

Megalokonomou, Rigissa; Silva Goncalves, Juliana; van Veldhuizen, Roel

Working Paper

Gender differences in self-promotion and career advice

Working Paper, No. 2025:11

Provided in Cooperation with:

Department of Economics, School of Economics and Management, Lund University

Suggested Citation: Megalokonomou, Rigissa; Silva Goncalves, Juliana; van Veldhuizen, Roel (2025) : Gender differences in self-promotion and career advice, Working Paper, No. 2025:11, Lund University, School of Economics and Management, Department of Economics, Lund

This Version is available at:

<https://hdl.handle.net/10419/339056>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



LUND UNIVERSITY

Gender Differences in Self-Promotion and Career Advice

Megalokonomou, Rigissa ; Silva Goncalves, Juliana; van Veldhuizen, Roel

2025

Document Version:
Other version

[Link to publication](#)

Citation for published version (APA):

Megalokonomou, R., Silva Goncalves, J., & van Veldhuizen, R. (2025). *Gender Differences in Self-Promotion and Career Advice*. (pp. 1-97). (Working Papers; Vol. 2025:11).

Total number of authors:

3

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Working Paper 2025:11

Department of Economics
School of Economics and Management

Gender Differences in Self-Promotion and Career Advice

Rigissa Megalokonomou
Juliana Silva-Goncalves
Roel van Veldhuizen

Dec 2025



LUND
UNIVERSITY

Gender Differences in Self-Promotion and Career Advice*

Rigissa Megalokonomou¹, Juliana Silva-Goncalves², and Roel van
Veldhuizen³

¹Monash University

²The University of Sydney

³Lund University

December 12, 2025

Abstract

We study the role of self-promotion and career advice in sustaining gender differences in labor market outcomes. We conduct an online experiment in which “advisers” advise “workers” to choose between a more ambitious and a less ambitious task based on the worker’s subjective self-assessment. We find that women have lower self-assessments and receive less ambitious career advice as a result. We also show that these gender differences are similar for both quantitative and qualitative self-assessments, and that the gender difference in advice received can be mitigated by informing advisers of the worker’s true performance or of the gender gap in self-promotion.

Keywords: Advice, Gender, Self-Promotion, Randomized Experiment

JEL codes: C91, D91, J16

**Megalokonomou*: Monash University, Department of Economics, IZA, and CESifo (rigissa.megalokonomou@monash.edu); *Silva-Goncalves*: The University of Sydney, ARC Centre of Excellence for Children and Families over the Life Course and IZA (juliana.goncalves@sydney.edu.au); *van Veldhuizen*: Lund University (roel.van_veldhuizen@nek.lu.se).

We thank Adeline Delavande, Kirby Nielsen, and participants at seminars at Lyon (GATE), Jönköping, Berlin (BBE), and the UTS Business School (EDG), as well as participants at the 2025 Australia New Zealand Workshop on Experimental Economics (ANZWEE), the 2024 ESA European Meeting in Helsinki, and the 2025 Nordic Conference for Behavioral and Experimental Economics in Stavanger for helpful comments. The study was approved by The University of Sydney Human Research Ethics Committee (Project No. 2022/872). We pre-registered our experimental design, hypotheses, and analysis plan (AEARCTR-0013803). Rigissa Megalokonomou acknowledges financial support from the Monash Establishment Grant (No. 1755774). Roel van Veldhuizen gratefully acknowledges financial support from the Jan Wallander och Tom Hedelius foundation.

1 Introduction

A robust stylized fact established in recent years in many countries is that women obtain more education and are more likely to graduate from college than men. However, there are still large differences in the fields of specialization chosen by men and women, which leads to persistent gender gaps in labor market outcomes. Women are underrepresented in male-dominated fields, such as science, technology, engineering, and mathematics (STEM), which traditionally pay better (Blau & Kahn, 2017; Goldin, 2014; Goulas, Griselda, & Megalokonomou, 2024; Oreopoulos & Petronijevic, 2013). Women are also less likely to take leadership positions (Bertrand, Goldin, & Katz, 2010; Eagly & Carli, 2003). Previous research has highlighted the role of stereotypes, discrimination, and childbearing responsibilities as well as differences in beliefs and preferences (see Blau and Kahn, 2017 for a review). A recent line of work relates gender gaps in the labor market to gender differences in self-promotion strategies, whereby women may be less likely to advocate for their achievements compared with men (Exley & Kessler, 2022; Exley & Nielsen, 2024).

In this paper, we focus on the role of advice in shaping labor market outcomes for men and women. Advice from others can impact important economic decisions and shape people’s career trajectory. People turn to parents, teachers, peers, managers and others to seek advice when facing important decisions, such as what subjects to study in school, whether to go to university and which degree to choose and whether to apply for a specific job or promotion. When we lack information, advice can be invaluable for resolving uncertainty and lead to better decisions. However, advice may contribute to gender gaps in the labor market if it is given in a gender-biased manner. Advisers may encourage men toward more competitive, risk-taking, or historically lucrative careers and women toward careers that align with traditional gender norms. Such biases may emerge even among well-meaning or experienced advisers, if they do not have perfect information on the achievement potential of their advisee, and therefore, must rely on their own subjective judgment. In such cases, the advisee’s judgment is likely to matter as well, which is important since women have been shown to be less likely to advocate for their own skills and accomplishments than men (Exley & Kessler, 2022). Therefore, a gender gap in advice, driven by a gender bias in subjective

judgments, may help explain the persistence of gender differences in labor market outcomes.

Two main empirical challenges arise while examining the role of career advice in shaping labor market outcomes for men and women. First, in many contexts, advice is difficult to measure, as it is often private. Second, even if advice can be measured, any gender differences in advice could be driven by gender differences in unobserved characteristics. These issues explain why, despite the widespread scholarly and policy importance of understanding the role of advice in explaining gender gaps in labor market outcomes, well-identified empirical estimates are rare.

We overcome these challenges by studying gender differences in advice received using a pre-registered online experiment conducted on Prolific. As a first step, 350 participants (the “workers”) work on an initial STEM task, after which they are asked to evaluate their own performance in several ways, both qualitative and quantitative. These self-evaluations are then sent to a second set of 1,750 participants in the role of “advisers,” who are tasked with recommending a more or less ambitious follow-up task to five different workers. Upon receiving the advice, workers decide between the two follow-up tasks and are paid according to their performance. What makes the advice invaluable is that workers have no knowledge of what the task entails, whereas advisers have full information about the task and payments involved, as well as an incentive to maximize the worker’s earnings. Our design allows us to measure the advice itself and any information available to advisers, and thus overcomes both identification challenges described above. Since workers are aware that their self-evaluations will be sent to a third party (the advisers), we follow Exley and Kessler (2022) and refer to them as *self-promotions*.

Our main interest lies in studying gender differences in advice. In our baseline treatment, we provide advisers with workers’ answers to one verbal and five quantitative self-evaluation questions. We implement three additional treatments that vary the advisers’ information set. Two treatments focus on the role of subjective judgment. In the “subjective” treatment, we reduce the advisers’ information set by providing them with only verbal self-assessments. In the “objective” treatment, we instead increase the information set by also providing advisers with the workers’ actual performance. These two treatments allow us to examine whether the gender gap in advice depends on the extent to which they need to rely on subjective

information. Finally, we also consider a “nudge” treatment that allows us to test whether informing advisers about prior evidence of a gender gap in self-promotion reduces the gender gap in advice.

We report four main results. First, we find a substantial gender gap in advice, with women 12 percentage points less likely to be advised to opt for the more ambitious task compared with men in the baseline treatment. Second, we find that self-promotion is a key driver of the gender gap in advice. Women submit less optimistic self-evaluations, and this in turn explains why they receive less ambitious advice than men. Our findings demonstrate that, even among equally performing men and women, women communicate their performance less favorably than men and, consequently, receive less ambitious advice. Third, we find that the gender gap in advice is remarkably robust to changes in the advisers’ information set. Fourth, informing advisers about the gender gap in self-promotion removes and even reverses the gender gap in advice in favor of women. This suggests that information interventions or nudges can be an effective way to eliminate gender gaps in advice.

This paper contributes to several research fields. First, it advances the literature on the drivers of gender differences in the labor market (Blau & Kahn, 2017) by focusing on the role of and channels that drive a gender gap in advice. Previous work has examined the role of gender stereotypes, gender differences in earnings expectations, family trade-offs and childbearing responsibilities, and behavioral traits such as overconfidence, competitiveness, and risk preferences in shaping gender gaps in career decisions.¹ We also contribute to a recent branch of this literature that studies gender differences in self-promotion (Chang, Saccardo, & Gallus, 2025; Exley & Kessler, 2022; Exley & Nielsen, 2024; Römer & Schröder, 2025). We build on this work and show that the gender gap in self-promotion may affect third-party advice, which in turn may affect individuals’ choices and outcomes.

Second, this paper contributes to a broader literature on advice. One branch of this literature uses laboratory experiments to study whether the advice of other participants or

¹The literature covers gender gaps in career aspirations and expectations (Bohren, Imas, & Rosenberg, 2019; Carlana, 2019; Lavy & Megalokonomou, 2024); earnings expectations (Blau & Kahn, 2017; Reuben, Wiswall, & Zafar, 2017); family trade-offs and childbearing responsibilities (Angelov, Johansson, & Lindahl, 2016; Bertrand, Goldin, & Katz, 2010; Cortés & Pan, 2023); overconfidence (Bandiera, Parekh, Petrongolo, & Rao, 2022); competitiveness (Niederle & Vesterlund, 2007; Niederle, 2016); and risk preferences (Croson & Gneezy, 2009).

experimenters impacts behavior in, for example, economic games (Çelen, Kariv, & Schotter, 2010; Chaudhuri, Schotter, & Sopher, 2009; Schotter & Sopher, 2003, 2007); matching markets (Guillen & Hakimov, 2018); and tournament entry (Kessel, Mollerstrom, & Van Veldhuizen, 2021). Another branch relies on exogenous information, mentoring, or role model interventions to proxy for the effect of advice in the field (Falk, Kosse, & Pinger, 2025; Goulas, Gunawardena, Megalokonomou, & Zenou, 2024; Hoxby & Turner, 2015; Porter & Serra, 2020). When it comes to gender, Brandts, Groenert, and Rott (2015) show in the lab that receiving advice from a more experienced participant does not eliminate the gender gap in tournament entry, and Brandts and Rott (2021) show that male and female advisers give similar advice. In the field, Gallen and Wasserman (2021) demonstrate that professionals acting as career advisers focus on different aspects of the job depending on the gender of their advisee and in line with gender stereotypes.² We contribute to this body of research by studying gender gaps in advice in an experimental context that relates to career decisions and while allowing self-promotion to influence advice, in line with many real-world applications.

Third, we also contribute to the literature that examines how ambiguity or subjectivity about personal qualities can generate biased judgments. Previous studies provide mixed evidence on the role of subjective judgment in generating gender biases (Bohren, Imas, & Rosenberg, 2019; Krawczyk & Smyk, 2016; Mengel, Sauermann, & Zölitz, 2019; Özgümüş, Rau, Trautmann, & König-Kersting, 2020; Sarsons, Gërkhani, Reuben, & Schram, 2021). For instance, Bohren, Imas, and Rosenberg (2019) found that questions asked by inexperienced female users of an online STEM forum receive lower subjective ratings than those asked by inexperienced male users, but the gender gap is reversed among experienced users. Özgümüş, Rau, Trautmann, and König-Kersting (2020) find a gender gap in subjective evaluations of teaching materials, with female teachers receiving lower ratings relative to male teachers, but only from evaluators with tertiary education. In these studies, gender gaps were driven by the evaluators' personal beliefs or preferences. In contrast, we find no evidence in our setting that advisers have gender-biased beliefs or preferences per se, regardless of how subjective the information about the worker's performance is.

²Also related are Carlana (2019) and Lavy and Megalokonomou (2024) and Wolter and Zöllner (2024), who investigate gender stereotypes among teachers and parents, respectively. While some of the effects they document likely operate through advice, they do not directly observe this.

Our results have important policy implications. The substantial gender gap in self-promotion indicates that women communicate their achievements less favorably than men. This self-promotion gap is critical, because it influences both the advice women receive and the decisions they make about their own career paths. Policies such as mentorship programs could help mitigate this gap and encourage more women to pursue ambitious career goals, particularly in STEM fields. Our findings suggest that increasing awareness of the confidence gap between women and men can be a useful policy tool. Policies that incorporate such nudges—such as information interventions—could further help mitigate gender gaps in male-dominated fields.

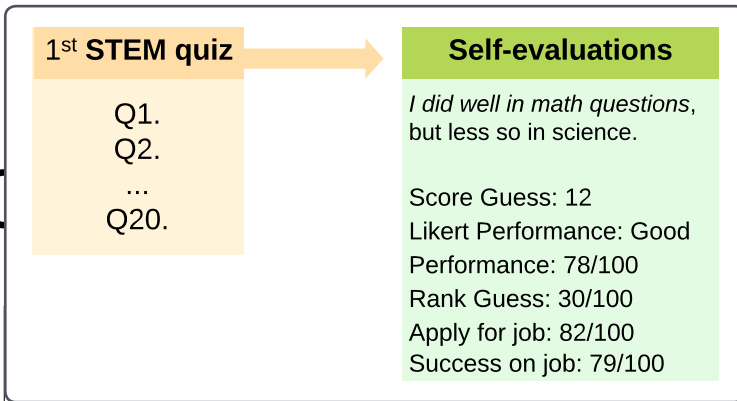
2 Experimental Design

We conducted an online experiment in which participants randomly assigned to the role of workers received advice from other participants randomly assigned to the role of advisers. The experiment consisted of three parts. In part one, workers completed a STEM quiz and subsequently evaluated their own performance. In part two, advisers received the workers' self-evaluations and submitted their advice on whether workers should perform a more or less ambitious follow-up STEM task. In part three, workers received the advice and decided in which of the two follow-up STEM tasks they would perform.

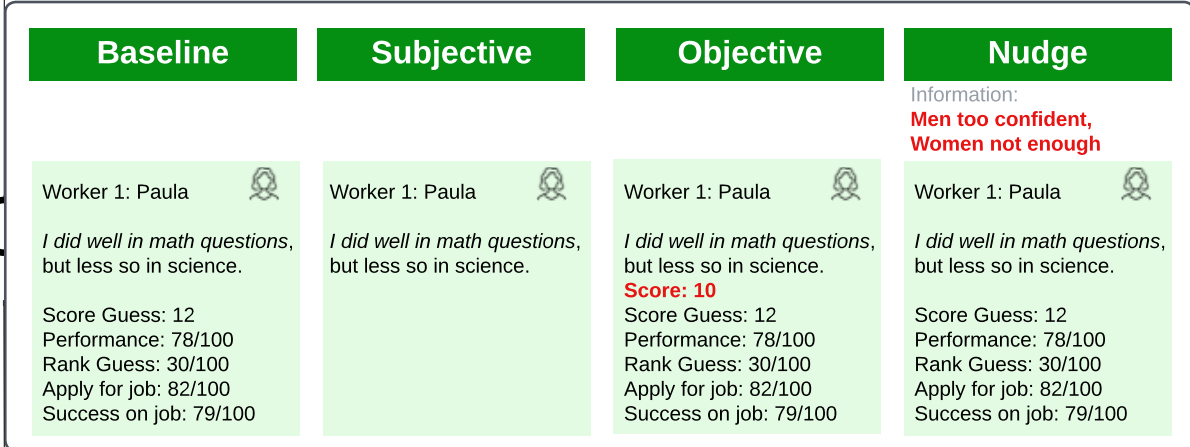
In what follows, we describe our experiment in sequential order. Figure 1 illustrates our experimental design, and all instructions and materials are presented in Appendix A.11. We previously conducted a pilot experiment, the results of which we report in Silva-Goncalves and van Veldhuizen (2020). The design and results of additional online pilots are described in Appendix A.8.

Figure 1: Overview of the Experimental Design

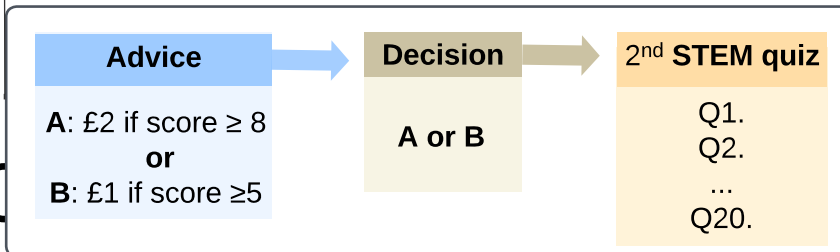
Part 1 Workers



Part 2 Advisers



Part 3 Workers



2.1 Self-Promotion Measurement (Workers)

Part one featured only participants assigned to the role of workers, who were informed that they would participate in two separate sessions. These participants started with a brief sociodemographic survey, which included questions on gender, age, education, employment, and income (see Appendix A.11 for the full survey). We then instructed participants to pick a name and avatar to represent themselves during the rest of the experiment. Available names and avatars were based on the person’s reported gender (identifiably female for women and identifiably male for men).³ Unbeknownst to the workers, we included this feature to give advisers a way to identify the gender of the workers. A similar approach has been used in experimental research; for example, by Gangadharan, Jain, Maitra, and Vecci (2016) and Mengel (2020).

Workers then completed a 20-question math and science quiz with multiple choice questions. They had 30 seconds for each question and were paid £0.10 per correct answer. Following Exley and Kessler (2022), we selected 20 questions from practice tests for the ASVAB (an aptitude test of the US Armed Services), and conducted several pilots to ensure that the questions were neither too easy nor too difficult. Upon completing the quiz, we asked workers to guess their performance. In order to encourage careful responses, we informed them that they would receive £0.50 for a correct guess and £0.20 if they were within 3 of their true score.⁴

As our next step, we informed workers that in the following week they would once again have to complete some kind of math and science task (Part 3 in Figure 1). We also told workers that they would be able to choose between two versions of the task (job A and job B) and that their payment would depend on their chosen task. However, we did not tell them anything about the task’s specific content or incentives. Instead, we informed workers that they would receive advice from another participant—the adviser—with knowledge of the nature of the two jobs and an incentive to give good advice. This feature captures the

³Participants who identified as non-binary in the sociodemographic survey could have chosen a set of identifiably male or female or gender-neutral names and avatars. In our experiment, none of the workers identified as non-binary. The full list of aliases and avatars is presented in Appendix 11.1.

⁴In order to reduce cheating, we presented the questions as images in order to prevent, e.g., copy-pasting questions into search engines.

information asymmetry between advisers and workers that makes advice invaluable in many applications.

Next, we informed workers that they would be asked to answer several self-evaluation questions. Importantly, they were told that one or more of these questions would be sent to the adviser to help them give the best possible advice. The questions we used are summarized in Table 1. The first five questions are taken from Exley and Kessler (2022), who find that men give more optimistic answers on each of these questions despite being worse at the task. The rank guess provides a measure of relative ability beliefs. Following Exley and Kessler (2022), the fact that the responses are sent to the adviser implies that they can be seen as measuring self-promotion. Part one concluded with two questions to measure general self-confidence and risk-aversion (as in Dohmen, Falk, Huffman, Sunde, Schupp, & Wagner, 2011), followed by a reminder to return for the next week's session.

Table 1: Self-Evaluation Questions

Measure	Question	Scale
Verbal Assessment	Describe how well you think you performed on the quiz.	Open-ended
Likert Performance	Indicate how good your performance was on the quiz	7-point Likert scale (“terrible” to “exceptional”)
Performed Well	I performed well on the quiz	0–100 scale (“entirely disagree” to “entirely agree”)
Would Apply for STEM Job	I would apply for a job that required me to perform well on the quiz.	0–100 scale (“entirely disagree” to “entirely agree”)
Would Succeed in STEM Job	I would succeed in a job that required skills tested on the quiz.	0–100 scale (“entirely disagree” to “entirely agree”)
Guessed Rank	How well do you think you performed compared with 100 other participants?	1 (best)–101 (worst) rank
Guessed Score	How many quiz questions do you think you answered correctly?	0-20

2.2 Advice Measurement (Advisers)

Part two featured only those participants assigned to the role of advisers. As a first step, we informed advisers that they would give advice to a worker about choosing between two jobs (A and B). Advisers then learned the specifics of the jobs, both of which involved a math and science quiz with twenty 30-second multiple-choice questions. The only difference between the two jobs was the payment scheme. Job A required workers to solve 8 questions to receive a bonus of £2. Job B had a lower threshold (5 questions) and lower bonus (£1). Workers who failed to reach their chosen threshold would receive no bonus. Advisers also learned that workers were unaware of these details, which rendered their advice invaluable.

For each adviser, there was a one in five chance that their advice would be delivered to a worker. When this happened, the adviser would receive an additional payment equal to half of the worker’s bonus. We introduced a monetary incentive to give good advice in order to increase advisers’ attention to the information provided to them when giving advice. A risk-neutral adviser should advise any worker with at least a 62.5% chance of making the higher threshold to pick the more ambitious job A.

On the next page, advisers were then presented with the name and avatar chosen by the worker. This introduced advisers to workers in a natural way and allowed advisers to clearly identify the gender of the worker, without making it too apparent that gender was the focus of our study. On the following page, advisers saw the worker’s answers to the self-evaluation questions. On this page we used gendered pronouns to reinforce the signal of the worker’s gender (all workers in our sample identified as either male or female). The information available to advisers differed by treatment, as we will explain later in this section.

Advisers saw the workers’ self-evaluation and were then given the opportunity to provide advice on the same page. The different types of advice are presented in Table 2. The first question asked the adviser for their advice on a Likert scale including “definitely job B”, “most likely job B”, “no advice”, “most likely job A”, and “definitely job A”. The second question was similar, but used a binary scale. Third, we asked advisers for their confidence in the binary question on a scale from 0 to 10. If the adviser’s advice was selected to be sent to a worker, one of these two pieces of advice (Likert scale advice or binary advice with confidence level) would randomly be chosen to be sent to the worker. On the next page, we then also asked advisers to imagine that instead of having to choose between two jobs, workers would be able to select their own threshold, and would earn a bonus of £0.20 times the threshold if they reached it. Advisers were asked which performance goal they would recommend the worker choose in this scenario. This item provides us with an additional, more continuous—but hypothetical—measure of advice.

Table 2: Advice Questions

Measure	Question	Scale
Likert Advice	Which advice do you want to send to [worker’s name]?	5-point Likert scale (“definitely B” to “definitely A”)
Binary Advice	Which job do you advise that [worker’s name] chooses?	Binary (0-B, 1-A)
Binary Advice Confidence	On a scale from 0 to 10, how confident are you that your advice is the best for [worker’s name]?	0-10
Performance Goal	<p>Now imagine that instead of having to choose between Job A (Advanced) and Job B (Basic), [worker’s name] would choose [his/her/their] performance goal as follows:</p> <ul style="list-style-type: none"> • Goal 1: £0.20 for 1 or more correct answers • (...) • Goal 20: £4.00 for 20 correct answers <p>Which goal would you advise [worker’s name] to choose in this case?</p>	1-20

After giving advice to their initial worker, advisers learned that they would be giving advice to four additional workers. We included these additional (randomly selected) workers to increase power and only revealed this to advisers at this point to allow us to treat the first worker as a one-shot observation in robustness checks. We informed advisers that each of the five workers was equally likely to receive their advice. Advisers read the information on each of the four additional workers sequentially, and gave their advice by answering the same questions as they had for the first worker.

At the end of the session, the advisers completed a sociodemographic survey. They

also answered questions about their general level of confidence and risk attitudes (as in Dohmen, Falk, Huffman, Sunde, Schupp, & Wagner, 2011) and their beliefs in regard to gender differences in performance and confidence. Finally, we asked advisers to choose a name, which was used when communicating their advice to workers.

The procedures detailed so far were followed in our four adviser-level treatments. Below, we outline the procedures specific to each treatment, which differ in terms of the information the adviser receives prior to giving advice to the workers.

2.2.1 Baseline Treatment

In the baseline treatment, advisers see workers' answers to all self-evaluation questions (listed in Table 1), with the exception of the Likert performance measure, which we left out to avoid duplicate information. This treatment captures the usual situation, in which objective performance cannot be perfectly observed by advisers. This forces them to rely on subjective indicators that may be influenced by the workers' self-promotion skills and the adviser's gender attitudes.

2.2.2 Subjective Treatment

The subjective treatment removes all quantitative self-evaluation measures from the information given to advisers. Advisers only see the workers' answer to the verbal open-ended question when determining their advice. We use this treatment to study whether providing advisers with only subjective, verbal information attenuates or increases the gender gap in advice. Attenuation is possible if gender differences in self-promotion are less pronounced or less easily detected in brief self-evaluation verbal statements than in quantitative ones. Yet the gender gap in advice may also increase if any gender biases held by advisers, driven by their beliefs or preferences, become more important by increasing subjectivity in self-evaluations, as suggested by previous work on motivated reasoning and subjective judgment (e.g., Bohren, Imas, & Rosenberg, 2019).

2.2.3 Objective Treatment

The objective treatment takes the opposite approach, by decreasing subjectivity in the worker portfolio. It presents advisers with the same information as in the baseline treatment and an additional objective indicator: the worker’s performance in part one. Access to this indicator will reduce the role of the subjective indicators, which may limit both the role of gender differences in self-promotion and the role of potential underlying gender biases among advisers. Providing more objective indicators has also often been suggested as an equity-improving policy.⁵

2.2.4 Nudge Treatment

The purpose of the nudge treatment is to test whether a gender gap in advice could be mitigated by making advisers aware of prior results on gender gaps in self-confidence just before they give advice. Advisers see the same information in the worker’s portfolio, as in the baseline treatment. In addition, prior to seeing the portfolio of their first worker, we presented advisers with three statements that describe situations in which a gender gap in self-confidence in a STEM setting led to unequal gender outcomes. Two of those statements summarize the results from two recent research studies, as follows:

1. Using data from more than 4,000 participants in online labor markets, researchers at Harvard University found that women rated their performance in a math and science quiz lower than men, despite similar performance levels. As a result, women who were hired performed better than men on average.
2. Using data from more than 3,600 people, researchers at University College London report that men are more confident in the self-assessment of their abilities than women. This gender gap in self-assessment contributes to the gender gap in who gets hired in a top job.

⁵For example, Heilman and Haynes (2005) and Cecchi-Dimeglio (2017) point out a potential role for subjectivity in generating gender bias in evaluations, and propose using objective measures to address this bias. Objective information has also long been proposed as a primary method to reduce statistical discrimination (Autor & Scarborough, 2008).

The first statement was based on Exley and Kessler (2022) and the second on Adamecz-Völgyi and Shure (2022). The third statement summarized a claim made by a major IT firm as follows:

3. An internal study by the IT firm Hewlett Packard found that men who apply for a job meet only 60% of the qualifications, but women who apply meet 100% of them.⁶

We asked two comprehension questions to ensure that advisers read the statements. We also asked advisers whether they thought the information might affect their advice to encourage them to reflect on the information. Although there was no explicit recommendation that advisers adjust their advice based on this information, this recommendation was implicit. Effectively, the information provided by the nudge encouraged advisers to discount self-evaluations from male workers and inflate those of female workers. In this sense, the nudge may capture both a pure “informational” effect and an “encouragement” effect (see, e.g., Kessel, Mollerstrom, and Van Veldhuizen, 2021). Both aspects are likely to be present in similar information campaigns in the real world as well, where the encouragement effect would come from the authorities providing the information rather than experimenters.

2.2.5 Treatment Assignment

We included the baseline, subjective, and nudge treatments as between-subject treatments. The objective treatment was implemented as a within-subject treatment. For this purpose, the advisers in the baseline treatment (but not the other treatments) were given the chance to re-evaluate their five original workers while knowing the workers’ actual score. We used a within-subject design for the objective treatment to increase statistical power, as detailed in our pre-analysis plan.

2.3 Second Quiz (Workers)

For part three, all workers who completed part one were invited to take part in a second session approximately one week after the first session. Workers received the advice from

⁶Source: “Why Women Don’t Apply for Jobs Unless They’re 100% Qualified,” Harvard Business Review (2014).

their matched adviser in binary or Likert-scale form and were then asked to make their decision (job A or B). After making their decision, workers learned both the nature of their task (twenty 30-second math and science questions for both tasks) and the incentives (A: £2 for solving at least 8 questions; B: £1 for solving at least 5 questions) of their chosen task. They were not told anything about the task they had not chosen. After completing the quiz, workers were given the option to provide general feedback.

2.4 Procedures

We conducted our experiment online on Prolific with 2,100 participants from the United Kingdom in July-August 2024. The experiment was programmed in oTree (Chen, Schonger, & Wickens, 2016). Throughout the experiment participants answered multiple comprehension questions. We recruited 350 participants (175 men and 175 women) to serve as workers. Our aim was to minimize the rate of worker attrition by making their bonus payments conditional on showing up a second time and by asking them to make a written commitment to complete the second part of the study (referenced as part three above). In the end, 320 out of 350 workers (91%) completed both parts of the study. Workers only received feedback on their performance on the two quizzes at the end of their second session. Workers earned an average of £8.60, which includes a fixed payment of £6 (£3 paid out after completing each part of the study), for a median time commitment of about 25 minutes.

The remaining 1,750 participants were assigned to the role of advisers. Of these, 700 advisers were assigned to the baseline/objective treatment, 525 to the subjective treatment, and 525 to the nudge treatment. We oversampled the baseline/objective treatment to increase power, since this treatment is used for all of our statistical tests. Advisers earned a fixed payment of £3 for completing the study, for a median time commitment of 15 minutes. Three hundred and fifty advisers, whose advice was randomly selected to be sent to workers, also received an additional payment of either 0, £0.50 cents or £1, depending on the workers' performance in part three.

To ensure that the adviser-level treatment effects are not confounded with order, day of the week, time of day, or any other effects, the treatments were randomly assigned to participants. In all treatments, advisers gave advice to the same group of 350 workers, such

that treatment differences in advice are not driven by random variation in workers across treatments. In each treatment, each adviser gave advice to 5 randomly selected workers. Our random assignment of workers to advisers mechanism ensured the following. In the baseline/objective treatment, each of the 350 workers was assessed by 10 advisers. Therefore, in the baseline treatment, we have a total of 3,500 adviser-worker pairings and observations. In the subjective and nudge treatments, 175 workers were assessed by 7 advisers and the other 175 workers were assessed by 8 advisers. In these treatments, we therefore have a total of 2,625 adviser-worker pairings and observations.

2.5 Main Hypotheses

We will test three pre-registered hypotheses.

Hypothesis 1 (Gender Gap in Self-promotion) *Female workers are less confident than male workers in their self-evaluations of performance on the STEM quiz.*

This auxiliary hypothesis replicates Exley and Kessler (2022)'s study, in which men revealed greater confidence than women in five self-evaluation questions about performance on a math and science quiz. This was true both when the evaluations were sent to a third party (and thus measure self-promotion) and when they were private. We hypothesize that we will find a similar gender gap in self-promotion in our data.

Hypothesis 2 (Gender Gap in Advice) *Female workers receive less ambitious advice than male workers in the baseline treatment.*

This is our first main hypothesis. We expect women to receive less ambitious advice for two potential reasons. The first directly relates to the gender gap in self-promotion: If women promote themselves less, we also expect advisers (who likely rely heavily on workers' self-evaluations) to give less ambitious advice to female workers. Second, advisers may give less ambitious advice to women, even when controlling for self-promotion, if advisers are influenced by stereotypes.

Hypothesis 3 (Treatment Differences) *The gender gap in advice is significantly smaller in the subjective, objective, and nudge treatments than in the baseline.*

This is our second main hypothesis. We expect all three treatments to reduce the gender gap in advice. For the subjective treatment, the gap in advice could either increase, due to increased scope for gender biases with less information, or decrease, due to reduced influence of the gender gap in self-promotion. We expect the latter effect to dominate and lead to a smaller gender gap in advice in this treatment relative to the baseline. For the objective treatment, we expect that making advice less subjective by informing advisers of workers' past performance will lead to a smaller gender gap in advice, relative to the baseline. For the nudge treatment, we expect that providing advisers with information on the general gender gap in self-confidence will encourage advisers to give more ambitious advice to women and less ambitious advice to men, relative to the baseline.

3 Results

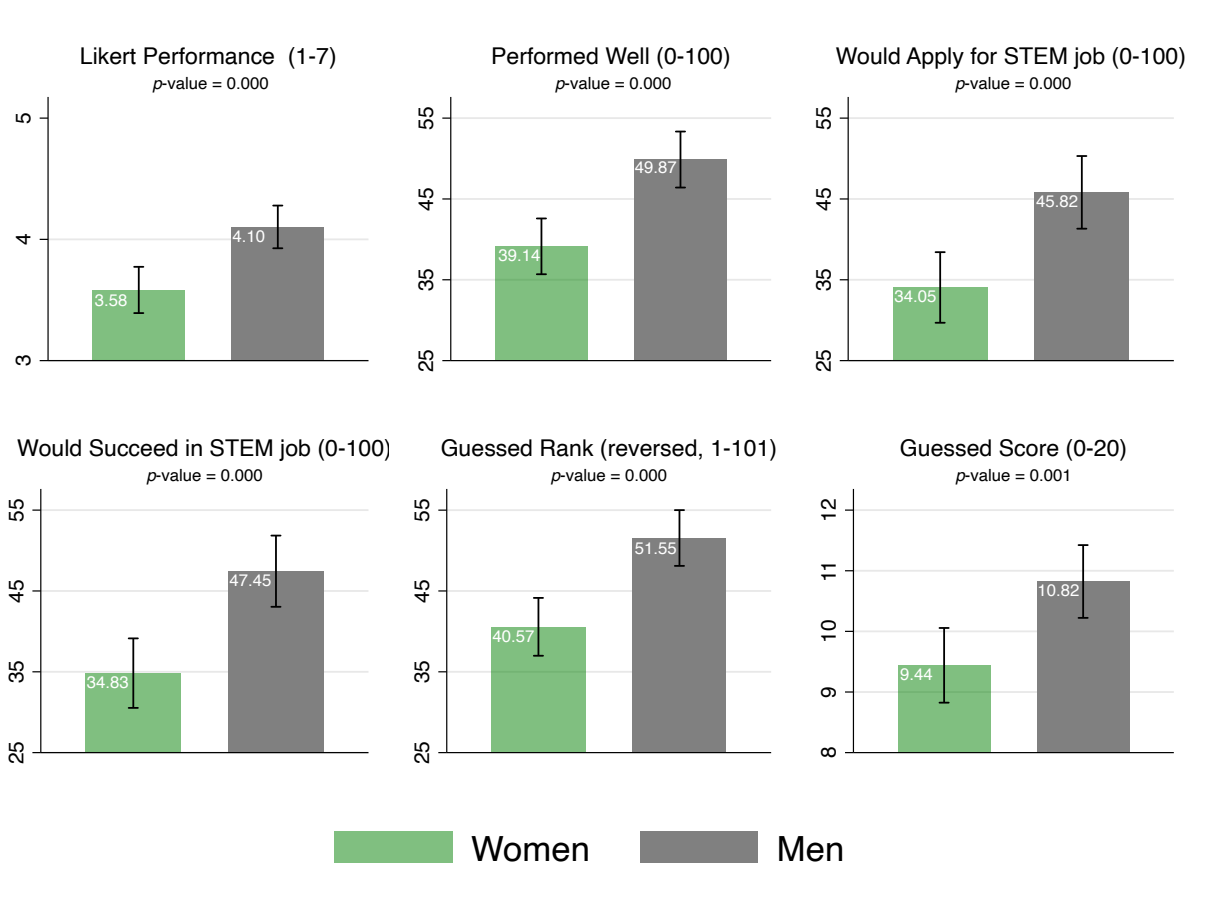
In this section we will present the pre-registered tests for our main hypotheses. We start with the gender gap in self-promotion, followed by the gender gap in advice and finally our treatment effects. We also include additional tests that serve to identify mechanisms, check robustness, or investigate other outcomes. All of our tests are pre-registered unless otherwise indicated. We present our pre-analysis plan and power calculations in Appendix A.9, alongside a brief description of where each of the pre-registered tests can be found in the paper. We use one-sided tests for every pre-registered test with a directional hypothesis, and two-sided tests in all other cases.

3.1 Gender Gap in Self-promotion

Figure 2 plots the average answer to each quantitative self-evaluation question, separately for men and women. In line with Exley and Kessler (2022) and our first hypothesis, women are less confident in all six measures (see also Table 3). Following Exley and Kessler (2022), we refer to this gender gap in self-evaluations as a gender gap in self-promotion because workers are aware that the evaluations are sent to an interested third party (in our case, the

adviser).⁷ In Table A.1 and Figure A.1 in Appendix A.1 we show that there is a positive correlation between a worker’s score and their self-evaluation and that the correlation does not differ by gender.

Figure 2: Gender Gap in Self-promotion



Notes: Mean self-promotion scores for women and men are shown for each of the six measures. Error bars denote 95% confidence intervals. p-values come from the corresponding tests in Table 3.

The magnitudes of the gender gap in self-promotion in our sample are similar to those of Exley and Kessler (2022), although slightly smaller in most variables. For example, the gender gap in “Would Apply for STEM job” in our sample is 12 percentage points, whereas in Exley and Kessler (2022) it was between 13 and 19 percentage points across their various samples. Similar gender differences in self-promotion are reported by Exley and Nielsen (2024) and Römer and Schröder (2025), as well as in studies with different math or science

⁷The exception is the score guess variable, which was elicited before workers were informed that their evaluations would be shared with an adviser.

tasks (e.g., Niederle and Vesterlund, 2007).

With the notable exception of Exley and Kessler (2022), these studies also tend to find that men outperform women on math and science quizzes, which is reasonable given that men tend to have more training in STEM. We also find that men outperform women by an average of 1 point (Table A.2). For our purpose, it is important that the gender difference in self-promotion persists when controlling for score fixed effects (Table A.3), and that advisers did not anticipate the gender gap in performance (Table A.19). In other words, the gender difference in self-evaluations captures both a gender difference in underlying performance and a gender difference in self-promotion.

Table 3: Gender Gap in Self-promotion

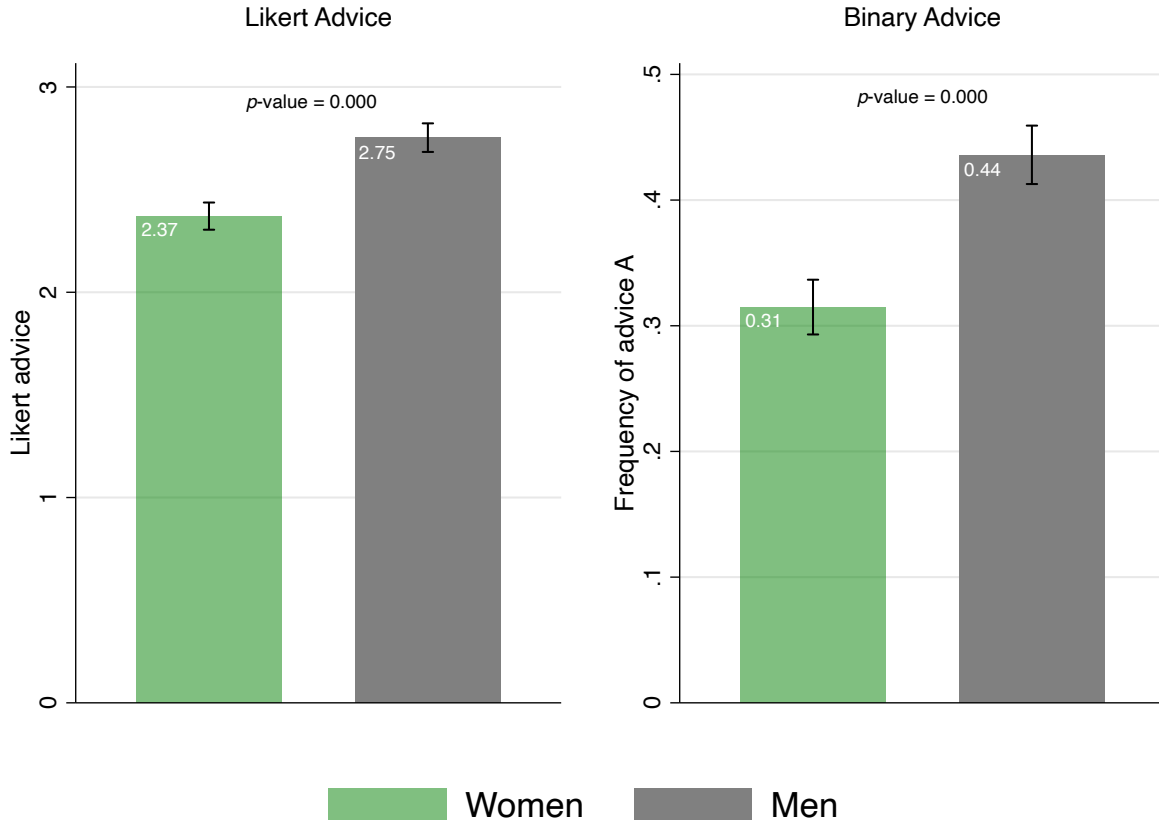
	Women		Men		Diff.	<i>p</i> -value
	Mean	SD	Mean	SD		
Likert Performance (1-7)	3.583	1.283	4.103	1.185	-0.520	0.000
Performed Well (0-100)	39.137	23.208	49.874	23.303	-10.737	0.000
Would Apply for STEM Job (0-100)	34.051	29.374	45.817	30.229	-11.766	0.000
Would Succeed in STEM Job (0-100)	34.834	28.901	47.446	29.629	-12.611	0.000
Gussed Rank (reversed, 1-101)	40.566	24.069	51.554	23.198	-10.989	0.000
Gussed Score (0-20)	9.440	4.143	10.823	4.037	-1.383	0.001

Notes: *p*-values come from pre-registered one-sided t-tests. The *Gussed Rank* variable is reversed, such that a higher value corresponds to a higher rank. The sample includes a total of 175 women and 175 men.

3.2 Gender Gap in Advice

We now turn to our main analysis and examine the extent to which there is a gender gap in advice in the baseline treatment. In line with our pre-analysis plan, we focus on the two main measures of advice (Likert and binary scale). Figure 3 plots the results. In line with our first hypothesis, women receive less ambitious advice on both the Likert scale (0.38 units on a 5-point scale) and on the binary measure (13 percentage points). Both differences are significant at the 1% level; the pre-registered test is presented in columns 1 and 3 of Table 4.

Figure 3: Gender Gap in Advice in the Baseline Treatment



Notes: 95% CIs shown. p -values derived from the analysis in columns 1 and 2 of Table 5. A higher value on the Likert scale means that advice is closer to A.

What explains the gender gap in advice? To answer this question, columns 2 and 4 in Table 4 control for all five quantitative self-evaluation questions.⁸ Two results stand out. First, the self-evaluation questions are very predictive of advice received. This suggests that advisers relied heavily on workers’ self-evaluations when giving advice. Second, the gender gap in advice vanishes after controlling for self-evaluations. In other words, the gender gap appears to have been driven entirely by gender differences in self-evaluations. The lack of a residual gender gap in advice also suggests that there was no intrinsic gender bias (e.g., due to stereotypes) among advisers in our population. Had such a bias existed, advisers

⁸We deviate from our pre-analysis plan by controlling for all five self-evaluation questions as opposed to only the guessed score. This allows us to capture all information available to advisers. When we instead control only for the guessed score, a significant residual gender gap does remain; see Table A.5.

would have given more ambitious advice to men even between men and women with similar self-evaluations.

Table 4: Drivers of the Gender Gap in Advice

<i>Advice Type (1-A)</i>	Likert		Binary	
	(1)	(2)	(3)	(4)
Female	-0.382*** (0.048)	0.022 (0.038)	-0.121*** (0.016)	0.004 (0.013)
Performed Well (0-100)		0.003* (0.002)		0.000 (0.001)
Gussed Rank (1-101)		0.006*** (0.002)		0.002*** (0.001)
Would Apply for STEM Job (0-100)		0.004** (0.002)		0.002*** (0.001)
Would Succeed in STEM Job (0-100)		0.008*** (0.002)		0.002*** (0.001)
Gussed Score (0-20)		0.113*** (0.008)		0.034*** (0.003)
Constant	2.704*** (0.058)	0.446*** (0.073)	0.408*** (0.020)	-0.285*** (0.024)
N	3,500	3,500	3,500	3,500
Clusters (Advisers)	700	700	700	700
R ²	0.019	0.432	0.017	0.382

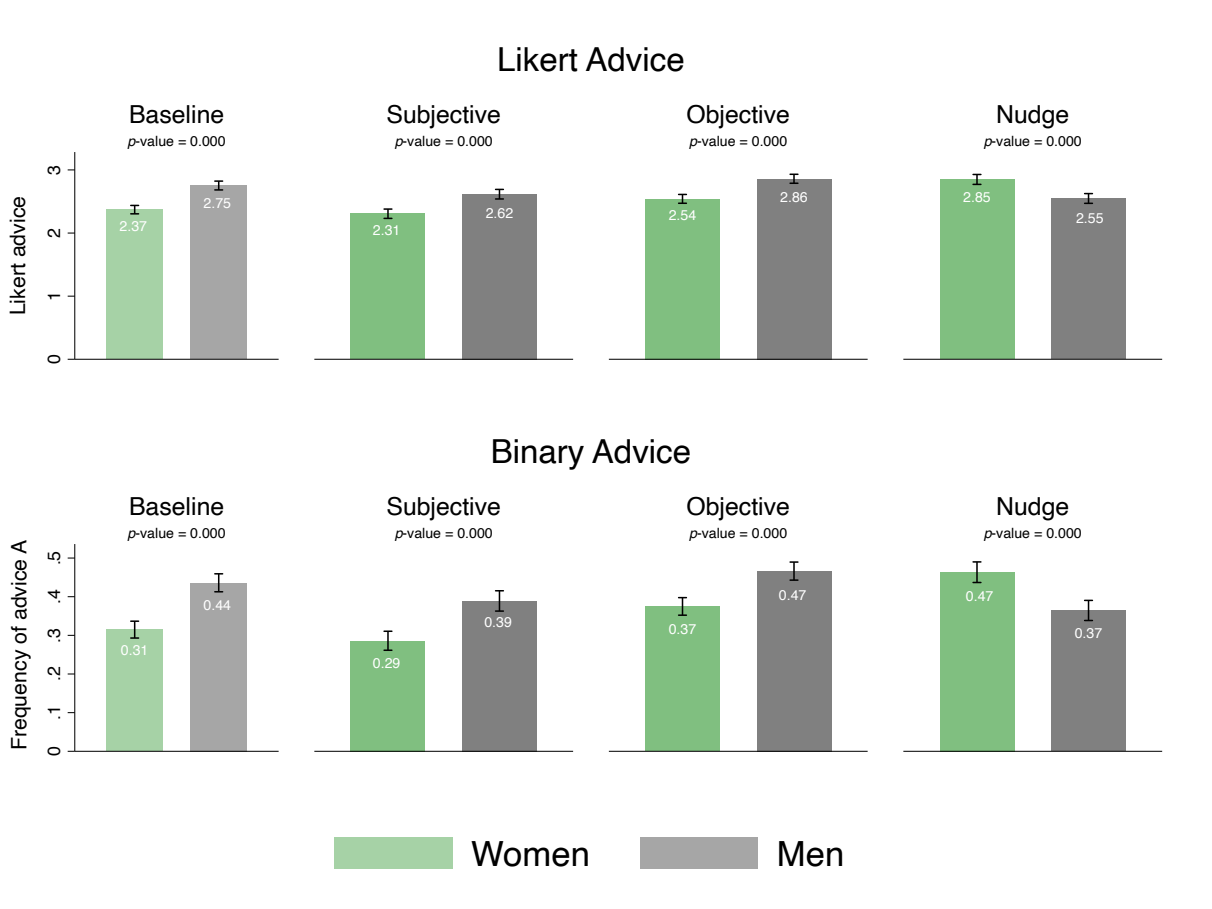
Notes: All specifications control for a worker’s order (1 to 5). *Female* is a binary indicator for the worker’s gender (0-male, 1-female). *Performed Well* indicates how much the worker agreed with having performed well in part 1 (from 0 to 100%). *Would Apply for STEM job* expresses willingness to apply for jobs that utilize similar skills as the STEM quiz (from 0 to 100%). *Would Succeed in STEM job* expresses agreement with expecting to perform well on jobs featuring similar tasks (from 0 to 100%). *Gussed Rank* indicates the expected rank compared with 100 other individuals, reverse coded such that 101 is the highest rank. *Gussed Score* is the worker’s gussed score in part 1. Standard errors are clustered at adviser level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3.3 Treatment Differences in Advice

We now turn to our third hypothesis and test whether the gender gap in advice can be attenuated by providing advisers with more or less information or through a nudge. Figure 4 presents the advice given to men and women separately for the four treatments and the two

types of advice. The biggest difference emerges in the nudge treatment: Informing advisers of prior results on gender gaps in confidence eliminates and even reverses the gender gap in advice. Indeed, women in the nudge treatment are 10 percentage points *more likely* than men to receive ambitious advice. Columns 7 and 8 in Table 5, which formally compare the gender gap in advice across treatments, confirm that these differences are statistically significant. In other words, making advisers aware of the gender gap in self-confidence is a highly effective way to remove and even reverse the gender gap in advice in our context.

Figure 4: Gender Gap in Advice by Treatment



Notes: Mean advice for women and men is shown for each treatment, separately for the Likert (upper panel) and binary (lower panel) advice measures. A higher value on the Likert scale means that the advice is closer to A. Vertical bars indicate 95% confidence intervals. p -values correspond to gender-difference tests and are derived from the analysis in Table 5.

Table 5: Gender Gap in Advice, by Treatment

<i>Treatment</i>	Baseline vs. <i>Treatment</i>							
	Baseline		Subjective		Objective		Nudge	
<i>Advice Type (1-A)</i>	Likert (1)	Binary (2)	Likert (3)	Binary (4)	Likert (5)	Binary (6)	Likert (7)	Binary (8)
Female	-0.382*** (0.048)	-0.121*** (0.016)	-0.382*** (0.048)	-0.121*** (0.016)	-0.382*** (0.048)	-0.121*** (0.016)	-0.382*** (0.048)	-0.121*** (0.016)
Subjective			-0.138** (0.056)	-0.047** (0.019)				
Objective					0.107*** (0.032)	0.030*** (0.011)		
Nudge							-0.205*** (0.058)	-0.068*** (0.019)
Female×Subjective			0.074 (0.072)	0.019 (0.024)				
Female×Objective					0.064* (0.043)	0.030** (0.014)		
Female×Nudge							0.683*** (0.077)	0.221*** (0.026)
Constant	2.704*** (0.058)	0.408*** (0.020)	2.710*** (0.050)	0.399*** (0.017)	2.732*** (0.056)	0.414*** (0.018)	2.722*** (0.050)	0.412*** (0.017)
Gender Gap in <i>Treatment</i>								
Coefficient			-0.308*** (0.053)	-0.102*** (0.018)	-0.318*** (0.049)	-0.091*** (0.016)	0.301*** (0.060)	0.099*** (0.021)
Standard error								
N	3,500	3,500	6,125	6,125	7,000	7,000	6,125	6,125
Clusters (Advisers)	700	700	1,225	1,225	700	700	1,225	1,225
R ²	0.019	0.017	0.018	0.018	0.017	0.015	0.018	0.017

Notes: All specifications control for worker order (1 to 5). *Female* is a binary indicator for the workers' gender (0-male, 1-female). The other variables are binary indicators for the respective treatments. Standard errors are clustered at the adviser level. Statistical significance of the coefficient estimates for the interaction terms is based on a one-sided test for a positive coefficient, in line with our pre-analysis plan. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Results for the other two treatments are less clear-cut. Decreasing subjectivity in advice by informing advisers of the workers' true performance reduces the gender gap in advice by 0.06 units (p -value=0.06) or 3 percentage points (p -value=0.02) on the Likert and binary scale, respectively (see columns 5 and 6 in Table 5). Still, the gender gap in the objective treatment remains sizable and significant. This residual gap is driven by gender differences in performance; when we control for performance differences, the gender gap in advice is small and no longer significant (see Table A.4 in Appendix). In other words, while advisers rely heavily on objective performance indicators, doing so does not greatly reduce the gender

gap in advice due to men’s superior performance. Our regression results also imply that the objective treatment would have been more effective in a setting with no gender gap in performance (or a superior performance among women, as in Exley and Kessler, 2022).

Finally, we find no evidence that the gender gap in advice is reduced by increasing subjectivity in advice, when advisers are only provided with qualitative self-evaluation measures (Table 5, columns 3-4). To shed further light on this result, we used Microsoft Copilot (version 365) to rate the confidence level expressed by the qualitative performance question (see Appendix A.3 for more details). Table A.7 shows that the gender gap in advice vanishes when controlling for these AI ratings. This suggests that women also promoted themselves less on the qualitative question, which in turn fully explains why women receive less ambitious advice in this treatment. Importantly, it demonstrates that gender gaps in advice driven by gender differences in self-promotion can arise even when self-promotion is limited to brief (generally, one or two sentences in our sample) self-evaluation narratives. These results also imply that increased subjectivity does not generate an increased intrinsic gender bias among our advisers.

In summary, we find that the gender gap in advice is remarkably robust to increasing subjectivity in advice, by providing advisers only with workers’ brief verbal self-evaluation statements (subjective treatment). The gender gap in advice is also robust to decreasing subjectivity in advice, by providing advisers with information on workers actual performance (objective treatment). However, decreasing subjectivity in advice by providing advisers with factual information about the general gender gap in confidence affects the gender gap in advice, eliminating and even reversing the gap.

3.4 Worker Earnings

What do these differences in advice imply for worker earnings? As a first step, it is important to note that, as expected, the vast majority of workers (97%) followed the advice received (see Appendix A.5 for more details). Rather than looking at realized earnings (which depend on the adviser randomly matched with each worker), we examine *expected* earnings, based on all the advice generated for a particular worker (across all advisers), assuming that workers would have followed the advice. We compute expected earnings using both their initial (part

one) and realized (part three) scores. Table 6 presents results based on workers' initial score (part one) using either (1) a dummy for whether the advice maximized the worker's expected earnings (our pre-registered test, columns 1 and 2); (2) expected earnings when following the advice (columns 3 and 4, not pre-registered); and (3) the gap between expected earnings based on following the advice and what a risk-neutral worker would have done based on their part one score guess, that is, choose job A if their score guess in part one was 8 or greater (columns 5 and 6, not pre-registered).

Our results show that in all three measures, expected earnings were highest when advisers knew the worker's part one score (objective treatment) and lowest when they only had access to brief open-ended self-assessments of performance (subjective treatment). In other words, the more information advisers had, the better their advice, which is intuitive. Columns 2, 4, and 6 show that these effects are quite similar for men and women. These columns also show that the nudge treatment reduces the expected earnings of men (although not statistically significant in column 6) while increasing those of women. This result is driven by a tendency for advisers to recommend job A at suboptimally low rates (see Appendix A.8). The nudge offsets this tendency for female workers, which increases their expected earnings and hence the quality of advice. These results are similar when we calculate expected earnings based on the part three score instead (Table A.9, not pre-registered).

Table 6: Expected Earnings Based on Part 1 Score

	Advice Maximized Expected Earnings		Expected Earnings Based on Advice		Gap in Expected Earnings With vs. Without Advice	
	(1)	(2)	(3)	(4)	(5)	(6)
Female		-0.027* (0.017)		-0.161*** (0.018)		0.034* (0.018)
Subjective	-0.032** (0.013)	-0.046** (0.019)	-0.032** (0.015)	-0.045** (0.020)	-0.035** (0.016)	-0.047** (0.022)
Objective	0.087*** (0.008)	0.073*** (0.011)	0.083*** (0.008)	0.073*** (0.011)	0.083*** (0.008)	0.073*** (0.011)
Nudge	0.022 (0.013)	-0.038** (0.019)	0.024* (0.015)	-0.040** (0.020)	0.027* (0.016)	-0.035 (0.022)
Female×Subjective		0.028 (0.026)		0.025 (0.027)		0.025 (0.029)
Female×Objective		0.029** (0.014)		0.022 (0.014)		0.022 (0.014)
Female×Nudge		0.118*** (0.026)		0.129*** (0.029)		0.123*** (0.028)
Constant	0.469*** (0.014)	0.482*** (0.016)	1.262*** (0.015)	1.343*** (0.018)	-0.270*** (0.016)	-0.287*** (0.018)
N	12,250	12,250	12,250	12,250	12,250	12,250
Clusters (Advisers)	1,750	1,750	1,750	1,750	1,750	1,750
R ²	0.009	0.011	0.007	0.021	0.007	0.012

Notes: All specifications control for a worker’s order (1 to 5). *Female* is a binary indicator for the worker’s gender (0-male, 1-female). The other variables are binary indicators for the respective treatments. Standard errors are clustered at adviser level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3.5 Robustness Checks

We report the results of our pre-registered robustness checks in Appendix A.6. First, we show that the gender gap in self-promotion is not affected by controlling for the worker’s sociodemographic characteristics (Table A.10). We also find that workers who are less educated and have low income satisfaction are less confident in their math and science abilities. Second, we test for a gender gap in advice (in the baseline treatment), controlling for adviser sociodemographic characteristics (Table A.11). We find no differences relative to our main

results, although more risk-averse advisers are less likely to recommend job A. Finally, we test for treatment differences in the gender gap in advice (1) using alternative measures of advice (Table A.12) and (2) restricting our analysis to only the first worker advised (Table A.13). The main difference relative to our main results is that the reduction in the gender gap in advice in the objective treatment is no longer statistically different from the baseline gender gap when using alternative measures of advice or when restricting to the first worker.

3.6 Additional Results

We report several additional results in Appendix A.7. First, we show that baseline advisers tend to be less confident in their advice when recommending job A, and less confident about their advice to women when recommending job A (Table A.14), driven by women’s lower confidence (Table A.15). Second, we show using two questionnaire variables that advisers in the baseline, objective, and subjective treatments expect women, on average, to be less confident than men in math and science tasks but to have similar performance. However, we do not see that advisers who believe men to be more confident adjust their advice in a meaningful way (Table A.16). Moreover, we observe that the gender gap in advice is largest among advisers who believe that men perform better than women, and non-significantly different from zero among advisers who believe that women perform better than men (Table A.17). We also see that the nudge treatment successfully shifted these views in the expected direction: The fraction of advisers who believe men to be more confident increases by 25 percentage points, and the fraction who believe women to be superior performers increases by 9 percentage points (Table A.18 and Table A.19). Third, we show that male and female advisers tend to give similar advice, with the exception of the nudge treatment, in which female advisers are more likely to give more ambitious advice to women. This suggests that the nudge was particularly effective in convincing female advisers to change their advice (Table A.20, not pre-registered).

4 Discussion

In this study, we examined the role of self-promotion and career advice in sustaining gender differences in labor market outcomes. We conducted an online experiment in which “advisers” advise “workers” to choose between a more ambitious and a less ambitious task based on the worker’s subjective self-promotion. We find that women have lower self-promotion and receive less ambitious career advice as a result. We also show that these gender differences are similar for both quantitative and verbal self-assessments, and that the gender difference in advice received can be mitigated by informing advisers of the workers’ true performance or of the gender gap in self-promotion.

Our results suggest that gender differences in self-promotion may play a key role in generating gender differences in career advice. This may also help explain differences between our results and those of Brandts, Groenert, and Rott (2015) and Gallen and Wasserman (2021). Those studies find that men and women receive similarly ambitious advice in a tournament entry laboratory experiment and a professional career advice setting, respectively. Yet whereas advisers in those papers relied on fixed messages written by the experimenter (Gallen & Wasserman, 2021) or on the worker’s true performance (Brandts, Groenert, & Rott, 2015; Brandts & Rott, 2021), our advisers had access to workers’ self-evaluations, which differed significantly by gender. Consistent with this, the gender gap in advice in our study vanishes entirely when controlling for gender differences in self-promotion. This highlights a potential key role for self-promotion in potentially inducing even unbiased, well-meaning, and paternalistic advisers to give different advice to men and women.

By contrast, we find no evidence that advisers are intrinsically biased in the sense of giving different advice to men and women with similar levels of stated confidence. Interestingly, this occurs despite the fact that the average adviser believes that men are more confident than women, which implies that men “should” get *less* ambitious advice. This latter result is consistent with Exley and Nielsen (2024), who find that evaluators incorporate male and female confidence signals similarly even when they are aware that men are more overconfident. The limited role of an intrinsic bias may also explain why giving advisers access to more or less objective information did not impact the gender gap in advice.

One result that distinguishes our study from Exley and Kessler (2022) (but not Exley and Nielsen, 2024, Römer and Schröder, 2025, or Chang, Saccardo, and Gallus, 2025) is that male workers significantly outperform women. This result implies that greater male self-promotion is partially justified in our setting. At the same time, the gender gap in self-promotion remains significant when we control for performance. More importantly, advisers did not anticipate the gender performance gap, and therefore evaluated workers “as-if” there were no gender gap in performance. This implies that our setting was functionally similar to a setting with no gap in performance, as might be expected in most real-world applications.

Although we constructed our setting to capture key aspects involved in generating gender gaps in career advice, our setting may differ from important applications in key ways. For example, real-world advisers may have more deeply ingrained gender stereotypes than our advisers, based on many years of work experience. This may limit the role of self-promotion and increase the role of intrinsic gender bias among these advisers. At the same time, their experience may also increase the quality of advice, and thus limit the role of gender bias. An interesting follow-up would implement our paradigm with experienced advisers and test whether they are more or less susceptible to gender differences in self-promotion when forming their advice.

Our results also have potentially important policy implications. The substantial gender gap in self-promotion indicates that women underestimate their abilities compared with their male counterparts, even when their performance is equal. This self-perception gap is critical, because it influences both the advice women receive and the decisions they make about their own career paths. We show that in our context these differences can be neutralized by informing advisers of prior research on gender gaps in self-confidence and self-promotion. Another approach would be to boost women’s self-promotion through mentorship programs, skill-building workshops, and related programs.

References

- Adamecz-Völgyi, A., & Shure, N. (2022). The gender gap in top jobs—the role of overconfidence. *Labour Economics*, *79*, 102283.
- Angelov, N., Johansson, P., & Lindahl, E. (2016). Parenthood and the gender gap in pay. *Journal of Labor Economics*, *34*(3), 545–579. <https://doi.org/10.1086/684851>
- Autor, D. H., & Scarborough, D. (2008). Does job testing harm minority workers? Evidence from retail establishments. *The Quarterly Journal of Economics*, *123*(1), 219–277.
- Bandiera, O., Parekh, N., Petrongolo, B., & Rao, M. (2022). Men are from Mars, and women too: A Bayesian meta-analysis of overconfidence experiments. *Economica*. <https://doi.org/10.1111/ecca.12407>
- Bertrand, M., Goldin, C., & Katz, L. F. (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, *2*(3), 228–255.
- Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, *55*(3), 789–865.
- Bohren, J. A., Imas, A., & Rosenberg, M. (2019). The dynamics of discrimination: Theory and evidence. *American Economic Review*, *109*(10), 3395–3436.
- Brandts, J., Groenert, V., & Rott, C. (2015). The impact of advice on women’s and men’s selection into competition. *Management Science*, *61*(5), 1018–1035.
- Brandts, J., & Rott, C. (2021). Advice from women and men and selection into competition. *Journal of Economic Psychology*, *82*, 102333. <https://doi.org/10.1016/j.joep.2020.102333>
- Carlana, M. (2019). Implicit stereotypes: Evidence from teachers’ gender bias. *The Quarterly Journal of Economics*, *134*(3), 1163–1224.
- Cecchi-Dimeglio, P. (2017). How gender bias corrupts performance reviews, and what to do about it. *Harvard Business Review*.
- Çelen, B., Kariv, S., & Schotter, A. (2010). An experimental test of advice and social learning. *Management Science*, *56*(10), 1687–1701.

- Chang, J., Saccardo, S., & Gallus, J. (2025). Leaning in softly: On gender and self-promotion. *Available at SSRN 5240652*.
- Chaudhuri, A., Schotter, A., & Sopher, B. (2009). Talking ourselves to efficiency: Coordination in inter-generational minimum effort games with private, almost common and common knowledge of advice. *The Economic Journal*, *119*(534), 91–122.
- Chen, D. L., Schonger, M., & Wickens, C. (2016). Otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97.
- Cortés, P., & Pan, J. (2023). Children and the remaining gender gaps in the labor market. *Journal of Economic Literature*, *61*(4), 1359–1409.
- Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, *47*(2), 448–474.
- 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.
- Eagly, A. H., & Carli, L. L. (2003). The female leadership advantage: An evaluation of the evidence. *The leadership quarterly*, *14*(6), 807–834.
- Exley, C. L., & Kessler, J. B. (2022). The gender gap in self-promotion. *The Quarterly Journal of Economics*, *137*(3), 1345–1381.
- Exley, C. L., & Nielsen, K. (2024). The gender gap in confidence: Expected but not accounted for. *American Economic Review*, *114*(3), 851–885.
- Falk, A., Kosse, F., & Pinger, P. (2025). Mentoring and schooling decisions: Causal evidence. *Journal of Political Economy* (Forthcoming).
- Gallen, Y., & Wasserman, M. (2021). Informed choices: Gender gaps in career advice.
- Gangadharan, L., Jain, T., Maitra, P., & Vecci, J. (2016). Social identity and governance: The behavioral response to female leaders. *European Economic Review*, *90*, 302–325.
- Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, *104*(4), 1091–1119.
- Goulas, S., Griselda, S., & Megalokonomou, R. (2024). Comparative advantage and gender gap in stem. *Journal of Human Resources*, *59*(6), 1937–1980.

- Goulas, S., Gunawardena, B., Megalokonomou, R., & Zenou, Y. (2024). Gender role models in education. *CESifo Working Paper No. 11330*.
- Guillen, P., & Hakimov, R. (2018). The effectiveness of top-down advice in strategy-proof mechanisms: A field experiment. *European Economic Review, 101*, 505–511.
- Heilman, M. E., & Haynes, M. C. (2005). No credit where credit is due: Attributional rationalization of women’s success in male-female teams. *Journal of Applied Psychology, 90*(5), 905.
- Hoxby, C. M., & Turner, S. (2015). What high-achieving low-income students know about college. *American Economic Review, 105*(5), 514–517.
- Kessel, D., Mollerstrom, J., & Van Veldhuizen, R. (2021). Can simple advice eliminate the gender gap in willingness to compete? *European Economic Review, 138*, 10377.
- Krawczyk, M., & Smyk, M. (2016). Author’s gender affects rating of academic articles: Evidence from an incentivized, deception-free laboratory experiment. *European Economic Review, 90*, 326–335.
- Lavy, V., & Megalokonomou, R. (2024). The short- and the long-run impact of gender-biased teachers. *American Economic Journal: Applied Economics, 16*(2), 176–218. <https://doi.org/10.1257/app.20210052>
- Mengel, F., Sauermann, J., & Zölitz, U. (2019). Gender bias in teaching evaluations. *Journal of the European Economic Association, 17*(2), 535–566.
- Mengel, F. (2020). Gender differences in networking. *The Economic Journal, 130*(630), 1842–1873.
- 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.
- Niederle, M. (2016). Gender. In J. Kagel & A. E. Roth (Eds.), *Handbook of Experimental Economics, 2nd ed.* (pp. 481–553). Princeton University Press.
- Oreopoulos, P., & Petronijevic, U. (2013). Making college worth it: A review of the returns to higher education. *The Future of Children, 23*(1), 41–65.
- Özgümüs, A., Rau, H. A., Trautmann, S. T., & König-Kersting, C. (2020). Gender bias in the evaluation of teaching materials. *Frontiers in Psychology, 11*, 1074.

- Porter, C., & Serra, D. (2020). Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics*, 12(3), 226–254.
- Reuben, E., Wiswall, M., & Zafar, B. (2017). Preferences and biases in educational choices and labour market expectations: Shrinking the black box of gender. *The Economic Journal*, 127(604), 2153–2186.
- Römer, N., & Schröder, M. (2025). *Effects of written self-promotion on gender bias and decision quality* (tech. rep.). Hannover Economic Papers (HEP).
- Sarsons, H., Gërxhani, K., Reuben, E., & Schram, A. (2021). Gender differences in recognition for group work. *Journal of Political Economy*, 129(1), 101–147.
- Schotter, A., & Sopher, B. (2003). Social learning and coordination conventions in intergenerational games: An experimental study. *Journal of Political Economy*, 111(3), 498–529.
- Schotter, A., & Sopher, B. (2007). Advice and behavior in intergenerational ultimatum games: An experimental approach. *Games and Economic Behavior*, 58(2), 365–393. <https://doi.org/10.1016/j.geb.2006.03.005>
- Silva-Goncalves, J., & van Veldhuizen, R. (2020). Subjective judgment and gender bias in advice: Evidence from the laboratory. *Life Course Centre Working Paper*, (2020-29).
- Wolter, S. C., & Zöllner, T. (2024). *Are parents an obstacle to gender-atypical occupational choices?* (Tech. rep.). IZA Discussion Papers.

A Appendix

A.1 Gender Gap in Performance

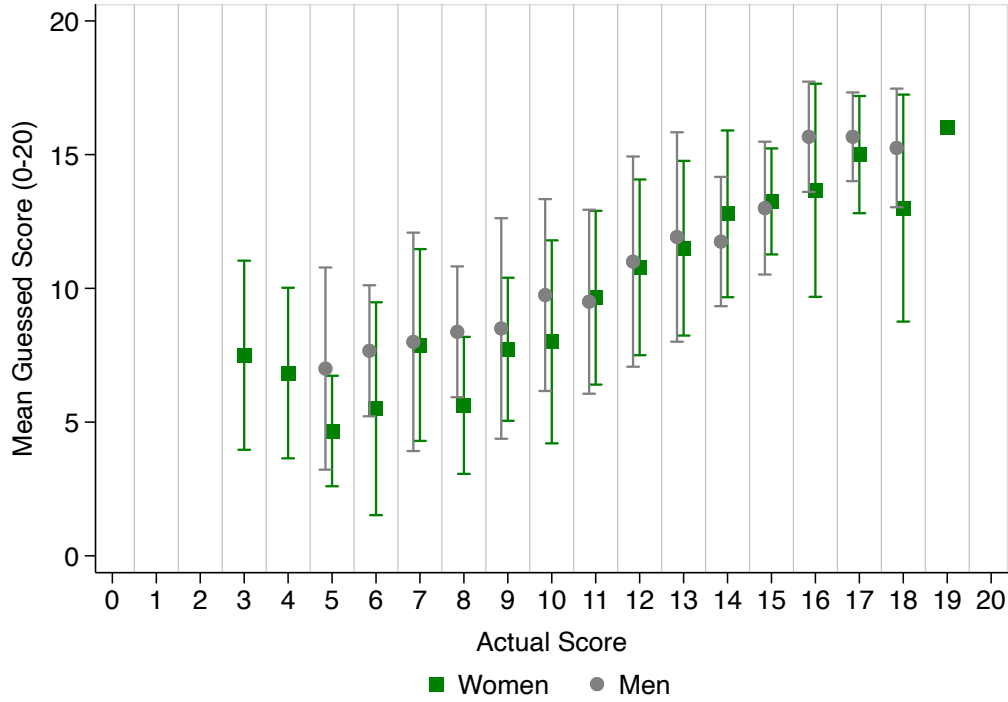
Table A.1 and Figure A.1 (not pre-registered) show that the relationship between the actual score and self-promotion does not differ by gender. The estimated coefficients of the interaction terms between gender and (actual) score for the different outcomes are small and statistically indistinguishable from 0. Table A.2 presents the means, standard deviations, differences of the means, and P-values of differences in worker scores on the 1st and 2nd quizzes for men and women. The scores range from 0 to 20. Men outperform women in both cases by about 1 point (out of 20). Table A.3 (pre-registered) shows that the gender gap in self-promotion remains significant when we include score fixed effects. This is shown by the negative and statistical significant coefficients of the variable “Female” across the different outcomes. Table A.4 (not pre-registered) shows that the gender gap in advice in the baseline treatment remains significant when controlling for score fixed effects, and that the gender gap in advice in the objective treatment vanishes when controlling for workers’ performance in the STEM quiz in part 1.

Table A.1: Relationship Between Actual Score and Self-promotion

	Likert Performance (1-7)	Performed Well (0-100)	Would Apply for STEM job (0-100)	Would Succeed in STEM job (0-100)	Gussed Rank (1-101)	Gussed Score (0-20)
	(1)	(2)	(3)	(4)	(5)	(6)
Score	0.164*** (0.024)	3.352*** (0.460)	3.346*** (0.629)	3.073*** (0.610)	2.845*** (0.479)	0.713*** (0.075)
Female	-0.801** (0.388)	-10.646 (7.310)	-12.532 (9.991)	-21.014** (9.696)	-17.204** (7.609)	-0.860 (1.198)
Female × Score	0.043 (0.033)	0.327 (0.631)	0.408 (0.862)	1.106 (0.836)	0.876 (0.656)	0.022 (0.103)
Constant	2.208*** (0.293)	11.106** (5.530)	7.119 (7.559)	11.903 (7.336)	18.654*** (5.756)	2.573*** (0.906)
N	350	350	350	350	350	350
R ²	0.300	0.304	0.197	0.222	0.271	0.381

Notes: *Female* is a binary indicator for the worker’s gender (0-male, 1-female). *Score* is a continuous variable in the range [0,20] and captures the worker’s actual score. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A.1: Mean and Standard Deviation of Guessed Score by Actual Score by Gender



Notes: Markers show mean guessed scores by actual score, with vertical lines indicating the corresponding standard deviations. Circles correspond to means for men, and squares correspond to means for women.

Table A.2: Gender Gap in Performance

	Women		Men		Diff.	<i>p</i> -value
	Mean	SD	Mean	SD		
Score in 1 st Quiz, <i>Before Advice</i>	10.514	3.510	11.566	3.291	-1.051	0.004
Score in 2 nd Quiz, <i>After Advice</i>	12.062	3.592	13.044	3.654	-0.981	0.016

Notes: The sample “*Before Advice*” is based on part 1, and includes 175 women and 175 men. The sample “*After Advice*” is based on part three, and includes 160 women and 160 men.

Table A.3: Gender Gap in Self-promotion Controlling for Performance

	Likert Performance (1-7)	Performed Well (0-100)	Would Apply for STEM job (0-100)	Would Succeed in STEM job (0-100)	Gussed Rank (1-101)	Gussed Score (0-20)
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.359*** (0.118)	-7.357*** (2.202)	-8.361*** (2.977)	-8.942*** (2.903)	-7.725*** (2.298)	-0.733** (0.361)
Constant	8.233*** (2.341)	3.359*** (0.763)	46.857*** (14.280)	48.775*** (14.899)	56.861*** (19.304)	56.942*** (18.824)
Score FE	✓	✓	✓	✓	✓	✓
N	350	350	350	350	350	350
R ²	0.319	0.333	0.248	0.264	0.298	0.407

Notes: *Score FE* (fixed effects) controls for actual performance. *Female* is a binary indicator for the worker’s gender (0-male, 1-female). Statistical significance of the coefficient estimate for the binary indicator *Female* is based on a one-sided test. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Gender Gap in Advice Controlling for Performance

<i>Treatment</i>	Baseline		Subjective		Objective		Nudge	
	Likert (1)	Binary (2)	Likert (3)	Binary (4)	Likert (5)	Binary (6)	Likert (7)	Binary (8)
Female	-0.225*** (0.047)	-0.075*** (0.016)	-0.192*** (0.052)	-0.063*** (0.018)	-0.025 (0.040)	-0.005 (0.014)	0.433*** (0.055)	0.140*** (0.019)
Constant	2.226*** (0.306)	0.297*** (0.100)	1.700*** (0.243)	0.088 (0.072)	1.131*** (0.083)	0.088 (0.069)	2.366*** (0.385)	0.309** (0.132)
Score FE	✓	✓	✓	✓	✓	✓	✓	✓
N	3,500	3,500	2,625	2,625	3,500	3,500	2,625	2,625
Clusters (Advisers)	700	700	525	525	700	700	525	525
R ²	0.203	0.186	0.125	0.115	0.457	0.388	0.160	0.142

Notes: All specifications control for worker order (1 to 5). *Female* is a binary indicator for the workers’ gender (0-male, 1-female). Standard errors are clustered at the adviser level. *Score FE* (fixed effects) controls for actual performance. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.2 Gender Gap in Advice in Baseline Treatment

In line with the pre-analysis plan, we report the analysis of the drivers of the gender gap in advice in the baseline treatment, controlling for only one self-evaluation measure—namely,

participants’ guessed score. The results show that the gender gap in advice is strongly attenuated, with a small gender gap persisting. Since this gap is less than one-third of the original gap and since it vanishes completely when we also control for the other self-evaluation variables, we decided to report the results that control for all variables available to advisers in the main text.

Table A.5: Drivers of the Gender Gap in Advice

<i>Advice Type (1-A)</i>	Likert (1)	Binary (2)
Female	-0.082** (0.041)	-0.029** (0.014)
Guessed Score (0-20)	0.217*** (0.005)	0.067*** (0.002)
Constant	0.356*** (0.075)	-0.316*** (0.024)
N	3,500	3,500
Clusters (Advisers)	700	700
R ²	0.384	0.334

Notes: All specifications control for a worker’s order (1 to 5). *Female* is a binary indicator for the worker’s gender (0-male, 1-female). *Guessed Score* is the worker’s guessed score in part 1. Standard errors are clustered at adviser level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.3 Gender Gap in Qualitative Self-promotion

In line with the pre-analysis plan, we use AI to quantify responses to the open-ended self-evaluation question. Most workers answered this question with one or two sentences. We explore the extent to which there are gender differences in the qualitative measure of self-promotion using AI tools (Copilot for Microsoft 365), which allow us to quantify the expressed level of confidence in the statements. Specifically, we asked Copilot to evaluate the statements according to how confident the worker is about their performance (on a scale from 0 to 10), to guess the likelihood that the worker gave more than 8 correct answers (the threshold for success in job A), to guess the actual score of the worker, and to rate the statement based on the quality of the writing (on a scale from 0 to 10). The results show that (a)

women were also significantly less confident in the open-ended question (Table A.6, pre-registered) and (b) these gender differences can fully explain the gender gap in advice in the subjective treatment (Table A.7, not pre-registered). Note that we did not pre-register the exact question, or the specific AI used to quantify the qualitative self-promotions.

Table A.6: Gender Gap in Confidence Based on AI Ratings of the Qualitative Performance Statements

	AI Confidence in Performance (0-10)	AI More Than 8 Correct	AI Score Guess (0-20)	AI Writing Quality (0-10)
	(1)	(2)	(3)	(4)
Female	-0.367*** (0.134)	-0.421*** (0.145)	-0.888*** (0.266)	-0.226*** (0.084)
Constant	3.367*** (0.870)	3.421*** (0.939)	6.888*** (1.727)	5.226*** (0.545)
Score FE	✓	✓	✓	✓
N	350	350	350	350
R^2	0.296	0.247	0.285	0.281

Notes: *Female* is a binary indicator for the worker's gender (0-male, 1-female). *Score FE* (fixed effects) controls for actual performance. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: Drivers of the Gender Gap in Advice in Subjective Treatment

Advice A Measure:	Likert		Binary	
	(1)	(2)	(3)	(4)
Female	-0.308*** (0.053)	-0.020 (0.047)	-0.102*** (0.018)	-0.014 (0.016)
AI Confidence in Performance		0.566*** (0.059)		0.197*** (0.021)
AI More Than 8 Correct		0.028 (0.044)		0.011 (0.016)
AI Score Guess		-0.043 (0.041)		-0.027* (0.015)
AI Writing Quality		0.067 (0.049)		0.030* (0.017)
Constant	2.579*** (0.065)	-0.005 (0.199)	0.339*** (0.022)	-0.496*** (0.066)
N	2,625	2,625	2,625	2,625
Clusters (Advisers)	525	525	525	525
R^2	0.015	0.307	0.016	0.269

Notes: All specifications control for worker order (1 to 5). *Female* is a binary indicator for the worker's gender (0-male, 1-female). Standard errors are clustered at the adviser level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.4 Advice and Worker Guessed Score

We found in our main results (Table 4) that self-evaluations strongly predict the advice received by the worker in the baseline treatment. Table A.8 (pre-registered) re-examines this result by examining the role of the guessed score in a more flexible way. We observe a monotonically increasing relationship between the worker's guessed score and the chance of receiving advice A. Moreover, we observe that generally advisers only advised A to workers with a guess well above the performance threshold for job A (8 correct answers).

Table A.8: Effect of Worker Score Guess on Advice in Baseline Treatment

Advice A measure:	Likert (1)	Binary (2)
Guessed Score=2	-0.800*** (0.127)	-0.118*** (0.044)
Guessed Score=3	-0.768*** (0.102)	-0.139*** (0.029)
Guessed Score=4	-0.556*** (0.106)	-0.112*** (0.029)
Guessed Score=5	-0.568*** (0.096)	-0.091*** (0.030)
Guessed Score=6	-0.450*** (0.114)	-0.073** (0.035)
Guessed Score=7	-0.200** (0.095)	-0.029 (0.031)
Guessed Score=9	0.000 (0.110)	0.007 (0.036)
Guessed Score=10	0.235** (0.100)	0.083** (0.034)
Guessed Score=11	0.700*** (0.111)	0.236*** (0.039)
Guessed Score=12	0.798*** (0.108)	0.305*** (0.038)
Guessed Score=13	0.742*** (0.143)	0.282*** (0.052)
Guessed Score=14	1.461*** (0.107)	0.514*** (0.037)
Guessed Score=15	1.689*** (0.119)	0.597*** (0.039)
Guessed Score=16	1.983*** (0.104)	0.713*** (0.036)
Guessed Score=17	2.323*** (0.109)	0.765*** (0.033)
Guessed Score=18	2.375*** (0.154)	0.757*** (0.049)
Guessed Score=19	2.650*** (0.220)	0.832*** (0.026)
Constant	2.001*** (0.091)	0.140*** (0.030)
N	3,500	3,500
Clusters (Advisers)	700	700
R^2	0.400	0.363

Notes: The omitted category is Guessed Score=8. All specifications control for worker order (1 to 5). Standard errors are clustered at the adviser level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.5 Worker Behavior and Expected Earnings

Of the 320 workers who completed the second session, 121 were advised to choose job A on either the binary or Likert scale, 198 were advised to choose job B, and 1 worker received no advice (the middle point on the Likert scale). Of the 319 workers who received advice,

308 chose to follow it (96.6%), whereas 11 workers (5 men and 6 women) did not. The high rate of advice-following is understandable, as workers had minimal information beyond the adviser’s recommendation. In total, 121 workers were advised to choose job A and 198 were advised to choose job B.

How did advice impact earnings and how do earnings differ by gender and treatment? We can answer this question in two ways. First, we can look at expected earnings based on all the advice generated for a particular worker (across all advisers). The results of this approach are discussed in Table 6 in Section 3.4 in the main text based on part 1 score. Results are similar if we re-do this analysis using part three score instead (Table A.9). Second, we can examine realized earnings in part three. This has the benefit of not requiring us to assume that workers follow advice. However, it comes at the cost of a much smaller sample size (320 workers rather than 8,750 pieces of advice), which reduces power and increases the scope for random adviser draws to impact the results. Across all treatments, women are 6.3 percentage points less likely to choose job A (N=320, p=0.122, one-sided t-test). The gender gap is also insignificant in the baseline treatment (10.8 pp, N=88, p=0.150, one-sided t-test); nudge treatment (12.1 pp, N=62, p=0.166, one-sided t-test); subjective treatment (5.2 pp, N=71, p=0.328, one-sided t-test); and objective treatment (-0.7 pp, N=99, p=0.527, one-sided t-test) individually. Due to the small sample size, these estimates are much noisier than the adviser-level estimates presented in Section 3.4 in the main text. The fact that women also make less ambitious choices in the nudge treatment is driven by female workers randomly receiving much less ambitious advice from their assigned advisers (28% job A) than the average advice sent by all advisers in this treatment (47% job A for women).

Table A.9: Expected Earnings Based on Part 3 Score

	Advice Maximized		Expected Earnings		Gap in Expected Earnings	
	Expected Earnings	Expected Earnings	Based on Advice	Based on Advice	With vs. Without Advice	With vs. Without Advice
	(1)	(2)	(3)	(4)	(5)	(6)
Female		-0.047*** (0.017)		-0.102*** (0.019)		0.013 (0.019)
Subjective	-0.036*** (0.014)	-0.048** (0.019)	-0.034** (0.015)	-0.048** (0.021)	-0.040** (0.017)	-0.050** (0.022)
Objective	0.068*** (0.008)	0.066*** (0.011)	0.065*** (0.008)	0.061*** (0.011)	0.060*** (0.008)	0.054*** (0.011)
Nudge	0.028** (0.014)	-0.054*** (0.019)	0.029* (0.015)	-0.061*** (0.021)	0.034** (0.017)	-0.067*** (0.022)
Female×Subjective		0.023 (0.026)		0.026 (0.028)		0.020 (0.029)
Female×Objective		0.002 (0.015)		0.008 (0.015)		0.011 (0.015)
Female×Nudge		0.162*** (0.027)		0.179*** (0.030)		0.202*** (0.030)
Constant	0.403*** (0.014)	0.426*** (0.016)	1.287*** (0.015)	1.338*** (0.018)	-0.341*** (0.016)	-0.348*** (0.018)
N	11,197	11,197	11,197	11,197	10,531	10,531
Clusters (Advisers)	1,750	1,750	1,750	1,750	1,750	1,750
R ²	0.007	0.011	0.006	0.012	0.006	0.016

Notes: All specifications control for a worker's order (1 to 5). *Female* is a binary indicator for the worker's gender (0-male, 1-female). The other variables are binary indicators for the respective treatments. Standard errors are clustered at adviser level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.6 Robustness Checks

Table A.10 replicates Table 3 while also controlling for worker sociodemographics. We find that less educated participants are less confident in all six dimensions, as are participants who report low income satisfaction. Table A.11 replicates Table 4 while controlling for adviser-level demographics. Older and more risk-averse participants give less ambitious advice. Both the gender gap in self-promotion and the gender gap in advice remain similar when controlling for demographics.

Table A.10: Gender Gap in Self-promotion Controlling for Worker Performance and Sociodemographics

	Likert Performance (1-7)	Performed Well (0-100)	Guessed Rank (rev)	Would Apply for STEM job (0-100)	Would Succeed in STEM (0-100)	Score Guess (0-20)
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.374*** (0.118)	-7.549*** (2.193)	-7.680*** (2.323)	-9.244*** (2.986)	-9.751*** (2.871)	-0.703** (0.367)
Age	-0.011* (0.006)	-0.211* (0.110)	0.042 (0.116)	-0.142 (0.150)	-0.147 (0.144)	0.003 (0.018)
<i>Education (ref: Postgraduate degree)</i>						
High School	-0.550** (0.237)	-10.379** (4.393)	-12.024** (4.652)	-18.324*** (5.982)	-20.118*** (5.751)	-1.173 (0.735)
Vocational training	-0.465*** (0.175)	-9.151*** (3.242)	-8.952*** (3.433)	-15.144*** (4.415)	-19.600*** (4.244)	-1.132** (0.543)
Undergraduate degree	-0.248 (0.153)	-6.557** (2.845)	-4.464 (3.013)	-6.304 (3.874)	-9.493** (3.725)	-0.427 (0.476)
<i>Occupation (ref: Employed)</i>						
Studying	0.069 (0.194)	3.015 (3.595)	2.787 (3.807)	1.843 (4.895)	2.088 (4.706)	0.819 (0.602)
Retired	0.445 (0.299)	4.810 (5.539)	0.741 (5.866)	0.698 (7.542)	3.795 (7.251)	1.094 (0.927)
Unemployed	0.265 (0.268)	2.983 (4.981)	2.692 (5.275)	8.336 (6.782)	2.806 (6.520)	0.331 (0.834)
Not in Labor Force	0.179 (0.323)	-0.385 (5.991)	-1.307 (6.344)	2.963 (8.158)	1.728 (7.843)	-1.134 (1.003)
<i>Income satisfaction (ref: Live comfortably)</i>						
Find it difficult	-0.380** (0.175)	-8.084** (3.243)	-7.499** (3.434)	-7.279 (4.416)	-7.189* (4.246)	-0.620 (0.543)
Coping	-0.114 (0.133)	-2.580 (2.470)	-2.111 (2.616)	-1.729 (3.364)	-1.401 (3.234)	-0.469 (0.414)
Constant	4.488*** (0.816)	68.009*** (15.140)	64.596*** (16.032)	83.039*** (20.615)	86.020*** (19.819)	9.370*** (2.534)
Score FE	✓	✓	✓	✓	✓	✓
N	350	350	350	350	350	350
R ²	0.370	0.394	0.342	0.306	0.339	0.437

Notes: Score FE (fixed effects) controls for actual performance. Statistical significance of the coefficient estimates for the variable *Female* is based on a one-sided test. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.11: Gender Gap in Advice With Adviser Sociodemographics

Advice A Measure:	Likert		Binary	
	(1)	(2)	(3)	(4)
Worker characteristics:				
Female	-0.387*** (0.047)	0.020 (0.038)	-0.123*** (0.016)	0.004 (0.013)
Performed Well (0-100)		0.003* (0.002)		0.000 (0.001)
Gussed Rank (1-101)		0.006*** (0.002)		0.002*** (0.001)
Would Apply for STEM job (0-100)		0.004** (0.002)		0.002*** (0.001)
Would Succeed in STEM job (0-100)		0.008*** (0.002)		0.002*** (0.001)
Gussed Score (0-20)		0.114*** (0.009)		0.034*** (0.003)
Adviser characteristics:				
Female	0.028 (0.057)	0.090* (0.054)	0.009 (0.018)	0.029* (0.017)
Age	-0.003 (0.003)	-0.005* (0.003)	-0.001 (0.001)	-0.002** (0.001)
<i>Education (ref: Postgraduate degree)</i>				
High School	0.065 (0.122)	0.162 (0.122)	0.021 (0.039)	0.051 (0.039)
Vocational training	0.077 (0.088)	0.131 (0.086)	0.029 (0.027)	0.045* (0.026)
Undergraduate degree	-0.025 (0.074)	-0.001 (0.072)	-0.006 (0.023)	0.000 (0.022)
<i>Occupation (ref: Employed)</i>				
Studying	0.108 (0.099)	0.019 (0.100)	0.023 (0.032)	-0.004 (0.033)
Retired	0.159 (0.158)	0.102 (0.148)	0.061 (0.050)	0.042 (0.046)
Unemployed	-0.064 (0.101)	-0.091 (0.086)	-0.024 (0.035)	-0.034 (0.030)
Not in Labor Force	-0.108 (0.103)	-0.024 (0.107)	-0.044 (0.037)	-0.016 (0.035)
<i>Income satisfaction (ref: Live comfortably)</i>				
Find it difficult	-0.059 (0.087)	-0.033 (0.084)	-0.056** (0.026)	-0.048* (0.025)
Coping	-0.040 (0.064)	-0.016 (0.063)	-0.017 (0.020)	-0.010 (0.020)
Self-confidence	-0.005 (0.017)	-0.003 (0.016)	-0.004 (0.005)	-0.003 (0.005)
Risk attitude	0.032** (0.016)	0.036** (0.015)	0.009** (0.004)	0.010** (0.004)
Constant	2.657*** (0.166)	0.386** (0.171)	0.425*** (0.052)	-0.272*** (0.056)
N	3,500	3,500	3,500	3,500
Clusters (Advisers)	700	700	700	700
R ²	0.026	0.441	0.024	0.390

Notes: All specifications control for a worker's order (1 to 5). Standard errors are clustered at the adviser level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We next investigate whether our results on advice are robust to using two alternative measures of advice: the performance goal and a version of binary advice that adjusts for adviser confidence. In the latter case, in line with our pre-analysis plan, we combine the binary variable with the confidence variable in the following way: We code advice given with confidence 1-3 as “no advice,” confidence 4-8 as “probably A/B,” and 9+ as “definitely A/B,” and encode this as a 5-point scale from 1-definitely B to 5-definitely A, to approximate the distribution of responses on the Likert scale. Our results, reported in Table A.12, (a) replicate the gender gap in advice in the baseline, (b) replicate the effect of the nudge, and (c) replicate the null effect of the subjective treatment on the gender gap in advice. The main difference from our main results is that we no longer find a significant effect of the objective treatment on the gender gap in advice.

Table A.12: Treatment Differences in Gender Gap in Secondary Advice Measures

<i>Treatment</i>	Baseline vs. <i>Treatment</i>							
	Baseline		Subjective		Objective		Nudge	
<i>Advice Measure:</i>	Goal (1)	Confidence (2)	Goal (3)	Confidence (4)	Goal (5)	Confidence (6)	Goal (7)	Confidence (8)
Female	-1.062*** (0.134)	-0.342*** (0.046)	-1.062*** (0.134)	-0.342*** (0.046)	-1.062*** (0.134)	-0.342*** (0.046)	-1.062*** (0.134)	-0.342*** (0.046)
Subjective			-0.307 (0.191)	-0.136** (0.054)				
Objective					0.306*** (0.083)	0.118*** (0.031)		
Nudge							-0.515*** (0.187)	-0.199*** (0.055)
Female×Subjective			0.092 (0.204)	0.033 (0.069)				
Female×Objective					0.122 (0.105)	0.033 (0.040)		
Female×Nudge							1.547*** (0.216)	0.668*** (0.073)
Constant	9.890*** (0.120)	2.648*** (0.056)	10.186*** (0.140)	2.640*** (0.048)	10.129*** (0.152)	2.662*** (0.054)	10.257*** (0.143)	2.670*** (0.048)
Gender Gap in <i>Treatment</i>								
Coefficient			-0.970*** (0.154)	-0.308*** (0.052)	-0.940*** (0.120)	-0.309*** (0.049)	0.485*** (0.170)	0.327*** (0.057)
Standard error								
N	3,500	3,500	6,125	6,125	7,000	7,000	6,125	6,125
Clusters (Advisers)	700	700	1,225	1,225	700	700	1,225	1,225
R^2	0.015	0.018	0.017	0.019	0.018	0.017	0.014	0.020

Notes: All specifications control for a worker's order (1 to 5). *Female* is a binary indicator for the worker's gender (0-male, 1-female). The other variables are binary indicators for the respective treatments. Standard errors are clustered at adviser level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We next investigate whether our results are robust to restricting our analysis of the gender gap in advice (Table 5) to the first worker assessed by each adviser. We only informed advisers about workers 2-5 after they finished evaluating worker 1, which allowed us to treat the first worker as a one-shot observation. In addition, restricting ourselves to the first worker removes the potential for spillovers from previous workers. The results (Table A.13) are similar to our main results, in that (a) female workers receive less ambitious advice (columns 1 and 2); (b) the nudge reverses the gender gap (columns 3 and 4); and (c) there is no difference between the gender gaps in the baseline and subjective treatments (columns 7 and 8). The main difference is that the gender gap in the objective treatment no longer differs from the baseline (columns 5 and 6).

Table A.13: Treatment Differences in Gender Gap in Advice, 1st worker only

<i>Treatment</i>	Baseline vs. <i>Treatment</i>							
	Baseline		Subjective		Objective		Nudge	
Advice A measure:	Likert (1)	Binary (2)	Likert (3)	Binary (4)	Likert (5)	Binary (6)	Likert (7)	Binary (8)
Female	-0.289*** (0.106)	-0.106*** (0.036)	-0.289*** (0.106)	-0.106*** (0.036)	-0.289*** (0.106)	-0.106*** (0.036)	-0.289*** (0.106)	-0.106*** (0.036)
Subjective			-0.124 (0.113)	-0.087** (0.039)				
Objective					0.203*** (0.062)	0.046** (0.022)		
Nudge							-0.288*** (0.111)	-0.094** (0.039)
Female×Subjective			0.072 (0.159)	0.056 (0.053)				
Female×Objective					-0.017 (0.085)	0.023 (0.031)		
Female×Nudge							0.920*** (0.158)	0.287*** (0.055)
Constant	2.657*** (0.076)	0.400*** (0.026)	2.657*** (0.076)	0.400*** (0.026)	2.657*** (0.076)	0.400*** (0.026)	2.657*** (0.076)	0.400*** (0.026)
Gender gap in <i>Treatment</i>								
Coefficient			-0.217** (0.118)	-0.050 (0.040)	-0.306*** (0.112)	-0.083*** (0.037)	0.631*** (0.117)	0.182*** (0.042)
Standard error								
N	700	700	1,225	1,225	1,400	1,400	1,225	1,225
Clusters (Advisers)	700	700	1,225	1,225	700	700	1,225	1,225
R^2	0.010	0.012	0.010	0.013	0.015	0.013	0.032	0.025

Notes: *Female* is a binary indicator for the worker's gender (0-male, 1-female). The other variables are binary indicators for the respective treatments. Statistical significance of the coefficient estimates for the variable *Female*×*Treatment* is based on a one-sided test. Standard errors are clustered at adviser level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.7 Additional Results

Table A.14 reports the stated confidence of advisers giving binary advice. The key results are that advisers are less confident when advising job A, and they are less confident when giving ambitious advice (job A) to female workers than when giving this advice to male workers. Table A.15 repeats this analysis while controlling for worker guessed score fixed effects. The results are similar, though the gender gap for advice A falls from 0.209 to 0.131 and is no longer significant ($p=0.125$, two-sided Wald test). Table A.16 shows that the gender gap in advice in the baseline treatment is similar for advisers who anticipate that men are more confident than women. Table A.17 shows that the gender gap in advice is largest when advisers believe that men perform better than women and non significantly different from zero when advisers believe that women perform better than men. Table A.18 shows that the nudge treatment increases the share of advisers who expect men to be more confident relative to the baseline/objective treatment, whereas the Subjective treatment reduces it. The nudge treatment also increases the share of advisers who expect women to outperform men. Table A.19 presents the full distribution of answers to the expected gender performance gap question in each treatment. Finally, Table A.20 looks at heterogeneity by the gender of the adviser. There are no differences in the baseline and subjective treatments. Female advisers appear to give less ambitious advice to female workers relative to male workers in the objective treatment, though this is only significant when looking at the Likert scale. In the nudge treatment, by contrast, female advisers are significantly more likely to advise female workers to go for the ambitious task (job A). This suggests that female advisers internalized the nudge to a greater extent than their male counterparts.

Table A.14: Gender Gap in Adviser Confidence

	Advice		Worker gender		Diff.	<i>p</i> -value
	Job A	Job B	Female	Male		
Confidence in Binary Advice	7.467	7.986			-0.520	0.000
Confidence in Binary Advice			7.786	7.796	-0.010	0.870
Confidence when Advising A			7.345	7.554	-0.210	0.019
Confidence when Advising B			7.989	7.983	0.006	0.935

Notes: The sample includes 700 advisers in the baseline treatment.

Table A.15: Gender Gap in Adviser Confidence With Worker Guessed Score Fixed Effects

	Confidence in Advice		
	(1)	(2)	(3)
Female	-0.065 (0.057)		-0.058 (0.074)
Job A		-0.572*** (0.094)	-0.542*** (0.111)
Female×Job A			-0.073 (0.112)
Constant	9.001*** (0.280)	8.948*** (0.263)	9.009*** (0.285)
Guessed Score FE	✓	✓	✓
N	3,500	3,500	3,500
Clusters (Advisers)	700	700	700
R^2	0.048	0.064	0.065

Notes: All specifications control for a worker's order (1 to 5). *Female* is a binary indicator for the worker's gender (0-male, 1-female). *Job A* is a binary indicator that takes the value 1 if the adviser advised the worker to choose Job A and 0 otherwise. Standard errors are clustered at the adviser level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.16: Gender Gap in Advice by Advisers Beliefs about the Gender Gap in Confidence in Baseline

Advice A measure:	Men are more confident				Men and women equally confident				Women more confident			
	Likert		Binary		Likert		Binary		Likert		Binary	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Female	-0.456*** (0.058)	-0.010 (0.048)	-0.147*** (0.020)	-0.008 (0.017)	-0.261*** (0.091)	0.093 (0.067)	-0.091*** (0.030)	0.016 (0.023)	-0.230 (0.196)	-0.003 (0.173)	-0.018 (0.063)	0.055 (0.052)
Performed Well (0-100)		0.005** (0.002)		0.001 (0.001)		-0.002 (0.004)		-0.001 (0.001)		0.009 (0.007)		0.000 (0.002)
Guessed Rank (1-101)		0.006*** (0.002)		0.002*** (0.001)		0.005 (0.003)		0.001 (0.001)		0.012** (0.006)		0.006*** (0.002)
Would Apply for STEM Job (0-100)		0.005** (0.002)		0.002*** (0.001)		0.002 (0.003)		0.002 (0.001)		-0.004 (0.007)		0.002 (0.002)
Would Succeed in STEM Job (0-100)		0.007*** (0.002)		0.002*** (0.001)		0.010*** (0.004)		0.003** (0.001)		0.011 (0.007)		0.001 (0.002)
Guessed Score (0-20)		0.105*** (0.010)		0.029*** (0.004)		0.148*** (0.015)		0.048*** (0.005)		0.066* (0.034)		0.028** (0.011)
Constant	2.723*** (0.072)	0.461*** (0.093)	0.417*** (0.025)	-0.263*** (0.031)	2.604*** (0.106)	0.268** (0.118)	0.381*** (0.035)	-0.335*** (0.041)	2.984*** (0.246)	0.964*** (0.334)	0.438*** (0.084)	-0.278*** (0.081)
N	2,240	2,240	2,240	2,240	1,015	1,015	1,015	1,015	245	245	245	245
Clusters (Advisers)	448	448	448	448	203	203	203	203	49	49	49	49
R ²	0.025	0.442	0.023	0.383	0.008	0.435	0.009	0.384	0.006	0.401	0.000	0.442

Notes: All specifications control for a worker's order (1 to 5). *Female* is a binary indicator for the worker's gender (0-male, 1-female). Standard errors are clustered at the adviser level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.17: Gender Gap in Advice by Advisers Beliefs about the Gender Gap in Performance in Baseline

Advice A measure:	Men perform better				Men and women perform equally well				Women perform better			
	Likert		Binary		Likert		Binary		Likert		Binary	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Female	-0.458*** (0.058)	-0.010 (0.048)	-0.147*** (0.020)	-0.008 (0.017)	-0.257*** (0.090)	0.093 (0.067)	-0.090*** (0.030)	0.016 (0.023)	-0.237 (0.189)	-0.003 (0.173)	-0.019 (0.060)	0.055 (0.052)
Performed Well (0-100)		0.005** (0.002)		0.001 (0.001)		-0.002 (0.004)		-0.001 (0.001)		0.009 (0.007)		0.000 (0.002)
Guessed Rank (1-101)		0.006*** (0.002)		0.002*** (0.001)		0.005 (0.003)		0.001 (0.001)		0.012** (0.006)		0.006*** (0.002)
Would Apply for STEM Job (0-100)		0.005** (0.002)		0.002*** (0.001)		0.002 (0.003)		0.002 (0.001)		-0.004 (0.007)		0.002 (0.002)
Would Succeed in STEM Job (0-100)		0.007*** (0.002)		0.002*** (0.001)		0.010*** (0.004)		0.003** (0.001)		0.011 (0.007)		0.001 (0.002)
Guessed Score (0-20)		0.105*** (0.010)		0.029*** (0.004)		0.148*** (0.015)		0.048*** (0.005)		0.066* (0.034)		0.028** (0.011)
Constant	2.806*** (0.046)	0.461*** (0.093)	0.455*** (0.015)	-0.263*** (0.031)	2.627*** (0.069)	0.268** (0.118)	0.398*** (0.022)	-0.335*** (0.041)	2.810*** (0.146)	0.964*** (0.334)	0.422*** (0.043)	-0.278*** (0.081)
N	670	670	670	670	2,330	2,330	2,330	2,330	500	500	500	500
Clusters (Advisers)	134	134	134	134	466	466	466	466	100	100	100	100
R ²	0.029	0.419	0.031	0.355	0.016	0.441	0.013	0.391	0.010	0.427	0.012	0.413

Notes: All specifications control for a worker's order (1 to 5). *Female* is a binary indicator for the worker's gender (0-male, 1-female). Standard errors are clustered at the adviser level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.18: Treatment Effects on Adviser’s Opinion About Gender Gap in Confidence and Performance

	Share answering that:	
	Men more confident (1)	Women perform better (2)
<i>Ref: Objective</i>		
Nudge	0.225*** (0.026)	0.253*** (0.023)
Subjective	-0.053** (0.026)	-0.010 (0.023)
Constant	0.640*** (0.017)	0.143*** (0.015)
N	1,750	1,750

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.19: Advisers Stated Beliefs About Gender Gaps (Final Questionnaire)

	% of sample in each treatment		
	Objective ($N = 700$)	Nudge ($N = 525$)	Subjective ($N = 525$)
Gender performance gap			
Women perform better / a lot better	4.71	14.10	5.33
Women perform slightly better	9.57	25.52	8.00
Women and men perform equally well	66.57	46.29	67.62
Men perform slightly better	14.71	11.24	14.86
Men perform better / a lot better	4.43	2.86	4.19

Notes: Since the question was asked at the end of the experiment and we use a within-subject design for the baseline and objective treatments, we do not have advisers’ stated beliefs for the baseline treatment.

Table A.20: Heterogeneity in Advice by Adviser Gender

<i>Treatment</i>	Baseline		Subjective		Objective		Nudge	
	Likert (1)	Binary (2)	Likert (3)	Binary (4)	Likert (5)	Binary (6)	Likert (7)	Binary (8)
Advice A measure:								
W Female	0.043 (0.057)	0.004 (0.019)	0.004 (0.066)	-0.002 (0.023)	0.088 (0.063)	0.021 (0.022)	0.509*** (0.067)	0.161*** (0.023)
A female	0.051 (0.065)	0.012 (0.022)	0.069 (0.080)	0.020 (0.026)	0.169** (0.075)	0.039 (0.025)	-0.030 (0.079)	-0.024 (0.026)
Wfemale × A female	-0.040 (0.076)	0.000 (0.026)	0.031 (0.095)	0.008 (0.032)	-0.187** (0.085)	-0.038 (0.029)	0.336*** (0.098)	0.112*** (0.034)
Performed Well (0-100)	0.003* (0.002)	0.000 (0.001)	0.017*** (0.002)	0.005*** (0.001)	0.003 (0.002)	0.001 (0.001)	0.005** (0.002)	0.002* (0.001)
Gussed Rank (1-101)	0.006*** (0.002)	0.002*** (0.001)	0.002 (0.002)	0.001 (0.001)	0.007*** (0.002)	0.002*** (0.001)	0.008*** (0.002)	0.002*** (0.001)
Would Apply for STEM job (0-100)	0.004** (0.002)	0.002*** (0.001)	-0.004* (0.002)	-0.001* (0.001)	-0.000 (0.002)	0.001 (0.001)	0.002 (0.002)	0.001 (0.001)
Would Succeed in STEM job (0-100)	0.008*** (0.002)	0.002*** (0.001)	0.011*** (0.003)	0.004*** (0.001)	0.000 (0.002)	-0.000 (0.001)	0.006*** (0.002)	0.002** (0.001)
Gussed Score (0-20)	0.113*** (0.009)	0.034*** (0.003)	0.029*** (0.010)	0.007* (0.004)	0.142*** (0.010)	0.044*** (0.003)	0.104*** (0.011)	0.034*** (0.004)
Constant	0.418*** (0.081)	-0.291*** (0.027)	0.983*** (0.090)	-0.152*** (0.030)	0.707*** (0.091)	-0.196*** (0.030)	0.407*** (0.089)	-0.318*** (0.031)
N	3,500	3,500	2,625	2,625	3,500	3,500	2,625	2,625
Clusters (Advisers)	700	700	525	525	700	700	525	525
R^2	0.432	0.383	0.271	0.238	0.280	0.241	0.382	0.330

Notes: All specifications control for a worker's order (1 to 5). **W** means worker; **A** means adviser. Standard errors are clustered at the adviser level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.8 Pilot Experiments

We conducted several pilots for our online experiment. We conducted our first pilot in April 2023 with 24 workers. Each worker answered 30 questions randomly selected from a set of 46. Based on the results, we dropped two easy questions (accuracy rates greater than 94%) and two hard questions (accuracy rates lower than 20%). To obtain set of 30 questions for our second pilot, we also dropped two questions in which women strongly outperformed men and nine questions in which men strongly outperformed women. We conducted our second pilot in May 2023 with 23 participants. Men and women performed equally well, but men were more confident in the quantitative measures. We tested two qualitative questions in this pilot, one rating their performance on the task and one on general math and science skills. We each evaluated the responses on these questions in a gender-blind way and found a larger gap in the performance question. We therefore decided to include this question in the main experiment.

Our third pilot took place in September 2023 and involved 20 workers and 20 advisers. Each worker answered 25 questions. We piloted the baseline treatment and found a gender gap in self-promotion and a gender gap in advice. However, we saw a reverse gender gap (with women more confident) in the guessed rank question, which we suspected was driven by this question's being reverse-coded to all other questions (most confident being on the left rather than right of the screen). We therefore decided to flip the scale for this question. We also included a second rank question in which participants were asked to compare themselves with similar individuals. However, the answers to the two rank questions were almost perfectly correlated, so we decided to drop this second (more complicated) question.

Our fourth pilot took place in early October 2023 with a further 20 workers and 20 advisers. This time we found no gender gap in self-evaluations, which led us to conduct a fifth pilot in late October 2023 with 100 workers, and we once again did find a gender gap in self-evaluations. We also introduced a qualitative question asking participants to reflect on their math and science skills, for which they needed to spend at least 1 minute and received 50 pence for answering. We also saw a gender gap in performance. To reduce the gender gap, we decided to remove the five questions with the strongest male advantage in

performance. We tested the set of questions in a sixth pilot in early December 2023 with 100 workers, and found no gender gap in performance. We then decided to use this set of 20 questions for our main experiment. We conducted a seventh pilot in late December 2023 with 100 workers to test the new qualitative question from the fifth pilot. After evaluating the answers using ChatGPT and finding no gender difference and little correlation with the quantitative question on our AI, we decided to remove this question.

The purpose of our eighth pilot (conducted in February 2024 with 100 workers and 100 advisers) was to test the final design of the baseline treatment and be used as the basis for our power calculations. Upon observing the results, we realized that the performance thresholds for task A might be a little too ambitious. In our ninth and final pilot (conducted in March 2024), we therefore lowered the threshold for job A from 10 to 8 correct answers. We used the data from pilots six to nine as the basis for our power calculations. These, in turn, convinced us to move away from a one-to-one matching of workers and advisers to the one-to-many matching we used in our main experiment. The time between the final pilot and the main sessions was spent on the pre-analysis plan and finalizing the design of the three other treatments (which we did not pilot).

A.9 Pre-Analysis Plan

In this section, we reprint our pre-analysis plan. We also added notes explaining where in the paper the respective analysis can be found. For any analysis not included in the paper, we will both (a) summarize its results and (b) explain why we did not include it in the paper. We also include the power calculation. The pre-analysis plan is also available from our pre-registration page at <https://www.socialscisceregistry.org/trials/13803>.

A Description of the Experiment

We conduct a three-part experiment. In session 1, participants in the role of workers work on a math and science quiz. Upon completing the quiz, they indicate their confidence in several ways, forming a worker “portfolio”. In session 2, a different set of participants in the role of advisers sees the portfolio of five workers and advises each worker whether they should go for an advanced (A) or a basic (B) version of the task. In session 3, each worker receives the advice from one adviser and decides which of the two versions of the task to pursue.

We consider three between-subject treatments that vary the information the adviser receives in session 2. In the “baseline” treatment, the adviser receives the full portfolio of worker confidence measures from session 1. In the “nudge” treatment, advisers also receive a brief message informing them about the gender gap in self-assessment. In the “subjective” treatment, the advisers instead only receive the worker’s verbal self-evaluation (a qualitative and subjective self-assessment). In addition, participants in the baseline treatment (but not the other treatments) are also asked to evaluate the same five workers a second time while also having access to those workers’ true performances on the task. We refer to this second set of advice as the (within-subject) “objective” treatment.

B Main Outcomes

1. Worker Confidence (Session 1)
 - a) Guessed number of correct answers
 - b) I performed well on the quiz (0-100)
 - c) I would apply for a job that would require me to perform well on the quiz (0-100)
 - d) I would succeed in a job requiring me to perform well on the quiz (0-100)
 - e) Relative performance rank (compared to 100 other participants)
 - f) Subjective performance Likert scale (1-terrible to 5-great).
 - g) Standardized sum of the previous measures
 - h) Verbal self-evaluation
2. Advice (Session 2)
 - a) Binary advice (A or B)
 - b) Five-point scale advice (Definitely A to Definitely B)

C Secondary Outcomes

3. Worker Task Choice (Session 3)
4. Other measures of advice (Session 2)

D Main Explanatory Variable

5. Worker Gender: dummy variable comparing males to females.

E Other Variables

6. Workers (session 1): general confidence/risk aversion measures, score, demographics
7. Advisors: gender, other demographics, confidence/risk aversion
8. Advisor beliefs about gender differences:
 - a) Are men or women better at the task?
 - b) Are men or women more confident in their self-assessment?
9. Worker score in session 3.

Demographics include household income, income satisfaction, age, occupation, job title, and education.

F Main analysis

Our experiment aims to test the following hypotheses:

1. Gender Gap in Self-Promotion: Female workers are less confident than male workers.
 - (a) Main tests: one-sided t-test for each of the quantitative worker confidence variables listed in part B1 above (a-g).
Authors' note: We included these tests in Figure 2
2. Gender Gap in Advice: Female workers receive less ambitious advice than male workers in the baseline treatment.
 - (a) Main tests: linear regression of the (a) binary advice measure and (b) Likert scale advice measure on a dummy for the worker's gender. We use a one-sided test on the gender dummy.
Authors' note: We included these tests in Table 4 (columns 1 and 3) and reprint the results in Table 5 (columns 1 and 2).
3. Treatment Differences: The gender gap in advice is significantly smaller in the objective, nudge and subjective treatments than in the baseline.
 - (a) Main tests: a linear regression of advice on gender, treatment and a gender*treatment interaction. The key test is a one-sided test on the significance of the gender*treatment interaction. We do three regressions comparing each of nudge, subjective and objective treatment to the baseline.
Authors' note: We included these tests in Table 5 (columns 3, 5 and 7).

We will use robust standard errors clustered at the referee level for the regression analysis (F2- F3). In addition to testing these three hypotheses, we will also investigate the drivers of the gender gap in advice:

4. Drivers of the Gender Gap in Advice: Is the gender gap in advice (F2) driven by gender differences in self-promotion or by a gender bias among advisers?
 - (a) Main tests: linear regression of the (a) binary advice measure and (b) Likert scale advice measure on (i) a dummy for the worker's gender and (ii) the worker's guessed score (B1a). We use a one-sided test for both variables (less ambitious advice for women, positive effect of the guessed score).
Authors' note: The pre-registered specification can be found in Table A.5. In our main analysis in Table 4 we decided to control for all six self-evaluation questions instead (as per analysis G11 below). The main difference between the two tables is that the residual gender gap is no longer significant in Table A.5. Upon reflection, Table A.5 seemed to us like a fairer reflection of our results since it controls for all information available to advisers.

G Additional Analysis

Here is an overview of additional tests we intend to include in our paper to shed further light on our results, but that do not directly address our main hypotheses. We start with some

further analysis of worker confidence.

5. Is there a gender difference in performance?
 - a) Two-sided t-test for a gender difference in worker score (session 1).
 - b) Two-sided t-test for a gender difference in worker score (session 3), after receiving advice.
Authors' note: We present these results in Table A.2.

6. Are any gender differences in confidence significant after controlling for performance and demographics?
 - a) Regression of confidence on the worker gender dummy (1-female, 0-male) while controlling for (a) worker score from session 1 and (b) also controlling for worker demographics. The relevant test is a one-sided test on the regression coefficient on the female dummy (as per main analysis F1).
Authors' Note: We present the results (a) in Table A.3 and (b) in Table A.10.

7. Are there any gender differences in the subjective self-evaluations?
 - a) For this purpose, we intend to explore the use of AI methods to quantify the expressed levels of confidence.
Authors' Note: We present these results in A.6.

Next we plan to present robustness checks for our analysis on gender differences in advice, as well as some additional outcomes.

8. Is the gender difference in advice F2 robust to only looking at the first worker evaluated?
 - a) Replication of main analysis F2 while including only the first worker evaluated.
Authors' note: We present this analysis in Table A.13.
9. Do gender differences and treatment differences in advice (F2 and F3) also appear in our secondary measures (C4)?
 - a) Replication of main analysis F2 and F3 for the performance target C4a.
 - b) Replication of main analysis F2 and F3 using the confidence in advice variable (C4b). For this purpose, we combine the binary variable with the confidence variable as follows: confidence 1–3 → “no advice”, confidence 4–8 → “probably A/B”, confidence 9+ → “definitely A/B”, and encode this as a 5-point scale from 1 = definitely B to 5 = definitely A.
Authors' note: This analysis is present in Table A.12
10. Does the adviser's confidence (C4b) depend on the worker's gender and advice sent?
 - a) Two-sided t-test investigating whether confidence differs significantly by (i) worker gender and (ii) binary advice (A or B). Also two t-tests studying whether confidence differs by worker gender *within* each advice category.
 - b) A linear regression of adviser confidence on (i) worker gender, (ii) binary advice (A or B), or (iii) gender within the sample of a given type of advice, always controlling for the worker's guessed score.
Authors' note: This analysis can be found in Tables A.14 and A.15.

11. What are the main drivers of advice?

- a) Linear regressions of advice (B2a and B2b) in the baseline on (i) all quantitative measures of confidence available to advisers (B1a–e) and (ii) referee demographics.
- b) Regression exploring whether the effect of confidence is non-linear by including guessed-score fixed effects.

Authors' note: We present the analysis controlling for all measures of confidence in Table 4, and the one also controlling for demographics in Table A.11. Results with guessed score fixed effects are in Table A.8.

12. What is the quality of advice and how does it differ by treatment and gender?

- a) We create a binary indicator for optimal advice—whether advice would maximize a worker's expected earnings based on session-1 performance. Regress this indicator on treatment dummies. Also replicate including a gender dummy and interaction terms.

Authors' note: This analysis is presented in Table 6.

13. Do we observe heterogeneous effects based on adviser bias?

- a) Replicating main analysis F2 while splitting the sample based on adviser bias E8.
- b) We split the sample into three groups: no difference, men at least slightly better/more confident and women at least slightly better/more confident. We only analyze a given group if at least 5% of our sample are in this group.
- c) We analyze variables E8a and E8b separately.

Authors' note: This analysis is presented in Table A.16 and Table A.17.

14. Do any of our treatments reduce any gender bias that exists in variables E8a and E8b?

- a) Regression of each variable on treatment dummies (two-sided tests).
- b) Conduct this test only if the baseline shows significant bias (one-sided test: women perceived as worse/less confident).

Authors' note: This analysis is presented in Table A.18.

We also consider several tests of behavior in Session 3. These are of lesser importance because we (a) expect nearly all workers to follow advice and (b) we have less power due to the smaller sample size per treatment.

15. Do female workers choose less ambitious jobs in session 3?

- a) One-sided t-test for worker choice in session 3 by gender.

Authors' note: This test is presented in Appendix A.5.

16. An analysis of the efficiency consequences of advice.

- a) Create a binary indicator of the optimal choice (maximizing expected earnings based on session-1 performance). Regress on treatment dummies, and replicate including gender and interaction terms.
- b) Perform a similar analysis for whether advisers' advice is itself optimal given session-3 performance.

Authors' note: This analysis is presented in Table 6 and Table A.9.

17. Are gender differences in worker choice in Session 3 driven by advice, and do they differ by treatment?

- a) Split the sample by whether workers received binary or Likert advice.
- b) For binary advice: regress choice on gender and advice. Also use an ordinal confidence-weighted scale (1–3 no advice, 4–8 probably A/B, 9+ definitely A/B).
- c) For Likert advice: regress choice on gender and the 1–5 Likert recommendation.
- d) Test whether the gender gap in choices differs by treatment using regressions with gender, treatment dummies, and gender×treatment interactions.

Authors’ note: Due to the small sample size for these tests, we do not report the full pre-registered tables on the influence of advice on worker decisions. Instead, we analyze the quality of advice (Table 6 and Table A.9).

A.10 Power Calculations

Authors’ note: In what follows we reproduce the power calculation from our pre-analysis plan.

We conducted power calculations to check our power to detect (a) a gender gap in worker confidence (self-promotions), (b) a gender gap in advice and (c) treatment differences in the gender gap in advice. Based on these power calculations, we aim to collect data from 350 workers and 1750 advisers (700 in the baseline, and 525 in the Nudge and Subjective treatments). With the intended sample size, we will have a power greater than 0.90 to detect (a) gender differences in worker confidence in line with the pilot and previous literature, (b) a gender gap in advice of half the size of the gap implied by our pilots and (c) treatment effects that eliminate 70% or more of the baseline gender gap in advice. In the remainder of this document, we will present the power calculations in greater detail.

H1: Gender Differences in Self-Promotion

Our main interest in this experiment is in testing three sets of hypotheses. Our first hypothesis relates to the gender gap in self-confidence (self-promotion) among workers. We can do power calculations using the confidence data from three pilot sessions (N=301 participants). Table 1 presents the pilot results for our quantitative confidence variables (B1a-g). The gender gap was significant at the 1% level for all variables, and the point estimates for the first five variables are consistent with Exley and Kessler (2022) albeit slightly smaller.⁹ We use the averages and standard deviations for men and women to compute the power associated with a one-sided t-test testing whether men are more confident than women. The results suggests that a sample size of 210 workers (105 men and 105 women) is sufficient to achieve a power of 0.80 or greater for all six of our confidence variables, as well as when summing over all variables. A sample size of 290 workers would be sufficient for achieving a power of 0.90.

⁹The gender gaps in Exley and Kessler (2022) for the first five variables are 2.28 (p. 1359), 0.67, 13.83, 17.28 and 16.12 (Table II). They do not elicit the rank guess variable.

Table 1: Power Calculations for the Gender Gap in Self-Promotions

	Score Guess	Perform (Likert)	Perform (0-100)	Apply	Perf Job	Rank Guess	Sum
Men	10.14 (3.95)	3.06 (1.02)	44.82 (22.4)	38.63 (28.1)	39.96 (27.3)	54.44 (24.3)	0.29 (0.98)
Women	8.81 (3.75)	2.64 (1.02)	35.89 (22.5)	25.95 (25.1)	27.66 (26.1)	62.67 (22.8)	-0.29 (0.94)
Diff	1.33 (3.91)	0.42 (1.04)	8.93 (22.9)	12.68 (27.3)	12.30 (27.3)	-8.23 (23.9)	0.58 (1.0)
Samp (0.80)	210	150	158	112	118	206	70
Samp (0.90)	290	206	220	154	164	284	96

Notes: The cells report the means of the relevant variables and standard deviations are reported in parentheses. “Score Guess” asks participants to guess the number of questions they solved correctly. “Perform” elicits subjective performance on a five-point Likert scale and a 0-100 scale respectively. The next two questions ask, on a 0-100 scale, whether they would apply for a job that requires them to perform well on the quiz or be successful in such a job respectively. “Rank Guess” variable asks participants to guess their rank compared to 100 other participants (1-best, 101-worst). The final variable takes the sum of the standardized first six variables (reverse coding the rank guess variable) and then standardizes this sum. Required sample sizes (“Samp”) are based on “power twomeans mu_men mu_women, sd1(sd_men) sd2(sd_women) onesided”, where the four parameters are the ones presented in the first two rows of the table (means and sd). We present results for both power calculation with a power of 0.80 (“Samp (0.80)”) and for a power of 0.90.

H2: Gender Differences in Advice

Our second hypothesis relates to the gender gap in advice. We conducted pilot sessions with a total of 200 advisers who each evaluated 5 worker portfolios for a total of 1000 adviser-level observations. The results revealed that (a) the various confidence measures were highly correlated with each other ($r > 0.6$ in all cases) and (b) advisers relied heavily on the confidence variables in determining their advice. For example, a regression of binary advice on the worker’s guessed quiz score yielded a coefficient of 0.067 (se: 0.003, $p < 0.0001$). The resulting gender gap in advice received was also significant at 5.6pp ($p = 0.015$, one-sided t-test). This gap might understate the true expected gender gap in the main sessions, however, because the sample of workers the advisers evaluated had a lower gender gap in confidence (0.42 points on the score guess variable) than the one we saw when combining worker data from all pilots (1.33 points).

Hence, we compute power in the following way. First, we randomly pick a sample of workers from the pilot sessions with replacement. We then simulate advice data using the pilot coefficient estimates of a regression of advice on workers’ guessed quiz score. We also add noise to the simulated advice to the point that the guessed quiz score explains the same fraction of the total variance as in the pilot. We ensure that the Likert scale advice has five values by winsorizing at 1 and 5 and rounding to the nearest integer. Similarly, we binarize the binary advice measure by converting, for example, a predicted advice of 0.75 into either advice A (with 75% chance) or advice B (with 25% chance). Power is then equal to the fraction of simulated samples in which women receive less ambitious advice with $p < 0.05$ (one-sided).

Table 2 presents the results. The power calculations predict an approximately 8.6pp gender gap in the binary measure and a 0.25-point gap in the Likert scale measure. Noting that each adviser evaluates five workers, a sample size of 120-150 advisers would already yield power above the 0.80 threshold, where power is slightly higher for the Likert scale

measure. Note that this assumes that each piece of advice is an independent observation, which seems reasonable in the present context (and given that the data were simulated from data in an identical environment). Note also that we conservatively assume that advisers are not intrinsically biased against women; adding such a bias would raise the predicted gender gap and thereby further increase our power.

Table 2: Power Calculations for a Gender Difference in Advice

	Gender Gap (Binary)	Power	Gender Gap (Likert)	Power
600 evals / 120 adv	-0.085	0.724	-0.249	0.816
800 evals / 150 adv	-0.088	0.876	-0.263	0.930
1000 evals / 200 adv	-0.087	0.929	-0.254	0.964
2000 evals / 400 adv	-0.086	0.993	-0.251	1.000

Notes: The table presents results from four sets of simulations that vary the sample size from 600 to 2000 portfolio evaluations (evenly split across the two genders). Binary advice is a dummy variable for whether advice A is given, and Likert advice is a five-point scale ranging from “definitely B” to “definitely A”. Each row is based on 1000 simulated samples. “Gender Gap” is the difference between men and women on the relevant outcome averaged across all simulation samples. Power is the fraction of simulated samples in which an OLS regression of the outcome on gender produced a significant result at the 5% level. Since we predict that women receive less ambitious advice, we use a one-sided test.

H3a: Treatment Effects in Advice in the Subjective Treatment

Our third set of hypotheses relates to treatment differences in the gender gap in advice. A first between-subject treatment examines what happens if referees only have access to a subjective verbal self-evaluation (and no longer see the worker’s guessed score or other quantitative self-evaluations). For our power calculations, we assume that advice in this treatment is no longer connected to (guessed) performance. This implicitly assumes that subjective self-evaluations are not informative or, if they are, that male and female workers write similar subjective self-evaluations. We also assume that advisers are not directly influenced by the gender of the worker, i.e., they are not intrinsically gender-biased. In practice, this implies that we randomly generate advice in such a way that its distribution resembles the distribution of advice given in the pilot in terms of advice, while removing any connection to performance and gender. We then test whether the gender difference in advice is significantly smaller in this treatment (where it is zero in expectation) than in the baseline treatment (where it is approximately 8.6pp). Table 3 presents the results: having 400 advisers in each treatment gives us a power greater than 0.90.

Two further remarks are in order. First, since we use the baseline treatment in multiple comparisons, it may be prudent to oversample this treatment. Table 4 shows that shifting observations from the Subjective to the baseline treatment only starts reducing power when more than 30% of observations are shifted this way. Second, the true treatment effect may be larger than assumed here (e.g., if advisers are gender-biased) or smaller (e.g., if men write

Table 3: Power Calculations for the Subjective Treatment

Evaluations	Advisers	Power (Binary)	Power (Likert)
1000	200	0.682	0.706
2000	400	0.910	0.923
3000	600	0.977	0.990
3500	700	0.989	0.994
4000	800	0.995	0.996

Notes: The table presents results from simulations. The first column presents the number of evaluations included in each simulation sample for the baseline and (in parentheses) the subjective treatment. The second column presents the corresponding number of advisers (one per five evaluations). The two subsequent rows summarize the results of 1000 simulations, where power is calculated as the proportion of simulated samples in which the gender gap in the subjective treatment is significantly smaller than in the baseline as per a difference-in-difference test (5% level). Since we predict a smaller gender gap in the Subjective treatment, we use a one-sided test.

more optimistic subjective self-evaluations). We will discuss minimum detectable effect sizes at the end of this document.

Table 4: Power Calculations for Overrepresented Baseline in the Subjective Treatment

Per Treatment		Binary Advice	Likert Advice
#Evaluations	#Advisers	Power	Power
2000 (2000)	400 (400)	0.902	0.917
2200 (1800)	440 (360)	0.913	0.912
2400 (1600)	480 (320)	0.904	0.917
2600 (1400)	520 (280)	0.883	0.911
2800 (1200)	560 (240)	0.860	0.842
3000 (1000)	600 (200)	0.837	0.819
3200 (800)	640 (160)	0.773	0.782
3400 (600)	680 (120)	0.710	0.686
3600 (400)	720 (80)	0.578	0.555
3800 (200)	760 (40)	0.377	0.375

Notes: See the notes to Table 3.

H3b: Treatment Effects in Advice in the Nudge Treatment

Our second between-subject treatment manipulation concerns a nudging/information intervention that attempts to lower the gender gap in advice by nudging the advisers to take the existing gender gap in self-promotion into account in their advice. For our power calculations, we assume that the intervention completely eliminates the gender gap in self-confidence/self-promotion (which is equivalent to assuming that the advisers remove the gender bias in

worker confidence). We implement this in the power calculation by randomly drawing observations (i.e., confidence elicitation) from both male and female workers in the pilot, and then assigning a gender randomly. This removes the gender gap in worker self-promotions by construction, while ensuring that the confidence distribution is the same as in the pilot. Table 5 presents the results, which are very similar to the Subjective treatment (Table 3).

Table 5: Power Calculations for the Nudge Treatment

Evaluations	Advisers	Power (Binary)	Power (Likert)
1000	200	0.682	0.759
2000	400	0.909	0.951
3000	600	0.910	0.993
3500	700	0.993	0.996
4000	800	0.994	0.999

Notes: See the notes to Table 3.

H3c: Treatment Effects in Advice in the Objective Treatment

As a final step, we include a within-subject treatment testing whether providing objective information (the worker’s actual score) decreases the gender gap. For our power calculations, we simulate the data for this treatment in the same manner as in the baseline treatment above (see the section H2), except that we replace the “guessed score” with the “actual score” in the process generating simulated advice. Since we use a within-subject design, we assume that the noise term used in the simulations is identical for the baseline and Objective treatment. This essentially assumes that the adviser’s reading of the worker’s portfolio does not change other than through replacing the worker’s subjective performance with their true performance. Table 6 shows that in this treatment the gender gap reduces by about two thirds (approximately 5.3pp on the binary scale and 0.15 points on the Likert scale), which is because the gender gap in performance (while not zero) is smaller than the gender gap in confidence/self-promotions.

Table 6 also suggests that a relatively small sample size may be sufficient to detect these effects. This is because we use a within-subject design and because we assume that all advisers update in line with the new information (e.g., less ambitious advice if the true objective performance was less than the subjective performance). In practice, power may be lower if advisers choose not to incorporate the objective information (e.g., because they are happy with their original advice), or if the noise term is not perfectly correlated across the baseline and objective evaluations (e.g., if advisers change their mind about the interpretation of the portfolio from the baseline to objective choice, or if they are making mistakes or are just noisy). Nevertheless, even in such cases we can expect our power to be high for the intended sample size of our experiment, as we discuss in the next section.

Table 6: Power Calculations for the Objective Treatment

Evaluations	Advisers	Treat Effect (Binary)	Power	Treat Effect (Likert)	Power
600	120	-0.054	0.505	-0.148	0.785
800	150	-0.054	0.584	-0.153	0.882
1000	200	-0.051	0.641	-0.151	0.939
2000	400	-0.052	0.901	-0.150	0.993
3000	600	-0.053	0.971	-0.153	1.000

Notes: The table presents results from five sets of simulations that vary the sample size from 600 to 3000 portfolio evaluations. Binary advice is a dummy variable for whether advice A is given, and Likert advice is the five-point scale ranging from “definitely B” to “definitely A”. Each row presents the averages from 1000 simulated samples. “Treat Effect” is the estimated difference-in-difference between the gender gap in the baseline and the gender gap in the objective treatment averaged across all simulation samples. Power is the fraction of simulated samples in which an OLS of the difference between the outcome in the two treatments on gender produced a significant result at the 5% level (this is equivalent to a difference-in-difference test). Since we predict a smaller gender gap in the objective treatment, we use a one-sided test.

Final Sample Size and Cost Calculations

The previous sections suggest that a sample of around 300 workers would be sufficient to detect a gender gap in confidence with a power 0.90 (Table 1). To detect a gender gap in advice with a power of 0.90 a sample of 200 advisers would be sufficient (Table 2). And for studying treatment effects in the gender gap in advice, approximately 400 advisers per treatment would also achieve a power of 0.90. A larger sample size may still be helpful, however, in particular if the treatment effects are smaller than assumed in the power calculations. In addition, it is useful to oversample the baseline treatment, since it will be used in all of our tests (in contrast to the other treatments, which are used in a single comparison only).

Also keeping in mind budget constraints, we therefore aim to collect data from 350 workers and 1750 advisers (700 in the baseline, and 525 in the Nudge and Subjective treatments). The total costs are estimated to be around 9 AUD per adviser and 21.5 AUD per worker for a total of 23,275 AUD. This also implies that each adviser has a 20% chance of being matched to a worker.

With the intended sample size, we have a power greater than 0.90 to detect a gender difference in confidence among workers based on the pilot data (H1). For the gender gap in advice (H2) the present sample size would give us a power greater than 0.90 even if the gender gap in confidence is half the size in the pilot (or equivalently: the advisers put only half as much weight on worker confidence in their final decision as they did in the pilot). For our between-subject treatments, we expect to still have a power of 0.84 on the Likert scale variable to detect a decrease in the gender gap by 60%, assuming that the gender gap in the baseline is in line with the pilot (and a power greater than 0.90 if the gender gap decreases by 70% or more).

A.11 Experimental Instructions/Screenshots

11.1 Part 1 Worker Session

Welcome!

Overview: Your participation in this study involves taking part in **two sessions**. **Today, you will complete Session 1. Next week you will complete Session 2.**

Today's session is expected to take 25 minutes to complete. You will start by completing a survey, followed by a math and science quiz that will take about half of the session's time. You will get instructions for the remainder of today's session after you completed the quiz.

Your payment: You will receive a participation fee of £3 for both Session 1 and Session 2 within 24 hours of completing each session. You may also receive a bonus depending on your performance in each session. Your bonus will only be paid out to you next week once you have completed Session 2. On average, someone who completes both sessions can expect to earn £9 in total.

Please download and read through the information about the study and the consent form [here](#).

Based on the instructions provided above, please indicate whether each statement below is True or False:

Your participation in this study means you will take part in two sessions, today's session and another session next week.

- True
 False

You can only take part in Session 2 next week if you complete Session 1 today.

- True
 False

You will receive both your bonus and your participation fee of £3 within 24 hours of completing today's session.

- True
 False

It is important for the study that you only complete today's session if you can commit to also completing Session 2 next week. You can complete Session 2 during any day and time of next week. You will receive a reminder to complete Session 2 at the beginning of next week.

If you agree to participate, please write the following sentence "I commit to completing session 2 next week" in the space provided below.

Initial survey

1) Please indicate your age:

 ▾

2) Are you:

- Female Male Other

If you chose *Other*, please specify:

3) What is the highest degree or level of education you have completed?

 ▾

4) What is your current occupation?

- Studying
 Employed (full-time or part-time)
 Currently looking for work
 Not in the labour force
 Retired
 Other
 Prefer not to say

5) What is (or was) your full job title? *Examples: Retail assistant, Civil engineer, Cleaner, Nurse, Primary School Teacher*

6) How much do you get paid for your job, before tax and other deductions are taken out? If your pay before tax changes from week to week, think about what you earn on average.

 ▾

7) Now please consider your **household income** (*your household includes everyone who lives with you regularly and with whom you share money*). How do you feel about your **household income** nowadays?

 ▾

During this study, you will be represented by an "in-study" name and avatar. Please choose your name and avatar now.

Name:

- Gabriella
- Erica
- Felicia
- Caroline
- Stephanie
- Philippa
- Alana
- Georgia
- Olivia
- Louisa
- Victoria
- Jennifer
- Paula
- Alexandra
- Lucy
- Christina
- Samantha
- Julia
- Marcia
- Danielle

Avatar:






- 
- 
- 
- 
- 

During this study, you will be represented by an "in-study" name and avatar. Please choose your name and avatar now.

Name:

- Alexander
- Charles
- Eric
- Alan
- Philip
- George
- Stephen
- Marcus
- Felix
- Victor
- John
- Gabriel
- Christopher
- Louis
- Oliver
- Daniel
- Julian
- Paul
- Luke
- Samuel

Avatar:











- 
- 
- 
- 
- 

During this study, you will be represented by an "in-study" name and avatar. Please choose your name and avatar now.

Name:

- Freddie
- Alexander
- Charlie
- Louis
- Felicia
- Felix
- Sam
- Philippa
- Julian
- Stephen
- Philip
- Alex
- Louisa
- Leo
- Alexandra
- Julia
- Theo
- Stephanie

Avatar:

- 
- 
- 
- 
- 
- 
- 
- 
- 
- 

Math and science quiz

You will complete a **20 multiple choice questions quiz** that will test your skills in **math and science**.

Here is an example of a question:

If $5^5 = 5^2 \times 5^m$, what is the value of m ?

- A. 3
- B. 5.5
- C. 9
- D. 12.5

You will see each question on a separate page, and you will have **30 seconds to answer each question**.

You will earn a bonus of **£0.10 per correct answer**. (There is no penalty for wrong answers.)

Based on the instructions provided above, please indicate whether each statement below is True or False:

The quiz will test your math and science skills.

- True
- False

You will earn a bonus of £0.10 per correct answer on the quiz.

- True
- False

Math and science quiz

When you are ready to start the quiz, please press Start.

Quiz (1/20)

Time left to complete this page: 0:25

If n is a positive integer divisible by 7, and if $n < 70$, what is the greatest possible value of n ?

- A. 49
- B. 56
- C. 63
- D. 69

Quiz (2/20)

Time left to complete this page: 0:25

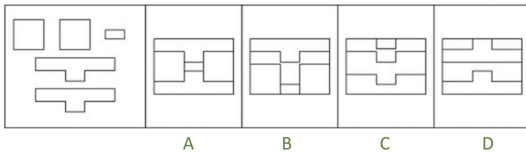
On the pH scale, any substance below what number is considered acid?

- A. 0
- B. 5
- C. 7
- D. 15

Quiz (3/20)

Time left to complete this page: 0:25

Which of the images below *best* solves the problem in the first picture?



Quiz (4/20)

Time left to complete this page: 0:25

There are 4 red marbles and 8 green marbles in a box. If Emma randomly selects a marble from the box, what is the probability of her selecting a red marble?

- A. $\frac{1}{3}$
- B. $\frac{1}{2}$
- C. $\frac{2}{3}$
- D. $\frac{4}{5}$

Quiz (5/20)

Time left to complete this page: 0:25

In the field of genetics, an organism's appearance is referred to as its?

- A. Genes
- B. Alleles
- C. Genotype
- D. Phenotype

Quiz (6/20)

Time left to complete this page: 0:24

Air is less dense than water because

- A. It is lighter.
- B. Its molecules are further apart.
- C. Its molecules are closer together.
- D. It moves quicker.

Quiz (7/20)

Time left to complete this page: 0:25

If $5^{11} = 5^2 \times 5^m$, what is the value of m ?

- A. 3
- B. 5.5
- C. 9
- D. 12.5

Quiz (8/20)

Time left to complete this page: 0:24

In vertebrates, which substance gives red blood cell their colour?

- A. carotene
- B. dopamine
- C. melanin
- D. hemoglobin

Quiz (9/20)

Time left to complete this page: 0:26

How much greater is the value of $3x + 5$ than the value of $3x - 7$?

- A. 8
- B. 10
- C. 12
- D. 14

Quiz (10/20)

Time left to complete this page: 0:25

Which of the following is **not** a blood type?

- A. A
- B. AB
- C. O
- D. X

Quiz (11/20)

Time left to complete this page: 0:26

If $(3.2 + 3.3 + 3.5)w = w$, what is the value of w ?

- A. 0
- B. $\frac{1}{10}$
- C. 1
- D. 10

Quiz (12/20)

Time left to complete this page: 0:22

Which of the following is a characteristic of the automatic nervous system?

- A. It regulates the voluntary control of body movements through the skeletal muscles.
- B. It regulates involuntary activity in the heart, stomach, lungs, and intestines.
- C. It is thought to be the centre of intelligence.
- D. It is part of the nervous system consisting of the brain and spinal cord.

Quiz (13/20)

Time left to complete this page: 0:22

DNA has which of the following structures?

- A. Double helix
- B. Helix
- C. Parabola
- D. Hexagon

Quiz (14/20)

Time left to complete this page: 0:25

What is the latitude of any point on earth's equator?

- A. 0
- B. 90
- C. 180
- D. 270

Quiz (15/20)

Time left to complete this page: 0:25

Which of the following planets is located farthest from the sun?

- A. Neptune
- B. Saturn
- C. Jupiter
- D. Uranus

Quiz (16/20)

Time left to complete this page: 0:26

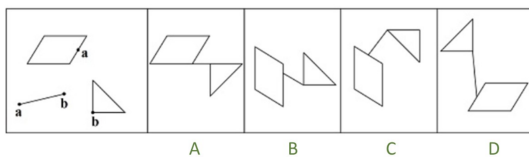
Which layer of the atmosphere is closest to the earth?

- A. Mesosphere
- B. Stratosphere
- C. Troposphere
- D. Thermosphere

Quiz (17/20)

Time left to complete this page: 0:21

Which figure best shows how the objects in the left box will touch if the letters for each object are matched?



Quiz (18/20)

Time left to complete this page: 0:25

Which of the following is **not** a factor of 90?

- A. 5
- B. 6
- C. 12
- D. 15

Quiz (19/20)

Time left to complete this page: 0:25

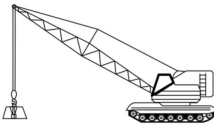
If b does not equal zero and $ab = \frac{b}{4}$, what is the value of a ?

- A. $\frac{1}{8}$
- B. $\frac{1}{4}$
- C. $\frac{1}{3}$
- D. $\frac{1}{2}$

Quiz (20/20)

Time left to complete this page: 0:23

Extending the reach of this crane will shift its:



- A. Total weight
- B. Allowable speed
- C. Centre of gravity
- D. Centre of buoyancy

End of the Quiz

Thank you for completing the quiz.

Before proceeding to the next part of today's session, you will have to **guess how many questions (out of 20) you answered correctly** in the quiz.

You will receive a bonus of £0.50 if your guess is exactly correct, and a bonus of £0.20 if your guess lies within 3 of the correct number.

How many quiz questions do you think you answered correctly (out of 20)?

 ▼

What happens next week

Before proceeding with today's session tasks, you will get information about what will happen in Session 2 (next week).

In Session 2, **you will have to decide whether to work on one of two jobs: Job A or Job B.** Both jobs will last for approximately 10 minutes and involve another math and science task. Apart from this, you will not be told anything else about the two jobs until you have to decide which job you will work on.

Prior to making your decision (in Session 2), **you will receive advice from an Adviser regarding which job you should take.** The Adviser will be another Prolific participant. **Unlike you, the Adviser knows what Job A and Job B are.** Hence, **their advice may help you make an informed decision.**

Your bonus in Session 2 will depend on your performance on your chosen job. The Adviser will earn a bonus equal to 50% of your earnings on the job. This means that it is in the Adviser's best interest to advise you to take the job that will maximize your earnings, because this will maximize their earnings too.

Based on the instructions provided above, please indicate whether each statement below is True or False:

The Adviser will decide whether you work on Job A or Job B.

- True
- False

The Adviser will advise you on whether to opt for Job A or Job B.

- True
- False

The Adviser will have more information than you about what Job A and Job B are when giving you advice.

- True
- False

Self-assessment questionnaire

You will now answer a **self-assessment questionnaire**, which will include several questions **about how well you think you did on the math and science quiz** today.

To help the Adviser give you the best possible advice about which job to choose in Session 2, the **Adviser will see one or more of your answers** in this questionnaire.

Based on the instructions provided above, please indicate whether the statement below is True or False:

The Adviser will see one or more of your answers in the self-assessment questionnaire.

- True
- False

Self-assessment questionnaire (1/4)

In the space provided below, please describe (using 2 to 3 sentences) how well you think you performed today on the math and science quiz.

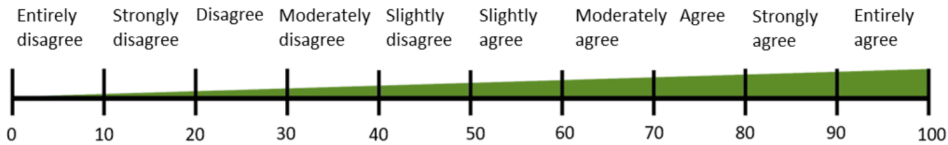
Self-assessment questionnaire (2/4)

Please indicate how good you think your performance was on the quiz.

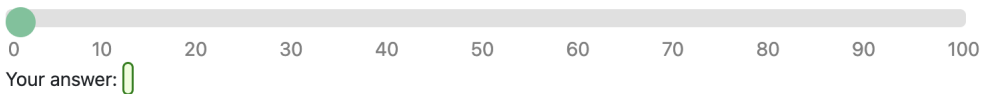
Terrible	Very poor	Poor	So-so	Good	Very good	Exceptional
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Self-assessment questionnaire (3/4)

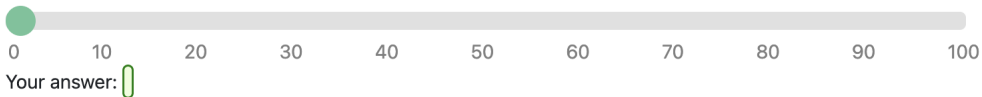
Please indicate the extent to which you agree with each of the following statements, on a scale from 0 (entirely disagree) to 100 (entirely agree).



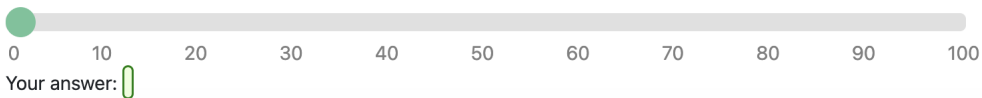
I performed well on the quiz.



I would apply for a job that required me to perform well on the quiz.



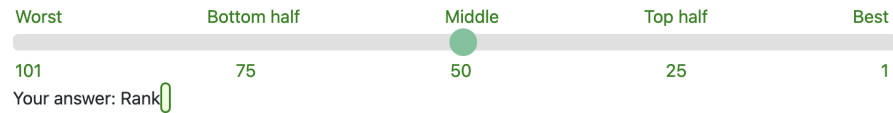
I would succeed in a job that required me to perform well on the quiz.



Self-assessment questionnaire (4/4)

For the next question, you will **guess how well you performed on the math and science quiz compared to 100 other participants in Prolific**. You can choose any number **between 1 (better than everyone else) and 101 (worse than everyone else)** as your best guess for your rank.

How well do you think you performed on the quiz compared to 100 other participants?



Final questionnaire

Thank you for completing the self-assessment questionnaire.

We will now ask you to answer two **final questions about how you see yourself in general**.

Q1. Are you in general a person who is very confident or someone who is not confident at all? Please tick a box on the scale, where the value 0 means: *not confident at all* and the value 10 means: *very confident*

0 1 2 3 4 5 6 7 8 9 10

Q2. Are you in general a person who is fully prepared to take risks or do you try to avoid taking risks? Please tick a box on the scale, where the value 0 means: *not at all willing to take risks* and the value 10 means: *very willing to take risks*

0 1 2 3 4 5 6 7 8 9 10

End of Session 1

Thank you for completing Session 1. You will receive your participation fee of £3 within 24 hours.

Sometime next week, you will receive a reminder through Prolific to complete Session 2 of this study. Session 2 is expected to take 15 minutes to complete. You can choose to complete Session 2 any day and time next week, after you receive the invitation to complete Session 2.

You will receive the bonus payment for today's session after completing Session 2 next week.

To confirm completion of Session 1 and receive your payment, please press the button below before exiting the screen.

Finish and return to the Prolific website

11.2 Part 2 Adviser Session

Baseline treatment

Welcome

Overview: In this study, you will be taking the role of an Adviser. Your main task will be to advise a Worker (another Prolific participant), on which job to complete in this study. You will also be asked to answer a short survey. The study is expected to take 20 minutes to complete.

Your payment: You will receive a participation fee of £3 within 24 hours of completing the study. In addition, there is a 20% chance that your advice will be sent to the Worker. If so, you will receive a bonus, depending on the performance of the Worker who received your advice. Your bonus will be paid out to you after the Worker has completed their job (within two weeks from today).

Please download and read through information about the study and the consent form [here](#).

Based on the instructions provided above, please indicate whether each statement below is True or False:

You will advise the Worker on which job to choose.

- True
 False

You will receive any bonus payment within 24 hours of completing the study.

- True
 False

You will receive any bonus payment with a delay, after the Worker has performed their job.

- True
 False

The Worker's Job

In this study, **you will be matched to a Worker** who is another participant on Prolific. The **Worker will have to choose between two jobs: Job A (Advanced) and Job B (Basic)**. Both jobs require the Worker to answer **20 multiple choice math and science questions**, with a limited time of 30 seconds per question. You can see example questions below.

Example 1

If $5^5 = 5^2 \times 5^m$, what is the value of m ?

- A. 3
B. 5.5
C. 9
D. 12.5

Example 2

Which layer of the atmosphere is closest to the earth?

- A. Mesosphere
B. Stratosphere
C. Troposphere
D. Thermosphere

Job A (Advanced) and Job B (Basic) require the Worker to answer the same 20 questions within the same time limit but **their bonuses differ**, as follows:

- **Job A (Advanced):** Worker receives a bonus of **£2** if they get **8 or more correct answers**, and no bonus otherwise.
- **Job B (Basic):** Worker receives a bonus of **£1** if they get **5 or more correct answers**, and no bonus otherwise.

Advice

Your task is to give advice to the Worker regarding which job to take. When the Worker has to decide whether to take Job A (Advanced) or Job B (Basic), they do not know what task each job entails nor the associated bonus. Hence, your advice may help the Worker make an informed decision. If your advice is sent to the Worker, they will see it before making their decision.

A few days ago, the Worker took part in another study in which they completed a math and science quiz that is similar to the job they will have to work on. They **then answered self-assessment questions regarding their performance** in that quiz, without having received any feedback about how well they actually did.

To help you give the Worker the best possible advice, **you will receive the Worker's answers to the self-assessment questions**, where the Worker describes how well they think they did in the math and science quiz.

Your Bonus

There is a 20% chance that your advice is sent to the Worker. If so, you will receive a bonus equal to 50% of the earnings of the Worker. Here are the possible outcomes:

- Worker chooses **Job A (Advanced)**: you earn a bonus of **£1** if they have **8 or more correct answers**, and no bonus otherwise.
- Worker chooses **Job B (Basic)**: you earn a bonus of **£0.50** if they have **5 or more correct answers**, and no bonus otherwise.

It is in your best interest to advise the Worker to take the job that will maximize their earnings, because this will maximize your earnings too.

Based on the instructions on the previous pages, please indicate whether each statement below is True or False:

You decide whether the Worker works on Job A (Advanced) or Job B (Basic).

- True
 False

You give advice to the Worker on whether to opt for Job A (Advanced) or Job B (Basic).

- True
 False

Your bonus depends on the Worker's performance on their chosen job.

- True
 False

Nudge Treatment Screen

Information

On the next page, you will be able to give advice to the Worker. Below are some facts that you may want to keep in mind when giving advice:

- 1 *Using data from more than 4,000 participants in online labour markets, researchers at Harvard University found that women rated their performance in a math & science quiz lower than men, despite similar performance levels. As a result, women who were hired performed better than men on average.*
- 2 *Using data from more than 3,600 people, researchers at the University College London report that men are more confident in the self-assessment of their abilities compared to women. This gender gap in self-assessment contributes to the gender gap in who ends up in a top job.*
- 3 *An internal study by the IT firm Hewlett Packard found that men who apply for a job meet only 60% of the qualifications, but women who apply meet 100% of them.*

Based on the information provided above, please answer the following questions:

Men were more confident in their self-assessment in the previous studies.

- True
 False

Suppose you are comparing a male and female worker with the same self-assessment of their performance. Based on previous research, which of the two workers is likely to have performed better?

- Male
 Female
 Neither

Do you think the information you received will affect your advice?


- Yes
 No

In one or two sentences, please briefly justify your answer to the previous question.

Advice

You will now advise the following Worker:

Name: **Lucy**

A simple black line-art icon of a woman's head and shoulders. She has short, wavy hair and is wearing a collared shirt. The icon is centered within a light green rounded rectangular border.

Advice

Below are Lucy's answers to the self-assessment questions. Please read her answers carefully before giving your advice.

Name: **Lucy**



Q1. Please describe how well you think you performed on the math and science quiz.

Lucy's answer:

I don't think I did very well. I didn't have time to think for very long what the answers are so I was a bit under pressure and out of time on quite a few questions

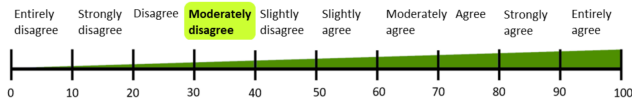
Q2. How many quiz questions do you think you answered correctly (out of 20)?

Lucy's answer: **15**

Q3. To which extent do you agree with each of the following statements? Please indicate your answer on a scale from 0 (entirely disagree) to 100 (entirely agree).

I performed well on the quiz.

Lucy's answer: **30**



I would apply for a job that required me to perform well on the quiz.

Lucy's answer: **40**



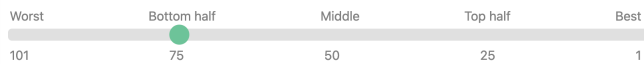
I would succeed in a job that required me to perform well on the quiz.

Lucy's answer: **30**



Q4. How well do you think you performed on the quiz compared to 100 other participants? Please give an answer between 1 (better than everyone else) and 101 (worse than everyone else).

Lucy's answer: Rank **75**



Please submit your advice to Lucy by answering the questions below. Your answer to one of these questions will be sent to Lucy and help her decide between Job A (Advanced) and Job B (Basic).

Q1. Which advice do you want to send to Lucy?

Definitely Job B	Most likely Job B	No advice	Most likely Job A	Definitely Job A
------------------	-------------------	-----------	-------------------	------------------

Please indicate your answer:

Q2. Which job do you advise that Lucy chooses?

Job A (Advanced) Job B (Basic)

On a scale from 0 to 10, how confident are you that your advice in Q2 is the best for Lucy?

0 1 2 3 4 5 6 7 8 9 10

Advice

Now imagine that instead of having to choose between Job A (Advanced) and Job B (Basic), Lucy would choose her performance goal as follows:

- Goal 1: £0.20 for 1 or more correct answers
- Goal 2: £0.40 for 2 or more correct answers
- Goal 3: £0.60 for 3 or more correct answers
- (...)
- Goal 9: £1.80 for 9 or more correct answers
- Goal 10: £2.00 for 10 or more correct answers
- Goal 11: £2.20 for 11 or more correct answers
- (...)
- Goal 18: £3.60 for 18 or more correct answers
- Goal 19: £3.80 for 19 or more correct answers
- Goal 20: £4.00 for 20 correct answers

This means that Lucy would earn a bonus of £0.20 times the selected performance goal if her performance met or exceeded her goal, and no bonus otherwise. You would always earn a bonus equivalent to 50% of Lucy's earnings.

Which goal would you advise Lucy to choose in this case?

[See Lucy's answers to self-assessment questions again](#)

Advice

Thank you for submitting your advice.

Next, you will be asked to give advice to four additional Workers. Just like for your previous Worker (Lucy), there is a chance that these workers will receive your advice. Overall, there is a 20% chance that your advice is sent to a Worker. Each of the five Workers (including the previous Worker, Lucy) is equally likely to be selected.

[Adviser sees 4 additional worker portfolios, sequentially]

Objective Treatment

Advice

Thank you for submitting your advice.

Next, you will see the portfolios of each of the Workers again. This time, in addition to their answers to the self-assessment questions, you will also **learn about their actual score in the previous math and science quiz**, and you will have the possibility to **change your advice** about which job they should choose.

Based on the instructions provided above, please indicate whether each statement below is True or False:

You will see the actual score of the Workers in the math and science quiz.

- True
- False

You will be able to change your advice.

- True
- False

Advice to 1st Worker

Below are Lucy's answers to the self-assessment questions and **new information about her score** in the previous math and science quiz.

Name: Lucy



NEW: Lucy got a score of 6/20 in the previous math and science quiz.

Q1. Please describe how well you think you performed on the math and science quiz.

Lucy's answer:

I don't think I did very well. I didn't have time to think for very long what the answers are so I was a bit under pressure and out of time on quite a few questions

Q2. How many quiz questions do you think you answered correctly (out of 20)?

Lucy's answer: 15

Q3. To which extent do you agree with each of the following statements? Please indicate your answer on a scale from 0 (entirely disagree) to 100 (entirely agree).

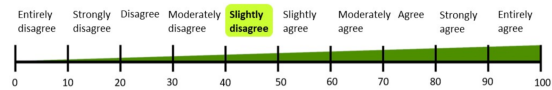
I performed well on the quiz.

Lucy's answer: 30



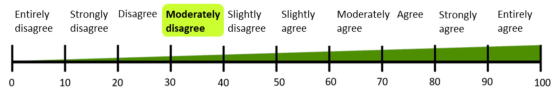
I would apply for a job that required me to perform well on the quiz.

Lucy's answer: 40



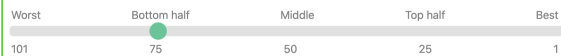
I would succeed in a job that required me to perform well on the quiz.

Lucy's answer: 30



Q4. How well do you think you performed on the quiz compared to 100 other participants? Please give an answer between 1 (better than everyone else) and 101 (worse than everyone else).

Lucy's answer: Rank 75



Below is the advice you previously gave to Lucy. You can either change or maintain your previous advice. The only new information you received is that Lucy's actual score on the previous math and science quiz was 6/20. If you wish to change your advice, please select your new advice, and then press Next. If you do not wish to change your advice, please press Next.

Q1. Which advice do you want to send to Lucy?

- Definitely Job B
 Most likely Job B
 No advice
 Most likely Job A
 Definitely Job A

Please indicate your answer:

-

Q2. Which job do you advise that Lucy chooses?

- Job A (Advanced)
 Job B (Basic)

On a scale from 0 to 10, how confident are you that your advice in Q2 is the best for Lucy?

- 0
 1
 2
 3
 4
 5
 6
 7
 8
 9
 10

Q3. Which performance goal would you advise Lucy to choose?

Recall that Lucy would earn a bonus of £0.20 times the selected performance goal if her performance met or exceeded her goal, and no bonus otherwise.

1

Final questionnaire (1/2)

Thank you for submitting your advice. We will now ask you to answer a **final questionnaire**.

1) Please indicate your age:

2) Are you:

Female Male Other

If you chose *Other*, please specify:

3) What is the highest degree or level of education you have completed?

4) What is your current occupation?

- Studying
- Employed (full-time or part-time)
- Currently looking for work
- Not in the labour force
- Retired
- Other
- Prefer not to say

5) What is (or was) your full job title? *Examples: Retail assistant, Civil engineer, Cleaner, Nurse, Primary School Teacher*

6) How much do you get paid for your job, before tax and other deductions are taken out? If your pay before tax changes from week to week, think about what you earn on average.

7) Now please consider your **household income** (*your household includes everyone who lives with you regularly and with whom you share money*). How do you feel about your **household income** nowadays?

8) Are you in general a person who is very confident or someone who is not confident at all? Please tick a box on the scale, where the value 0 means: *not confident at all* and the value 10 means: *very confident*

0 1 2 3 4 5 6 7 8 9 10

9) Are you in general a person who is fully prepared to take risks or do you try to avoid taking risks? Please tick a box on the scale, where the value 0 means: *not at all willing to take risks* and the value 10 means: *very willing to take risks*

0 1 2 3 4 5 6 7 8 9 10

10) Who do you think, if anybody, would perform better on the math and science quiz?

Women a lot better	Women better	Women slightly better	Both equally well	Men slightly better	Men better	Men a lot better
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

11) Who do you think, if anybody, would be more confident in their self-assessment?

Women a lot more confident	Women more confident	Women slightly more confident	Both equally confident	Men slightly more confident	Men more confident	Men a lot more confident
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

12) Is there any feedback you would like to give us about the study?

Please **pick a name** for yourself from the list below. (They will be used when communicating your advice to the Worker.)

- Stephanie
- Stephen
- Felicia
- Sam
- Louis
- Louisa
- Alexander
- Philip
- Julian
- Julia
- Freddie
- Leo
- Charlie
- Alexandra
- Alex
- Philippa
- Theo
- Felix

End of the Study

Thank you for completing this study. Your payment is as follows:

Participation fee: £3. You will receive this payment within 24 hours.

In addition to the participation fee, if your advice is sent to a Worker, you may receive a bonus, depending on your worker's performance in their chosen job, which will be paid out to you within 2 weeks.

Thank you for your participation.

**To confirm completion and receive your payment,
please press the button below before exiting the screen.**

Finish and return to the Prolific website

11.3 Part 3 Worker Session

Welcome back!

Thank you for returning to the study!

Last week you did a math and science quiz and answered questions about your performance in the quiz.

Today you will have to decide whether to work on one of two jobs: Job A (the advanced job) and Job B (the basic job).

Both jobs will last for approximately 10 minutes and require you to work on a math and science task. Apart from this, you will not be told anything else about the two jobs until after you have decided which job you will work on.

Prior to making your decision, you will receive advice from an Adviser regarding which job you should take. The Adviser is another Prolific participant. The Adviser has seen some of your answers to last week's self-assessment questions and knows what the two jobs are. Because the Adviser earns a bonus equal to 50% of your earnings in the job, it is in the Adviser's best interest to help you make an informed decision.

Your payment: You will receive a participation fee of £3, your bonus for Session 1 and your bonus for today's session, within 24 hours of completing today's session.

Based on the instructions provided above, please indicate whether each statement below is True or False:

The Adviser will advise you on whether to take Job A (Advanced) or Job B (Basic).

- True
 False

The Adviser has more information than you about what Job A (Advanced) and Job B (Basic) are.

- True
 False

You will learn what your chosen job (A or B) entails after you have made your choice and before you start your chosen job.

- True
 False

Advice & Decision

Your Adviser, who chose the name John, sends you the following advice:

The best job for you is definitely Job A.

Please indicate **your decision** about which job you will work on:

- Job A (Advanced)
 Job B (Basic)

Advice & Decision

Your Adviser, who chose the name John, sends you the following advice:

You should choose Job A.

He rated his **confidence** level in his advice as **9 out of 10**.

Please indicate **your decision** about which job you will work on:

- Job A (Advanced)
 Job B (Basic)

Job A

You decided to work on Job A (Advanced). The job consists of a **quiz with 20 multiple choice questions testing your skills in math and science**. You will have up to 30 seconds to answer each question.

For this job, you receive a bonus of **£2 if you solve at least 8 questions correctly**. If your score is lower than 8, you will not receive any bonus.

When you are ready, please press the Next button below to start the task.

Job B

You decided to work on Job B (Basic). The job consists of a **quiz with 20 multiple choice questions testing your skills in math and science**. You will have up to 30 seconds to answer each question.

For this job, you receive a bonus of **£1 if you solve at least 5 questions correctly**. If your score is lower than 5, you will not receive any bonus.

When you are ready, please press the Next button below to start the task.

Quiz (1/20)

Time left to complete this page: 0:25

If $(3.2 + 3.3 + 3.5)w = w$, what is the value of w ?

- A. 0
B. $\frac{1}{10}$
C. 1
D. 10

Quiz (2/20)

Time left to complete this page: 0:25

In the field of genetics, an organism's appearance is referred to as its?

- A. Genes
- B. Alleles
- C. Genotype
- D. Phenotype

Quiz (3/20)

Time left to complete this page: 0:25

Which layer of the atmosphere is closest to the earth?

- A. Mesosphere
- B. Stratosphere
- C. Troposphere
- D. Thermosphere

Quiz (4/20)

Time left to complete this page: 0:24

Air is less dense than water because

- A. It is lighter.
- B. Its molecules are further apart.
- C. Its molecules are closer together.
- D. It moves quicker.

Quiz (5/20)

Time left to complete this page: 0:24

What is the latitude of any point on earth's equator?

- A. 0
- B. 90
- C. 180
- D. 270

Quiz (6/20)

Time left to complete this page: 0:25

DNA has which of the following structures?

- A. Double helix
- B. Helix
- C. Parabola
- D. Hexagon

Quiz (7/20)

Time left to complete this page: 0:19

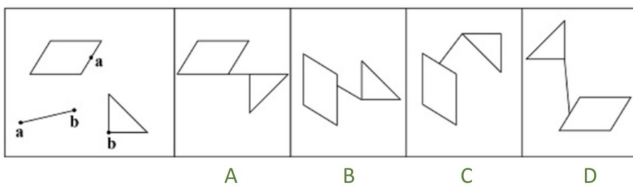
If b does not equal zero and $ab = \frac{b}{4}$, what is the value of a ?

- A. $\frac{1}{8}$
- B. $\frac{1}{4}$
- C. $\frac{1}{3}$
- D. $\frac{1}{2}$

Quiz (8/20)

Time left to complete this page: 0:25

Which figure best shows how the objects in the left box will touch if the letters for each object are matched?



Quiz (9/20)

Time left to complete this page: 0:24

If n is a positive integer divisible by 7, and if $n < 70$, what is the greatest possible value of n ?

- A. 49
- B. 56
- C. 63
- D. 69

Quiz (10/20)

Time left to complete this page: 0:25

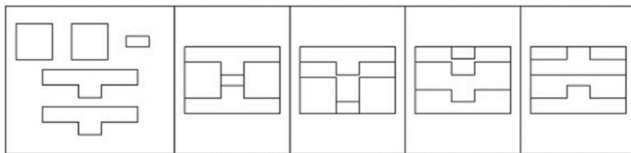
There are 4 red marbles and 8 green marbles in a box. If Emma randomly selects a marble from the box, what is the probability of her selecting a red marble?

- A. $\frac{1}{3}$
- B. $\frac{1}{2}$
- C. $\frac{2}{3}$
- D. $\frac{4}{5}$

Quiz (11/20)

Time left to complete this page: 0:25

Which of the images below *best* solves the problem in the first picture?



A

B

C

D

Quiz (12/20)

Time left to complete this page: 0:20

Which of the following is **not** a factor of 90?

- A. 5
- B. 6
- C. 12
- D. 15

Quiz (13/20)

Time left to complete this page: 0:26

How much greater is the value of $3x + 5$ than the value of $3x - 7$?

- A. 8
- B. 10
- C. 12
- D. 14

Quiz (14/20)

Time left to complete this page: 0:26

On the pH scale, any substance below what number is considered acid?

- A. 0
- B. 5
- C. 7
- D. 15

Quiz (15/20)

Time left to complete this page: 0:26

Which of the following is a characteristic of the automatic nervous system?

- A. It regulates the voluntary control of body movements through the skeletal muscles.
- B. It regulates involuntary activity in the heart, stomach, lungs, and intestines.
- C. It is thought to be the centre of intelligence.
- D. It is part of the nervous system consisting of the brain and spinal cord.

Quiz (16/20)

Time left to complete this page: 0:24

Which of the following is **not** a blood type?

- A. A
- B. AB
- C. O
- D. X

Quiz (17/20)

Time left to complete this page: 0:24

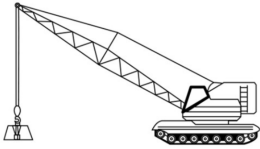
If $5^{11} = 5^2 \times 5^m$, what is the value of m ?

- A. 3
- B. 5.5
- C. 9
- D. 12.5

Quiz (18/20)

Time left to complete this page: 0:24

Extending the reach of this crane will shift its:



- A. Total weight
- B. Allowable speed
- C. Centre of gravity
- D. Centre of buoyancy

Quiz (19/20)

Time left to complete this page: 0:23

In vertebrates, which substance gives red blood cell their colour?

- A. carotene
- B. dopamine
- C. melanin
- D. hemoglobin

Quiz (20/20)

Time left to complete this page: 0:25

Which of the following planets is located farthest from the sun?

- A. Neptune
- B. Saturn
- C. Jupiter
- D. Uranus

End of the quiz

Thank you for completing the quiz.

Before ending today's session, is there any feedback you would like to give us about the study?

End of the study!

Thank you for completing this study. Your payment for today's session is as follows:

Bonus from Session 1: You answered 18 correct answers in the quiz in Session 1, for a bonus of £1.80. In addition, you will receive £0.00 based on the accuracy of your guess regarding your performance in Session 1.

Bonus from Session 2: You answered 5 correct answers in the quiz, for a bonus of £1.00.

Including your £3 participation fee, your total payment for today's session is: £5.80. You will receive this payment within 24 hours.

**To confirm completion of the study and receive your payment,
please press the button below before exiting the screen.**

Finish and return to the Prolific website