings further suggest that

when it is costly to identify "merit" ex ante, non-merit-based defaults can be just as effective in encouraging individuals to participate.

3 Experimental Design and Procedure

3.1 Experimental design

Our experiment draws from previous research on willingness to contribute ideas when subjects face tasks with varying gender stereotypes (Bordalo et al., 2016; Chen and Houser, 2019; Coffman, 2014). We also draw on previous research exploring the role of defaults in influencing gender gaps (Erkal et al., 2022).

Subjects were randomly assigned to one of four treatments: Merit-Default (MD), Random Default with No Feedback (RD_NF), Random Default with Feedback (RD_F), and Control (i.e., no default and no feedback). All treatments followed a similar structure as detailed below.

The experiment was comprised of three incentivized parts (Part A, B, and C) and a post-experimental survey. Participants received a three-dollar completion fee. All participants received general instructions informing them that one part of the experiment would be selected for experimental payment and announced at the conclusion of the experiment. Points were used in the experiment as the currency with the exchange rate of 1 point = 0.05 dollar. Participants faced multiple-choice questions (MCQ) from six categories: Arts and Literature (Art), cars (Car), Disney movies (Disney), the Kardashians (Kard), Sports and Games (Sport), and Video Games (Video). Each question included five possible answers and was labeled with its corresponding category. Those six categories varied in their perceived gender stereotypeness as revealed in the answers to the post-experimental survey questions. With the exception of MD and RD_F, no feedback was provided during the experiment.

Part A: individual task Participants answered 30 MCQ (five from each category) on their own (see Figure A1 for an example question). The data from this part of the experiment provided us with a baseline measurement of individual ability for each category. Subjects received four points for each correct answer.

Part B: elicitation of willingness to contribute answers to the group As the subjects proceeded to Part B, they were informed that they would be working with one randomly selected participant in the study as a group for this part. Each group was given new 30 MCQ. Each participant's payment in this part depended on the submitted group answers. Each group member received four points for each correct answer and lost one point for each incorrect answer.

To decide the group answers, participants made two decisions for each of the *new* 30 MCQ: (i) their answer to the question; and (ii) their position in line to submit their answer as the group answer. There were four positions in line (1, 2, 3 and 4). In each pair, the participant who selected the lowest number would have their answer submitted as the group's answer. If both members selected the same position in line, the computer randomly selected one member's answer as the group answer. Thus, the *lower* the position in line, the *more* willing the subject was to contribute their answers to the group. Our four treatments vary on whether any positions in line were pre-selected. Figure A2 provides an example of the decision screen in the Control.

Before subjects made the two above decisions, they were asked to make incentivized guesses about their own rank within the newly formed group for each of the categories from Part A. They receive an additional point for each correct guess. We used the guess as a measure of belief.

Part C: Belief elicitation In Part C, we measured participants' confidence in their own answers and the answers of their group members, for each question answered in Part B. A simplified Becker-DeGroot-Marschak (BDM) method is used in Part C. Participants were asked to state the probabilities that their own and their group member's answer were correct. The BDM method ensured that participants truthfully reported their beliefs.[1] We use these belief measures to examine whether defaults affect willingness to contribute through changes in confidence. The analysis, reported in Appendix D., provides little evidence for this mechanism.

Post-experimental questionnaire We asked our subjects for demographic and attitudinal information, including, e.g., gender, age, where they attended high school, and which question categories they liked or disliked. As the last question in the questionnaire, we asked the subjects to evaluate the male- or female-typeness of each category. For each of the categories, the subjects were asked to indicate their answers by selecting a number from -5 to 5, where -5 was labeled as "women know more," and 5 was labeled "men know more". We use a normalized value – maleness score - to represent the maleness of each category (for details see Section 4).

3.2 Treatments

Treatments varied only in Part B, where we implemented a between-subject design with two factors: feedback (with or without) and default implementation (with default or no default). This resulted in four treatments: Control, Merit-Default (MD), Random Default with Feedback (RD_F), and Random Default with No Feedback (RD_NF). Table 1 summarizes the features of all treatments.

[1]For implementation details, please refer to Figure A3.

Table 1: Summary of Treatment Conditions

Treatment	Part A Feedback	Part B Default Rule
Control	No	No default
Merit Default (MD)	Yes	Default based on past performance in the corresponding category
Random Default with No Feedback (RD_NF)	No	Default position randomly selected
Random Default with Feedback (RD_F)	Yes	Default position randomly selected

Control (no default + no feedback) There were no pre-selected positions for Part B, nor was any Part A performance feedback provided to the subjects in Control.

Merit Default (MD) Before the start of the Part B questions, subjects received Part A performance feedback information as compared to their paired group member (see Figure A4). They were also informed that Position "**1**" was pre-selected for all the questions in the categories that they were at least as good as their group member and "**4**" was preselected for all questions in the categories that they were worse than their group member. However, the subjects were free to choose other positions in line (see Figure A5). While our focus is the case when subjects are defaulted to Position 1, we include in the design default to Position 4. We think this feature can help reduce experimenter demand effect, which is probably more likely when we only allow default to Position 1. Our design also allows us to learn the effect of opt-out when one is defaulted to a "passive position" (i.e., not to contribute).

Random Default with No Feedback (RD_NF) Contrary to the MD treatment, participants were advised that Position "**1**" was pre-selected for all the questions in some *randomly selected* categories and Position "**4**" was pre-selected for all questions in the remaining categories. Moreover, the subjects did NOT receive any Part A performance feedback. Thus, this treatment informs the pure default effect.

Random Default with Feedback (RD_F) The only difference between this treatment and RD_NF is that subjects received Part A performance feedback before they started on the questions in Part B.

3.3 Experiment procedures

Our study was pre-registered on OSF registries (osf.io/gbqv5). The experiment was programmed in oTree (Chen et al., 2016) and conducted online using Prolific (`https://www.prolific.co`)

between July and August 2023.[2] [3] Only those who reside in the United States were eligible to participate, given the familiarity with the question categories (e.g., questions about the Kardashians).

To minimize cheating in the experiment, a 15-second time limit was imposed on each of the MCQ questions in Part A and Part B. No time limit was imposed on the position in line decisions in Part B. Attention check questions were included in both Part A and Part B.

A total of 804 participants attended our experiment, with 200 in Control, 200 in MD, 202 in RD_F, and 200 in RD_NF. Ages ranged from 18 to 77 (Mean = 37.82, Median = 35); 47.89% female; 69.9% White, 7.46% Hispanic or Latino and 9.20% Black or African American; the highest attained degree ranged from "Some high school" to "Doctorate degree" (with the median being an "Associated degree"). Each subject participated in exactly one treatment. A typical session lasted about 20 mins. The average payment was $6.11.[4]

Due to the nature of the experiment and the possibility of significant attrition in data collection, we chose not to live match subjects into groups.[5] Instead, we sequentially conducted two waves of data collection per treatment condition. After data collection in Wave 1, we recruited participants in Wave 2. We matched randomly the subjects in Wave 2 with those (or their decisions) in Wave 1 (one-to-one matching). To ensure common information, the same instructions were used for subjects in both waves for the same treatment condition.

In Control, subjects in both waves made the exact same decisions. In the MD, RD_NF, and RD_F treatments, subjects in both waves were told that one person would have a pre-selected position when answering questions in Part B. Subjects in Wave 1 were informed that no position in line was pre-selected for them in Part B and that position in line was pre-selected for their matched partner. The explanations as to how the pre-selected position was determined vary among the three default treatments as described above.[6] In the MD and RD_F treatments, the Part A performance of each participant in Wave 2 was compared with that of the matched participant in Wave 1 and the feedback was provided to the Wave 2 participant at the beginning of Part B. In the MD treatment,

[2]This 2023 experiment followed a similar earlier experiment in 2019, which needed to be discarded due to a technical glitch in how the data were recorded. The delay in detecting the problem and re-running the experiment was due to one of the authors' job transition.

[3]The Prolific platform is widely used in social sciences in recent years and has been compared favorably against Amazon Mechanical Turk (see, e.g., Peer et al., 2017). We choose the balanced-gender sample for each wave and each treatment in Prolific.

[4]The Prolific platform enforces an average hourly payment of $8 per hour. Our hourly payment is 1.5 times than the suggested amount.

[5]The subjects on the Prolific platform may join a study any time while the study is active. It can be the case that one subject starts the experiment and the next subject does not show up until 5 – 10 mins later. If we opt for live matching, the first subjects may have to wait up to 20 mins to be matched. The prolonged waiting time will inevitably create a significant drop out.

[6]To make the instructions simple and easy to follow, we did not explain the logistic details of the procedure of the matching and default assignment. Subjects were only told that one person in their pair will be randomly assigned to a pre-selected position, which is truthful description given subjects were randomly assigned to different treatments and thus has a random chance to be assigned to a pre-selected position.

the Wave 2 participant's pre-selected position (1 or 4) was determined according to their relative performance in Part A compared to the matched Wave 1 participant. The participants in Wave 2 in RD_NF received no feedback. but were randomly defaulted to either Position 1 or 4. After Wave 2 participants finished the experiments, participants in both waves received their payments.

3.4 Hypotheses

To investigate how default options affect gender differences in willingness to contribute ideas, we distinguish between two components of contribution behavior: *baseline willingness to contribute* and *stereotype sensitivity*. Baseline willingness to contribute refers to the overall tendency to choose Position 1, holding constant the gender-stereotyped nature of the category, which we measure using the maleness score. Stereotype sensitivity refers to the extent to which willingness to contribute changes as the category becomes more male-typed or female-typed. In other words, stereotype sensitivity captures whether participants become less willing to contribute as the category becomes less gender-congruent. Gender differences in contribution may arise because men and women differ in their baseline willingness to contribute, their stereotype sensitivity, or both. In turn, the overall gender gap may reflect a baseline gender difference in willingness to contribute, a gender difference in stereotype sensitivity, or the combination of the two. The pattern documented in the literature (e.g., Coffman, 2014; Chen and Houser, 2019) is most consistent with a case in which both components contribute to the observed gender gap. Accordingly, we hypothesize that a default can reduce the overall gender gap in two ways. First, it may increase baseline willingness to contribute, especially among women. Second, it may reduce stereotype sensitivity, so that willingness to contribute depends less on whether the category is male- or female-typed.

Because our treatments vary in their default mechanisms, we first isolate the pure default effect by comparing RD_NF with the Control, where the only difference is the presence of a default. We then examine the merit effect by comparing MD with RD_NF and RD_F treatment.

Our hypothesis on the pure default effect build on prior evidence that individuals tend to remain with pre-selected positions. That is, when Position 1 is pre-selected, we expect more people will choose 1, regardless of gender congruent categories. This corresponds to a positive default effect on baseline willingness to contribute.[7] Prior literature, however, does not inform whether defaults also impact stereotype sensitivity - that is, whether defaults weaken the relationship between category stereotype and willingness to choose the leading position. If defaults reduce stereotype sensitivity, then willingness to choose Position 1 should vary less with category stereotype under default

[7]The gender difference in baseline willingness to contribute (and, consequently, the overall gender gap) may be reduced if the default has a larger positive effect on women's baseline willingness to contribute than on men's (He et al., 2021). Even if the default has a similar positive effect on both women and men, the gender difference in baseline willingness to contribute may still become statistically insignificant if the increase is sufficiently large relative to the original difference.

conditions.

Hypothesis 1a *(Default effect on baseline willingness to contribute):* Holding maleness of the questions constant, participants in the RD_NF are more likely to select Position 1 when that position is pre-selected for them than are participants in the Control.

Hypothesis 1b *(Default effect on stereotype sensitivity):* The negative (positive) effect of the maleness of the questions on women's (man's) willingness to choose leading position is weaker when subjects in RD_NF are defaulted to Position 1 than when subjects are in the Control.

These two default effects may independently or jointly reduce the gender gap observed in the Control. Our experiment is designed to identify which of these channels accounts for any effect of defaults on the gender gap.

Next, we consider the effect of merit-based default. We expect any merit-based default effect to operate through the channels specified in Hypotheses 1. Furthermore, mechanically, assigning only the high-ability member to default Position 1 should improve group performance under the merit-default relative to the random default. Our next hypotheses concern the potential behavioral effect of the merit-based defaults: individuals may be more likely to remain with the default when the default position is assigned on the basis of merit rather than randomly. There are two possible reasons for this.

First, being assigned to Position 1 under MD indicates that the subject performed better than the other group member. We refer to this as the information effect. Previous literature suggests that feedback can affect willingness to contribute. For example, Chen and Houser (2019) show that public feedback increases female leaders' willingness to contribute. In our setting, feedback indicating that one performed better than one's partner may likewise increase the likelihood of remaining with the default leading position. If so, high performers should be more likely to contribute in both MD and RD_F than in RD_NF.

Second, the assignment of Position 1 may be viewed as more legitimate under MD. We refer to this as the legitimacy effect. Previous research has shown that people are more likely to comply with a request when it is legitimate (see, e.g., Jackson et al., 2012). In our setting, Position 1 may carry greater legitimacy under MD because participants know that it is assigned on the basis of relative performance in the relevant category. Under RD_F, by contrast, participants also learn whether they performed better than their partner, but the assignment of positions is random. Thus, if legitimacy strengthens compliance with the default, high performers should be more likely to remain with pre-selected Position 1 in MD than in RD_F.

Hypothesis 2a *(Information effect of the merit default):* For high performers in Part A who are defaulted to Position 1 in Part B, the likelihood of remaining with the pre-selected position is higher in MD and RD_F than in RD_NF.

Hypothesis 2b *(Legitimacy effect of the merit default):* For high performers in Part A who are

defaulted to Position 1 in Part B, the likelihood of remaining with the pre-selected position is higher in MD than in RD_F.

4 Results

4.1 Summary Statistics and Category Stereotypes

We ran eight online sessions in total, with two sessions per treatment, and obtained observations from 804 subjects. Table 2 reports summary statistics for task performance in Parts A and B. Men answered slightly more questions correctly than women on average, but the difference is small and not statistically significant (men vs. women, 12.40 vs. 11.96, $p = 0.08$, two-sided t-test). We also do not observe significant differences in the number of correct answers across treatments ($p = 0.37$, ANOVA).[8]

Table 2: Part A and Part B Performance Summary Statistics

	Part A			Part B			
Treatment	Women	Men	Overall	Women	Men	Overall	Obs.
Control	11.74	12.28	12.01	15.23	15.95	15.59	200
	(0.31)	(0.42)	(0.26)	(0.45)	(0.46)	(0.31)	
MD	11.96	12.65	12.31	16.49	17.14	16.82	200
	(0.32)	(0.31)	(0.22)	(0.39)	(0.40)	(0.28)	
RD_F	11.92	13.04	12.48	16.30	17.65	16.97	204
	(0.36)	(0.36)	(0.25)	(0.40)	(0.39)	(0.28)	
RD_NF	12.21	11.62	11.92	16.11	15.69	15.90	200
	(0.37)	(0.43)	(0.28)	(0.46)	(0.50)	(0.24)	
Overall	11.96	12.40	12.18	16.02	16.61	16.32	804
	(0.17)	(0.19)	(0.13)	(0.17)	(0.19)	(0.15)	

As explained in Section 3.3, in the three default treatments, subjects in Wave 2 were assigned a pre-selected position - either 1 or 4. Subjects in Wave 1 in all three default treatments were not assigned any pre-selected positions.[9] Our analysis focuses on comparing choices under default conditions to those in the Control condition. Therefore, while we include all participants in the Control group, for the default treatments, we only include Wave 2 participants. As a result, the

[8]Given that everyone has multiple observations, we calculate individual averages first. We treat the averages for each individual as independent observations. We do the same for the remaining tests.

[9]Unsurprisingly, their choices of Position 1 and average chosen positions are statistically indistinguishable from those in Control (see Appendix Table B1).

number of subjects used in the main analysis is 200, 100, 102, and 100 in Control, MD, RD_F, and RD_NF, respectively.

We measure the category stereotype of each question category using a standardized maleness score, constructed from responses to the post-experimental questionnaire as described in Section 3.1. Participants were asked to rank each category according to whether, in general, men or women know more about it. They used a sliding $[-5, 5]$ scale, where -5 was labeled "women know more," 0 was labeled "no gender difference," and 5 was labeled "men know more." In Table 3, we present the maleness score of each category, by gender. Men and women generally agree on the femaleness/maleness of the categories. We provide the average perception for each category in the third column; each average is significantly different from 0 ($p < 0.001$). In the last column, we provide the normalized z-score for maleness of the category: we rescale the average perceptions to have mean 0 and standard deviation 1. Going forward, we refer to Kardashians, Disney, and Arts & Literature as female-typed categories, and Video Games, Sports, and Car as male-typed categories.

Table 3: Perceived Maleness of Categories

Category	Avg. maleness given by men	Avg. maleness given by women	Overall average	Normalized maleness score
Kardashians	-3.26	-3.19	-3.23	-1.11
Disney	-1.77	-2.35	-2.06	-0.72
Art & Literature	-1.13	-1.74	-1.44	-0.52
Video Games	2.42	1.70	2.06	0.64
Sports	2.81	2.32	2.57	0.80
Car	2.93	2.83	2.88	0.90
Obs.	402	402	804	804

4.2 Pure Default Effects on Contribution and Stereotype Sensitivity

4.2.1 Descriptive Evidence

We first examine whether the Control treatment replicates prior findings on gender differences in willingness to contribute ideas. We then compare RD_NF with the Control treatment to test the pure default effect (Hypothesis 1a and Hypothesis 1b). Because RD_NF differs from Control only by the presence of a randomly assigned default, this comparison isolates the effect of making contribution the default option. Because our focus is on encouraging idea contribution, we concentrate on cases in which subjects are assigned to Position 1 in the default treatments. The corresponding results for

cases in which subjects are assigned to Position 4 are reported in Appendix C.[10]

In the Control treatment, we replicate the standard gender pattern in willingness to contribute. Women are significantly less likely than men to choose the leading position overall (27.87% vs. 34.54%, two-sided t-test, $p = 0.01$). Willingness to contribute also varies systematically with category stereotype: as categories become more male-typed, women become less willing to choose Position 1, whereas men become more willing to do so.

We first test Hypothesis 1a, which predicts that the default increases baseline willingness to contribute. The descriptive evidence strongly supports this hypothesis. As shown in Table 4, the likelihood of choosing Position 1 increases from 27.87% to 52.99% for women (two-sided t-test, $p < 0.01$) and from 35.54% to 55.54% for men (two-sided t-test, $p < 0.01$). As a result, the overall gender gap in choosing Position 1 is no longer statistically significant in RD_NF (52.99% vs. 55.54%, two-sided t-test, $p = 0.67$). Thus, the random default without feedback increases baseline willingness to contribute for both women and men, and this increase is large enough to eliminate the overall gender gap.

As supporting descriptive evidence, Appendix Table B3 reports the same comparison separately by category. The increase in choosing Position 1 under RD_NF is not concentrated in a single category: for both women and men, the likelihood of choosing Position 1 is higher in RD_NF than in Control in most categories. The gender gap is also reduced in most categories, with the only marginal exception being Video Games (two-sided t-test, $p = 0.06$). These category-level patterns are consistent with the aggregate evidence that the random default without feedback increases baseline willingness to contribute. [11]

We next test Hypothesis 1b, which predicts that the default reduces stereotype sensitivity. As shown in Table 4, the descriptive evidence provides a more mixed pattern. In the Control treatment, women are less likely to choose Position 1 in male-typed categories than in female-typed categories (19.41% vs. 36.38%, two-sided t-test, $p < 0.01$), while men are more likely to choose Position 1 in male-typed categories than in female-typed categories (38.46% vs. 30.66%, two-sided t-test, $p = 0.02$). In RD_NF, women continue to display stereotype sensitivity: they remain less likely to choose Position 1 in male-typed categories than in female-typed categories (44.02% vs. 62.24%, two-sided t-test, $p < 0.01$). By contrast, men no longer exhibit this pattern under the default (54.93% vs. 58.55%, two-sided t-test, $p = 0.61$). These results suggest that the random default increases women's willingness to contribute across categories, but does not remove women's stereotype

[10] In the Control treatment, women are significantly more likely than men to choose Position 4 overall (31.48% vs. 23.49%, two-sided t-test, $p = 0.01$). When participants are defaulted to Position 4 in RD_NF, both men and women are more likely to choose Position 4, and the overall gender gap is no longer statistically significant (women: 55.85% vs. men: 53.77%, two-sided t-test, $p = 0.72$).

[11] Appendix Table B2 reports the corresponding intensive-margin evidence using the average chosen position number. The overall pattern is similar, although category-level changes vary.

Table 4: Likelihood of Choosing Position 1 by Treatment, Gender, and Category Type

Gender	Treatment	Female-Typed Categories	Male-Typed Categories	All Categories	Category Difference p-value
Women	Control	36.38%	19.41%	27.87%	$< 0.01^{***}$
	RD_NF	62.24%	44.02%	52.99%	$< 0.01^{***}$
Men	Control	30.66%	38.46%	35.54%	0.02^{**}
	RD_NF	58.55%	54.93%	55.54%	0.61
Gender gap p-value: Control		0.08	$< 0.01^{***}$	$< 0.01^{***}$	
Gender gap p-value: RD_NF		0.60	0.10	0.67	

Note: The table reports the percentage of decisions in which participants chose Position 1. Female-typed categories are Kardashians, Disney, and Arts & Literature. Male-typed categories are Video Games, Sports, and Car. The category-difference p-value tests the difference between female-typed and male-typed categories within each gender-treatment cell. The gender-gap p-values test the difference between women and men within each treatment-category type cell. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

sensitivity. By contrast, for men, the default promotes both baseline willingness to contribute and reduces their stereotype sensitivity.

4.2.2 Regression Evidence

The descriptive evidence suggests that the random default without feedback increases baseline willingness to contribute, but does not eliminate women's stereotype sensitivity. We next examine whether these patterns hold in regression models that control for individual characteristics, task performance, and question-level differences. The regression analysis proceeds in three steps. The results are reported in Table 5. Column (1) establishes the benchmark pattern in the Control treatment. Column (2) shows how this pattern changes under RD_NF. Column (3) pools the Control and RD_NF treatments and tests the treatment differences directly.

We begin by estimating the following specification separately for the Control and RD_NF treatments:

$$Y_{ij} = \alpha + \beta_1 Women_j + \beta_2 MS_i + \beta_3 Women_j \times MS_i + \Psi Controls_{ij} + \epsilon_{ij}. \quad (1)$$

The dependent variable, Y_{ij}, equals one if participant j chooses Position 1 for question i, and zero otherwise. $Women_j$ is an indicator for female participants, and MS_i is the standardized maleness score of the question category. Because the maleness score is standardized, the coefficient on $Women_j$ captures the gender difference in willingness to choose Position 1 at the average level of category maleness. The interaction term captures whether willingness to choose Position 1 changes

differently with category maleness for women and men.

Table 5: Regression Analysis of Pure Default Effect

	DV: Choosing Position 1 (=1)		
	(1) Control	(2) RD_NF	(3) Pooled
Women	−0.04**	−0.03	−0.04**
	(0.02)	(0.05)	(0.02)
Maleness Score	0.02**	−0.07**	0.02**
	(0.01)	(0.03)	(0.01)
Women × Maleness Score	−0.09***	−0.02	−0.10***
	(0.02)	(0.04)	(0.02)
RD_NF			0.22***
			(0.04)
RD_NF × Women			0.01
			(0.06)
RD_NF × Maleness Score			−0.11***
			(0.03)
RD_NF × Women × Maleness Score			0.09**
			(0.04)
Constant	0.11	0.14	0.07
	(0.08)	(0.14)	(0.09)
Controls	Yes	Yes	Yes
R^2	0.22	0.19	0.24
Adj. R^2	0.22	0.19	0.24
Num. obs.	5882	1472	7354

Note: Coefficients of OLS regressions. The signs of coefficients are identical to those in Probit regressions. The unit of observation is question i answered by participant j. Each participant answered 30 questions. Standard errors, clustered at the individual level, are reported in parentheses. Controls include a dummy for whether question i was answered correctly, Part A score for the category, Part B score for the category, race dummy, US high school dummy, education level dummies, and the overall probability of a correct answer for question i in Part B.
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Column (1) confirms the benchmark pattern in the Control treatment. Women are significantly less likely than men to choose Position 1. The coefficient on MS is positive, while the coefficient on $Women \times MS$ is negative. Thus, men become more likely to choose Position 1 as categories become more male-typed, whereas women become less likely to do so. These estimates reproduce the gender gap and stereotype sensitivity documented in the descriptive evidence.

Column (2) shows how this pattern changes under RD_NF. The coefficient on $Women$ is no longer statistically significant, indicating that the overall gender gap in baseline willingness to contribute disappears when participants are randomly defaulted to Position 1. At the same time,

the estimates do not indicate that stereotype sensitivity disappears for women. The coefficient on MS is negative, and women's estimated stereotype sensitivity remains negative. Thus, the separate regressions suggest that the random default without feedback increases baseline willingness to contribute, but does not remove women's lower willingness to contribute in more male-typed categories. To test the treatment differences directly, Column (3) pools the Control and RD_NF treatments and estimates the following specification:

$$\begin{aligned} Y_{ij} = \alpha &+ \beta_1 Women_j + \beta_2 MS_i + \beta_3 Women_j \times MS_i + \beta_4 RD\_NF_j \\ &+ \beta_5 RD\_NF_j \times Women_j + \beta_6 RD\_NF_j \times MS_i + \beta_7 RD\_NF_j \times Women_j \times MS_i \\ &+ \Psi Controls_{ij} + \epsilon_{ij}. \end{aligned} \tag{2}$$

This pooled specification directly tests the two channels identified in the hypotheses. The coefficient on RD_NF captures the effect of RD_NF on men's baseline willingness to contribute. The estimated effect for women is calculated as the sum of the coefficients on RD_NF and $RD\_NF \times Women$. Similarly, the coefficient on $RD\_NF \times MS$ captures how RD_NF changes men's stereotype sensitivity. The estimated change in stereotype sensitivity for women is calculated as the sum of the coefficients on $RD\_NF \times MS$ and $RD\_NF \times Women \times MS$.

Column (3) provides strong evidence for Hypothesis 1a. The coefficient on RD_NF is positive and statistically significant, indicating that the random default without feedback increases men's baseline willingness to contribute by 22 percentage points. The estimated effect for women, calculated as the sum of the coefficients on RD_NF and $RD\_NF \times Women$, is also positive and statistically significant ($p < 0.01$). Moreover, the coefficient on $RD\_NF \times Women$ is not statistically significant, indicating that the increase in baseline willingness to contribute is not significantly different for women and men. Thus, the regression evidence confirms that RD_NF increases baseline willingness to contribute for both genders and eliminates the overall gender gap.

The evidence for Hypothesis 1b is more limited. For women, the default does not significantly reduce stereotype sensitivity: the estimated change in stereotype sensitivity for women, calculated as the sum of the coefficients on $RD\_NF \times MS$ and $RD\_NF \times Women \times MS$, is not statistically different from zero ($p = 0.61$). For men, the negative coefficient on $RD\_NF \times MS$ is statistically significant, indicating that the default changes the relationship between category maleness and willingness to contribute. Indeed, the estimated stereotype sensitivity for men in RD_NF, calculated as the sum of the coefficients on MS and $RD\_NF \times MS$, is negative ($H_0 : \beta_2 + \beta_6 = 0$, $p < 0.01$), suggesting a reversal of the stereotype pattern observed in the Control treatment. Taken together, the regression results show that RD_NF operates primarily by increasing baseline willingness to contribute, rather than by reducing women's stereotype sensitivity.

Result 1: The random default without feedback increases baseline willingness to contribute for both women and men, eliminating the overall gender gap in choosing Position 1. However, its effect on stereotype sensitivity differs by gender. For women, stereotype sensitivity persists, so Hypothesis 1b is not supported. For men, the default significantly changes—and reverses—the relationship between category maleness and willingness to contribute, so Hypothesis 1b is supported for men. Overall, Hypothesis 1a is supported for both women and men.

4.3 Merit-Based Assignment and the Default Effect

Having established the pure default effect, we next examine whether merit-based assignment strengthens the effect of the default. In the MD treatment, the default position is assigned according to relative performance in Part A. This design can affect behavior through two channels. First, it provides performance feedback: participants who are defaulted to Position 1 learn that they performed at least as well as their partner in the corresponding category. This is the information effect of the merit default (Hypothesis 2a). Second, the default position may be perceived as more legitimate when it is assigned on the basis of performance rather than randomly. This is the legitimacy effect of the merit default (Hypothesis 2b).

4.3.1 Descriptive Evidence

We begin with descriptive comparisons across the three default treatments. Because the hypotheses concern whether high performers are more likely to remain with a pre-selected leading position, we focus on participants who were defaulted to Position 1. For comparison, in the random-default treatments we report results separately for high and low performers.

Figure 1 reports the likelihood of choosing Position 1. Among high performers, participants in MD are more likely to choose Position 1 than participants in RD_NF (64.07% vs. 54.41%, $p = 0.03$, two-sided t-test). Taken alone, this comparison is consistent with the possibility that merit-based assignment increases the tendency to stay with the default. Futher analysis suggests that this difference cannot be attributed to either the information effect of the merit alone or the legitimacy effect alone. To see this, first we note that High performers in RD_F, who receive the same performance feedback but are assigned the default position randomly, are not significantly more likely to choose Position 1 than high performers in RD_NF (59.68% vs. 54.41%, $p = 0.26$, two-sided t-test). Low performers also do not appear to respond to negative performance feedback: their likelihood of choosing Position 1 is not significantly different in RD_F than in RD_NF (51.69% vs. 54.46%, $p = 0.60$, two-sided t-test). These comparisons do not support the information effect of the merit default (Hypothesis 2a).

The descriptive evidence also does not support the legitimacy effect of the merit default

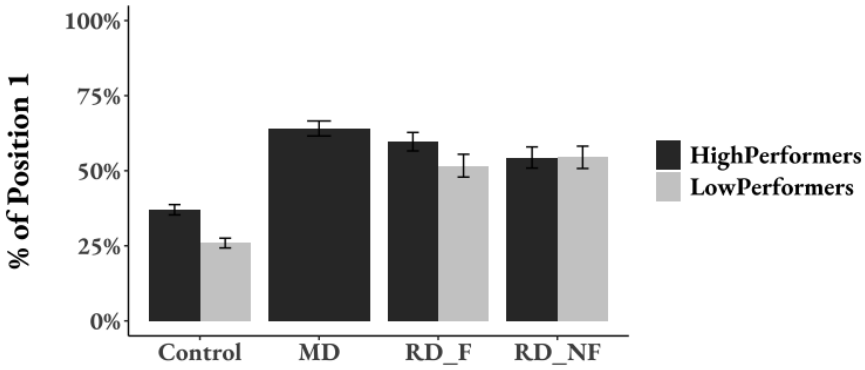

Figure 1: Likelihood of Choosing Position 1 among High and Low Performers

(Hypothesis 2b). High performers in MD are not significantly more likely to choose Position 1 than high performers in RD_F (64.07% vs. 59.68%, two-sided t-test, $p = 0.27$). Thus, participants do not appear more willing to stay with the default when the default position is explicitly assigned on the basis of merit rather than randomly.[12]

4.3.2 Regression Evidence

We next examine whether these descriptive patterns hold in regression models. The regression analysis focuses on high performers in a category, because Hypothesis 2a and Hypothesis 2b concern whether high performers are more likely to remain with the pre-selected Position 1 when the default is accompanied by performance feedback or assigned on the basis of merit.

We estimate the following specification:

$$Y_{ij} = \alpha + \beta_1 MD_j + \beta_2 RD\_F_j + \beta_3 RD\_NF_j + \beta_4 Women_j + \beta_5 MS_i + \beta_6 Women_j \times MS_i + \Psi Controls_{ij} + \epsilon_{ij}. \quad (3)$$

The dependent variable, Y_{ij}, equals one if participant j chooses Position 1 for question i, and zero otherwise. The treatment indicators MD_j, RD_F_j, and RD_NF_j capture the effect of being defaulted to Position 1 under each default treatment, relative to the Control treatment. Column (1) reports the specification without controls, while Column (2) adds the full set of controls used in the previous regressions.

[12] Nor do we find that defaulting according to merit significantly improves correct group answers relative to the other treatments. While the results are in the right direction, average group correct answers in Part B are 19.21 in MD and around 18.85 in Control and the two random-default treatments. The explanation may be the general improvement in performance in Part B relative to Part A, as shown in Table 2.

Because the merit-related hypotheses concern whether feedback or merit-based assignment increases high performers' tendency to remain with the pre-selected Position 1, the main specification focuses on treatment differences among high performers rather than on gender heterogeneity in those treatment effects. Fully interacting treatment, gender, and maleness score would substantially complicate the specification and is not central to Hypothesis 2a or Hypothesis 2b. As a robustness check, Appendix Tables B5 and B6 report regressions that allow the default effect to vary by gender and maleness score; these results do not provide evidence that MD or RD_F reduces women's stereotype sensitivity.

This specification allows us to test the two merit-related hypotheses through comparisons across treatment coefficients. The information effect of the merit default compares treatments with performance feedback against RD_NF. Thus, if Hypothesis 2a is supported, the estimated effects of MD and RD_F should be larger than the estimated effect of RD_NF. The legitimacy effect of the merit default compares MD against RD_F. Thus, if Hypothesis 2b is supported, the estimated effect of MD should be larger than the estimated effect of RD_F.

Table 6 shows that all three default treatments significantly increase high performers' likelihood of choosing Position 1 relative to Control. However, the comparisons across default treatments provide little evidence that feedback or merit-based assignment strengthens the default effect. The estimated effect of MD is larger than that of RD_NF without controls, but this difference is only weakly significant in Column (1) and is no longer statistically significant after controls are added (MD vs. RD_NF, $p = 0.08$ in Column (1), $p = 0.27$ in Column (2)). The comparison between RD_F and RD_NF provides no evidence that performance feedback increases the tendency to choose Position 1 (RD_F vs. RD_NF, $p = 0.20$ in Column (1), $p = 0.37$ in Column (2)). Thus, the regression evidence does not support the information effect of the merit default (Hypothesis 2a).

The regression evidence also does not support the legitimacy effect of the merit default (Hypothesis 2b). If legitimacy increased compliance with the default, high performers in MD should be more likely to choose Position 1 than high performers in RD_F. We do not observe such a difference (MD vs. RD_F, $p = 0.33$ in Column (1), $p = 0.46$ in Column (2)). Thus, conditional on being defaulted to Position 1, high performers do not appear more likely to follow the default when the position is assigned on the basis of merit rather than randomly.

The coefficients on $Women$ and $Women \times MS$ provide additional evidence on gender differences among high performers. The negative coefficient on $Women$ indicates that, at the average level of category maleness, female high performers are less likely than male high performers to choose Position 1. This gender difference is statistically significant both without controls ($p < 0.01$) and with controls ($p < 0.05$). The negative coefficient on $Women \times MS$ further indicates that female high performers' likelihood of choosing Position 1 decreases relative to male high performers as categories become more male-typed. This interaction is statistically significant in both specifications

Table 6: Regressions on Merit-Based Assignment for High Performers

	DV: Choosing Position 1 (=1)	
	(1) No controls	(2) Controls
MD	0.26^{***}	0.25^{***}
	(0.03)	(0.03)
RD_F	0.22^{***}	0.22^{***}
	(0.03)	(0.03)
RD_NF	0.17^{***}	0.18^{***}
	(0.04)	(0.03)
Women	-0.06^{***}	-0.05^{**}
	(0.02)	(0.02)
Maleness Score	0.00	0.00
	(0.01)	(0.01)
Women × Maleness Score	-0.15^{***}	-0.08^{***}
	(0.02)	(0.02)
Controls	No	Yes
R^2	0.08	0.24
Adj. R^2	0.08	0.24
Num. obs.	6592	6592

Note: Coefficients of OLS regressions. The signs of the coefficients are identical to those in Probit regressions. The unit of observation is question i answered by participant j. Each participant answered 30 questions. Standard errors, clustered at the individual level, are reported in parentheses. Controls include a dummy for whether question i was answered correctly; Part A score for the category; Part B score for the category; race dummy; US high school dummy; education level dummies; and the overall probability of a correct answer for question i in Part B.
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

($p < 0.01$). This pattern is consistent with the earlier finding that defaults increase willingness to contribute but do not eliminate women's stereotype sensitivity.

Result 2: Merit-based assignment does not significantly strengthen the default effect. Neither the information effect of the merit default nor the legitimacy effect of the merit default is supported: high performers are not more likely to choose Position 1 in MD than in RD_F or RD_NF once the relevant comparisons are made.

5 Conclusion and Discussion

In this study, we use online experiments to develop and test an easily implementable default approach to reducing the gender gap in willingness to contribute ideas in environments with varying gender stereotypes. We consider the impact of randomly defaulting people to contribute ideas, as well as merit-based assignment, where an individual's default position is determined by their skill and ability. We find that both random and merit-based defaults substantially increase baseline willingness to contribute ideas. The effect is larger in magnitude for women, although the difference is not statistically significant. However, we find little evidence that defaults reduce women's stereotype sensitivity. In particular, although defaults increase women's likelihood of contributing across all areas, women still contribute relatively less in male-stereotyped environments than in female-stereotyped ones.

This research highlights the persistent challenge of the under-representation of qualified women in leading and contributing ideas to groups, particularly in male-stereotyped STEM fields. Despite substantial effort to mitigate this gap, including DEI initiatives such as quotas and training, the gender gap remains. Our results suggest that changing the default to an opt-out system can be a simple yet effective strategy to encourage more diverse idea contributions, potentially leading to improved organizational performance and greater job satisfaction.

Importantly, we find no evidence that merit-based assignment significantly strengthens the default effect relative to random assignment. Nor do we find clear evidence that merit-based assignment significantly improves group performance. This finding indicates that the simplicity and ease of random defaults might be just as effective, if not more so, in promoting participation without the complexity of evaluating merit, which could introduce its own set of challenges or biases. In this context, random defaults may offer a more equitable and efficient approach to encouraging contributions from a broader and more diverse group of individuals.

The impact of default options in reducing gender gaps is significant, but not without limitations. While defaults are effective in encouraging contributions, they do not completely eliminate the influence of gender stereotypes. This finding suggests that default options are not a panacea for deeply ingrained biases that may shape individuals' behaviors and perceptions in certain contexts.

More work is needed to explore how defaults can interact with other mechanisms to address these biases more comprehensively.

The findings of this study offer important implications for policymakers, organizations, and leaders committed to reducing gender disparities in workplace participation and leadership. First, the use of default mechanisms serves as a practical, low-cost intervention to increase participation and idea contribution, particularly among women. For example, in workplace meetings, academic settings, or collaborative projects, automatically designating individuals as contributors - unless they opt out - can help overcome initial hesitations and foster a more inclusive environment.

Importantly, the finding that merit-based assignment does not significantly strengthen the default effect suggests that policy mechanisms need not be complex to be effective. Random defaults, which are easier to implement and administer, can be just as effective in promoting participation. This simplicity makes defaults an attractive option for organizations with limited resources or those seeking quick and scalable solutions.

Moreover, while default mechanisms can play a valuable role, they should be integrated into broader systemic efforts to challenge and reduce gender stereotypes. These could include targeted training programs to raise awareness of implicit biases, mentorship opportunities to support women in male-stereotyped fields, and efforts to promote diverse role models and inclusive leadership practices.

One limitation of our research is its reliance on a particular task with categories of gender-stereotyped questions. Although we employed a representative sample, the task's particular design may constrain the generalizability of our results. Future research should explore the effectiveness of defaults in broader contexts, including different tasks, industries, and cultural settings. Additionally, further investigation is needed to identify complementary strategies that can more effectively reduce stereotype sensitivity and promote gender equity in participation and leadership, particularly in STEM disciplines. For example, combining default options with other interventions, such as bias training, mentorship programs, or structural changes in organizational culture, might more effectively reduce the impact of stereotypes and encourage greater gender equity in leadership and participation.

A second limitation concerns the observed alignment between perceived stereotypes and actual performance in our setting. Specifically, in categories perceived as more male-typed, men did outperform women on average. As a result, our merit-based default intervention may be less effective in contexts where stereotype perceptions strongly reflect underlying performance. Future studies may benefit from exploring settings where stereotypes are less aligned or misaligned with actual ability, particularly those involving "inaccurate" or "harmful" stereotypes that deviate substantially from empirical reality.

In conclusion, our findings highlight the potential of simple, easily implementable default options to promote more equitable participation in idea generation and leadership roles. While defaults alone

may not fully eliminate the influence of gender stereotypes, they offer a practical and cost-effective strategy for advancing diversity and inclusion, especially in environments where conventional interventions have yielded limited results. By leveraging the power of defaults, organizations and policymakers can take meaningful steps toward narrowing gender gaps and fostering more inclusive, innovative, and high-performing environments.

References

Afridi, F., V. Iversen, and M. R. Sharan, "Women Political Leaders, Corruption and Learning: Evidence from a Large Public Program in India," *Economic Development and Cultural Change*, 2017, *66* (1), 1–30.

Almås, I., A. Cappelen, B. Tungodden, A. Armand, B. Bartling, R. Benabou, and R. Weber, "Cutthroat Capitalism versus Cuddly Socialism: Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians?," *Journal of Political Economy*, 2020, *128* (5), 1753–1788.

Ariely, Dan and Michael I. Norton, "How Actions Create—Not Just Reveal—Preferences," *Trends in Cognitive Sciences*, 2008, *12* (1), 13–16.

Babcock, L. and S. Laschever, *Why Women Don't Ask: The High Cost of Avoiding Negotiations—and Positive Strategies for Change*, New York: Bantam Books, 2007.

Bénabou, Roland and Jean Tirole, "Identity, Morals, and Taboos: Beliefs as Assets," *The Quarterly Journal of Economics*, 2011, *126* (2), 805–855.

Benjamin, D., J. Choi, and A. J. Strickland, "Social Identity and Preferences," *American Economic Review*, 2010, *100* (4), 1913–1928.

Bohnet, I., *What Works: Gender Equality by Design*, Cambridge: Harvard University Press, 2016.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer, "Stereotypes," *The Quarterly Journal of Economics*, 2016, *131* (4), 1753–1794.

_ , _ , _ , and _ , "Beliefs about Gender," *American Economic Review*, 2019, *109* (3), 739–773.

Cappelen, A. W., A. D. Hole, E. Ø. Sørensen, and B. Tungodden, "The Pluralism of Fairness Ideals: An Experimental Approach," *American Economic Review*, 2007, *97* (3), 818–827.

Chen, D. L., M. Schonger, and C. Wickens, "oTree—An Open-Source Platform for Laboratory, Online, and Field Experiments," *Journal of Behavioral and Experimental Finance*, 2016, *9*.

Chen, J. and D. Houser, "When Are Women Willing to Lead? The Effect of Team Gender Composition and Gendered Tasks," *Leadership Quarterly*, 2019, *30* (6), 101340.

Cherry, T. L., P. Frykblom, and J. F. Shogren, "Hardnose the Dictator," *American Economic Review*, 2002, *92* (4), 1218–1221.

Coffman, K. B., "Evidence on Self-Stereotyping and the Contribution of Ideas," *The Quarterly Journal of Economics*, 2014, *129* (4), 1625–1660.

Cooper, D. and J. Kagel, "A Failure to Communicate: An Experimental Investigation of the Effects of Advice on Strategic Play," *European Economic Review*, 2016, *82*, 24–45.

Davidai, S., T. Gilovich, and L. D. Ross, "The Meaning of Default Options for Potential Organ Donors," *Proceedings of the National Academy of Sciences*, 2012, *109* (38), 15201–15205.

DeJong, William, "An Examination of Self-Perception Mediation of the Foot-in-the-Door Effect," *Journal of Personality and Social Psychology*, 1979, *37* (12), 2221–2239.

Dinner, I., E. J. Johnson, D. G. Goldstein, and K. Liu, "Partitioning Default Effects: Why People Choose not to Choose," *Journal of Experimental Psychology: Applied*, 2011, *17* (4), 332–341.

Erkal, N., L. Gangadharan, and E. Xiao, "Leadership Selection: Can Changing the Default Break the Glass Ceiling?," *The Leadership Quarterly*, 2022, *33* (2), 101563.

Everett, J., L. Caviola, G. Kahane, J. Savulescu, and N. Faber, "Doing Good by Doing Nothing? The Role of Social Norms in Explaining Default Effects in Altruistic Contexts," *European Journal of Social Psychology*, 2015, *45* (2), 230–241.

Festinger, Leon and James M. Carlsmith, "Cognitive Consequences of Forced Compliance," *Journal of Abnormal and Social Psychology*, 1959, *58* (2), 203–210.

Gangadharan, L., T. Jain, P. Maitra, and J. Vecci, "Social Identity and Governance: The Behavioral Response to Female Leaders," *European Economic Review*, 2016, *90*, 302–325.

Gravert, C., "How Luck and Performance Affect Stealing," *Journal of Economic Behavior & Organization*, 2013, *93*, 301–304.

He, Joyce C., Sonia K. Kang, and Nicola Lacetera, "Opt-out Choice Framing Attenuates Gender Differences in the Decision to Compete in the Laboratory and in the Field," *Proceedings of the National Academy of Sciences*, 2021, *118* (42), e2108337118.

Jachimowicz, J. M., S. Duncan, E. U. Weber, and E. J. Johnson, "When and Why Defaults Influence Decisions: A Meta-Analysis of Default Effects," *Behavioural Public Policy*, 2019, *3* (2), 159–186.

Jackson, J., B. Bradford, M. Hough, A. Myhill, P. Quinton, and T. R. Tyler, "Why Do People Comply with the Law? Legitimacy and the Influence of Legal Institutions," *The British Journal of Criminology*, 2012, *52* (6), 1051–1071.

Johnson, E. J. and D. Goldstein, "Do Defaults Save Lives?," *Science*, 2003, *302* (5649), 1338–1339.

_ , **S. B. Shu, B. G. Dellaert, C. Fox, D. G. Goldstein, G. Häubl, and E. U. Weber**, "Beyond Nudges: Tools of a Choice Architecture," *Marketing Letters*, 2012, *23* (2), 487–504.

Kahneman, D. and A. Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 1979, *47* (2), 263–292.

Kajackaite, A., "Lying about Luck versus Lying about Performance," *Journal of Economic Behavior & Organization*, 2018, *153*, 194–199.

Konow, J., "Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions," *American Economic Review*, 2002, *90* (4), 1072–1091.

Madrian, B. C. and D. F. Shea, "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior," *The Quarterly Journal of Economics*, 2001, *116* (4), 1149–1187.

Matsa, D. and A. Miller, "A Female Style in Corporate Leadership? Evidence from Quotas," *American Economic Journal: Applied Economics*, 2013, *5* (3), 136–169.

McKenzie, C., M. J. Liersch, and S. R. Finkelstein, "Recommendations Implicit in Policy Defaults," *Psychological Science*, 2006, *17*, 414.

Peer, E., L. Brandimarte, S. Samat, and A. Acquisti, "Beyond the Turk: Alternative Platforms for Crowdsourcing Behavioral Research," *Journal of Experimental Social Psychology*, 2017, *70.*

Tannenbaum, D., C. Fox, and T. Rogers, "On the Misplaced Politics of Behavioral Policy Interventions," *Nature Human Behavior*, 2017, *1* (7), 130.

Urminsky, O., "What Should the Ask Be a Nudge? The Effect of Default Amounts on Charitable Donations," *Journal of Marketing Research*, 2016, *53* (5), 829–846.

Weber, Elke U. and Eric J. Johnson, "Constructing Preferences from Memory," in Sarah Lichtenstein and Paul Slovic, eds., *The Construction of Preference*, Cambridge University Press, 2006.

Appendix A. Experiment Screenshots

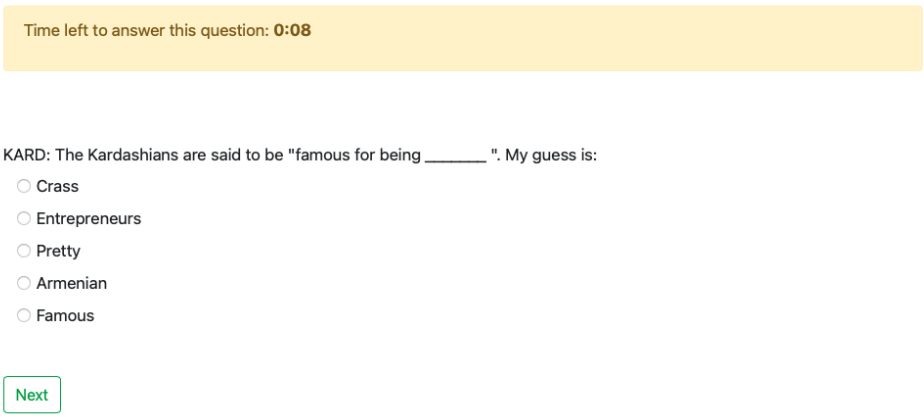

Figure A1: An example question from Part A

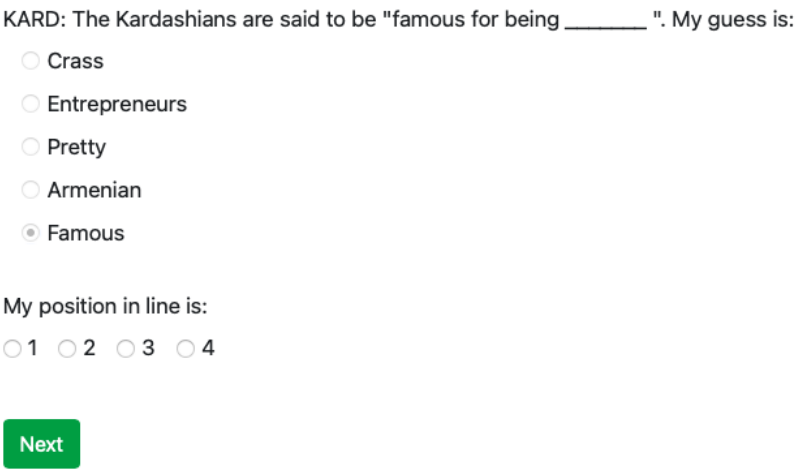

Figure A2: An example question in Part B

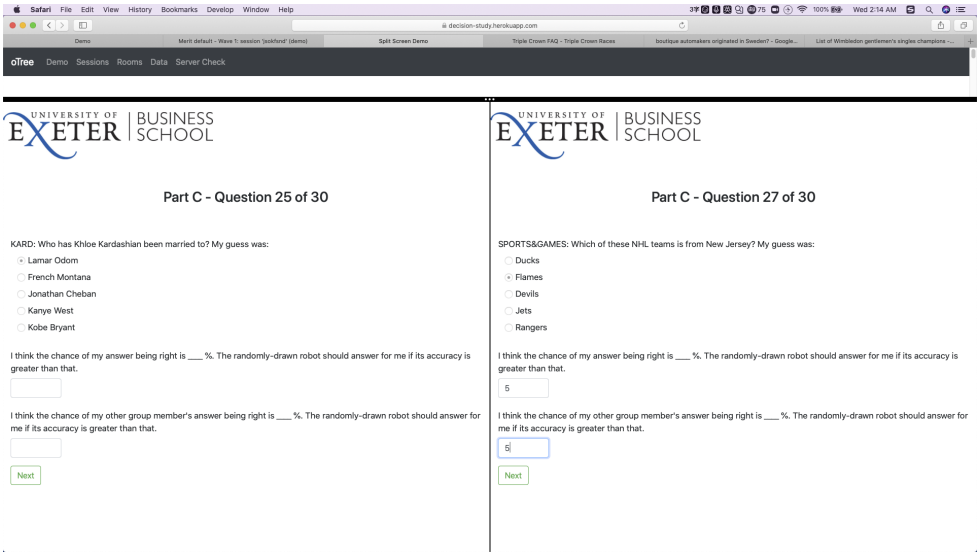

Figure A3: Example questions in Part C

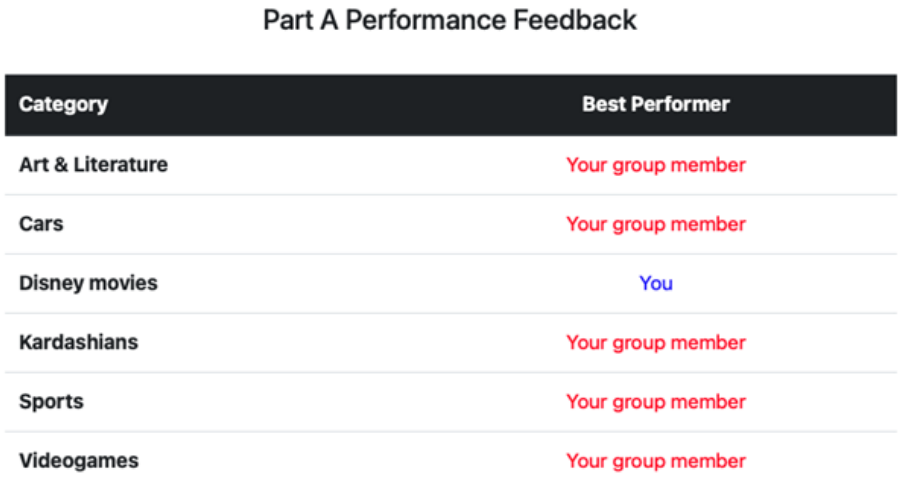

Figure A4: Example of feedback information in MD and RD_F

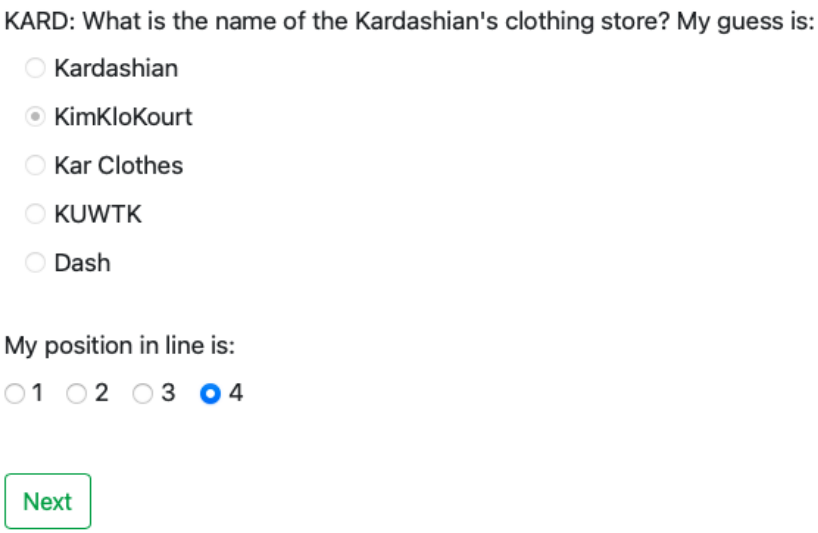

Figure A5: Pre-selected position in line example

Appendix B. Additional Tables and Figures for Default to Position 1

Table B1 shows the number of correct answers in Part A and Part B, the likelihood of choosing Position 1, and the average position number for Control and Wave 1 subjects across three default treatments. Wave 1 subjects in these default treatments were statistically indistinguishable from those in the Control group in terms of Part A performance and willingness to contribute ideas. The only exception is that in Part B, Wave 1 subjects in the RD_F answered significantly more questions correctly compared to the Control (RD_F vs. Control: 17.27 vs. 15.59, $p < 0.01$). Wave 1 subjects in the other two treatments showed no statistically significant differences from the Control group (MD vs. Control: 16.27 vs. 15.59, $p = 0.19$; RD_NF vs. Control: 15.89 vs. 15.59, $p = 0.63$).

Table B1: Control and Default Treatments in Wave 1

	No. of Questions Correct				
Treatment	Part A	Part B	Likelihood of Choosing 1	Average Position Number	Obs.
Control	12.01 (0.04)	15.59 (0.32)	31.20% (0.01)	2.44 (0.04)	200
MD (Wave 1)	12.03 (0.31)	16.27 (0.40)	32.49% (0.02)	2.37 (0.06)	100
RD_F (Wave 1)	12.66 (0.35)	17.61 (0.41)	35.76% (0.02)	2.34 (0.06)	102
RD_NF (Wave 1)	11.75 (0.44)	15.89 (0.52)	31.07% (0.02)	2.36 (0.06)	100
p-value (ANOVA)	0.34	$< 0.01^{***}$	0.23	0.39	

Standard errors are reported in parentheses where applicable.
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

The increase in the likelihood of choosing Position 1 is also reflected in the average chosen position number. As shown in Table B2, the average chosen position is lower in RD_NF than in Control for both men and women in most categories (for women: $p < 0.01$ for each female-typed category and $p > 0.05$ for each male-typed category; for men: $p < 0.1$ for each female-typed category). These patterns provide additional descriptive evidence that the random default without feedback increases baseline willingness to contribute.

No significant differences were observed among treatments (ANOVA, $p = 0.75$). Specifically, 62.83% of groups in the Control answered the question correctly, compared to 64.14% in MD ($p = 0.45$), 64.31% in RD_F ($p = 0.44$), and 62.49% in RD_NF ($p = 0.87$), with all p-values showing comparisons with Control. In Task B, subjects were presented with a new set of multiple-choice questions, which may help explain why no improvement in group performance is observed in the default treatments.[13]

[13]Individual performance in Part A and Part B is positively correlated: categories with higher correct response rates

Table B2: Average Chosen Position Number in Control and RD_NF

	Female-Typed Categories			Male-Typed Categories			
	Kardashian	Disney	Art & Lit.	Video Games	Sports	Car	All
Men							
Control	2.45	2.08	2.65	1.96	2.57	2.26	2.33
RD_NF	1.89	1.71	1.95	1.90	2.27	2.17	1.98
p-value	$< 0.01^{***}$	0.06^{*}	$< 0.01^{***}$	0.79	0.13	0.67	$< 0.01^{***}$
Women							
Control	2.17	2.04	2.70	2.73	3.02	2.70	2.56
RD_NF	2.53	1.51	2.19	3.33	2.50	2.11	2.05
p-value	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$	0.04^{**}	0.02^{**}	0.01^{***}	$< 0.01^{***}$
Gender Gap (p-value)							
Control	0.02^{**}	0.76	0.61	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$
RD_NF	0.12	0.34	0.31	0.12	0.37	0.81	0.63

Figures B1–B2 visualize the results reported above by showing the likelihood that women and men choose Position 1 in Part B across categories and treatments. Figure B1 compares the Control treatment with RD_NF, while Figure B2 compares the Control treatment with RD_F and MD.

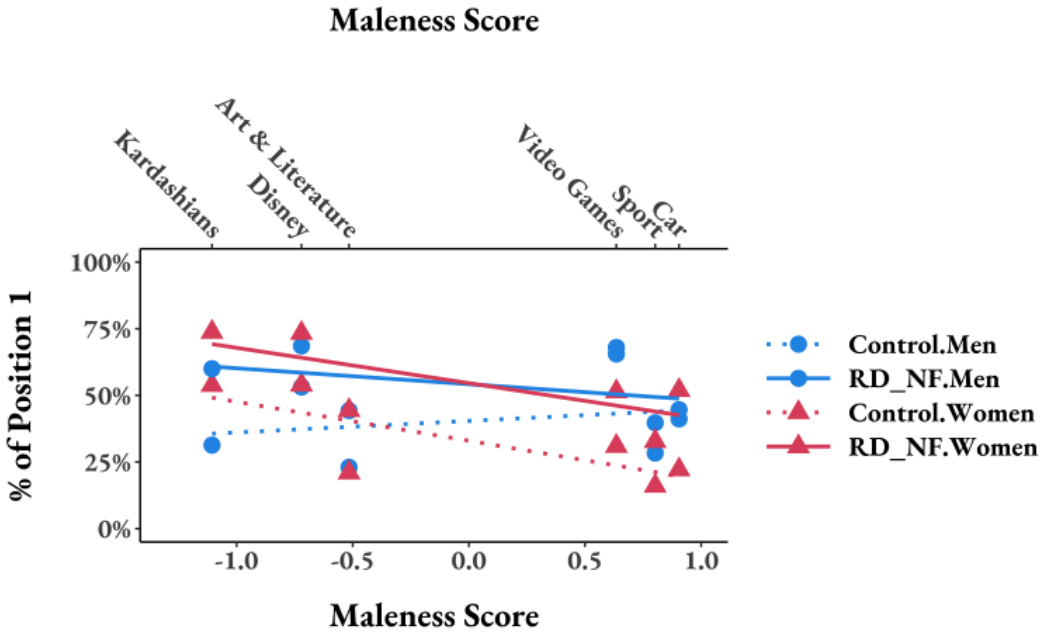

Figure B1: Likelihood of choosing Position 1 among high performers in Control and RD_NF

Tables B3–B6 report additional category-level and regression evidence for high-performing participants. Columns (1) estimate the baseline model for high performers in the Control treatment,

in Part A also tend to have higher correct response rates in Part B ($p < 0.01$). However, the explanatory power of this relationship is limited ($R^2 < 0.1$).

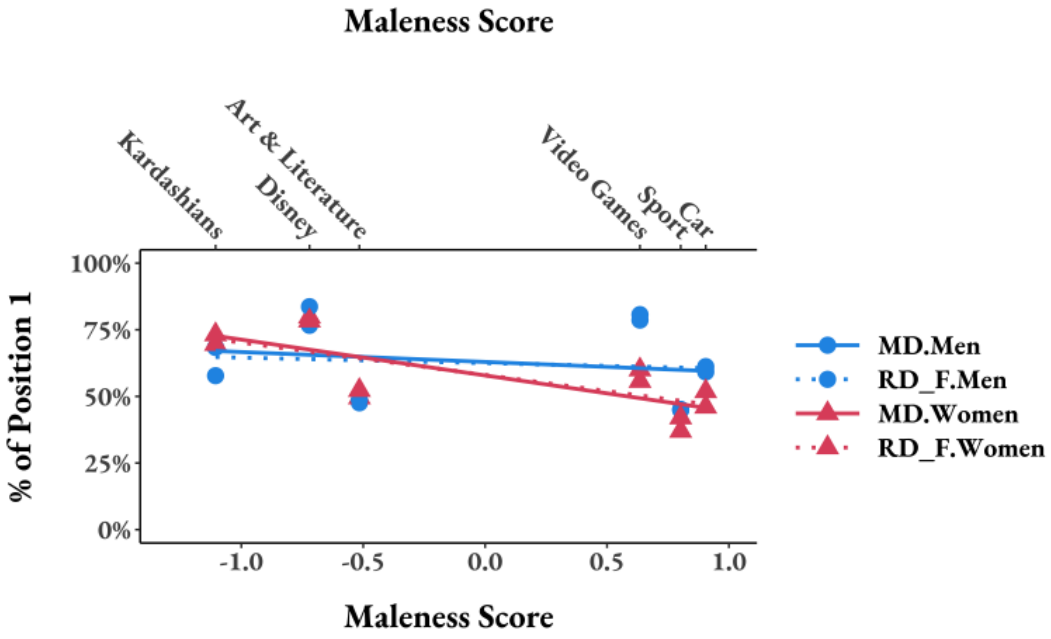

Figure B2: Likelihood of choosing Position 1 among high performers in RD_F and MD

Columns (2) estimate the same model for high performers in each treatment, and Columns (3) pool the Control with the corresponding treatment to examine treatment effects on baseline willingness to contribute and stereotype sensitivity. Except for RD_NF, we find no evidence that RD_F or MD reduces stereotype sensitivity.

Table B3: Likelihood of Choosing Position 1 in Control and RD_NF, by Category

	Female-Typed Categories				Male-Typed Categories				
	Kardashian	Disney	Art & Lit.	Overall	Video Games	Sports	Car	Overall	All
Men									
Control	28.60%	42.70%	20.70%	30.66%	53.57%	24.15%	37.75%	38.46%	35.54%
RD_NF	61.59%	67.86%	51.82%	58.55%	63.00%	41.35%	46.67%	54.93%	55.54%
p-value	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$	0.30	0.02^{**}	0.26	$< 0.01^{***}$	$< 0.01^{***}$
Women									
Control	43.90%	47.35%	17.75%	36.38%	24.25%	11.60%	22.32%	19.41%	27.87%
RD_NF	73.86%	71.43%	44.42%	62.24%	43.10%	38.45%	52.11%	44.02%	52.99%
p-value	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$
Gender Gap (p-values)									
Control	$< 0.01^{***}$	0.33	0.37	0.08	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$
RD_NF	0.19	0.70	0.46	0.60	0.06^{*}	0.77	0.59	0.10	0.67

Table B4: Effect of RD_NF for High Performers

	DV: Choosing Position 1 (=1)		
	(1) Control	(2) RD_NF	(3) Pooled
Women	-0.04^{*}	-0.00	-0.05^{*}
	(0.03)	(0.06)	(0.03)
MS	0.03	-0.09^{**}	0.03
	(0.02)	(0.04)	(0.02)
Women × MS	-0.11^{***}	-0.01	-0.11^{***}
	(0.02)	(0.06)	(0.02)
RD_NF			0.17^{***}
			(0.05)
RD_NF × Women			0.03
			(0.07)
RD_NF × MS			-0.13^{***}
			(0.05)
RD_NF × Women × MS			0.14^{**}
			(0.06)
Constant	0.09	-0.10	0.09
	(0.10)	(0.19)	(0.10)
Controls	Yes	Yes	Yes
R^2	0.26	0.22	0.26
Adj. R^2	0.26	0.21	0.25
Num. obs.	2943	864	3807

Note: Coefficients of OLS regressions. The signs of coefficients are identical to those in Probit regressions. The unit of observation is question i answered by participant j. Standard errors, clustered at the individual level, are reported in parentheses. Controls include a dummy for whether question i was answered correctly, Part A score for the category, Part B score for the category, race dummy, US high school dummy, education level dummies, and the overall probability of a correct answer for question i in Part B.
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Table B5: Effect of RD_F for High Performers

	DV: Choosing Position 1 (=1)		
	(1) Control	(2) RD_F	(3) Pooled
Women	-0.04^{*}	-0.11^{*}	-0.05^{*}
	(0.03)	(0.06)	(0.03)
MS	0.03	0.02	0.03
	(0.02)	(0.04)	(0.02)
Women × MS	-0.11^{***}	−0.06	-0.11^{***}
	(0.02)	(0.05)	(0.02)
RD_F			0.25^{***}
			(0.05)
RD_F × Women			-0.06
			(0.06)
RD_F × MS			-0.01
			(0.04)
RD_F × Women × MS			0.05
			(0.06)
Constant	0.09	0.20	0.04
	(0.10)	(0.17)	(0.07)
Controls	Yes	Yes	Yes
R^2	0.26	0.17	0.26
Adj. R^2	0.26	0.16	0.26
Num. obs.	2943	907	3850

Note: Coefficients of OLS regressions. The signs of coefficients are identical to those in Probit regressions. The unit of observation is question i answered by participant j. Standard errors, clustered at the individual level, are reported in parentheses. Controls include a dummy for whether question i was answered correctly, Part A score for the category, Part B score for the category, race dummy, US high school dummy, education level dummies, and the overall probability of a correct answer for question i in Part B.
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Table B6: Effect of MD for High Performers

	DV: Choosing Position 1 (=1)		
	(1) Control	(2) MD	(3) Pooled
Women	−0.04*	-0.02	−0.05**
	(0.03)	(0.05)	(0.03)
MS	0.03	-0.02	0.03*
	(0.02)	(0.02)	(0.02)
Women × MS	−0.11***	-0.07	−0.12***
	(0.02)	(0.12)	(0.03)
MD			0.23***
			(0.04)
MD × Women			0.03
			(0.06)
MD × MS			-0.05
			(0.03)
MD × Women × MS			0.07
			(0.04)
Constant	0.09	0.27**	0.10
	(0.10)	(0.12)	(0.09)
Controls	Yes	Yes	Yes
R^2	0.26	0.16	0.26
Adj. R^2	0.26	0.16	0.26
Num. obs.	2943	1878	4821

Note: Coefficients of OLS regressions. The signs of coefficients are identical to those in Probit regressions. The unit of observation is question i answered by participant j. Standard errors, clustered at the individual level, are reported in parentheses. Controls include a dummy for whether question i was answered correctly, Part A score for the category, Part B score for the category, race dummy, US high school dummy, education level dummies, and the overall probability of a correct answer for question i in Part B.
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Appendix C. The Effect of Defaulting to Position 4

Following the hypotheses for defaulting participants to Position 1, we summarize the corresponding hypotheses for defaulting participants to Position 4.

Hypothesis D1a. *(Default effect on choosing Position 4)* For questions in each gender category, participants in RD_NF are more likely than participants in the Control to select Position 4 when this position is pre-selected for them.

Hypothesis D1b. *(Default effect on stereotype sensitivity)* The relationship between category maleness and the willingness to choose Position 4 is weaker when subjects in RD_NF are defaulted to Position 4 than when subjects are in the Control.

Hypothesis D2a. *(Information effect of the merit default)* For low performers in Part A who are defaulted to Position 4 in Part B, the likelihood of remaining with the pre-selected position is higher in MD and RD_F than in RD_NF.

Hypothesis D2b. *(Legitimacy effect of the merit default)* For low performers in Part A who are defaulted to Position 4 in Part B, the likelihood of remaining with the pre-selected position is higher in MD than in RD_F.

For all three default treatments, we focus on cases where subjects are assigned to Position 4. Table C1 presents the percentage of choosing Position 4 in both the Control and RD_NF conditions, broken down by gender and category. The results show that the default effect is symmetric: across all categories, both men and women are significantly more likely to choose Position 4 in RD_NF compared to Control ($p < 0.05$ for all categories). This effect also reduces the gender gap in choosing Position 4. In the Control condition, women are significantly more likely than men to choose Position 4 (31.48% vs. 23.49%, $p = 0.01$). However, in the RD_NF condition, this gender difference is no longer significant (women: 55.85% vs. men: 53.77%, $p = 0.72$). This result supports Hypothesis D1a.

Figure C1 visualizes the likelihood of choosing Position 4 by gender, category, and treatment. We observe stereotype sensitivity in women's choices of Position 4: in Control, women are more likely to choose Position 4 in male-typed categories than in female-typed categories (23.03% vs. 39.69%, two-sided t-test, $p < 0.01$). We find that in the RD_NF treatment, women remain more likely to choose Position 4 in male-typed categories than in female-typed categories (47.00% vs. 63.17%, two-sided t-test, $p = 0.01$). This result does not support Hypothesis D1b.[14]

We use a regression framework to further investigate the effect of merit-based assignment on choosing Position 4. Table C2 shows the regression results (Column (1) is for high performers; Column (2) is for low performers). The significant positive coefficients on MD, RD_F, and RD_NF

[14] In Control, men are not more likely to choose Position 4 in female-typed categories than in male-typed categories (23.80% vs. 23.32%, two-sided t-test, $p = 0.88$). Men's decisions do not vary by category type in RD_NF (54.57% vs. 53.37%, two-sided t-test, $p = 0.87$).

Table C1: Likelihood of Choosing Position 4 in Control and RD_NF

	Female-Typed Categories			Male-Typed Categories			
	Kardashian	Disney	Art & Lit.	Video Games	Sports	Car	All
Men							
Control	26.40%	15.95%	28.85%	16.56%	29.50%	23.72%	23.49%
RD_NF	58.70%	40.00%	64.26%	38.33%	67.17%	51.20%	53.77%
p-value	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$	0.02^{**}	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$
Women							
Control	23.10%	17.90%	28.60%	40.10%	44.10%	35.08%	31.48%
RD_NF	40.00%	51.30%	52.40%	73.18%	64.55%	58.28%	55.85%
p-value	0.02^{**}	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$	$< 0.01^{***}$
Gender Gap (p-value)							
Control	0.41	0.61	0.95	$< 0.01^{***}$	$< 0.01^{***}$	0.01^{**}	0.01^{**}
RD_NF	0.05^{*}	0.34	0.24	$< 0.01^{***}$	0.77	0.45	0.72

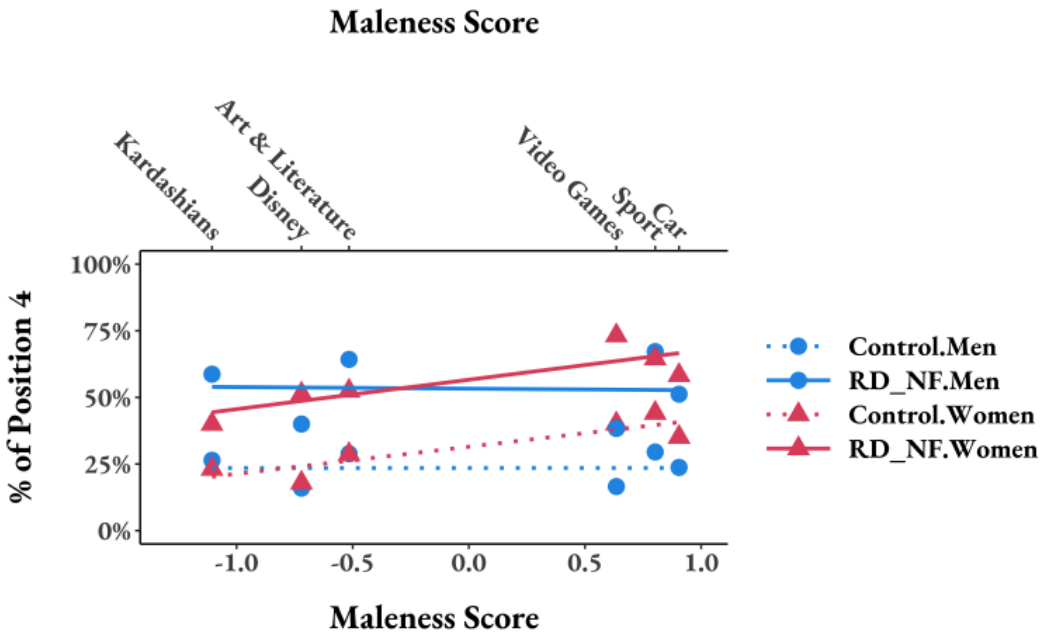

Figure C1: Percentage choosing Position 4 by treatment and category

show the default effect: both high performers and low performers are more likely to choose Position 4 when they are defaulted to Position 4. We do not find evidence for the information effect of the merit default, as there is no magnitude difference between RD_F and RD_NF ($p = 0.30$ in Column (1), $p = 0.50$ in Column (2)). The legitimacy effect of the merit default is not significant (MD vs. RD_F, $p = 0.78$ in Column (2)). The coefficients on Women and Women × MS provide additional evidence on gender differences and stereotype sensitivity. We omitted coefficients on interactions among treatment condition, MS, and Female in both regressions, as they are all insignificant. Neither Hypothesis D2a nor Hypothesis D2b is supported.

Table C2: Regressions on Default Effect on Choosing Position 4

	DV: Choosing Position 4 (=1)	
	(1) High Performers	(2) Low Performers
MD		0.28***
		(0.04)
RD_F	0.25***	0.28***
	(0.03)	(0.04)
RD_NF	0.30***	0.25***
	(0.03)	(0.05)
Women	0.04*	0.02
	(0.03)	(0.03)
MS	-0.00	0.03**
	(0.02)	(0.01)
Women × MS	0.08***	0.05**
	(0.02)	(0.02)
Constant	0.33***	0.38***
	(0.09)	(0.13)
Controls	No	Yes
R^2	0.23	0.14

Note: Coefficients of OLS regressions. The signs of the coefficients are identical to those in Probit regressions. The unit of observation is question i answered by participant j. All regressions use cluster-robust standard errors at the individual level. Controls include a dummy for whether question i was answered correctly; Part A score for the category; Part B score for the category; race dummy; US high school dummy; education level dummies; and the overall probability of a correct answer for question i in Part B.
Tests of coefficient differences: RD_F vs. RD_NF, $p = 0.30$ in Column (1) and $p = 0.50$ in Column (2); MD vs. RD_F, $p = 0.78$ in Column (2).
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

The results of defaulting to Position 4 demonstrate the consistency of the default effect. Similar to the findings for defaulting to Position 1, the default changes the likelihood of choosing the pre-selected position but does not provide evidence that merit-based assignment strengthens the

default effect.

Result D1: Hypothesis D1a is supported, while Hypotheses D1b, D2a, and D2b are not supported.

Appendix D. Confidence and Beliefs

The main text shows that default options directly affect willingness to contribute ideas. We now consider beliefs as a potential mechanism. Prior work suggests that beliefs play an important role in contribution decisions (Coffman, 2014). Following Coffman (2014) and Bordalo et al. (2019), we elicit beliefs in Part C of the experiment. In principle, defaults without feedback should not affect beliefs about performance. Subjects who are defaulted to Position 1 in RD_NF should not be more confident than subjects in the Control. Likewise, high performers who are defaulted to Position 1 should not be more confident in MD than in RD_F. Nevertheless, prior studies suggest that people may shape their beliefs and attitudes based on their previous actions (Festinger and Carlsmith, 1959; DeJong, 1979; Bénabou and Tirole, 2011; Ariely and Norton, 2008; Weber and Johnson, 2006). In our setting, participants who are nudged to choose Position 1 may subsequently form the belief that their performance is superior to their partner's. This possibility may be especially relevant in the MD treatment, where both the default position and one's qualification to answer are made salient.

Following Coffman (2014), we use two belief measures to assess confidence: one measuring confidence in one's own answers and one measuring confidence in one's partner's answers. For each question in Part B, participants state the probability that their own answer is correct and the probability that their partner's answer is correct. These two confidence measures are incentivized using a variant of the BDM mechanism. In particular, participants are introduced to a robot that answers questions with a random probability of being correct. A participant's answer is replaced with the robot's answer if the robot's accuracy rate exceeds the participant's stated probability of being correct.

The belief mechanism requires two steps. First, defaults must affect participants' beliefs. Second, beliefs must affect participants' willingness to choose Position 1. We focus on high performers in a category, consistent with the analysis of merit-based assignment in the main text. We first test whether defaults affect beliefs using the following specification:

$$\begin{aligned} Y_{ij} = \alpha + \beta_1 MD_j + \beta_2 RD\_F_j + \beta_3 RD\_NF_j + \beta_4 Women_j + \beta_5 MS_i \\ + \beta_6 Women_j \times MS_i + \Psi Controls_{ij} + \epsilon_{ij}. \end{aligned} \tag{D1}$$

Table D1 reports the results. The dependent variable is self-confidence in Column (1) and confidence in one's partner in Column (2). The treatment indicators capture whether defaults affect beliefs relative to the Control condition. If defaults increase participants' confidence in their own answers, the coefficients on MD, RD_F, and RD_NF should be positive in Column (1). If defaults reduce participants' confidence in their partner's answers, these coefficients should be negative in Column (2).

The results provide little evidence that defaults affect participants' beliefs about their own

answers. The coefficients on MD, RD_F, and RD_NF are not statistically significant in Column (1). In Column (2), the coefficient on MD is negative and statistically significant ($p < 0.05$), indicating that MD reduces participants' confidence in their partner's answers. The coefficients on RD_F and RD_NF are not statistically significant. Thus, we find no evidence that defaults generally increase self-confidence, and only limited evidence that the merit-based default changes beliefs about the partner.

Table D1: Regressions on Beliefs

	DV: Belief Measure	
	(1) Self-confidence (%)	(2) Partner confidence (%)
MD	0.01	-0.05^{**}
	(0.01)	(0.02)
RD_F	0.02	-0.02
	(0.02)	(0.02)
RD_NF	0.00	-0.03
	(0.02)	(0.02)
Women	-0.03^{***}	-0.02
	(0.01)	(0.02)
MS	-0.02^{**}	-0.03^{***}
	(0.01)	(0.01)
Women × MS	-0.06^{***}	-0.00
	(0.01)	(0.01)
Constant	0.28^{***}	0.52^{***}
	(0.04)	(0.07)
Controls	Yes	Yes
R^2	0.38	0.10
Adj. R^2	0.38	0.10
Num. obs.	6593	6593

Note: Coefficients of OLS regressions. The signs of the coefficients are identical to those in Probit regressions. The unit of observation is question i answered by participant j. All regressions use cluster-robust standard errors at the individual level. Controls include a dummy for whether question i was answered correctly; Part A score for the category; Part B score for the category; race dummy; US high school dummy; education level dummies; and the overall probability of a correct answer for question i in Part B.
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

We next test whether beliefs predict participants' willingness to choose Position 1. We estimate the same regression framework used in the main analysis, but add self-confidence and partner confidence as explanatory variables. Table D2 reports the results.

Self-confidence strongly predicts willingness to choose Position 1. The coefficient on self-

confidence is positive and statistically significant in both specifications ($p < 0.01$), consistent with Coffman (2014). By contrast, the coefficient on partner confidence is not statistically significant in either specification. This is important because the only belief effect detected in Table D1 is the negative effect of MD on confidence in one's partner. Since partner confidence does not predict the decision to choose Position 1, the evidence does not support beliefs as the main mechanism through which defaults affect willingness to contribute.

Table D2: Regressions on Defaults and Beliefs

	DV: Choosing Position 1 (=1)	
	(1) No controls	(2) Controls
MD	0.24***	0.24***
	(0.03)	(0.03)
RD_F	0.21***	0.20***
	(0.03)	(0.03)
RD_NF	0.18***	0.18***
	(0.03)	(0.03)
Women	-0.01	-0.02
	(0.02)	(0.02)
MS	0.01	0.01
	(0.01)	(0.01)
Women × MS	−0.03*	-0.02
	(0.02)	(0.02)
Self-confidence	0.79***	0.72***
	(0.03)	(0.03)
Partner confidence	0.02	0.02
	(0.04)	(0.04)
Constant	−0.11***	−0.14**
	(0.02)	(0.07)
Controls	No	Yes
R^2	0.42	0.43
Adj. R^2	0.42	0.42
Num. obs.	6592	6592

Note: Coefficients of OLS regressions. The signs of the coefficients are identical to those in Probit regressions. The unit of observation is question i answered by participant j. All regressions use cluster-robust standard errors at the individual level. Controls include a dummy for whether question i was answered correctly; Part A score for the category; Part B score for the category; race dummy; US high school dummy; education level dummies; and the overall probability of a correct answer for question i in Part B.
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Overall, Appendix Appendix D. provides little support for beliefs as the primary mechanism

behind the default effect. Defaults do not generally increase self-confidence, and although MD reduces confidence in one's partner, partner confidence does not predict the choice of Position 1.