

Gender and preferences for performance feedback

Katherine Coffman* and David Klinowski†

January 2024

Abstract

Across multiple studies, we investigate whether there are gender differences in preferences for receiving performance feedback. We vary many features of the feedback context: whether the performance task is a cognitive test or a mock interview, whether the feedback is objective or subjective, and whether it is possible for the provider of the feedback to discriminate on the basis of gender. Consistent with past work, we find that women are less optimistic about their performance than men, and that, on average, more optimistic individuals have greater demand for feedback. Results like these have been hypothesized in the literature to imply that women will shy away from performance feedback more so than men. And, when we survey participants from a similar population, they also anticipate that women will demand feedback at lower rates than men. Yet, across our two incentivized studies, we find that women are no less eager to receive performance feedback than men. Understanding whether and how these results might generalize to broader contexts, particularly those with more social factors, is an important question for future work.

* Harvard Business School. kcoffman@hbs.edu.

† Katz Graduate School of Business, University of Pittsburgh. dklinowski@katz.pitt.edu.

1 Introduction

Performance feedback is prevalent in educational and professional settings. Peers, teachers, bosses, and mentors often have both the knowledge and the opportunity to offer feedback about an individual's strengths and weaknesses. This type of feedback may inform decisions about what educational tracks and career paths to pursue. Of course, in many important contexts, opportunities for feedback are not exogenously assigned: individuals can seek out, or avoid, performance feedback. These decisions about whether, when, and how intensively to pursue performance feedback may have implications for the quantity and quality of information an individual has about her own talents. Furthermore, preferences over whether to receive or avoid feedback may shape what opportunities and paths an individual pursues in the first place; a feedback-averse individual may choose to avoid careers that entail a lot of feedback, driving sorting into certain educational tracks and careers. For instance, someone who enjoys feedback may opt into more competitive, results-oriented industries or roles. Moreover, the management literature has found that "feedback-seeking behavior" (Ashford, Blatt, and Vandewalle, 2003) might in itself improve a person's labor market outcomes, as "[f]eedback-seeking behavior [...] has been linked to higher job satisfaction, greater creativity on the job, faster adaptation in a new organization or role, and lower turnover" (Stone and Heen, 2014 p. 9). For these reasons, understanding demand for performance feedback may be valuable for understanding economic outcomes.

In this paper, we explore preferences for feedback, focusing on gender differences. Cleanly identifying gender differences in demand for feedback is challenging using observational data, which is potentially plagued by selection and confounds. Instead, we conduct a series of controlled experiments in which participants complete a task, report their beliefs of their performance, and make incentivized decisions about whether to receive feedback on their performance. In addition,

we elicit beliefs from a separate sample about whether there are gender differences in demand for feedback, for our tasks and more broadly.

We explore demand for feedback across a range of settings, including examining different tasks, manipulating the difficulty of the task, allowing the possibility of gender-based discrimination by the provider of the feedback, and examining demand for both objective and subjective feedback. In each of these settings, we find that women demand performance feedback no less than men. This is notable given that our studies incorporate several aspects that ex-ante might be expected to lead women to disproportionately shy away from feedback, including receiving feedback on more stereotypically male-typed skills and in domains where they are less self-confident than men. In fact, when we ask different participants from a similar population to predict our results, they believe that women will demand performance feedback less than men. Yet, across a range of regression specifications with and without controls, we consistently find that women's demand for feedback is no less than that of men.

Our experiments share a common format: participants complete a task and we elicit their demand for performance feedback on that task. In Study 1, participants take a cognitive skills test featuring questions on math, general science, and mechanical comprehension, exogenously varying the difficulty of the test across participants. We elicit incentivized demand for learning their absolute and relative performance on the test. In Study 2, participants provide written answers to three questions about their own life achievements and personality that are commonly asked in job interviews. We then elicit incentivized demand for learning their relative performance in terms of assessed intellectual curiosity, tendency to strive for achievement, assertiveness, and tolerance to stress. These assessments are made by Human Resources (HR) professionals that we hire and are based upon the written answers participants provide. We exogenously vary whether the HR

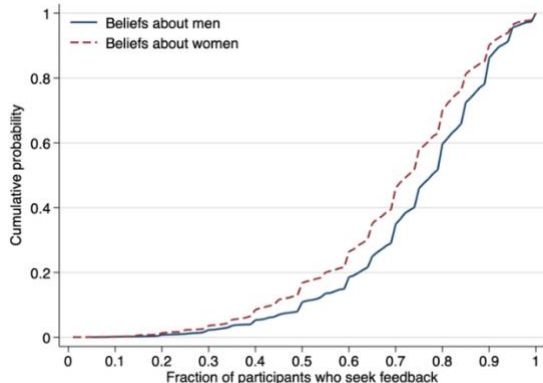
professionals observe the gender of the participants when evaluating their answers. This varies whether discrimination on the basis of gender is possible across treatments, allowing us to test whether the possibility of discrimination by the feedback provider impacts preferences for feedback.

In both studies, we collect individuals' ex-ante beliefs about their own performance, including both beliefs of how they performed and how certain they are about those beliefs. Across our two studies, we find that participants who have more optimistic beliefs of their performance, and who are more certain of their beliefs, have greater demand for feedback. While we see evidence of gender differences in beliefs about own performance (men more optimistic than women), this does not produce a gender gap in demand for feedback. On net, women are no more feedback averse than men.

We ask individuals from a separate population to predict our study participants' demand for feedback. We collect incentivized beliefs of the share of male and female participants that opt for feedback. Figure 1a shows that they do anticipate a gap: pooling across all opportunities to seek out feedback, the mean belief is that 70.2% of women and 74.7% of men would opt for feedback across our two studies (t-test $p < 0.001$), and beliefs about men's demand for feedback stochastically dominate beliefs about women's demand ($p < 0.001$ for both Kolmogorov-Smirnov and Somers' D tests). We also collect more qualitative beliefs of gender differences in willingness to seek feedback using unincentivized Likert-style questions. The belief that men have greater demand for feedback is approximately twice as common as the belief that women have greater demand for feedback, both in terms of our study tasks and in educational and professional settings more generally (Wilcoxon signed rank test $p < 0.001$ for both). Thus, while we see that women in

our studies demand performance feedback no less than men, others believe them to be more feedback averse than men.

a. Distribution of incentivized beliefs



b. Likert: Who demands feedback more?

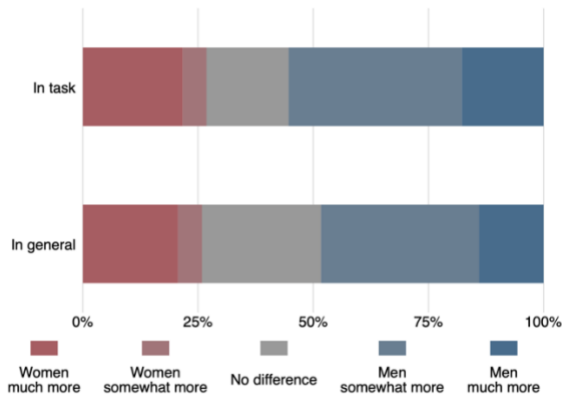


Figure 1 Beliefs of others' demand for feedback

Notes: Panel a plots the cumulative distribution of beliefs of what fraction of men and women choose to receive feedback on their performance on the task when there is no cost to receive or avoid feedback. Panel b shows responses to the 1-5 Likert questions “Overall, for the task participants completed in the previous stud, how would you describe differences in men’s and women’s preferences for finding out how they performed?” (In task), and “Thinking more generally—not just for the task the participants completed in the previous study—how would you describe gender differences in preferences for finding out how they performed in educational and professional settings, such as tasks in school and work?” (In general). Figures pool observations from the Cognitive Test and Interview studies.

Our main contribution is to test the hypothesis that women demand less feedback than men, using two well-powered, pre-registered experiments designed for this purpose. While this hypothesis has been discussed in the behavioral economics literature, e.g., in Niederle and Vesterlund (2007), it has been largely unexplored with a few important exceptions. In Eil and Rao (2011) and Mobius et al. (2022), subjects can pay to learn or avoid learning their relative performance on an ego-relevant domain after receiving a series of noisy signals. Both studies find no gender differences in the valuation for the information on average, although Mobius et al. (2022) find in addition that women are more likely than men to pay to avoid the information.³ In

³ Eil and Rao (2011) find some evidence that men have higher willingness to pay for the information than women among very confident participants and that women require a larger subsidy than men to receive the information among very underconfident participants, although these differences are not statistically significant.

Castagnetti and Schmacker (2022) and Sharma and Castagnetti (2023), subjects choose between different information structures that give noisy feedback on their relative performance on an ego-relevant domain. Sharma and Castagnetti (2023) find that women prefer more noisy structures than men, suggesting that women are more likely to avoid feedback, but Castagnetti and Schmacker (2022) find no such difference.

We build on this important work by conducting two well-powered, pre-registered studies specifically designed to explore gender differences in demand for feedback. Across our studies, we examine the role of confidence in explaining any gender gap in preferences for feedback. We also examine whether the possibility of gender-based discrimination by the provider of feedback impacts men's and women's demand for feedback differently, a factor that has not been studied previously. We complement our experiments with parallel studies investigating beliefs about gender differences in demand for feedback. Together, our studies bring new evidence to the important question of whether men and women vary in their preferences for receiving feedback, laying the groundwork for future investigations.

Finally, we should emphasize that our studies consider specific settings that intentionally shut down some interesting channels through which preferences for feedback may operate. The feedback provided in our study is not particularly actionable, nor does it have significant strategic value. It is provided within a minimalistic context, stripped of many social aspects that may loom large in other environments, such as relational, reputational, and power dynamics between the feedback-provider and feedback-recipient. Thus, we cannot rule out gender differences in receiving performance feedback in other settings. What we can say from our evidence is that, absent these other factors, it does not seem to be the case that women demand feedback less than

men. This finding may be important in interpreting and addressing gender gaps in demand for feedback in other settings, as it suggests there may be ways to alter contexts to reduce differences.

2 Hypotheses

This section presents the hypotheses that guide our analysis. Our primary interest in both studies is to test for gender differences in the demand for performance feedback.

Past literature on ego management and self-confidence provides clues as to why women might demand performance feedback less than men. In our studies, feedback has minimal instrumental value, at least within the study, as feedback is received at the end of the experiment and has no value for decision-making within the experiment. There is a large body of evidence that information is valued for reasons other than its instrumental use (for reviews see Golman, Hagmann, and Loewenstein, 2017, or the discussion of the literature by Masatlioglu, Orhun, and Raymond, 2021). One reason is ego management: if individuals derive utility from holding favorable beliefs about themselves, then information that affects these beliefs can affect utility directly.⁴ Indeed, Eil and Rao (2011), Burks et al. (2013), Masatlioglu, Orhun, and Raymond (2021), and Golman et al. (2022) show that individuals who are more optimistic about their performance on an ego-relevant task are more eager to receive performance feedback, indicating an intrinsic preference for news that is positive for self-image.^{5,6}

⁴ For reviews of the literature on belief-based utility, see e.g., Benabou and Tirole (2016) and Molnar and Loewenstein (2022).

⁵ Köszegi (2006) and Weinberg (2009) present models that make the opposite prediction. In these models, agents derive utility from believing they have high ability, thus low self-confidence agents are predicted to seek feedback on their ability in the hope of updating their beliefs upward, while high self-confidence agents are predicted to shun feedback to avoid the risk of revising their beliefs downward. Eliaz and Spiegel (2006) show that these models have difficulty explaining a positive relation between beliefs and demand for information.

⁶ Others have explored preferences for information in domains that are not necessarily ego-relevant, and have similarly found that individuals seek information that is expected to deliver good news and avoid information that is expected to deliver bad news even though avoiding it might be detrimental to decision making. Examples include investors looking up their portfolios more frequently in good market days (Karlsson, Loewenstein, and Seppi, 2009; Sicherman

Combine this insight with the findings of the literature on gender differences in self-confidence. Across a range of studies, researchers have found that women are more pessimistic than men about their own abilities and performance on tasks conditional on true ability (e.g., Niederle and Vesterlund, 2007; Grosse and Reiner, 2010; Shurchkov, 2012; Coffman, 2014; Buser, Niederle, and Oosterbeek, 2014; Bordalo et al., 2019; Klinowski, 2019; Exley and Kessler, 2022; Mobius et al., 2022; Exley and Nielsen, 2023). This evidence comes mostly from experiments in which participants perform male-typed, ego-relevant tasks, such as a cognitive skills test. Taken together, the implication would seem to be that women are more feedback averse than men: if more confident individuals are more eager to receive performance feedback, and women are less self-confident than men, they may demand less performance feedback.

Another reason men and women might differ in their demand for performance feedback could be anticipated responses to feedback. In studies that have considered exogenously-provided performance feedback, some researchers have found that women update their beliefs more conservatively than men do (Mobius et al., 2022; Coutts, 2018), particularly in male-typed domains (Coffman, Collis, and Kulkarni, 2023). And, a slew of recent work has documented that failure or negative feedback seems to deter women more so than men (Gill and Prowse, 2014; Ellison and Swanson, 2018; Brown et al., 2019; Buser and Yuan, 2019; Fang et al., 2021; Kang et al., 2021; Pereda et al., 2023; Wasserman, 2023). Similarly, Shastry et al. (2020) find that men are more likely to explain away negative feedback to luck; consistent with this finding, Coffman, Ugalde Araya, and Zafar (2023) show that women hold more pessimistic beliefs about their abilities after receiving negative feedback compared to men. These findings raise the question of

et al., 2016) or individuals avoiding medical tests for fear of receiving a positive result (Thornton, 2008; Oster, Shoulson, and Dorsey, 2013; Ganguly and Tasoff, 2017; Schünemann, Strulik, and Trimborn, 2023; and theoretical models Caplin and Eliaz, 2003, and Schweizer and Szech, 2018; see also Klinowski and Paulsen, 2013).

whether men and women, who differ in their responses to feedback, will also vary in their demand for feedback. If women anticipate more negative reactions to bad news, could this lead to lower demand for performance feedback?

Hypothesis 1: *Women demand performance feedback less than men.*

Hypothesis 1a: *Women are more pessimistic than men about their own performance.*

Hypothesis 1b: *Beliefs of own performance are positively related to demand for performance feedback.*

Our study designs allow us to test Hypothesis 1 as well as the sub-hypotheses that inform this prediction. Assuming evidence in favor of Hypothesis 1, 1a, and 1b, we expect that controlling for beliefs of own performance would help to explain any observed gender gap in the demand for performance feedback.

In addition, for both of our studies we are interested in testing whether a gender gap in the demand for feedback is expected. To do so, we will describe our study designs to a separate set of subjects and elicit their forecasts of men's and women's preferences for feedback. We hypothesize that these forecasters anticipate that women are more feedback averse than men.

Hypothesis 2: *Participants forecast that women demand performance feedback less than men.*

Finally, our last hypothesis is specific to Study 2. Given recent evidence that women may anticipate being discriminated against in evaluations of their capabilities and take action to prevent such discrimination (e.g., Exley et al., 2022; Lepage, Li, and Zafar, 2022; Alston, 2023), we are interested in studying whether women disproportionately shy away from performance feedback when feedback is provided by an individual who can discriminate on the basis of gender.

Hypothesis 3: *The possibility of gender discrimination in the performance evaluation leads women to decrease their demand for feedback more so than men.*

3 Demand for Feedback on a Cognitive Test

3.1 Experimental Design: Feedback Study

In the Feedback Study, participants take a cognitive test, report their beliefs of how they performed, and make decisions about whether they would like to receive accurate, objective feedback on how they performed.

3.1.1 Cognitive Test and Treatment Variation

Participants have 5 minutes to answer 30 questions on arithmetic reasoning, assembling objects, math, general science, and mechanical comprehension, drawn from the Armed Services Vocational Aptitude Battery (ASVAB).⁷ We intentionally chose questions from stereotypically male-typed domains. This increases the extent to which our results are likely to speak to real world settings of interest where gender gaps are largest, such as in STEM fields. Each question is multiple choice. Participants receive \$0.10 per correct answer, and \$0 for skipped or incorrect answers. As we detail later in this section, we add a stochastic component to the final payment so that participants cannot infer their performance from their earnings.

We randomize participants into either an easy or a hard version of the test in a between-subjects design. We exogenously manipulate the difficulty of the test in Study 1 to examine whether beliefs of own performance causally impact the demand for feedback. This rationale builds on previous findings that beliefs about ability depend upon the difficulty of the task;

⁷ We inform participants that this is a test of cognitive skills, but do not mention the term ASVAB.

specifically, individuals' confidence drops as the difficulty of the task increases (Moore and Healy, 2008; Bordalo et al., 2019).⁸ Assuming that randomly assigned difficulty level of the test does indeed impact beliefs of own performance, we can use this treatment assignment as an instrument for (over)confidence.

3.1.2 *Prior Beliefs*

After completing the cognitive test, participants report their beliefs about absolute and relative performance, and their degree of confidence in those beliefs. Participants first indicate how many questions they believe they answered correctly, receiving \$0.10 if their guess is correct and \$0 otherwise. Subjects then indicate how sure they feel about their guess, on a 1-5 scale ranging from "Not sure at all" to "Completely sure". We choose this qualitative scale in hopes that it produces less measurement error among this population than a fully incentivized probabilistic elicitation.⁹ Participants then indicate how they think they ranked relative to 9 other randomly-drawn study participants who completed the same test, receiving \$0.10 if their guess is correct and \$0 otherwise. Finally, participants indicate how sure they feel about their performance rank guess, again on a 1-5 scale.¹⁰

3.1.3 *Preferences over Feedback*

Following the belief reports, we elicit participants' preferences for receiving feedback on their performance. We do this in two parts. In the first part, we ask participants to indicate how interested they are in learning the number of questions they answered correctly and their rank

⁸ Previous work that induces variation in confidence by manipulating task difficulty include Dargnies, Hakimov, and Kübler, 2019; Barron and Gravert, 2022; and Möbius et al., 2022).

⁹ See Danz, Vesterlund, and Wilson (2022) and Healy and Kagel (2022) for recent work on how incentives may affect belief reports.

¹⁰ We see our measure of how sure participants feel about their guess as a measure of their certainty in their beliefs; essentially how tight or accurate is the prior belief, or what Moore and Healy (2008) refer to as *precision*.

relative to the other randomly-chosen 9 participants (as a single bundle of information), on a 1-5 scale ranging from "Not at all interested" to "Extremely interested".¹¹ This is an unincentivized report, since the answer to this question does not determine whether the participant receives the feedback. Our goal is to collect a simple, intuitive measure, unlikely to generate confusion, before continuing to incentivized measures.

In the second part of the elicitation, we inform participants that they have an opportunity to learn at the end of the session how many questions they answered correctly and how they ranked relative to the other randomly-chosen 9 participants, and that they will now be presented with three questions. Their answer to one randomly selected question determines whether they learn this information. For each of the three questions, participants must make a choice between two options: receiving or not receiving feedback. We vary the price associated with each option across question. We use real-effort task prices, rather than monetary prices, to avoid potential "house money" effects. In particular, we ask participants to complete sliders (Gill and Prowse, 2012; Araujo et al., 2016).¹² To familiarize participants with sliders, we required participants to complete two sliders before they advance to the three-question elicitation.

In Question 1, the choice is between receiving or not receiving the information, with no real-effort price attached to either choice. In Question 2, the choice is between receiving the information and completing 10 sliders, or not receiving the information. In Question 3, the choice is between receiving the information, or not receiving the information and completing 10 sliders. Therefore, relative to Question 1, Question 2 adds a real-effort cost to acquire the feedback, and

¹¹ We bundle together feedback on absolute and relative performance on the cognitive test to simplify the experiment design and analysis, and because we have no reason to expect that the factors that drive demand will differ dramatically for feedback about absolute and relative performance.

¹² A slider is completed by moving the indicator of a track bar to a target location in the range of integers from 0 to 100. It takes roughly 4 to 5 seconds to complete a slider.

Question 3 adds the same cost to avoid the feedback. The three questions are presented one at a time, on separate pages. Question 1 is always presented first, and the order of Questions 2 and 3 is randomized across participant.

3.1.4 Exit Questionnaire

After reporting their preferences for receiving feedback, participants provide their year of birth, gender, race, region of residence, and whether they attended high school in the US. Participants also indicate their beliefs about average gender differences in performance across all participants who completed the same test, by choosing one of the following options in an unincentivized manner: on average, (i) women answered at least 3 more questions correctly than men, (ii) women answered 1 or 2 more questions correctly than men, (iii) women and men answered correctly the same number of questions, (iv) men answered 1 or 2 more questions correctly than women, and (v) men answered at least 3 more questions correctly than women.

In this final part of the experiment, we also collect participant perceptions of how informative they expect feedback to be. We ask participants to imagine they were informed that they performed *better* than they expected and to indicate on a scale from 1-10 how much such feedback would: (i) influence their own evaluation of their performance, (ii) give them information on their cognitive ability generally, and (iii) give them information on their capabilities in other aspects of life. Participants also answer the same three questions under the assumption that they were informed that they performed *worse* than they expected. We randomize which block of three questions (better-than-expected or worse-than-expected) participants see first. While our design is intended to minimize the instrumental value of feedback, it is possible that subjects nevertheless anticipate that the feedback will be useful outside the study. We include these questions to capture participant perceptions of this form of instrumental value of the feedback.

3.1.5 Provision of Feedback

Following the exit questionnaire, participants who were selected to receive feedback based upon their choices learn their absolute and relative performance and complete any necessary sliders. We ask participants who receive feedback to type in the information they receive back to us on that same feedback screen; we inform them of this protocol at the time of their decisions. This ensures that participants make their choices over feedback knowing that they cannot avoid the feedback if they choose to receive it. Finally, all participants learn their total earnings in the study and the session concludes.

3.1.6 Implementation

We conducted the study in June 2020 on the Amazon MTurk platform. A total of 995 subjects completed the study, with 502 of them assigned to the easy version and 493 to the hard version of the test. All participants received a fixed payment of \$2.50 plus a bonus payment that was divided in two components. The first component corresponded to their performance on the cognitive test and the accuracy of their beliefs as detailed above. The second component was determined by a uniform random draw from [\$0, \$3] in increments of \$0.10. At the end of the session, we informed participants of their total earnings, but not of the breakdown of their earnings by components. We included the random earnings component to ensure that participants could not infer their absolute performance from their earnings, which would have diminished the value of receiving (or avoiding) feedback. We explained this feature to participants at the beginning of the study and again during the elicitation of preferences over feedback. The study lasted 15-20 minutes and was open only to MTurk workers 18 years of age or older, with IP addresses located in the US, with at least 100 previous HITs completed on MTurk, and with approval rating of at least 95%. Participants had to pass several comprehension and attention checks distributed throughout

the session to complete the study. We pre-registered the study before data collection (Coffman and Klinowski, 2020a).

3.2 Experimental Design: Forecast Study

After conducting the Feedback Study, we elicited beliefs about participants' demand for feedback from a separate sample of subjects. We call this the Forecast Study. We first elicit their demographic information (gender, age bracket, and region of residence), and then inform them about the Feedback Study. Subjects spend at least two minutes viewing the cognitive test, though they do not have to answer the test questions. Subjects are randomized into seeing either the easy or hard version of the test. We then describe to the subjects how we elicited the previous participants' demand for feedback on their performance on the test, and we familiarize subjects with the real-effort task prices by asking them to complete two sliders.

We elicit beliefs of the previous participants' demand for feedback. We elicit beliefs about male and female participants separately, asking in each case three questions. First, we ask subjects to guess how many out of 100 [male/female] participants chose to receive feedback when the price to receive feedback was zero. Then, to guess how many out of 100 [male/female] participants chose to receive feedback when it cost 10 sliders to receive feedback. Finally, to guess how many out of 100 [male/female] participants chose to receive feedback when it cost 10 sliders to avoid feedback. After subjects provide their guesses for these three questions for one gender, we ask the same three questions for the other gender. We randomize the order of the gender across subject, and show the set of three question always in the same order within gender. Subjects receive a bonus of \$0.25 if one of their six guesses, randomly chosen, is within 5 percentage points (pp) of the correct answer.

Finally, we directly elicit beliefs about gender differences in the previous participants' willingness to receive feedback on their performance on the cognitive test and on tasks more generally. We do this by asking two unincentivized, Likert-scale questions: (i) "*Overall, for the task participants completed in the previous study, how would you describe differences in men's and women's preferences for finding out how they performed?*", and (ii) "*Thinking more generally—not just for the task participants completed in the previous study—how would you describe gender differences in preferences for finding out how they performed in educational and professional settings, such as tasks in school and work?*" For each question, subjects must choose one of the following options: (i) men want to find out how they performed much more than women, (ii) men want to find out how they performed somewhat more than women, (iii) there is no gender difference in interest in finding out how they performed, (iv) women want to find out how they performed somewhat more than men, and (v) women want to find out how they performed much more than men.

We conducted the study in November 2022 on the Prolific platform. A total of 982 subjects completed the study, with 492 of them assigned to see the easy version and 490 to the hard version of the test. All subjects received a fixed payment of \$4 plus a bonus for guessing correctly as described above. The study lasted 10-15 minutes and was open only to Prolific participants 18 years of age or older, with IP addresses located in the US, with at least 100 completed studies on the platform, and with approval rate of at least 95%. We pre-registered the study before data collection (Coffman and Klinowski, 2022). Note that during this study, other subjects were randomized into treatments that elicited beliefs about demand for feedback on performance on the Interview, as described in Section 4.

3.3 Results: Feedback Study

3.3.1 Descriptive Statistics

The sample consists of 350 women and 645 men in the Feedback Study, and 463 women, 489 men, and 30 who identified as neither man or woman in the Forecast Study.¹³ Table A1 in Appendix A provides descriptive statistics, and Table A2 shows that demographic characteristics are largely balanced across treatments (easy and hard version of the test) within study. As expected, participants answer significantly more questions correctly on the easy version than the hard version of the test (9.4 vs. 7.2, $p < 0.001$). There are no gender differences in test scores in either treatment (Table A1).¹⁴

3.3.2 Demand for Feedback

We start by examining the participants' preferences for feedback. We present results in this section for the 860 subjects with monotonic preferences, which constitute 86.4% of the 995 subjects who completed the Cognitive Test Feedback Study. Failure of monotonicity does not vary significantly across treatment ($p = 0.869$) or gender ($p = 0.287$) and is positively correlated with finding the instructions difficult and with lower test score, suggesting that non-monotonicity is most likely related to confusion (Table A3).

Overall, we find substantial demand for feedback. Looking first at the qualitative measure asking how interested subjects were in learning about how they performed, the average is 3.78 on the 1–5 scale.¹⁵ Women report greater interest in receiving feedback: 3.73 for men and 3.87 for women ($p = 0.093$). This gender difference is concentrated within the easy version. On the easy version, men report on average 3.70 while women report on average 4.03 ($p = 0.004$). On the hard

¹³ This gender imbalance in the Feedback Study was not intentional and is unlikely to have been a product of men and women failing to complete the study at different rates, since 96% of participants who pass the initial bot and human checks and start the study go on to complete the study.

¹⁴ Throughout the paper, p -values come from two-sided t -tests except when noted otherwise.

¹⁵ There is no significant treatment difference: 3.73 on the easy version and 3.87 on the hard version ($p = 0.334$).

version, men report on average 3.75 while women report on average 3.71 ($p=0.702$). The difference-in-difference is significant in a regression ($p=0.026$ without demographic controls, $p=0.051$ with demographic controls).

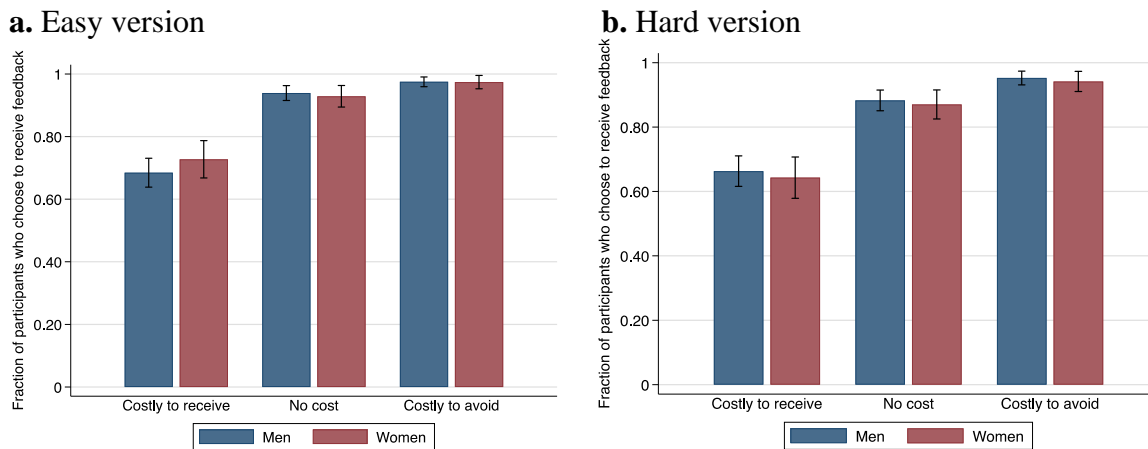


Figure 2. Demand for feedback on the cognitive test

Notes: Observations from the Cognitive Test Feedback Study. Sample restricted to subjects with monotonic preferences for feedback over all prices. Whiskers indicate 90-percent confidence intervals.

Figure 2 shows the share of men and women who choose to receive feedback for each of the three price-list items and two treatment conditions. Demand is high and there is no gender difference in this demand. More than two-thirds of our participants choose feedback even when it is costly. For both the easy and hard tests, demand for feedback declines as feedback becomes more costly to receive ($p<0.001$ for all pairwise comparisons), suggesting attentiveness and understanding among participants. More critically, there is no significant gender gap in demand for feedback at any price in either treatment.

Table 1. Demand for costly feedback on the cognitive test

	(1)	(2)	(3)	(4)	(5)
Female	0.011 (0.033)	0.027 (0.033)	0.031 (0.033)	0.041 (0.033)	0.053 (0.034)
Hard	-0.044 (0.032)	-0.048 (0.032)	-0.048 (0.032)	-0.033 (0.032)	-0.028 (0.032)
Actual rank (1:best , 10:worst)			0.014*** (0.006)	0.016*** (0.006)	0.015*** (0.006)
Belief of test score				0.005** (0.002)	
Certainty in test score				0.023 (0.014)	
Belief of rank (1:best , 10:worst)					-0.023**** (0.006)
Certainty in rank					0.029** (0.014)
Controls	N	Y	Y	Y	Y
N	860	860	860	860	860
R ²	0.0024	0.0262	0.0338	0.0448	0.0541

Notes: Observations from the Cognitive Test Feedback Study. Coefficient estimates from OLS regressions of the probability that the participant chooses to receive feedback when it is costly to receive it. Controls are age, race, region of residence, high school in the US, and the order of the three questions that elicit preferences for feedback. Sample restricted to subjects with monotonic preferences for feedback. Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

To examine in more detail the gender gap in the demand for feedback, we use OLS regressions to estimate the probability that the participant chooses to receive costly feedback on a female indicator and additional covariates as indicated in Table 1. We focus on the choice of receiving feedback when it is costly to do so because the variability is greatest for this outcome (as we note below, results are similar for the choices of receiving feedback when it is costly to avoid it and when there are no costs for either option). Table 1 columns 1-2 show that demand is not significantly different across treatment or gender, irrespective of the inclusion of demographic controls (there is no significant interaction effect between treatment and gender, Table A4 column

1).¹⁶ The result holds regardless of the order of the elicitation of preferences over feedback. The 95% confidence interval of the female coefficient in column 2 is [-0.038, 0.092]. Column 3 shows that these results are unchanged when we control for test performance. On average, performing better on the test (lower rank) is associated with being less likely to choose to receive feedback; this is seen by the positive coefficient on actual rank, which is a continuous variable from 1 to 10, where 1 is the best decile and 10 the worst decile of performance within treatment condition. In sum, across specifications 1-3 in Table 1 we find that women are no less likely than men to demand costly feedback. This result holds when there is no cost to receive or avoid feedback (Table A5) and when avoiding feedback is costly (Table A6). These robustness checks help to alleviate concerns that gender differences in the perceived or real costs of completing sliders drive our results.¹⁷

Result 1: *Contrary to Hypothesis 1, we cannot reject that men and women have the same demand for performance feedback.*

Next, we explore our sub-hypotheses, investigating the relationship between beliefs of performance and demand for costly feedback across gender.

¹⁶ Assuming a normal distribution for the standard error, to obtain 80% power for a 95% confidence interval the true effect size must be at least 2.8 standard errors away from zero (Gelman and Hill, 2006). Given our standard error for the female coefficient in column 1 of Table 1 equal to 0.033, we are powered to detect a male-female gap in demand for feedback of at least 9.24 pp. We can also use our estimates to identify the size of the male-female gap in demand for feedback that we can reject; using specification 3 from Table 1, we can reject men being 3 pp or more likely to demand feedback than women at the $p=0.05$ level.

¹⁷ We do not collect data on how long it takes participants to complete sliders. Murad et al. (2019) found that men perform better than women on the slider task under piece-rate compensation, suggesting that doing sliders might be less costly for men than for women. If this is true, we would be, if anything, under-estimating the extent to which women are more willing to pay for costly feedback than men, given their higher cost for completing the same number of sliders.

3.3.3 *The Role of Beliefs*

On average, participants overestimate their scores. Average beliefs of absolute score are 11.4 on the easy version and 9.6 on the hard version ($p < 0.001$). Figure A1 panels a-b plot beliefs about test score as a function of test score. Conditional on test performance, women's beliefs are significantly less optimistic than men's on the both the easy and hard versions of the test ($p = 0.056$ and $p = 0.023$, respectively, Table A7 columns 1-2).

In terms of beliefs of relative performance, participants on average rank themselves about in the middle of the pack, with an average rank of 5.8 on the hard test and 5.2 on the easy test, where rank of 1 is best and 10 is worst. On average, women are significantly less optimistic than men about their relative performance (Figure A1 panels c-d). Conditional on true rank, women's beliefs of rank are approximately 0.8 ranks lower than men's ($p < 0.001$, Table A8 column 3). Both the easy and the hard test produce this gender gap ($p < 0.001$ and $p = 0.005$, respectively, Table A8 columns 1-2), with no significant difference in the gap across treatments (n.s. interaction term in column 4 of Table A8).¹⁸ Overall, our evidence is consistent with sub-hypothesis 1a.

Result 1a: *Women hold significantly more pessimistic beliefs about their performance than men.*

Also consistent with past evidence, beliefs of both absolute and relative performance are positively predictive of demand for feedback. Returning to the estimation of the decision to receive costly feedback in Table 1, column 4 includes as covariates beliefs of absolute performance and certainty of these beliefs, and column 5 includes instead beliefs of relative performance and certainty of these beliefs. In both specifications, conditional on actual rank, more optimistic

¹⁸ We also ask participants how certain they are about these beliefs on a scale from 1 to 5. The average degree of certainty about absolute (relative) performance is 2.61 (2.86). Women express significantly less certainty in their absolute and relative beliefs than men, conditional on true performance and measured optimism (Tables A9-A10).

participants have significantly greater demand for feedback ($p=0.018$ for beliefs of absolute performance and $p<0.001$ for beliefs of relative performance). This is true for both men and women (Table A4 columns 3-4). Certainty of beliefs is also predictive of demand for feedback. Greater certainty is associated with greater demand, directionally so for beliefs of absolute performance ($p=0.107$, Table 1 column 4) and statistically significantly so for beliefs of relative performance ($p=0.046$, Table 1 column 5).¹⁹ A regression that drops actual rank from Table 1 column 5 does not change our conclusions for the remaining coefficients.

Finally, we examine whether self-confidence causally affects the demand for feedback by using random assignment to treatment to instrument for beliefs of relative performance. Treatment successfully manipulated confidence: beliefs of relative performance are significantly more pessimistic in the hard version conditional on true rank ($p<0.001$, Table A8 column 3). Directionally, we find that more optimistic participants are more likely to demand feedback, as the two-stage least squares (2SLS) estimate of the effect of beliefs on the demand for feedback is -0.064 ($p=0.165$, Table A12 column 2), which is more than twice as large as the OLS estimate of -0.024 ($p<0.001$, Table A12 column 1) but much less precisely estimated.²⁰ Taken as a whole, our evidence supports sub-hypothesis 1b.

¹⁹ The ego management hypothesis would suggest that it is those individuals who are most certain about performing well who are most eager to receive performance feedback. We can investigate this with our data. Looking at beliefs of absolute performance, the effect of certainty on demand is directionally larger for participants whose believed score is equal to or larger than the median belief within treatment condition: the effect is 0.014 ($p=0.504$, Table A11 column 1) for participants below the median belief and 0.031 ($p=0.141$, Table A11 column 2) for participants above the median belief; however, in a pooled regression the interaction between beliefs and certainty is not significant (Table A11 column 3). Similarly, for beliefs of relative performance, the effect of certainty is greater for participants whose believed rank is equal to or lower (i.e., better) than the median belief within treatment: the effect is 0.007 ($p=0.746$, Table A11 column 4) for participants below the median belief and 0.039 ($p=0.062$, Table A11 column 5) for participants above the median belief; however, in a pooled regression the interaction between beliefs and certainty is not significant (Table A11 column 6).

²⁰ Importantly, the 2SLS approach rests on the assumption that treatment assignment affects the demand for feedback only through its effect on self-confidence. This exclusion restriction would be violated if, for example, subjects considered the test to be systematically more, or less, informative of their abilities in the hard version of the test and this produced systematic differences in demand for feedback across treatment.

Result 1b: *Beliefs of own performance are positively related to demand for performance feedback.*

In sum, women are significantly less confident of their own performance than men, and on average more self-confident participants are more likely to demand feedback. Despite this, we do not observe a gender gap in demand for feedback.²¹

3.4 Results: Forecast Study

In this section, we examine the forecasts made by a separate set of subjects about the demand for feedback on the cognitive test. We restrict the forecaster sample to the 744 participants who have monotonic forecasts over nonzero prices (participants who forecasted that more subjects opt to receive feedback when it is free than when it is costly), which constitute 75.8% of the 982 participants in the Cognitive Test Forecast Study. Results are similar for the full sample of participants and for forecasts of the choice of receiving feedback at zero price (Tables A13-A14).²²

Table 2 estimates the forecasted probability that a participant opts for feedback when they must complete two sliders to receive it. Recall that we ask participants to make predictions about how many out of 100 men (women) would demand feedback at this price. We translate this into probability, with coefficients in hundreds of a percentage point, to allow for easier comparability with our results from the Feedback Study. Column 1 shows results from a within-subject analysis, regressing the forecasted probability of demanding feedback on an indicator that the forecast is about women and subject fixed effects. We estimate that participants forecast 58.5% of men and 55.0% of women opt for feedback, a difference of 3.5 pp ($p < 0.001$). Note that this forecasted

²¹ The results are similar when controlling for the participants' reports of how influential and generalizable they find the feedback to be.

²² Most subjects in the Forecast Study seem to have been confused about what it means to pay a cost to *avoid* feedback. In the Appendix, we discuss this issue in more detail and analyze forecasts of demand for feedback when it is costly to avoid feedback.

difference is outside of our 95% confidence interval around the estimated gender difference; comparing the analogous specification 2 from Tables 1 and 2, we estimate a forecasted gender difference of -0.044, outside of our confidence interval for the observed gender difference, [-0.038, 0.092].

Table 2. Forecasts of the demand for costly feedback on the cognitive test

	(1)	(2)	(3)
Female	-0.035 ^{****} (0.007)	-0.044 ^{**} (0.017)	-0.044 ^{**} (0.017)
Hard		-0.027 (0.017)	-0.027 (0.017)
Female forecaster			0.030 [*] (0.018)
Within subjects	Y	N	N
Across subjects	N	Y	Y
Controls	N	Y	Y
Mean	0.568	0.564	0.564
N	1,488	744	744
R ²	0.8286	0.0218	0.0257

Notes: Observations from the Cognitive Test Forecast Study. Coefficient estimates from OLS regressions of the forecasted probability of opting for feedback when feedback is costly. Controls are age bracket and region of residence. Sample restricted to the 744 subjects with monotonic forecasts over nonzero prices for feedback. Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

Column 2 shows results from an across-subjects analysis, restricting the sample to only the first forecasts provided by the participant (recall that subjects are randomized into being asked first about men or first about women) and controlling for demographics (the forecast provider's age bracket and region of residence). We continue to estimate a gender gap in this specification, with participants forecasting that women are 4.4 pp less likely to demand feedback (p=0.011). Column 3 shows that controlling for the forecaster's gender does not eliminate the forecasted gender gap. The forecasted gender gap in demand for feedback is indistinguishable across the hard and easy

versions of the cognitive test, and both men and women forecast that women have lower demand for feedback (Table A15).

Recall that participants also answered two qualitative Likert questions that elicited their beliefs of gender differences in demand for performance feedback on the cognitive test and on tasks more generally. For both questions, the median answer is that men seek feedback somewhat more than women. If we code the answers on a scale $\{-2, -1, 0, 1, 2\}$, with positive (negative) values indicating that men (women) seek relatively more feedback, and 0 indicating no gender difference, a Wilcoxon signed-rank test rejects equality of the distribution around zero ($p < 0.001$) for both questions. The mean answer is 0.39 ($p < 0.001$) for the cognitive test question and 0.28 ($p < 0.001$) for the general question.²³

Result 2: *Participants forecast that women demand performance feedback less than men, both in terms of the cognitive test study and more generally.*

Thus, women in our Feedback Study are no less eager for performance feedback than men, yet both men and women in the Forecast Study anticipate that women are more feedback averse.

In these studies feedback is objective and private. A computer scores the test, and feedback simply involves privately viewing that score and relative rank. However, in many settings of interest, feedback is likely to be substantially more subjective, and often provided by another individual. Is it possible that women become less eager to receive feedback in these more subjective settings, particularly when there is a possibility of gender discrimination in the feedback itself? We explore this in a second set of studies.

²³ Treatment, gender of the respondent, and order of the incentivized elicitation (forecast about men or women elicited first) are not significantly correlated with the answers to the qualitative Likert questions.

4 Demand for Feedback in an Interview Setting

In conducting a second study, we provide additional tests of our main hypotheses (hypotheses 1-2): do women demand performance feedback less than men, and do individuals believe that women demand performance feedback less than men? In addition, Study 2 varies across participants whether it is possible for the feedback provider—an HR professional in this case—to discriminate on the basis of gender.

4.1 Experimental Design: Feedback Study

The Feedback Study consists of two sessions that occur three weeks apart. In Session 1, participants answer common job interview questions about their life achievements and personality. We then hire HR professionals to rate the answers given by a random subset of participants. We manipulate whether the HR professionals observe the gender of the participants when evaluating their answers, which allows us to test whether the possibility of gender-based discrimination by the HR professional impacts preferences for feedback. After obtaining the ratings from the HR professionals, we invite participants back and conduct Session 2. In Session 2, participants are informed of the HR rating step that occurred between sessions and make decisions about whether they would like to receive feedback on how they ranked relative to other participants based on the ratings assigned by the HR professionals.

4.1.1 Session 1

Session 1 starts by asking participants for three pieces of demographic information: sex, age bracket (18 to 30, 31 to 50, or 51+), and region of residence (Northeast, South, Midwest, or West). We collect this information of the participant up front, so that we can reveal it later to the HR professional. While we are interested in the sex of the participant, we collect also the age bracket and the region of residence in order not to focus attention on sex and thus minimize any

potential priming or experimenter demand effects (De Quidt, Vesterlund, and Wilson, 2019). After providing these demographics, participants proceed to the task. They answer three questions that are commonly asked in job interviews to assess candidates. They have five minutes to answer each question, by typing their answers on their computer or devices. Participants are informed that they may be disqualified from participating in the second session of the study if they answer any of the interview questions in less than 60 words, if they submit an answer in two minutes or less, or if they navigate away from the screen during the interview. We also let participants know that copy-pasting has been disabled. This helps to ensure that participants give honest answers and to establish common knowledge of this fact, making the feedback about relative performance more meaningful to participants. The three interview questions are, "What is something you have achieved that you are proud of, and why?", "Describe a difficult task you were faced with and how you addressed it", and "What are you passionate about, and why?". The questions appear one at a time on separate screens. Following the three interview questions, participants are reminded that they will be invited to a second session in three weeks.

4.1.2 Evaluation by HR Professionals

In the three-week interim between Sessions 1 and 2, we hire two HR professionals from the platform Upwork to rate the answers to the interview questions of a subset of participants in Session 1. We randomly assign participants in Session 1 to either a blind or a nonblind condition in a between-subjects design, with a different HR professional assigned to each condition. In each condition, we randomly choose 10 participants to have their answers to the interview questions evaluated by the HR professional. In the blind condition, the HR professional is shown the answers to the interview questions for each of the 10 participants, but not their demographic information. In the nonblind condition, the HR professional is shown both the answers to the interview

questions and the demographic information (sex, age bracket, and region of residence) of the 10 participants evaluated. Other than this difference, the evaluation sessions are identical across conditions.

During the evaluation, we instruct the HR professional to rate each of the 10 participants on four different traits, based on the participant's answers to the three interview questions. The four traits are intellectual curiosity, a tendency to strive for achievement, assertiveness, and tolerance to stress. The HR professional assigns each participant a score from 1 to 10 on each trait that indicates the extent to which the participant's answers demonstrate the trait.

4.1.3 Session 2

Three weeks after completing Session 1, participants return for Session 2, in which we inform them that there was some chance that their answers to the interview questions have been evaluated by a HR professional, and we give them an opportunity to receive feedback on how they ranked on the evaluation relative to other participants.

Preliminary Instructions: At the beginning of Session 2, we remind participants that in Session 1 they provided their demographic information (sex, age bracket, and region of residence) and answered three questions commonly asked in job interviews. We inform participants that a HR professional with experience evaluating job candidates rated the answers given by 10 randomly-selected participants, assigning each of them a score from 1 to 10 on four traits that are generally valued by employers: intellectual curiosity, a tendency to strive for achievement, assertiveness, and tolerance to stress (we include a brief definition of each trait). We tell participants that, for each participant rated, we have constructed an "Interview Score" by averaging the subject's score across the four traits.

In the blind condition, we (truthfully) inform participants that the HR professional saw the answers to the interview questions and no other information about the participants. In the nonblind condition, we (truthfully) inform participants that the HR professional saw both the answers to the interview questions and the sex, age bracket, and region of residence of each participant. This treatment variation allows us to study whether anticipation of potential gender-based discrimination by the HR professional leads to gender differences in the demand for feedback. To sharpen the treatment, we mention several times throughout this preliminary information stage what information was seen by the HR professional. We also include understanding questions, one of which requires participants to indicate correctly what information was seen by the HR professional before they can advance to the next stage of the session.

Prior Beliefs: After participants receive the preliminary information, we elicit their beliefs about their rank on the Interview Score relative to the other 9 participants randomly selected to have their interview answers rated by the HR professional. Participants report their believed rank and, if indeed they were randomly-selected to be ranked by the HR professional, they receive \$10 if their guess is correct and \$0 otherwise. We then elicit the precision of this belief, by asking participants to indicate how sure they feel about their guess on a 1-5 scale.

Preferences over Feedback: We elicit participants' preferences for receiving feedback on their relative performance on the interview in two parts.²⁴ In the first part, we ask participants in an unincentivized fashion to indicate how interested they are in learning their rank on the Interview Score, on a 1-5 scale ranging from "Not at all interested" to "Extremely interested." In the second part, we measure participants' willingness to pay a real-effort cost (completing sliders) to receive

²⁴ We give subjects the opportunity to receive feedback only on their relative performance, not their absolute performance, since we felt that the Interview Score by itself was unlikely to convey much information on performance.

and avoid feedback, as in study 1. However, this time we include a more granular and wider-ranging real-effort price list.

After familiarizing participants with the slider task, we present participants with the price list as a sequence of 11 questions that appear one at a time on separate pages. We randomly select one of these 11 questions to determine their outcome, conditional on having indeed been ranked by the HR professional. The first question is always a choice between (A) being told the rank on the Interview Score or (B) not being told the rank on the Interview Score. This question is followed by a block of 5 questions that involve a choice between (A) being told the rank on the Interview Score and completing X sliders, or (B) not being told the rank on the Interview Score. This block is followed in turn by a second block of 5 questions that involve a choice between (A) being told the rank on the Interview Score, or (B) not being told the rank on the Interview Score and completing X sliders. Within a block of 5 questions, X always increases from 2, 5, 10, 50, to 100. (We estimate that it would take participants 7-8 minutes on average to complete 100 sliders, which corresponds to about 35-40% of the average session-2 completion time.) The order of the two 5-question blocks is randomized.

Exit Questionnaire: Following the preference elicitation, we instruct participants to imagine they received information that they ranked *better* than they expected, and ask them to indicate on a 1-10 scale how much such feedback would (i) influence their own evaluation of their abilities in terms of the traits of intellectual curiosity, a tendency to strive for achievement, assertiveness, and tolerance to stress, (ii) lead them to change their beliefs about their ability to perform well on a job interview, and (iii) lead them to change their beliefs about their capabilities in other aspects of life. Participants also answer the same three questions under the assumption that they received feedback

that they performed *worse* than they expected, and we randomize which block of three questions (better-than-expected or worse-than-expected) participants see first.

We then ask participants for race, educational attainment, current employment status, and whether they attended high school in the US. Following the demographics questions, participants indicate their agreement on a 1-7 scale with the statement, "In the past, I have worried whether I have been treated or evaluated unfairly because of my sex." They also indicate whether they think in the future, when trying to find or keep a job, employers will treat or evaluate them [substantially less, slightly less, equally, slightly more, substantially more] favorably than others because of their sex. These two questions are designed to elicit participants' beliefs of past and future sex-based discrimination. Finally, participants indicate their beliefs about average gender differences in the Interview Score across all participants in the treatment condition, by choosing one of the following options in an unincentivized manner: on average, (i) women obtained a much better Interview Score than men, (ii) women obtained a slightly better Interview Score than men, (iii) women and men obtain equal Interview Scores, (iv) men obtained a slightly better Interview Score than women, and (v) men obtained a much better Interview Score than women.

Provision of Feedback: At the end of the session, participants learn whether they had been randomly selected to have their answers to the interview questions evaluated, and, if so, they receive information on their rank on the Interview Score depending on their answer to the elicitation question that was implemented. They also complete the corresponding number of sliders, if applicable. As in the cognitive test study, we require participants who receive feedback to type in the feedback they receive back to us on the same feedback screen, and we inform participants of this feature before they respond to the elicitation mechanism.

4.1.4 Implementation

For Sessions 1 and 2, we advertised the study on Amazon MTurk as an academic study involving two sessions, 3 weeks apart from each other, and with a guaranteed payment of \$1 for completing Session 1 and an additional \$5 for completing Session 2. The larger fee for Session 2 was intended to discourage attrition between sessions. The study was open to MTurk workers 18 years of age or older, with IP addresses in the US, with at least 500 previous HITs completed on MTurk, and with approval rating of at least 95%.

We conducted Session 1 in November 2020. Session 1 lasted approximately 20 minutes and included several understanding and attention checks. Our pre-registered goal was to collect 1,500 observations that would be eligible for Session 2. To meet this goal, we invited a total of 2,451 participants to Session 1, yielding 1,515 participants who gave answers to the three interview questions that complied with the rules we established (at least 60 words per question, submitted in no less than 2 minutes, and typed without navigating away from the study page). Of these, we randomly assigned 759 to the blind condition and 756 to the nonblind condition. In each condition, we randomly selected 10 participants to have their answers evaluated by a HR professional. Note that Session 1 is identical across treatment; participants do not receive treatment-specific instructions until Session 2.

We collected the HR professional evaluations following Session 1. We used Upwork to recruit HR professionals. We advertised the rating task as a one-time, one-hour job involving evaluating answers to a mock job interview given by 10 participants of an academic study, for a fee of \$50. We opened the job to HR professionals with at least one year of experience in evaluating candidates in job interviews and we selected two. They completed the evaluation session by providing their ratings via a Qualtrics survey.

We opened Session 2 only to the 1,515 subjects who participated in Session 1 and gave valid answers to the interview questions. We sent these individuals an invitation to participate and reminders of Session 2 a few days prior to Session 2. A total of 1,350 subjects participated in Session 2. We included several attention and understanding checks throughout the session. Session 2 lasted approximately 20 minutes. We pre-registered the Interview Preferences for Feedback Study (see Coffman and Klinowski, 2020b).

4.2 *Experimental Design: Forecast Study*

After conducting the Feedback Study, we elicited forecasts about the participants' demand for feedback from a separate sample of subjects. We call this the Forecast Study. From these new subjects, we first elicit their demographic information (gender, age bracket, and region of residence), and then inform them about the Feedback Study in its entirety and familiarize subjects with the real-effort task prices by asking them to complete two sliders before they can proceed.

We then elicit the subjects' forecasts of the previous participants' demand for feedback. We elicit forecasts about male and female participants separately, asking in each case three blocks of questions. The first block of questions consists of only one question, in which we ask subjects to guess how many out of 100 [male/female] participants chose to receive feedback when the price to receive feedback was zero. Then, the second block of questions consists of five questions, in which we ask subjects to guess how many out of 100 [male/female] participants chose to receive feedback when it cost 2, 5, 10, 50, and 100 sliders to *receive* feedback. Finally, the third block of questions consists of five questions, in which we ask subjects to guess how many out of 100 [male/female] participants chose to receive feedback when it cost 2, 5, 10, 50, and 100 sliders to *avoid* feedback. After subjects provide their guesses in these three blocks of questions for one gender, we ask the same three blocks of questions for the other gender. These blocks always appear

in the same order (no price, positive price for receiving feedback, positive price for avoiding feedback). We randomize the order of which gender the participant provides guesses for first across subject. Subjects receive a bonus of \$0.25 if one randomly-selected guess (of the 22 total guesses, 11 for each gender) is within 5 pp of the correct answer.

Finally, we elicit forecasts about gender differences in demand for feedback using two unincentivized qualitative questions, one asking for beliefs about the gender gap in demand for performance feedback on the interview task specifically and the other asking for beliefs about the gender gap in demand for performance feedback more broadly.

This study was run as a branch of the same experiment used to elicit forecasts about demand for performance feedback on the cognitive test (see Section 3), conducted in November 2022 on the Prolific platform. A total of 970 subjects completed the interview version of the study, with 466 of them assigned to the HR-blind treatment and 504 to the HR-nonblind treatment.

4.3 Results: Feedback Study

4.3.1 Descriptive Statistics

Table B1 shows descriptive statistics of the sample for the Feedback Study, which consists of the 1350 subjects who completed Session 2 and thus provided their full set of demographic information and their demand for feedback. These subjects constitute 89.1% of the 1515 subjects who completed Session 1 and were invited to participate in Session 2. Observable demographics are balanced across treatment conditions (Table B2). Attrition from Session 1 to Session 2 was directionally smaller in the blind condition than the nonblind condition (10% vs. 12%, chi-squared test $p=0.153$), and directionally smaller for women than for men (10% vs. 12%, chi-squared test $p=0.203$). Since all participants underwent identical procedures up to the start of Session 2, any differential attrition across treatment is likely due to chance. Moreover, since subjects were not

informed of the opportunity to receive feedback in Session 1, attrition is unlikely to be directly related to preferences for feedback.

Unlike the cognitive test, the interview task does not produce an obvious, objective measure of performance. We constructed an objective measure of performance using the IBM Watson Personality Insights AI, a commercial artificial intelligence (AI) program that generates a personality profile from text. Using the participant's answers to the three interview questions, the AI outputs a personality profile consisting of a score from 0 to 1 on each Big Five personality trait and each facet component of each trait, where a higher score indicates that the participant exhibits the trait or facet to a larger extent. We extract the participant's scores on the facets of intellectual curiosity, assertiveness, striving for achievement, and tolerance to stress. We take the average of these four scores as the participant's objective score on the interview. We use this objective score as a control for performance in many of the analyses below, but note that participants are not made aware of this measure of performance, nor is it offered as feedback.²⁵

4.3.2 Demand for Feedback

We present results in this section for 1242 subjects with monotonic preferences for feedback over all prices, which constitute 92.0% of the 1350 participants who completed Session 2 of the Feedback Study. Failure of monotonicity is significantly correlated with being male, finding the instructions difficult, having lower AI score, and reporting lower interest in receiving feedback in the unincentivized elicitation (Table B3).

²⁵ We have some evidence that objective scores capture performance on the interview. For the 20 participants whose answers were rated by HR professionals, the correlation between objective scores and ratings assigned by the HR professional is 0.4 in the blind condition and 0.13 in the nonblind condition. Looking at the four facets that make up the objective scores, men score higher than women on intellectual curiosity and tolerance to stress, while women score higher on striving for achievement. There are no average gender differences in assertiveness, which is contrary to the stereotypical view that men are more assertive (Fiske, Cuddy, and Glick, 2007), but is consistent with Coffman, Flikkema, and Shurchkov (2021), who used text analysis of free-form conversation and found no gender differences in assertiveness as perceived by gender-blind coders, though the authors found a significant gender gap favoring men when coders were aware of conversant gender.

As in the cognitive test study, we see substantial demand for feedback. The average response to the qualitative measure of interest in receiving feedback on the 1-5 scale is 4.12. There is no significant treatment difference in this response: 4.08 in the nonblind condition and 4.16 in the blind condition ($p=0.184$). As in the cognitive test study, women report more interest in receiving feedback: 4.05 for men and 4.17 for women ($p=0.058$). There is no significant treatment-gender interaction ($p=0.615$ from an OLS regression with no controls).

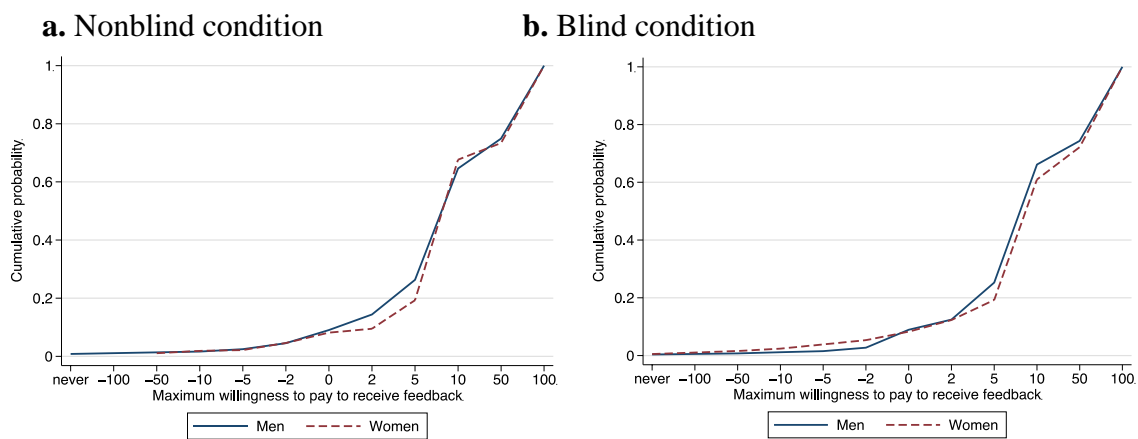


Figure 3. Maximum willingness to pay for feedback on the interview

Notes: Observations from the Interview Feedback Study. Sample restricted to 1242 subjects with monotonic preferences for feedback over all prices, which constitute 92.0% of the subjects who completed Session 2.

Figure 3 shows the cumulative distribution of the maximum a subject is willing to pay to receive feedback.²⁶ The majority of participants (91.5%) are willing to pay a strictly positive amount. The median and modal willingness to pay is 10 sliders. Strikingly, 26.5% of participants chose to receive feedback irrespective of the price. More central to our investigation, Figure 3 shows no large differences in the distributions across gender.

²⁶ A negative value for the maximum willingness to pay, $-X$, indicates that the subject is willing to complete up to X sliders to avoid receiving feedback.

Table 3. Maximum willingness to pay for feedback on the interview

	(1)	(2)	(3)	(4)	(5)
Female	3.523 (3.173)	0.713 (3.179)	-0.577 (4.478)	0.377 (3.159)	1.312 (3.146)
Blind	2.527 (3.118)	2.859 (3.074)	1.359 (4.837)	2.973 (3.065)	2.492 (3.043)
Female x Blind			2.544 (6.314)		
Average AI z-score				4.296* (2.365)	4.571* (2.349)
Word count				0.037** (0.016)	0.043*** (0.016)
Belief of rank (1:best , 10:worst)					-1.481 (0.931)
Certainty in rank					5.791**** (1.663)
Controls	N	Y	Y	Y	Y
N	1,309	1,309	1,309	1,309	1,309

Notes: Observations from the Interview Feedback Study. Coefficient estimates from interval regressions of the maximum willingness to pay to receive feedback when receiving feedback is costly (i.e., when the price to receive feedback is 2, 5, 10, 50, or 100 sliders). Controls are age, race, region of residence, educational attainment, high school in the US, currently looking for a job, and the order of the block of five questions that elicit willingness to pay to receive or to avoid feedback. Sample restricted to 1309 subjects with monotonic preferences for feedback over positive prices, which constitute 97.0% of the subjects who completed Session 2. Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

As in the cognitive test, we test for a gender gap in the demand for feedback using regressions and focusing on the choice of receiving feedback when it is costly to do so—that is, prices of 2, 5, 10, 50, and 100 sliders to receive feedback—since variability is greatest over positive prices. We use interval regressions to estimate the participant's maximum willingness to pay for feedback over positive prices on a female indicator and additional covariates as indicated in Table 3, restricting the sample to 1309 subjects with monotonic preferences for feedback over positive prices, representing 97.0% of the 1350 subjects who completed Session 2. Consistent with the cognitive test study, we find no significant evidence of gender differences in demand for feedback, and, if anything, directionally greater demand by women in most specifications. There are no

gender differences in the maximum willingness to pay regardless of inclusion of demographics controls (Table 3 columns 1-2), and regardless of the order of the elicitation of preferences over feedback.²⁷ We estimate that women are willing to pay 0.71 sliders more than men for feedback, with a 95% confidence interval of [-5.52, 6.94]. Put differently, we can reject that men are willing to pay 5.5 sliders or more for feedback than women at the $p=0.05$ threshold.

The total number of words written across the three interview questions is predictive of greater willingness to pay ($p=0.017$, Table 3 column 4), perhaps reflecting an association between effort in answering the questions and greater demand for feedback. Participants with higher objective scores as determined by the AI have greater willingness to pay ($p=0.069$, Table 3 column 4).²⁸ Conditioning on these measures continues to produce a null gender gap.

Result 1 (replication): *Contrary to Hypothesis 1, we cannot reject that men and women have the same demand for performance feedback.*

We do not find that the possibility of sex-based discrimination in the performance evaluation impacts participants' demand for feedback. The coefficient on *Blind* in Columns 1–2 in Table 3 reveals that, on average, there was no significant difference in demand for feedback across the two treatments. The insignificant interaction effect in Column 3 reveals that women were no more responsive to the treatment variation than men.²⁹

²⁷ Given our standard error for the female coefficient in column 1 of Table 3 equal to 3.173, we are 80% powered to detect a male-female gap in maximum willingness to pay for feedback of at least 8.88 sliders.

²⁸ This is contrary to the cognitive test study, for which we found that performance on the test is negatively correlated with demand for feedback. Comparing performance measures across studies is difficult, though, because the correlation between AI score and demand for feedback on the interview may reflect an effect of performance but also personality (intellectual curiosity, a tendency to strive for achievement, assertiveness, and tolerance to stress) on the demand for feedback.

²⁹ Is this null effect because the treatment was not received? We have suggestive evidence that the treatment was received: Looking at beliefs about average gender differences in the evaluations by HR professionals (a 1-5 variable where 5 is that HR professionals would rank men much better), in the blind condition the mean belief is 3.01 for male participants vs. 3.08 for female participants ($p=0.226$). But, when participants know that the HR evaluations will be non-blind, a gender gap emerges: the mean belief is 2.96 for men vs. 3.21 for women ($p<0.001$). In a regression, the difference-in-difference is significant ($p=0.058$ without controls, $p=0.091$ with controls). Thus, across subject, we

Result 3: *The possibility of sex-based discrimination in the performance evaluation does not significantly change demand for performance feedback among women nor men.*

The observed relationship between beliefs and demand is directionally consistent with the findings in the cognitive test study. Women's beliefs are on average more pessimistic than men's conditional on AI score rank (Table B5). The average degree of certainty is 2.54 on the 1–5 scale, with no significant gender or treatment difference (Table B6). Participants who are more optimistic of their relative performance are willing to pay more for feedback, although the relation is not significant ($p=0.112$, Table 3 column 5). And, as in the cognitive test study, greater certainty of beliefs is significantly associated with greater demand ($p<0.001$, Table 3 column 5).³⁰ Conditioning on beliefs, we continue to estimate a null gender difference in the demand for feedback (Table 3, Column 5).

Result 1a (replication): *Women hold significantly more pessimistic beliefs about their performance than men.*

Result 1b (replication): *Beliefs of own performance are (directionally) positively related to demand for performance feedback.*

In sum, we do not find that women demand feedback less than men.³¹ There are also no gender differences in the demand for feedback at zero price (Table B4). Furthermore, the possibility of sex-based discrimination does not significantly change demand for feedback among women nor men.

have some suggestive evidence that women seem to anticipate a modest amount of sex-based discrimination in evaluation.

³⁰ As in the cognitive test study, the positive relation between certainty of beliefs and the demand for feedback is concentrated among participants who are optimistic about their performance. The relation is much stronger and significant only for participants whose believed rank is equal to or lower (i.e., better) than the median belief within treatment (Table B7). This is another indication that individuals demand feedback in expectation of consuming good news.

³¹ These results are similar when controlling for the participants' reports of how influential and generalizable they find the feedback to be.

4.4 Results: Forecast Study

We now turn attention to the forecasts made by a separate set of subjects about the demand for feedback on the interview. To mirror the analysis in the previous section, we use the forecasts about how many of 100 men (women) demand feedback at different prices to construct implied forecasts of maximum willingness to pay. We do this using only observations from the 841 participants who provided monotonic forecasts over positive prices, which constitute 86.7% of the sample who completed the Interview Forecast Study. Table 4 presents the results from interval regressions that predict the implied forecasted maximum willingness to pay. Results are similar for the choice of receiving feedback at zero price in the full sample of participants (Table B8).

Table 4 column 1 performs a within-subject analysis, regressing the maximum willingness to pay on an indicator that the forecast is about women and subject fixed effects. We estimate that subjects forecast women are willing to pay 0.73 fewer sliders than men ($p < 0.01$). Column 2 performs an across-subjects analysis, using only the first set of forecasts the participant provided (recall that subjects are randomized into being asked first about men or first about women). In this specification we estimate a larger forecasted gender gap: subjects forecast women are willing to pay 4.16 fewer sliders than men ($p = 0.018$). Thus, just as in our cognitive test study, respondents forecast that women's demand for feedback will be less than men's, while our observed gap points in the other direction. But, in this case, the forecasted female-male difference of -4.16 does fall within our 95% confidence interval around the estimated female-male gap, [-5.52, 6.94]. Column 3 shows that controlling for the forecaster's gender does not eliminate the forecasted gender gap. The forecasted gender gap in demand for feedback is indistinguishable across the blind and nonblind conditions, and both men and women forecast that women have lower demand for feedback (Table B9).

Table 4. Forecasts of the maximum willingness to pay for feedback on the interview

	(1)	(2)	(3)
Female	-0.728*** (0.224)	-4.158** (1.764)	-4.151** (1.777)
Blind		2.007 (1.998)	1.949 (1.993)
Female forecaster			0.986 (1.945)
Within subjects	Y	N	N
Across subjects	N	Y	Y
Controls	N	Y	Y
Mean	28.677	25.353	25.353
N	168,200	84,100	84,100

Notes: Observations from the Interview Forecast Study. Coefficient estimates from interval regressions of the maximum willingness to pay to receive feedback when receiving feedback is costly (i.e., when the price to receive feedback is 2, 5, 10, 50, or 100 sliders). Controls are age bracket and region of residence. Sample restricted to the 841 subjects with monotonic forecasts about preferences for feedback over positive prices. Each subject generates a set of 100 observations corresponding to the subject's forecasts of the feedback choice of 100 males, and a set of 100 observations corresponding to the subject's forecasts of the feedback choice of 100 females. Column 1 uses the full set of 200 observations per subject. Columns 2-3 use 100 observations per subject, corresponding to the forecast about the gender first elicited from the subject. Robust standard errors in parentheses in column 1, and clustered at the subject level in columns 2-3. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

Finally, for our unincentivized, qualitative questions that elicited the forecasters' beliefs of gender differences in demand for performance feedback on the interview task and on tasks more generally, we find that forecasters believe women are less likely to demand performance feedback than men on the interview task specifically, as the median answer is that men seek feedback more than men, a Wilcoxon signed-rank test rejects equality of the distribution around zero ($p=0.057$), and the mean answer is 0.098 ($p=0.034$). However, for tasks more generally, forecasters anticipate no gender differences in demand for feedback; a Wilcoxon signed-rank test fails to reject equality of the distribution around zero ($p=0.221$), and the mean answer is 0.034 (t-test $p=0.435$). This is in contrast to what we found for the identical generalized question from forecasters assigned to the cognitive skills test version of the study. It is possible that the “broader tasks” that come to mind

for participants differ depending upon whether they have been thinking about the cognitive test versus the interview questions, but we cannot know for sure what drives this difference.

5 Discussion

Across two studies, we elicit individuals' demand for feedback on their performance on a task. We vary whether the task is a cognitive test or answering interview questions, whether the feedback is objective or subjective, and whether it is possible for the provider of the feedback to discriminate based on the participant's gender. Across these variations, we find that women demand performance feedback no less than men. And, when we elicit forecasts from other individuals, we observe that they anticipate women being less eager to receive performance feedback than men in our setting. These findings are summarized in Table 5, which presents an analysis pooling observations from our two studies. Column 1 shows the estimated probability that a participant chooses to receive feedback when receiving feedback costs 10 sliders, and column 2 shows the other individuals' forecasted probability of choosing to receive feedback when receiving feedback costs 10 sliders. Women are significantly more likely than men to choose feedback by 4.71 pp ($p=0.015$, Table 5 column 1), but they are forecasted to be significantly less likely than men to choose feedback by 4.27 pp ($p<0.001$, Table 5 column 2). The forecasted gender difference of -4.27 pp lies outside the 95% confidence interval for the observed gender difference of [0.91 pp, 8.57 pp]. In fact, we can reject any difference in the direction of men having greater demand for feedback than women.³²

³² In the pooled analysis in Table 5, we are 80% powered to detect an observed gender difference of 5.4 pp and a forecasted gender difference of 3.4 pp, which are 7% of the mean observed uptake of feedback when it costs 10 sliders to receive it (73.7%) and 6% of the mean forecasted uptake (54.6%), respectively.

Table 5. Observed and forecasted probability of choosing to receive feedback when it costs 10 sliders to receive feedback, both studies pooled

	Observed (1)	Forecasted (2)
Female	0.0471** (0.0194)	-0.0427**** (0.0121)
Cognitive Test Study	-0.0873**** (0.0204)	0.0359*** (0.0121)
N	2169	1585
R ²	0.0146	0.0127

Notes: Column (1) pools observations from the Cognitive Test Feedback Study and the Interview Feedback Study. Coefficient estimates from an OLS regression of the probability that the participant chooses to receive feedback when it costs 10 sliders to receive it, regressed on a female indicator and an indicator for the Cognitive Test Study. Sample restricted to the 860 subjects with monotonic preferences for feedback in the Cognitive Test Study and the 1309 subjects with monotonic preferences for feedback over positive prices in the Interview Study. Column (2) pools observations from the Cognitive Test Forecast Study and the Interview Forecast Study. Coefficient estimates from an OLS regression of the forecasted probability of choosing to receive feedback when it costs 10 sliders to receive it, regressed on an indicator that the forecast is about women's choices and an indicator for the Cognitive Test Study. Sample restricted to the 744 forecasters in the Cognitive Test Study with monotonic forecasts over nonzero prices for feedback and the 841 forecasters in the Interview Study with monotonic forecasts about preferences for feedback over positive prices, and always restricted to the forecast about the gender first elicited from the subject. Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

Consistent with past work, we find in our two studies that participants' expectations of receiving good news are positively associated with their demand for feedback and that women are less confident in their performance than men, results that have been hypothesized in the literature to imply that women will shy away from performance feedback more so than men. In this way, our results may be somewhat of a surprise. In our setting, more confident people demand more feedback, and women are less confident on average. Despite this, women demand no less feedback than men on average. Together, this suggests that there are likely factors beyond confidence that predict demand for feedback, and that, because of these other factors, controlling for confidence, women may even have a greater demand for feedback than men. Future work should investigate

this further, unpacking the factors beyond confidence that predict demand for largely non-instrumental performance feedback.³³

Investigations of this type may also help to further interpret some of the mixed findings in this literature. There are many dimensions that have been varied to some degree across studies, such as whether the feedback has instrumental or strategic value, whether feedback can be avoided entirely, whether the feedback is noisy, and what the costs of feedback are. It is worth noting that even given these substantial variations, most studies have found minimal evidence of gender differences in demand for feedback (Eil and Rao 2011, Mobius et al. 2022, and Castagnetti and Schmacker 2022). The one study that does point to sizable gender gaps in demand for feedback is Sharma and Castagnetti (2023), where participants choose the informativeness of noisy feedback. Work that explores the relationship between demand for deterministic information and demand for informativeness could enhance our understanding of the behavioral forces that shape information acquisition decisions and shed further light on these across-study differences.

Future work should also consider additional settings, such as opportunities for face-to-face performance feedback, to understand whether there are indeed contextual factors that might generate a gender gap. This would also help us to better understand what factors predict demand for feedback, and how these factors interact with gender. We consider only a few factors here, including overconfidence, difficulty of the task, and the possibility of bias in the feedback. Promising avenues for future research are investigating more social factors, including whether the

³³ More generally, our finding that overconfidence correlates with demand for feedback raises interesting questions for future work. When overconfident individuals receive feedback, they will be likely to receive (objectively) disappointing news on average. Viewed through the lens of long-run belief maintenance, this creates an apparent tension between remaining overconfident and continuing to demand more feedback. Zimmerman (2020) and others have begun to explore how individuals can maintain positive self-images in the face of negative feedback, through motivated reasoning and biased memory. Future work should continue to investigate this tension, linking it to demand for feedback over time.

feedback is provided publicly or privately and the relationship between the person providing the feedback and the recipient. One could also consider whether gender differences in demand emerge when feedback must be sought more proactively. In our setting, individuals respond to an offer to receive or avoid feedback, yet in many educational and labor-market scenarios the possibility of receiving feedback may not be as explicit. In these more ambiguous settings, might gender gaps in demand for feedback emerge?³⁴

Individual motivations for receiving feedback also deserve more attention. While our paper and others engage with ego-management and curiosity, understanding how a desire to learn (or the need to signal a desire to learn) impacts demand for feedback would be valuable. Finally, considering the type of feedback would also be worthwhile: how do preferences vary depending upon whether the feedback is about relative versus absolute talents, or whether the feedback is primarily evaluative as opposed to constructive.

We show that individuals expect there to be a gender gap in demand for feedback in our setting. With this we take a first step into connecting our findings to the question of whether gender affects the supply of feedback. A person's willingness to give feedback may depend on whether he thinks the other wishes to receive it; thus, inaccurate beliefs about preferences for feedback may prevent individuals from receiving the information they need to develop their skills and advance their careers.³⁵ Our study does not investigate the supply of feedback, including whether supply depends on beliefs about the preferences of the feedback recipient. Future work should explore whether and how the inaccurate beliefs we document impact the provision of feedback.

³⁴ This could be seen as analogous to results from the negotiation literature which suggest smaller gender differences when it is clear that negotiation is a possibility compared to situations with greater ambiguity about whether it is appropriate to negotiate (Bowles, Babcock, and McGinn, 2005).

³⁵ Closest to this question, Gallen and Wasserman (2022) show that a student's gender affects the information they receive about careers, with individuals providing more information about work-life balance to women. And, Dupas et al. (2021) find that female presenters receive more feedback during seminars. Understanding what beliefs – if any – underlie these differences is an important open question.

Importantly, our data do not allow us to speak to the question of whether men or women *should* demand more feedback: does more feedback make them better off? The answer to this question likely depends on a range of important factors, including how individuals update their beliefs about themselves in response to the feedback they receive and how valuable accurate beliefs about performance are in the context of interest. Our main contribution is instead to present evidence from multiple settings rejecting the hypothesis that women demand performance feedback less than men. We think this is a useful observation for researchers and practitioners looking to understand sources of gender gaps in self-confidence and career advancement. Future work should expand upon our studies to investigate the welfare implications of our findings, bringing together insights from the literatures on overconfidence and belief updating to paint a fuller picture of how the supply of feedback, demand for feedback, and responses to feedback together shape gender gaps.

REFERENCES

- Ahlbrecht, M., and Weber, M., 1997. Preference for gradual resolution of uncertainty. *Theory and Decision*, 43: 167-185.
- Alston, M., 2023. The (perceived) cost of being female: An experimental investigation of strategic responses to discrimination. Working paper.
- Ambuehl, S., and Li, S., 2018. Belief updating and the demand for information. *Games and Economic Behavior*, 109: 21-39.
- Araujo, F.A., Carbone, E., Conell-Price, L., Dunietz, M.W., Jaroszewicz, A., Landsman, R., Lame, D., Vesterlund, L., Wang, S.W., and Wilson, A.J., 2016. The slider task: An example of restricted inference on incentive effects. *Journal of the Economic Science Association*, 2: 1-12.
- Ashford, S.J., Blatt, R., and VandeWally, D., 2003. Reflections on the looking glass: A review of research on feedback-seeking behavior in organizations. *Journal of Management*, 29.6: 773-799.
- Barron, K., and Gravert, C., 2022. Confidence and career choices: An experiment. *Scandinavian Journal of Economics*, 124.1: 35-68.
- Benabou, R., and Tirole, J., 2016. Mindful economics: The production, consumption, and value of beliefs. *Journal of Economic Perspectives*, 30.3: 141-164.
- Bordalo, P., Coffman, K., Gennaioli, N., and Shleifer, A., 2019. Beliefs about gender. *American Economic Review*, 109.3: 739-773.
- Bowles, H.R., Babcock, L., and McGinn, K.L., 2005. Constraints and triggers: Situational mechanics of gender in negotiation. *Journal of Personality and Social Psychology*, 89.6: 951-965.
- Brown, A.L., and Kim, H., 2014. Do individuals have preferences used in macro-finance models? An experimental investigation. *Management Science*, 60.4: 805-1081.
- Brown, R., H. Mansour, S. D. O'Connell, and J. Reeves. 2019. Gender differences in political career progression: Evidence from u.s. elections. Discussion Paper 12569, IZA
- Burks, S.V., Carpenter, J.P., Goette, L., and Rustichini, A., 2013. Overconfidence and social signalling. *Review of Economic Studies*, 80.3: 949-983.
- Buser, T., Niederle, M., and Oosterbeek, H., 2014. Gender, competitiveness, and career choices. *Quarterly Journal of Economics*, 129.3: 1409-1447.

- Buser, T., Niederle, M., and Oosterbeek, H., 2020. Can competitiveness predict education and labor market outcomes? Evidence from incentivized choice and survey measures. Working paper.
- Buser, T. and H. Yuan. 2019. Do Women Give Up Competing More Easily? Evidence from the Lab and the Dutch Math Olympiad. *American Economic Journal: Applied Economics* 11 (3), 225–52.
- Caplin, A., and Eliasz, K., 2003. AIDS policy and psychology: A mechanism-design approach. *RAND Journal of Economics*, 34.4: 631-646.
- Caplin, A., and Leahy, J., 2001. Psychological expected utility theory and anticipatory feelings. *Quarterly Journal of Economics*, 116.1: 55-79.
- Castagnetti, A., and Schmacker, R., 2022. Protecting the ego: Motivated information selection and updating. *European Economic Review*, 142: 104007.
- Chad, I., Osun, E.B., and Ozbay, E.Y., 2023. Effect of feedback on beliefs about self-ability. Working paper.
- Chakraborty, P., and Serra, D., 2022. Gender and leadership in organizations: The threat of backlash. Working paper.
- Chen, S., and Heese, C., 2021. Fishing for good news: Motivated information acquisition. Working paper.
- Chew, S.H., and Ho, J.L., 1994. Hope: An empirical study of attitude toward the timing of uncertainty resolution. *Journal of Risk and Uncertainty*, 8: 267-288.
- Coffman, K., 2014. Evidence on self-stereotyping and the contribution of ideas. *Quarterly Journal of Economics*, 129.4: 1625-1660.
- Coffman, K., Collis, M., and Kulkarni, L. 2023. Stereotypes and belief updating. Accepted at *Journal of the European Economic Association*.
- Coffman, K., and Klinowski, D., 2020a. A study of feedback aversion. *AEA RCT Registry*, June 11. <https://doi.org/10.1257/rct.5990-1.0>
- Coffman, K., and Klinowski, D., 2020b. A study of feedback aversion on a job interview task. *AEA RCT Registry*. November 13. <https://doi.org/10.1257/rct.6735-1.0>
- Coffman, K., and Klinowski, D., 2022. Beliefs about receiving feedback. *AEA RCT Registry*. November 18. <https://doi.org/10.1257/rct.10413-1.0>
- Coffman, K., Ugalde Araya, P., and Zafar, B., 2023. A (Dynamic) Investigation of Stereotypes, Belief-Updating, and Behavior. *Working Paper*.

- Cortés, P., Pan, J., Reuben, E., Pilossoph, L., and Zafar, B., 2022. Gender differences in job search and the earnings gap: Evidence from the field and lab. *NBER Working Paper 28820*.
- Coutts, A. 2019. Good news and bad news are still news: Experimental evidence on belief updating. *Experimental Economics* 22 (2), 369 – 395.
- Dana, J., Weber, R.A., and Kuang, J.X., 2007. Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory*, 33: 67-80.
- Danz, D., Vesterlund, L., and Wilson A.J., 2022. Belief elicitation and behavioral incentive compatibility. *American Economic Review*, 112.9: 2851-2883.
- Dargnies, M.-P., Hakimov, R., and Kübler, D., 2019. Self-confidence and unraveling in matching markets. *Management Science*, 65.12: 5603-5618.
- De Quidt, J., Vesterlund, L., and Wilson, A.J., 2019. Experimenter demand effects. In *The Handbook of Research Methods and Applications in Experimental Economics*, eds. Schram, A., and Ule, A., Edward Elgar Publishing.
- Dillenberger, D., 2010. Preferences for one-shot resolution of uncertainty and Allais-type behavior. *Econometrica*, 78.6: 1973-2004.
- Drobner, C., 2022. Motivated beliefs and anticipation of uncertainty resolution. *American Economic Review: Insights*, 4.1: 89-105.
- Dupas, P., Modestino, A. S., Niederle, M., & Wolfers, J. 2021. Gender and the dynamics of economics seminars. NBER Working Paper w28494.
- Dweck, C.S., 2000. *Self-theories: Their role in motivation, personality, and development*. Psychology Press.
- Eil, D., and Rao, J.M., 2011. The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, 3.2: 114-138.
- Eliaz, K., and Schotter, A., 2010. Paying for confidence: An experimental study of the demand for non-instrumental information. *Games and Economic Behavior*, 70.2: 304-324.
- Eliaz, K., and Spiegel, R., 2006. Can anticipatory feelings explain anomalous choices of information sources? *Games and Economic Behavior*, 56: 87-104.
- Ellison, G. and A. Swanson. 2018. Dynamics of the gender gap in high math achievement. *NBER Working Paper 24910*.
- Ely, J., Frankel, A., and Kamenica, E., 2015. Suspense and surprise. *Journal of Political Economy*, 123.1: 215-260.

- Exley, C.L., Fisman, R., Low, C., Shan, X., Toma, M., 2022. The transparency gap.
- Exley, C.L., and Kessler, J.B., 2022. The gender gap in self-promotion. *Quarterly Journal of Economics*, 137.3: 1345-1381.
- Exley, C.L., and Nielsen, K., 2023. The gender gap in confidence: Expected but not accounted for. Working paper.
- Falk, A., and Zimmermann, F., 2022. Attention and dread: Experimental evidence on preferences for information. Working paper.
- Fang, C., E. Zhang, and J. Zhang. 2021. Do women give up competing more easily? Evidence from speedcubers. *Economics Letters* 205.
- Gallen, Y., and Wasserman, M., 2022. Informed choices: Gender gaps in career advice. Working paper.
- Ganguly, A., and Tasoff, J., 2017. Fantasy and dread: The demand for information and the consumption utility of the future. *Management Science*, 63.12: 4037-4060.
- Gelman, A., and Hill, J., 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.
- Gill, D., and Prowse, V., 2012. A structural analysis of disappointment aversion in a real effort competition. *American Economic Review*, 102.1: 469-503.
- Gill, D. and V. Prowse 2014. Gender differences and dynamics in competition: The role of luck. *Quantitative Economics* 5 (2), 351–376.
- Golman, R., Hagmann, D., and Loewenstein, G., 2017. Information avoidance. *Journal of Economic Literature*, 55.1: 96-135.
- Golman, R., and Loewenstein, G., 2018. Information gaps: A theory of preferences regarding the presence and absence of information. *Decision*, 5.3: 143-164.
- Golman, R., and Loewenstein, G., Molnar, A., and Saccardo, S., 2022. The demand for, and avoidance of, information. *Management Science*, 68.9: 6454-6476.
- Grant, S., Kajii, A., and Polak, B., 1998. Intrinsic preference for information. *Journal of Economic Theory*, 83.2: 233-259.
- Grosse, N.D., and Reiner, G., 2010. Explaining gender gaps in competitiveness: Gender-task stereotypes. Working paper.

- Grossman, Z., 2014. Strategic ignorance and the robustness of social preferences. *Management Science*, 60.11: 2659-2665.
- Grossman, Z., and van der Weele, J.J., 2017. Self-image and willful ignorance in social decisions. *Journal of the European Economic Association*, 15.1: 173-217.
- Healy, P.J., and Kagel J., 2022. Testing elicitation mechanisms via team chat. *Working paper*.
- Kang, L., Z. Lei, Y. Song, and P. Zhang. 2021. Gender differences in reactions to failure in high-stakes competition: evidence from the national college entrance exam retakes. *Working paper, SSRN*.
- Karlsson, N., Loewenstein, G., and Seppi, D., 2009. The ostrich effect: Selective attention to information. *Journal of Risk and Uncertainty*, 38.2: 95-115.
- Klinowski, D., 2019. Selection into self-improvement and competition pay: Gender, stereotypes, and earnings volatility. *Journal of Economic Behavior & Organization*, 158: 128-146.
- Klinowski, D., and Paulsen, J.S., 2013. Impaired awareness in Huntington disease: Medical evidence in relation to the optimal expectations model. *Working paper*.
- Kocher, M.G., Krawczyk, M., and van Winden, F., 2014. 'Let me dream on!' Anticipatory emotions and preference for timing in lotteries. *Journal of Economic Behavior & Organization*, 98: 29-40.
- Köszegi, B., 2006. Ego utility, overconfidence, and task choice. *Journal of the European Economic Association*, 4.4: 673-707.
- Kreps, D.M., and Porteus, E.L., 1978. Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, 185-200.
- Lepage, L.-P., Li, X., and Zafar, B., 2022. Anticipated gender discrimination and grade disclosure. *NBER Working Paper 30765*.
- Lovalló, D., and Kahneman, D., 2000. Living with uncertainty: Attractiveness and resolution timing. *Journal of Behavioral Decision Making*, 13.2: 179-190.
- Masatlioglu, Y., Orhun, A.Y., and Raymond, C. 2021. Intrinsic information preferences and skewness. *Working paper*.
- Mobius, M.M., Niederle, M., Niehaus, P., and Rosenblat, T.S., 2022. Managing self-confidence: Theory and experimental evidence. *Management Science*, 68.11: 7793-7817.
- Molnar, A., and Golman, R., 2023. Impatience for information: Curiosity is here today, gone tomorrow. *Working paper*.

- Molnar, A., and Loewenstein, G., 2022. Thoughts and players: An introduction to old and new economic perspectives on beliefs. In *The Cognitive Science of Belief: A Multidisciplinary Approach*, eds. Musolino, J., Sommer, J., and Hemmer, P., Cambridge University Press.
- Moore, D.A., and Healy, P.J., 2008. The trouble with overconfidence. *Psychological Review*, 115.2: 502-517.
- Murad, Z., Stavropoulou, C., and Cookson, G. 2019. Incentives and gender in a multi-task setting: An experimental study with real-effort tasks. *Plos one*, 14.3: e0213080.
- Murciano-Goroff, R., 2022. Missing women in tech: The labor market for highly skilled software engineers. *Management Science*, 68.5: 3262-3281.
- Niederle, M., and Vesterlund, L., 2007. Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*, 122.3: 1067-1101.
- Nielsen, K., 2020. Preferences for the resolution of uncertainty and the timing of information. *Journal of Economic Theory*, 189: 105090.
- Oster, E., Shoulson, I., and Dorsey, E.R., 2013. Optimal expectations and limited medical testing: Evidence from Huntington disease. *American Economic Review*, 103.2: 804-830.
- Pereda, P., Montoya Diaz, M. D., Rocha, F., Matsunaga, L., Borges, B. P., Mena-Chalco, J., ... & Brenck, C. 2023. Are women less persistent? Evidence from submissions to a nationwide meeting of economics. *Applied Economics*, 55(16), 1757-1768.
- Reuben, E., Wiswall, M., and Zafar, B., 2017. Preferences and biases in educational choices and labor market expectations: Shrinking the black box of gender. *Economic Journal*, 127.604: 2153-2186.
- Roberts, T.-A., and Nolen-Hoeksema, S., 1989. Sex differences in reactions to evaluative feedback. *Sex Roles*, 21: 11-12.
- Schünemann, J., Strulik, H., and Trimborn, T., 2023. Anticipation of deteriorating health and information avoidance. *Journal of Health Economics*, 89: 102755.
- Schweizer, N., and Szech, N., 2018. Optimal revelation of life-changing information. *Management Science*, 64.11: 5250-5262.
- Serra-Garcia, M., and Szech, N., 2022. The (in)elasticity of moral ignorance. *Management Science*, 68.7: 4815-4834.
- Sharma, K., and Castagnetti, A., 2023. Demand for information by gender: An experimental study. *Journal of Economic Behavior and Organization*, 207: 172-202.

- Shastry, G.K., Shurchkov, O., 2022. Reject or revise: Gender differences in persistence and publishing in economics. Working paper.
- Shastry, G.K., Shurchkov, O., and Xia L.L., 2020. Luck or skill: How women and men react to noisy feedback. *Journal of Behavioral and Experimental Economics*, 88: 101592.
- Shurchkov, O., 2012. Under pressure: Gender differences in output quality and quantity under competition and time constraints. *Journal of the European Economic Association*, 10.5: 1189-1213.
- Sicherman, N., Loewenstein, G., Seppi, D.J., and Utkus, S.P., 2016. Financial attention. *Review of Financial Studies*, 29.4: 863-897.
- Stone, D., and Heen, S., 2014. *Thanks for the Feedback: The Science and Art of Receiving Feedback Well*. Penguin Group, New York.
- Thornton, R.L., 2008. The demand for, and impact of, learning HIV status. *American Economic Review*, 98.5: 1829-1863.
- von Gaudecker, H.-M., van Soest, A., and Wengstrom, E., 2011. Heterogeneity in risky choice behavior in a broad population. *American Economic Review*, 101.2: 664-694.
- Wasserman, M. 2023. Gender differences in politician persistence. *The Review of Economics and Statistics*, 105(2), 275-291.2021
- Weinberg, B.A., 2009. A model of overconfidence. *Pacific Economic Review*, 14.4: 502-515.
- Zimmermann, F., 2015. Clumped or piecewise? Evidence on preferences for information. *Management Science*, 61.4: 707-929.
- Zimmermann, F., 2020. The dynamics of motivated beliefs. *American Economic Review*, 110.2: 337-363.

Appendix to

Gender and preferences for performance feedback

(For online publication)

Table of Contents

Appendix A. Supplementary Results, Cognitive Test	1
Appendix B. Supplementary Results, Interview	17
Appendix C. Forecasts of Demand for Feedback over Negative Prices.	26
Appendix D. Experimental Instructions, Cognitive Test Feedback Study.	27
Appendix E. Experimental Instructions, Interview Feedback Study.	55
Appendix F. Experimental Instructions, Forecast Study (Cognitive Test and Interview) . . .	103

Appendix A. Supplementary Results, Cognitive Test

Table A1 Descriptive statistics

	Pooled	By gender		
		Men	Women	p-value diff
Panel A: Feedback Study				
N	995	645	350	-
Age*	37.932	37.149	39.374	0.002
Race White	0.660	0.612	0.749	0.000
Race Black	0.243	0.290	0.157	0.000
Race Other/Mix	0.096	0.098	0.094	0.863
Region Northeast	0.267	0.281	0.243	0.199
Region South	0.346	0.338	0.360	0.486
Region Midwest	0.225	0.220	0.234	0.610
Region West	0.162	0.161	0.163	0.947
High school in US	0.970	0.972	0.966	0.574
Treatment Easy	0.505	0.512	0.491	0.543
Test score (Easy)*	9.386	9.155	9.831	0.161
Test score (Hard)*	7.247	7.305	7.146	0.629
Panel B: Forecast Study				
N	982	489	463	-
Age 18-30	0.263	0.270	0.231	0.167
Age 31-50	0.498	0.536	0.467	0.033
Age 51+	0.237	0.190	0.302	0.000
Region Northeast	0.209	0.223	0.203	0.454
Region South	0.375	0.342	0.417	0.017
Region Midwest	0.246	0.268	0.216	0.062
Region West	0.165	0.160	0.162	0.917

Notes: Observations from the Cognitive Test studies. In the Forecast Study, the *By gender* columns exclude observations from individuals who identify with a gender other than male or female. P-values from t-tests for starred variables, and Pearson Chi-squared tests otherwise.

Table A2 Randomization balance check across treatments

	Easy	Hard	p-value diff
Panel A: Feedback Study			
Female	0.343	0.361	0.543
Age*	37.863	38.002	0.837
Race White	0.667	0.653	0.637
Race Black	0.237	0.249	0.647
Race Other/Mix	0.096	0.097	0.926
Region Northeast	0.269	0.266	0.909
Region South	0.349	0.343	0.847
Region Midwest	0.221	0.229	0.760
Region West	0.161	0.162	0.969
High school in US	0.978	0.961	0.125
Panel B: Forecast Study			
Female	0.476	0.497	0.520
Age 18-30	0.285	0.241	0.119
Age 31-50	0.461	0.535	0.022
Age 51+	0.254	0.220	0.215
Region Northeast	0.224	0.194	0.252
Region South	0.339	0.410	0.022
Region Midwest	0.244	0.249	0.854
Region West	0.185	0.145	0.091

Notes: Observations from the Cognitive Test studies. P-values from t-tests for starred variables, and Pearson Chi-squared tests otherwise.

Table A3 Probability that the demand for feedback fails monotonicity

	(1)
Hard	0.020 (0.022)
Female	0.007 (0.023)
Difficulty with instructions	0.035**** (0.007)
Test score	-0.008*** (0.003)
Passed attention check	-0.028 (0.032)
Unincentivized interest	-0.004 (0.009)
Controls	Y
N	995
R ²	0.0686

Notes: Observations from the Cognitive Test Feedback Study. Coefficient estimates from an OLS regression of an indicator that the participant's demand for feedback is nonmonotonic over all prices. *Difficulty with instructions* is the answer to the question "From 1-7, how difficult did you find the instructions of the study?" (larger value indicates greater difficulty). Controls are age, race, region of residence, and high school in the US. Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01, **** p<0.001.

Table A4 Demand for costly feedback, differential effects by gender

	(1)	(2)	(3)	(4)
Female	0.070 (0.046)	0.141** (0.070)	0.016 (0.056)	0.089 (0.079)
Hard	-0.017 (0.040)	-0.046 (0.032)	-0.032 (0.032)	-0.028 (0.032)
Female x Hard	-0.087 (0.066)			
Actual rank (1:best , 10:worst)		0.022*** (0.007)	0.016*** (0.006)	0.015*** (0.006)
Female x Actual rank		-0.021* (0.011)		
Belief of test score			0.005* (0.003)	
Female x Belief of test score			0.003 (0.004)	
Certainty in test score			0.023 (0.014)	
Belief of rank (1:best , 10:worst)				-0.021*** (0.008)
Female x Belief of rank				-0.006 (0.013)
Certainty in rank				0.029** (0.014)
Controls	Y	Y	Y	Y
N	860	860	860	860
R ²	0.0281	0.0375	0.0451	0.0543

Notes: Observations from the Cognitive Test Feedback Study. Coefficient estimates from OLS regressions of the probability that the participant chooses to receive feedback when it is costly to receive it. Controls are age, race, region of residence, high school in the US, and the order of the three questions that elicit preferences for feedback. Sample restricted to subjects with monotonic preferences for feedback. Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

Table A5 Demand for feedback when there is no cost to receive or avoid feedback

	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.012 (0.021)	0.000 (0.021)	0.001 (0.025)	0.001 (0.021)	0.006 (0.021)	0.017 (0.021)
Hard	-0.057*** (0.020)	-0.056*** (0.020)	-0.056** (0.024)	-0.056*** (0.020)	-0.048** (0.020)	-0.042** (0.020)
Female x Hard			-0.001 (0.042)			
Actual rank (1:best , 10:worst)				0.003 (0.004)	0.004 (0.004)	0.004 (0.004)
Belief of test score					0.003*** (0.001)	
Certainty in test score					0.011 (0.009)	
Belief of rank (1:best , 10:worst)						-0.019**** (0.004)
Certainty in rank						0.004 (0.010)
Controls	N	Y	Y	Y	Y	Y
N	860	860	860	860	860	860
R ²	0.0100	0.0264	0.0264	0.0271	0.0359	0.0530

Notes: Observations from the Cognitive Test Feedback Study. Coefficient estimates from OLS regressions of the probability that the participant chooses to receive feedback when there is no cost to receive or avoid feedback. Controls are age, race, region of residence, high school in the US, and the order of the three questions that elicit preferences for feedback. Sample restricted to subjects with monotonic preferences for feedback. Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

Table A6 Demand for feedback when it is costly to avoid feedback

	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.006 (0.014)	0.001 (0.014)	0.005 (0.016)	0.000 (0.014)	0.003 (0.014)	0.007 (0.015)
Hard	-0.026** (0.013)	-0.024* (0.013)	-0.021 (0.015)	-0.024* (0.013)	-0.020 (0.013)	-0.018 (0.014)
Female x Hard			-0.008 (0.028)			
Actual rank (1:best , 10:worst)				-0.003 (0.002)	-0.003 (0.002)	-0.003 (0.002)
Belief of test score					0.001** (0.001)	
Certainty in test score					0.005 (0.006)	
Belief of rank (1:best , 10:worst)						-0.007** (0.003)
Certainty in rank						0.005 (0.005)
Controls	N	Y	Y	Y	Y	Y
N	860	860	860	860	860	860
R ²	0.0048	0.0338	0.0339	0.0362	0.0403	0.0447

Notes: Observations from the Cognitive Test Feedback Study. Coefficient estimates from OLS regressions of the probability that the participant chooses to receive feedback when it is costly to avoid feedback. Controls are age, race, region of residence, high school in the US, and the order of the three questions that elicit preferences for feedback. Sample restricted to subjects with monotonic preferences for feedback. Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

Table A7 Beliefs of absolute performance

	Easy (1)	Hard (2)	Both (3)	Both (4)
Female	-1.106* (0.578)	-1.490** (0.656)	-1.265*** (0.439)	-0.897 (0.585)
Hard			-1.013** (0.460)	-0.761 (0.572)
Female x Hard				-0.728 (0.880)
Test score	0.474**** (0.061)	0.272**** (0.081)	0.417**** (0.049)	0.416**** (0.049)
Controls	Y	Y	Y	Y
N	502	493	995	995
R ²	0.1566	0.1392	0.1523	0.1528

Notes: Observations from the Cognitive Test Feedback Study. Coefficient estimates from OLS regressions of the participant's belief of his or her own test score. Sample restricted to the easy version of the test in column (1) and the hard version of the test in column (2). Controls are age, race, region of residence, and high school in the US. Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01, **** p<0.001.

Table A8 Beliefs of relative performance (Rank 1 is best, 10 is worst)

	Easy (1)	Hard (2)	Both (3)	Both (4)
Female	0.931**** (0.227)	0.665*** (0.234)	0.780**** (0.162)	0.851**** (0.224)
Hard			0.607**** (0.153)	0.656**** (0.190)
Female x Hard				-0.140 (0.323)
Actual rank (1:best , 10:worst)	0.105*** (0.039)	0.053 (0.040)	0.083*** (0.028)	0.084*** (0.028)
Controls	Y	Y	Y	Y
N	502	493	995	995
R ²	0.0809	0.0998	0.0971	0.0973

Notes: Observations from the Cognitive Test Feedback Study. Coefficient estimates from OLS regressions of the participant's belief of his or her own rank in performance. Sample restricted to the easy version of the test in column (1) and the hard version of the test in column (2). Controls are age, race, region of residence, and high school in the US. Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

Table A9 Certainty in beliefs of absolute performance

	Easy (1)	Hard (2)	Both (3)	Both (4)
Female	-0.283*** (0.106)	-0.005 (0.109)	-0.137* (0.075)	-0.249** (0.103)
Hard			-0.219*** (0.072)	-0.295*** (0.087)
Female x Hard				0.222 (0.150)
Test score	-0.073*** (0.010)	-0.063*** (0.016)	-0.074*** (0.008)	-0.073*** (0.008)
Belief of test score	0.025*** (0.008)	0.051*** (0.008)	0.038*** (0.006)	0.039*** (0.006)
Controls	Y	Y	Y	Y
N	502	493	995	995
R ²	0.2660	0.2346	0.2352	0.2370

Notes: Observations from the Cognitive Test Feedback Study. Coefficient estimates from OLS regressions of the participant's expressed certainty in his or her belief of absolute performance. Sample restricted to the easy version of the test in column (1) and the hard version of the test in column (2). Controls are age, race, region of residence, and high school in the US. Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

Table A10 Certainty in beliefs of relative performance

	Easy (1)	Hard (2)	Both (3)	Both (4)
Female	-0.122 (0.112)	-0.114 (0.107)	-0.114 (0.077)	-0.109 (0.109)
Hard			-0.064 (0.071)	-0.060 (0.085)
Female x Hard				-0.010 (0.153)
Actual rank (1:best , 10:worst)	0.097**** (0.017)	0.061*** (0.019)	0.079**** (0.012)	0.079**** (0.012)
Belief of rank (1:best , 10:worst)	-0.035 (0.022)	-0.039* (0.023)	-0.037** (0.016)	-0.037** (0.016)
Controls	Y	Y	Y	Y
N	502	493	995	995
R ²	0.1627	0.1278	0.1382	0.1382

Notes: Observations from the Cognitive Test Feedback Study. Coefficient estimates from OLS regressions of the participant's expressed certainty in his or her belief of relative performance. Sample restricted to the easy version of the test in column (1) and the hard version of the test in column (2). Controls are age, race, region of residence, and high school in the US. Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

Table A11 Demand for feedback when it is costly to receive it

	Beliefs of absolute performance			Beliefs of relative performance		
	Less optimistic beliefs (1)	More optimistic beliefs (2)	All (3)	Less optimistic beliefs (4)	More optimistic beliefs (5)	All (6)
Female	0.053 (0.050)	0.026 (0.046)	0.041 (0.033)	0.025 (0.050)	0.072 (0.046)	0.054 (0.034)
Hard	-0.059 (0.055)	-0.031 (0.042)	-0.048 (0.032)	-0.067 (0.052)	-0.008 (0.042)	-0.027 (0.032)
Actual rank (1:best , 10:worst)	0.019* (0.010)	0.020** (0.008)	0.016*** (0.006)	0.030*** (0.009)	0.002 (0.007)	0.015*** (0.006)
Belief of test score	0.006 (0.014)	0.000 (0.003)	0.007 (0.006)			
Certainty in test score	0.014 (0.021)	0.031 (0.021)	0.024* (0.015)			
Belief x Certainty test score			-0.001 (0.002)			
Belief of rank (1:best , 10:worst)				-0.003 (0.021)	-0.028* (0.014)	-0.031* (0.017)
Certainty in rank				0.007 (0.022)	0.039* (0.021)	0.015 (0.029)
Belief x Certainty in rank						0.002 (0.005)
Controls	Y	Y	Y	Y	Y	Y
N	404	456	860	379	481	860
R ²	0.0440	0.0602	0.0449	0.0535	0.0600	0.0544

Notes: Observations from the Cognitive Test Feedback Study. Coefficient estimates from OLS regressions of the probability that the participant chooses to receive feedback when it is costly to receive it. The sample in column 1 (2) is the set of participants whose believed score on the test is below (at or above) the median belief of absolute performance within treatment. The sample is column 4 (5) is the set of participants whose believed rank on the test is above (at or below) the median belief of relative performance within treatment. Controls are age, race, region of residence, high school in the US, and the order of the three questions that elicit preferences for feedback. Sample restricted to subjects with monotonic preferences for feedback. Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01, **** p<0.001.

Table A12 Demand for feedback when it is costly to receive it, instrumental variables regression

	OLS	2SLS
Belief of rank (1:best , 10:worst)	-0.024**** (0.006)	-0.064 (0.046)
Female	0.054 (0.034)	0.086* (0.049)
Actual rank (1:best , 10:worst)	0.015*** (0.006)	0.019** (0.008)
Certainty in rank	0.029** (0.014)	0.020 (0.019)
<i>First stage instrument:</i>		
Hard		0.689**** (0.163)
F-stat for IV first stage		11.56
N	860	860

Notes: Observations from the Cognitive Test Feedback Study. Coefficient estimates from OLS and 2SLS regressions of the probability that the participant chooses to receive feedback when it is costly to receive it. 2SLS regressions use treatment assignment as an instrument for beliefs of rank. Additional controls in all regressions are age, race, region of residence, high school in the US, and the order of the three questions that elicit preferences for feedback. Sample restricted to subjects with monotonic preferences for feedback. Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

Table A13 Forecasts of demand for feedback when it is costly to receive it, full sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.038*** (0.007)	-0.048*** (0.015)	-0.047*** (0.015)	-0.035* (0.020)	-0.047*** (0.015)	-0.063*** (0.021)	-0.047* (0.028)
Hard		-0.028* (0.015)	-0.030** (0.015)	-0.017 (0.022)	-0.030** (0.015)	-0.029* (0.015)	-0.022 (0.031)
Female x Hard				-0.024 (0.030)			-0.030 (0.042)
Female forecaster					0.022 (0.015)	0.005 (0.022)	-0.000 (0.030)
Female x Female forecaster						0.034 (0.030)	0.025 (0.041)
Hard x Female forecaster							0.008 (0.043)
Female x Hard x Female forec.							0.020 (0.060)
Within subjects	Y	N	N	N	N	N	N
Across subjects	N	Y	Y	Y	Y	Y	Y
Controls	N	Y	Y	Y	Y	Y	Y
N	1,964	982	982	982	982	982	982
R ²	0.8104	0.0135	0.0164	0.0170	0.0186	0.0198	0.0209

Notes: Observations from the Cognitive Test Forecast Study. Coefficient estimates from OLS regressions of the forecasted probability of opting for feedback when feedback is costly. Controls are age bracket and region of residence. Sample is all subjects in the Feedback Study. Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

Table A14 Forecasts of demand for feedback at zero price, full sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.062*** (0.006)	-0.077*** (0.011)	-0.076*** (0.011)	-0.050*** (0.015)	-0.076*** (0.011)	-0.075*** (0.016)	-0.048** (0.021)
Hard		-0.048*** (0.012)	-0.048*** (0.012)	-0.021 (0.015)	-0.049*** (0.012)	-0.049*** (0.012)	-0.021 (0.021)
Female x Hard				-0.052** (0.023)			-0.055* (0.032)
Female forecaster					0.015 (0.012)	0.016 (0.015)	0.015 (0.022)
Female x Female forecaster						-0.002 (0.023)	-0.005 (0.031)
Hard x Female forecaster							-0.003 (0.030)
Female x Hard x Female forec.							0.009 (0.046)
Within subjects	Y	N	N	N	N	N	N
Across subjects	N	Y	Y	Y	Y	Y	Y
Controls	N	Y	Y	Y	Y	Y	Y
N	1,964	982	982	982	982	982	982
R ²	0.7472	0.0579	0.0655	0.0705	0.0670	0.0670	0.0717

Notes: Observations from the Cognitive Test Forecast Study. Coefficient estimates from OLS regressions of the forecasted probability of opting for feedback when there is no cost to receive or avoid feedback. Controls are age bracket and region of residence. Sample is all subjects in the Feedback Study. Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

Table A15 Forecasts of demand for feedback when it is costly to receive it

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.035**** (0.007)	-0.047*** (0.017)	-0.044** (0.017)	-0.026 (0.023)	-0.044** (0.017)	-0.060** (0.024)	-0.035 (0.032)
Hard		-0.028 (0.017)	-0.027 (0.017)	-0.008 (0.025)	-0.027 (0.017)	-0.025 (0.018)	-0.004 (0.036)
Female x Hard				-0.037 (0.035)			-0.051 (0.048)
Female forecaster					0.030* (0.018)	0.011 (0.025)	0.014 (0.035)
Female x Female forecaster						0.036 (0.035)	0.016 (0.047)
Hard x Female forecaster							-0.009 (0.049)
Female x Hard x Female forec.							0.043 (0.069)
Within subjects	Y	N	N	N	N	N	N
Across subjects	N	Y	Y	Y	Y	Y	Y
Controls	N	N	Y	Y	Y	Y	Y
N	1,488	744	744	744	744	744	744
R ²	0.8286	0.0127	0.0218	0.0234	0.0257	0.0271	0.0290

Notes: Observations from the Cognitive Test Forecast Study. Coefficient estimates from OLS regressions of the forecasted probability of opting for feedback when feedback is costly. Controls are age bracket and region of residence. Sample restricted to subjects with monotonic forecasts of demand over nonzero prices for feedback. Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

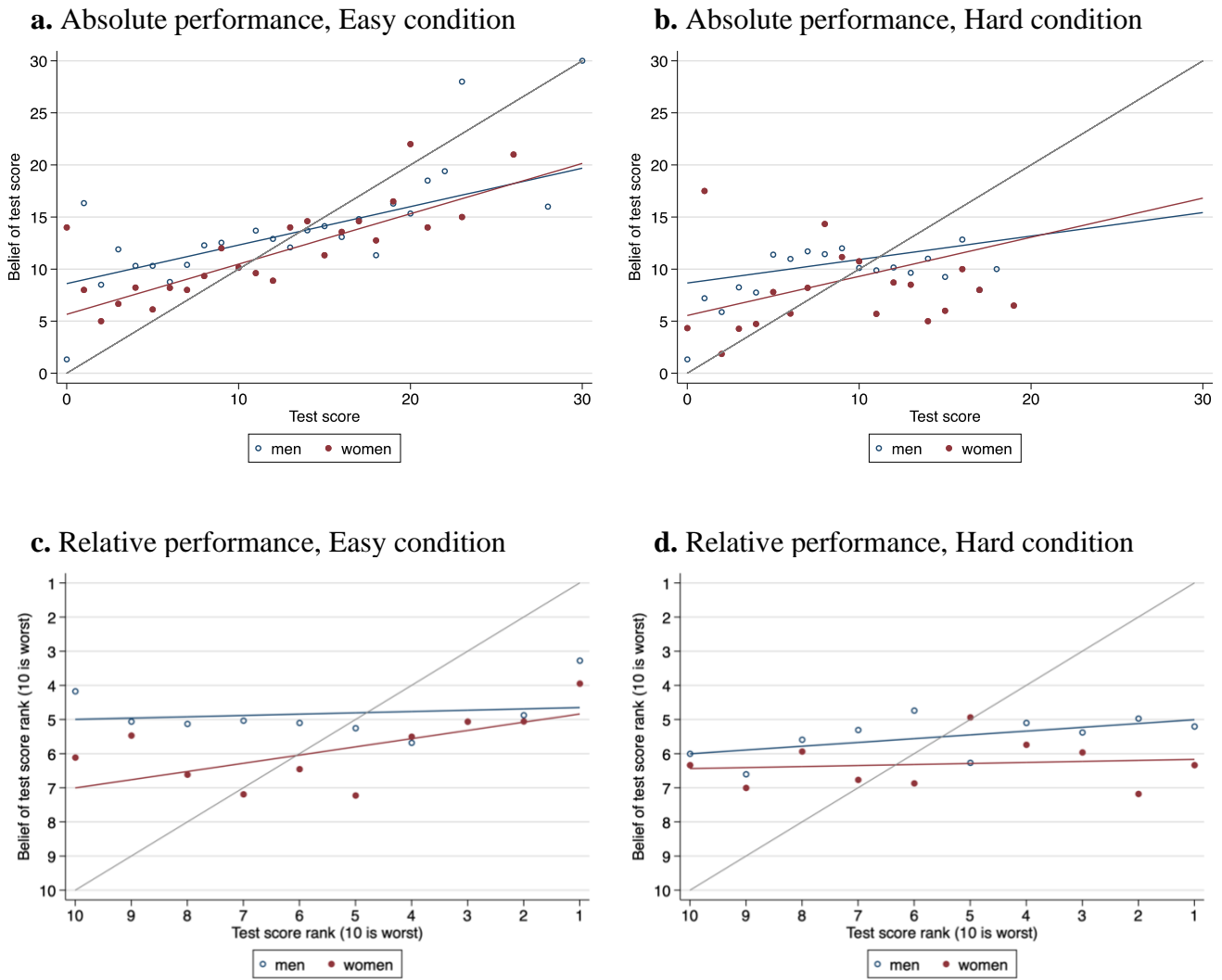


Figure A1 Beliefs of absolute and relative performance

Notes: Observations from the Cognitive Test Feedback Study. Dots are average beliefs conditional on gender and actual performance.

Appendix B. Supplementary Results, Interview

Table B1 Descriptive statistics

	Pooled	By gender		
		Men	Women	p-value diff
Panel A: Feedback Study				
N	1,350	560	790	-
Age 18-30	0.191	0.220	0.171	0.025
Age 31-50	0.529	0.562	0.505	0.037
Age 51+	0.280	0.218	0.324	0.000
Race White	0.787	0.752	0.813	0.007
Race Black	0.093	0.100	0.087	0.429
Race Other/Mix	0.120	0.148	0.100	0.007
Region Northeast	0.215	0.232	0.203	0.192
Region South	0.357	0.325	0.380	0.039
Region Midwest	0.224	0.259	0.200	0.011
Region West	0.204	0.184	0.218	0.129
High school in US	0.958	0.964	0.953	0.317
Educ No HS	0.004	0.005	0.004	0.671
Educ HS	0.070	0.054	0.081	0.051
Educ Some college	0.181	0.168	0.191	0.274
Educ Associates	0.111	0.086	0.129	0.012
Educ Bachelor	0.430	0.489	0.387	0.000
Educ Graduate	0.204	0.198	0.208	0.673
Empl Unemployed	0.097	0.070	0.116	0.004
Empl Less than FT	0.216	0.166	0.252	0.000
Empl FT	0.505	0.604	0.435	0.000
Empl Self-employed	0.158	0.138	0.172	0.085
Empl Looking	0.063	0.052	0.071	0.155
Empl Retired	0.073	0.066	0.077	0.437
Empl Unable	0.019	0.014	0.023	0.263
Treatment Blind	0.507	0.518	0.500	0.518
Panel B: Forecast Study				
N	970	494	462	-
Age 18-30	0.259	0.277	0.232	0.105
Age 31-50	0.505	0.538	0.472	0.040
Age 51+	0.235	0.184	0.297	0.000
Region Northeast	0.225	0.241	0.208	0.221
Region South	0.399	0.377	0.422	0.150
Region Midwest	0.223	0.215	0.232	0.527
Region West	0.151	0.164	0.136	0.233

Notes: Observations from the Interview studies. In the Feedback Study, observations are from Session 2 participants. In the Forecast Study, the *By gender* columns exclude observations from individuals who do not identify as male or female. P-values from Pearson Chi-squared tests.

Table B2 Balance check across treatments

	Nonblind	Blind	p-value diff
Panel A: Feedback Study			
Age 18-30	0.197	0.185	0.588
Age 31-50	0.516	0.542	0.342
Age 51+	0.287	0.273	0.561
Race White	0.783	0.791	0.727
Race Black	0.104	0.082	0.163
Race Other/Mix	0.113	0.127	0.421
Region Northeast	0.206	0.223	0.438
Region South	0.371	0.343	0.277
Region Midwest	0.217	0.232	0.493
Region West	0.206	0.201	0.835
High school in US	0.950	0.965	0.183
Educ No HS	0.005	0.004	0.971
Educ HS	0.069	0.070	0.948
Educ Some college	0.170	0.193	0.278
Educ Associates	0.119	0.104	0.376
Educ Bachelor	0.432	0.428	0.887
Educ Graduate	0.206	0.201	0.835
Empl Unemployed	0.107	0.088	0.234
Empl Less than FT	0.198	0.234	0.118
Empl FT	0.511	0.499	0.659
Empl Selfemployed	0.146	0.169	0.237
Empl Looking	0.065	0.061	0.800
Empl Retired	0.080	0.066	0.321
Empl Unable	0.018	0.020	0.749
Instruction how difficult*	1.883	1.823	0.307
Panel B: Forecast Study			
Female	0.463	0.505	0.187
Age 18-30	0.258	0.260	0.951
Age 31-50	0.510	0.500	0.758
Age 51+	0.230	0.240	0.709
Region Northeast	0.236	0.212	0.378
Region South	0.383	0.416	0.289
Region Midwest	0.234	0.210	0.373
Region West	0.143	0.159	0.488

Notes: Observations from the Interview studies. In the Feedback Study, observations are from participants of Session 2. P-values from t-tests for starred variables, and Pearson Chi-squared tests otherwise.

Table B3 Probability that the demand for feedback fails monotonicity

	(1)
Blind	0.001 (0.014)
Female	-0.036** (0.015)
Difficulty with instructions	0.024**** (0.007)
Average AI z-score	-0.029** (0.011)
Unincentivized interest	-0.022*** (0.007)
Controls	Y
N	1,350
R ²	0.0615

Notes: Observations from the Interview Feedback Study. Coefficient estimates from an OLS regression of an indicator that the participant's demand for feedback is nonmonotonic. *Difficulty with instructions* is the answer to the question "From 1-7, how difficult did you find the instructions of the study?" (larger value indicates greater difficulty). Controls are age, race, region of residence, educational attainment, high school in the US, and currently looking for a job. Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

Table B4 Demand for feedback at zero price

	(1)	(2)	(3)	(4)	(5)
Female	-0.004 (0.012)	-0.004 (0.012)	0.016 (0.019)	-0.005 (0.012)	-0.001 (0.012)
Blind	0.010 (0.012)	0.009 (0.012)	0.033* (0.018)	0.010 (0.012)	0.009 (0.012)
Female x Blind			-0.040* (0.024)		
Average AI z-score				0.019** (0.010)	0.019** (0.010)
Word count				0.102** (0.052)	0.106** (0.052)
Belief of rank (1:best , 10:worst)					-0.009** (0.004)
Certainty in rank					0.001 (0.006)
Controls	N	Y	Y	Y	Y
N	1,309	1,309	1,309	1,309	1,309
R ²	0.0007	0.0134	0.0155	0.0194	0.0257

Notes: Observations from the Interview Feedback Study. Coefficient estimates from OLS regressions of the probability that the participant chooses to receive feedback when there is no cost to receive or avoid feedback. Word count is divided by 1000. Controls are age, race, region of residence, educational attainment, high school in the US, currently looking for a job, and the order of the block of five questions that elicit willingness to pay to receive or to avoid feedback. Sample restricted to subjects with monotonic preferences for feedback over positive prices. Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01, **** p<0.001.

Table B5 Beliefs of relative performance (Rank 1 is best, 10 is worst)

	Nonblind (1)	Blind (2)	Both (3)	Both (4)
Female	0.456*** (0.150)	0.267* (0.148)	0.350**** (0.104)	0.397*** (0.147)
Blind			-0.081 (0.099)	-0.028 (0.156)
Female x Blind				-0.092 (0.204)
Average AI z-score	-0.011 (0.111)	0.057 (0.113)	0.017 (0.080)	0.019 (0.080)
Controls	Y	Y	Y	Y
N	665	685	1,350	1,350
R ²	0.0924	0.0815	0.0743	0.0745

Notes: Observations from the Interview Feedback Study. Coefficient estimates from OLS regressions of the participant's belief of their rank in the evaluation by the HR professional. Sample restricted to the nonblind condition in column (1) and the blind condition in column (2). Controls are age, race, region of residence, educational attainment, high school in the US, and currently looking for a job. Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

Table B6 Certainty in beliefs of relative performance

	Nonblind (1)	Blind (2)	Both (3)	Both (4)
Female	-0.042 (0.084)	-0.111 (0.080)	-0.075 (0.057)	-0.023 (0.082)
Blind			0.068 (0.054)	0.128 (0.088)
Female x Blind				-0.103 (0.112)
Average AI z-score	-0.025 (0.065)	-0.069 (0.060)	-0.052 (0.044)	-0.050 (0.044)
Belief of rank (1:best , 10:worst)	-0.046* (0.028)	-0.107**** (0.027)	-0.076**** (0.019)	-0.076**** (0.019)
Controls	Y	Y	Y	Y
N	665	685	1,350	1,350
R ²	0.0649	0.0945	0.0717	0.0723

Notes: Observations from the Interview Feedback Study. Coefficient estimates from OLS regressions of the participant's expressed certainty in their rank belief. Sample restricted to the nonblind condition in column (1) and the blind condition in column (2). Controls are age, race, region of residence, educational attainment, high school in the US, and currently looking for a job. Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

Table B7 Maximum willingness to pay for costly feedback on the interview

	Beliefs of relative performance		
	Less optimistic beliefs (1)	More optimistic beliefs (2)	All (3)
Female	10.246** (4.509)	-5.764 (4.415)	1.566 (3.150)
Blind	0.337 (4.336)	2.982 (4.283)	2.307 (3.038)
Average AI z-score	7.071** (3.271)	2.515 (3.434)	4.771** (2.361)
Word count	0.032 (0.022)	0.059*** (0.022)	0.043*** (0.016)
Belief of rank (1:best , 10:worst)	2.149 (1.838)	-6.001** (2.643)	2.237 (2.207)
Certainty in rank	2.514 (2.329)	6.738*** (2.368)	11.792*** (3.835)
Belief x Certainty in rank			-1.310* (0.758)
Controls	Y	Y	Y
N	642	667	1,309

Notes: Observations from the Interview Feedback Study. Coefficient estimates from interval regressions of the maximum willingness to pay to receive feedback when receiving feedback is costly (i.e., when the price to receive feedback is 2, 5, 10, 50, or 100 sliders). The sample in column 1 (2) is the set of participants whose believed rank on the test is above (at or below) the median belief of relative performance within treatment. Controls are age, race, region of residence, educational attainment, high school in the US, currently looking for a job, and the order of the block of five questions that elicit willingness to pay to receive or to avoid feedback. Sample restricted to subjects with monotonic preferences for feedback over positive prices. Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01, **** p<0.001.

Table B8 Forecast of demand for feedback at zero price, full sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.028*** (0.006)	-0.040*** (0.010)	-0.038*** (0.010)	-0.037** (0.015)	-0.038*** (0.010)	-0.043*** (0.014)	-0.057*** (0.020)
Blind		0.011 (0.010)	0.010 (0.010)	0.010 (0.014)	0.010 (0.010)	0.010 (0.010)	-0.028 (0.019)
Female x Blind				-0.001 (0.021)			0.032 (0.028)
Female forecaster					-0.015 (0.010)	-0.020 (0.014)	-0.059 (0.020)
Female x Female forecaster						0.011 (0.021)	0.045 (0.030)
Blind x Female forecaster							0.082*** (0.028)
Female x Blind x Female forec.							-0.071* (0.042)
Within subjects	Y	N	N	N	N	N	N
Across subjects	N	Y	Y	Y	Y	Y	Y
Controls	N	N	Y	Y	Y	Y	Y
N	1,940	970	970	970	970	970	970
R ²	0.7081	0.0162	0.0336	0.0336	0.0356	0.0358	0.0441

Notes: Observations from the Interview Forecast Study. Coefficient estimates from OLS regressions of the forecasted probability of opting for feedback when there is no cost to receive or avoid feedback. Controls are age bracket and region of residence. Full sample included (subjects with monotonic and nonmonotonic beliefs). Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

Table B9 Forecasts of maximum willingness to pay for feedback

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.728*** (0.224)	-4.707** (1.900)	-4.158** (1.764)	-3.884 (2.404)	-4.151** (1.777)	-6.084** (2.391)	-5.985* (3.274)
Blind		2.076 (1.896)	2.007 (1.998)	2.279 (2.724)	1.949 (1.993)	2.025 (1.991)	-0.146 (3.523)
Female x Blind				-0.563 (3.782)			-0.071 (5.243)
Female forecaster					0.986 (1.945)	-0.961 (2.529)	-3.510 (3.425)
Female x Female forecaster						4.077 (3.669)	4.699 (5.059)
Blind x Female forecaster							5.212 (4.960)
Female x Blind x Female forec.							-1.246 (7.555)
Within subjects	Y	N	N	N	N	N	N
Across subjects	N	Y	Y	Y	Y	Y	Y
Controls	N	N	Y	Y	Y	Y	Y
N	168,200	84,100	84,100	84,100	84,100	84,100	84,100

Notes: Observations from the Interview Forecast Study. Coefficient estimates from interval regressions of the maximum willingness to pay to receive feedback when receiving feedback is costly (i.e., when the price to receive feedback is 2, 5, 10, 50, or 100 sliders). Controls are age bracket and region of residence. Sample restricted to subjects with monotonic beliefs about preferences for feedback over positive prices. Each subject generates a set of 100 observations corresponding to the subject's forecasts of the feedback choice of 100 males, and a set of 100 observations corresponding to the subject's forecasts of the feedback choice of 100 females. Column 1 uses the full set of 200 observations per subject. Columns 2-7 use 100 observations per subject, corresponding to the forecast about the gender first elicited from the subject. Robust standard errors in parentheses in column 1, and clustered at the subject level in columns 2-7. *p<0.1, **p<0.05, ***p<0.01, ****p<0.001.

Appendix C. Forecasts of Demand for Feedback over Negative Prices

For the majority of subjects in the Cognitive Test Forecast Study and the Interview Forecast Study, beliefs of demand for feedback over positive prices are monotonic in the direction of the law of demand; that is, beliefs are lower for higher prices. But the opposite is true for negative prices. Beliefs of demand for feedback over negative prices are monotonic in the direction *opposite* to the law of demand; that is, beliefs are lower for more negative prices. This is illustrated in Figure C1 below.

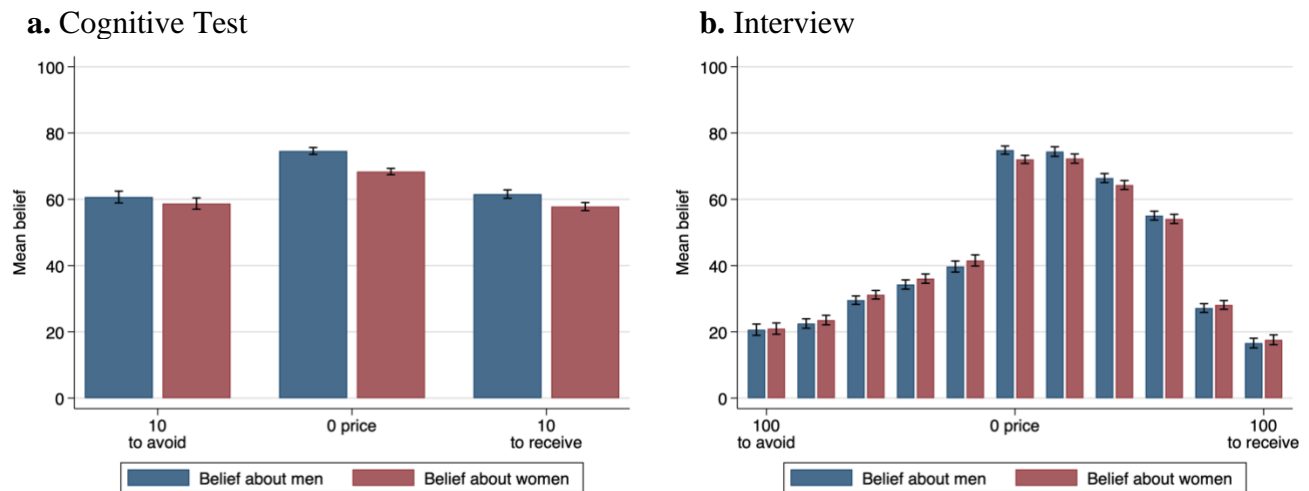


Figure C1 Average beliefs about demand for feedback

Notes: Observations from Cognitive Test Forecast Study in Panel A, and Interview Forecast Study in Panel B. Estimates from OLS regressions of the decision to opt for feedback on the price and subject fixed effects. Full sample in each study. Whiskers indicate 95-percent confidence intervals from robust standard errors.

We think this inconsistency in beliefs is due to confusion from subjects in the Forecast Studies about what it means to pay to avoid receiving feedback, since it is (at least to us) easier to understand what it means to "pay X sliders to receive feedback", than to "pay X sliders to avoid receiving feedback". Thus, we are hesitant to draw any conclusion from the data on beliefs over negative prices. However, if we take the beliefs data at face value, we estimate that subjects in the Cognitive Test Forecast Study guess 60.7% of men and 58.7% of women opt for feedback when the price is 10 sliders to avoid feedback, a difference of 1.9 percentage points ($p=0.020$ from a within-subject OLS regression). In the Interview Forecast Study, if we pool beliefs over all negative prices, we estimate that subjects guess 29.3% of men and 30.7% of women opt for feedback when the price is 2, 5, 10, 50 or 100 sliders to avoid feedback, a difference of 1.3 percentage points ($p<0.001$ from a within-subject OLS regression of the belief on an indicator for the gender which the guess is about, a control for the price, and subject fixed effects).

Appendix D. Experimental Instructions, Cognitive Test Feedback Study

Restart Survey

Place Bookmark

Mobile view off

Tools 

Before you proceed to the survey, please complete the captcha below.

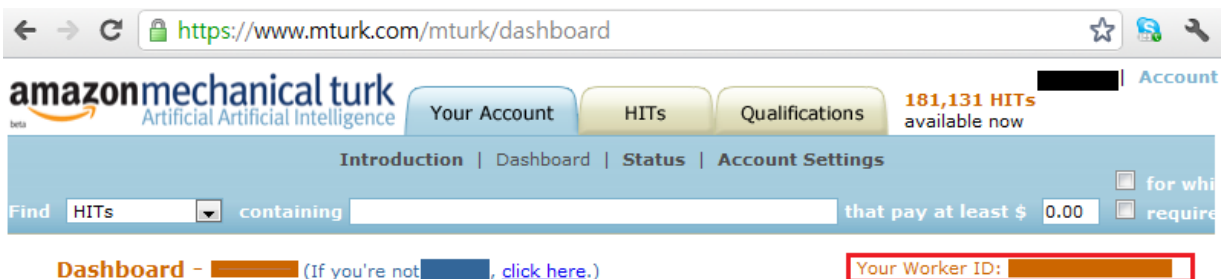
I'm not a robot reCAPTCHA
Privacy - Terms



Powered by Qualtrics 

To see if you are eligible to participate, please enter your Mechanical Turk Worker ID into the box below and then click NEXT.

Please see below for where you can find your Worker ID. Your Worker ID starts with the letter A and has 12-14 letters or numbers. It must be all CAPITAL letters and no spaces. It is NOT your email address.



→

Powered by Qualtrics 

You are eligible to participate. Please read the important consent information below.

Information Sheet

You are being asked to take part in a research study being done by Katherine B Coffman from Harvard University. The requester of this HIT is Manuela Collis, who is an approved Research Associate at Harvard University. The purpose of this research is to better understand decision-making processes. Your participation will take approximately 15 to 20 minutes to complete. As a participant, you will complete a test consisting of multiple-choice questions, including math and other topics, of varying degrees of difficulty, and answer a short follow-up questionnaire. You will receive clear instructions and will be provided with full and accurate information on how your answers and decisions will influence your payment. Also, while you will be fully informed about the version of this study that you have been randomly assigned to, you will not be informed about different versions of this study that other participants are in. None of the data you provide in this study will be linked to any information that could identify you to others, including the researchers. In the future, this non-identifiable data may be shared with other researchers or participants.

Your participation is completely voluntary. You can agree to take part and later change your mind. Your decision will not be held against you. All information that you provide will be kept entirely confidential, and only shared with members of this research team, which include Manuela Collis (Harvard University), Katherine B Coffman (Harvard University) and David Klinowski (University of Santiago, Chile). If you have questions, concerns, or complaints, or think the research has hurt you, talk to the research team at kcoffman@hbs.edu.

Compensation:

You will receive a guaranteed payment of \$2.50 upon completion and have the chance to earn additional payment.

You may print a copy of this information sheet for your own records.

If you would like to voluntarily participate in this survey and confirm that you are 18 years or older, please push the arrow to continue.

General Overview

This study consists of a test of cognitive skills and a number of related questions. We will explain them clearly as you move through the study. Note that in this study, there are several cognitive skills tests that we are using. While several MTurkers will see the same test you do, other MTurkers completing this HIT may see a different test.

→

Powered by Qualtrics [↗](#)

Your earnings

Following your participation, you also have a chance to earn a bonus payment. The amount of your bonus payment will depend on your performance on the test and your answers to the related questions. We will explain clearly how the bonus payment is determined as you move through the survey.

In addition, at the end of the survey we will randomly choose a value between \$0 and \$3.00, in increments of \$0.10. That is, we will randomly choose a value from the list:

{\$0, \$0.10, \$0.20, \$0.30, \$0.40, ... , \$2.60, \$2.70, \$2.80, \$2.90, \$3.00}

and we will add this value to your bonus. Each value is equally likely to be chosen.

Therefore, your total bonus will be an amount based on your answers to the test and related questions, plus the random value. At the end of the survey, you will be informed of this total sum amount, but will not be informed of how much of that sum comes from the payment based on your answers and how much comes from the random value.

For example, suppose your total bonus is \$1.20. At the end of the survey you would be informed that your bonus is \$1.20. But you would not be informed of how much of that amount corresponds to your answers and how much corresponds to the randomly-chosen value added to your bonus.

→

How to claim your earnings

At the end of the survey, you will see a completion code. You must enter this code into the space provided on the Amazon Mechanical Turk HIT in order to be paid. Your completion payment of \$2.50 will be transferred to you within 48 hours of completing the HIT. Any bonus payments you may have earned will be transferred to you within one week.

→

Powered by Qualtrics [↗](#)

Understanding and attention questions

This survey contains understanding and attention questions. If you answer an understanding question incorrectly, or if you fail to answer an attention question within the given time frame, you may be dismissed from the HIT and will not receive any completion payment or any additional payment. For that reason, you should read the instructions carefully and should not navigate away from the survey page during your participation.

→

Powered by Qualtrics 

Here's your first attention check.

In the space below, please spell the word "human" backwards. Please use all lowercase letters and insert no spaces between the letters.

→

Powered by Qualtrics 

Cognitive Skills Test

In the first part of this study, you will have 5 minutes to work on 30 cognitive skills questions. You will earn \$0.10 in bonus payment for every question you answer correctly. You will receive \$0 for any question you answer incorrectly or that you skip.

When you click next, the test will begin. The computer will time 5 minutes for you. All questions will appear on the same page. You may work on them in any order.

→

Powered by Qualtrics 

Timing

These page timer metrics will not be displayed to the recipient.

First Click 0 seconds
Last Click 0 seconds
Page Submit 0 seconds
Click Count 0 clicks

0447

ARITHMETIC REASONING: Aprille has \$50 to buy the items on her shopping list. Assuming there is no sales tax, about how much change will Aprille receive after buying all items on her list?

Aprille's List

ITEM	PRICE
Hammer	\$13.24
Screwdriver	\$11.99
Nails	\$4.27
Wrench	\$5.60

\$12

\$13

\$14

\$15

ARITHMETIC REASONING: At the grocery store, apples cost \$1.89 per pound and

Congratulations! You completed the test.

Now please answer the following questions about the test you just completed.

Please provide a guess of how many questions (from the total of 30) you answered correctly. If you guess correctly, you will earn an extra \$0.10 in bonus payment.

→

Powered by Qualtrics [↗](#)

How sure do you feel about your guess that you answered 3 out of the total of 30 questions correctly?

1 - Not at all sure

2 - A little sure

3 - Moderately sure

4 - Very sure

5 - Completely sure

→

Powered by Qualtrics [↗](#)

Other MTurkers have already completed this same test in a previous HIT under identical conditions to yours (same 30 questions you saw, with 5 minutes to answer, and same requirements to join the HIT). Suppose we compared you to 9 MTurkers drawn at random from those other MTurkers who completed the same test. Where would you rank in terms of your score on the test?

Please provide a guess from 1 to 10 of your rank relative to the other 9 MTurkers. If you guess correctly, you will earn an extra \$0.10 in bonus payment.

Selecting **1** means that you think you performed **better than** all 9 MTurkers.

Selecting **10** means that you think you performed **worse than** all 9 MTurkers.

Rank 1 - I performed better than all 9 MTurkers

Rank 2

Rank 3

Rank 4

Rank 5

Rank 6

Rank 7

Rank 8

Rank 9

Rank 10 - I performed worse than all 9 MTurkers

→

Powered by Qualtrics 

How sure do you feel about your guess that you ranked out of 10?

1 - Not at all sure

2 - A little sure

3 - Moderately sure

4 - Very sure

5 - Completely sure

→

Powered by Qualtrics [↗](#)

The computer has calculated the number of questions you answered correctly and your rank relative to the other randomly-chosen 9 MTurkers who already completed the same test you completed.

How interested are you in learning this information?

1 - Not at all interested

2 - Slightly interested

3 - Moderately interested

4 - Very interested

5 - Extremely interested

The NEXT button will take you to an additional, related question.

→

Powered by Qualtrics [↗](#)

As you will see, some questions in the following pages will involve the possibility of "completing sliders". Before you move to the next page, we'd like to explain now what it means to complete a slider.

To complete a slider, you must scroll the slider bar to a specific position. The specific position is a randomly-chosen value between 0 and 100, and is indicated with the slider. The slider is not completed successfully until its bar is scrolled exactly to the indicated position. If you are asked to complete sliders, you will have as much time as you would like to complete the sliders. But, you must complete all sliders successfully in order to advance to the next page.

To get a feel for what it means to complete sliders, below there are 2 sliders. Please scroll each of them to the indicated position. You will be able to advance to the next page only after you complete both sliders successfully.

Please complete the following sliders.

0 100

Scroll the slider to the number 29

Scroll the slider to the number 82



We wish to know more about your preferences for finding out how you performed on the test.

Your answers to the next three questions will determine whether you find out how you performed and may impact what you must do next in order to complete the HIT.

After you've answered all three questions, the computer will randomly select one question as the decision-that-counts. The computer will use your decision in only that question to determine what happens. Therefore, you should decide honestly in each question which option you prefer, because each question has a chance of determining what happens.

Note that if you choose not to learn the number of questions you answered correctly, you will not be able to infer this information by looking at the amount of your bonus at the end of the survey. This is because your bonus depends not only on your performance on the test, but also on your guesses to the other questions, and the randomly-chosen value between \$0 - \$3 that is added to your bonus.

Note also that this is the final part of this study. There will be no additional opportunity to learn this information and no further parts that depend upon this information.

If you choose to learn the information, you will see the information at the end of the survey and will be asked to enter it back to us to confirm that you received the information.

Please continue to question 1.


→

Question 1: Which option do you prefer?

Option A: You are told the information about how many problems you solved correctly on the test and how you performed relative to others.

Option B: You are NOT told the information about how many problems you solved correctly on the test and how you performed relative to others.

→


Powered by Qualtrics 

Question 2: Which option do you prefer?

Option A: You are told the information about how many problems you solved correctly on the test and how you performed relative to others AND you need to complete 10 sliders.

Option B: You are NOT told the information about how many problems you solved correctly on the test and how you performed relative to others.

→

Powered by Qualtrics 

Question 3: Which option do you prefer?

Option A: You are told the information about how many problems you solved correctly on the test and how you performed relative to others.

Option B: You are NOT told the information about how many problems you solved correctly on the test and how you performed relative to others AND you need to complete 10 sliders.

→

Powered by Qualtrics 

Attention Check

Instructions: For this question, we want to make sure you are reading everything. Please select the first option which indicates that you are unsure. Do not choose the option that claims to be the right answer.

What topics was the cognitive skills test on?

I am not sure.

General Science and Math among others

French literature and international business among others

→

Powered by Qualtrics 

Please answer a few short questions about yourself before seeing which question was chosen as the decision-that-counts.

Set 1 of Questions (out of 4)

What is your year of birth?

What is your sex?

Male

Female

Choose one or more races that you consider yourself to be:

White

Asian

Black or African American

Latinx

What region best describes where you live?

Northeast

South

Midwest

West

Did you attend high school in the United States?

Yes

No

How difficult did you find the instructions for this study?

Extremely easy 1	Moderately easy 2	Slightly easy 3	Neither easy nor difficult 4	Slightly difficult 5	Moderately difficult 6	Extremely difficult 7
------------------------	----------------------	--------------------	------------------------------------	-------------------------	------------------------------	-----------------------------

Difficulty of instructions

→

Set 2 of Questions (out of 4)

Considering the questions on the test, tell us whether you think men or women who completed this same HIT you completed answered more cognitive skills questions correctly, on average.

Women answered at least 3 more questions correctly than men on average

Women answered 1 or 2 more questions correctly than men on average

No gender difference

Men answered 1 or 2 more questions correctly than women on average

Men answered at least 3 more questions correctly than women on average

→

Set 3 of Questions (out of 4)

Suppose you learn that you performed **worse** than you believed. On a scale 1-10:


Not at all	Slightly	Moderately	Extremely
1	4	7	10

How much do you think this feedback would influence your own evaluation of your performance?

To what extent does this feedback give you information about your cognitive ability generally?

To what extent does this feedback give you information about your capabilities in other aspects of your life?

→

Powered by Qualtrics 

Set 4 of Questions (out of 4)

Suppose you learn that you performed **better** than you believed. On a scale 1-10:

Not at all	Slightly	Moderately	Extremely
1	4	7	10

How much do you think this feedback would influence your own evaluation of your performance?

To what extent does this feedback give you information about your cognitive ability generally?

To what extent does this feedback give you information about your capabilities in other aspects of your life?

→

Powered by Qualtrics 

Congratulations! You completed the study.

The computer selected Question 1 as the decision that counts.

In that question, you chose Option A. Therefore, you will receive information about your performance on the test. There are no sliders to complete based upon your answer.

Here is the information about your performance on the test:

You answered 0 questions correctly.

Your rank was 10 relative to the other 9 MTurkers. That means you performed better than 0 of the other 9 MTurkers.

Please answer the following questions to confirm that you received the information.

How many questions did you answer correctly?

What was your rank?

Please click next to receive your completion code (SECRET KEY) which lets you claim your earnings.

→

Appendix E. Experimental Instructions, Interview Feedback Study

Restart Survey

Place Bookmark

Mobile view off

Tools ▾

Click to write the question text

This question will not be displayed to the recipient.

Browser	Safari
Version	15.4
Operating System	Macintosh
Screen Resolution	1920x1080
Flash Version	-1
Java Support	0
User Agent	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7) AppleWebKit/605.1.15 (KHTML, like Gecko) Version/15.4 Safari/605.1.15

Overview and Consent

Note that the Requester of this HIT is John-Henry Pezzuto, who is an approved Research Associate at Harvard University. The following is a short summary of this study to help you decide whether or not to be a part of this study. More detailed information is listed later on in this form.

Summary

Why am I being invited to take part in a research study?

We invite you to take part in this research study because you fulfill the eligibility criteria based upon your profile information on Amazon Mechanical Turk.

What should I know about a research study?

- This research study will be explained to you.
- Whether or not you take part is up to you.
- Your participation is completely voluntary.
- You can choose not to take part.

Before you proceed to the survey, please complete the captcha below.

QaH6!-V

Please type in the characters that appear in the above picture.

→

Powered by Qualtrics A

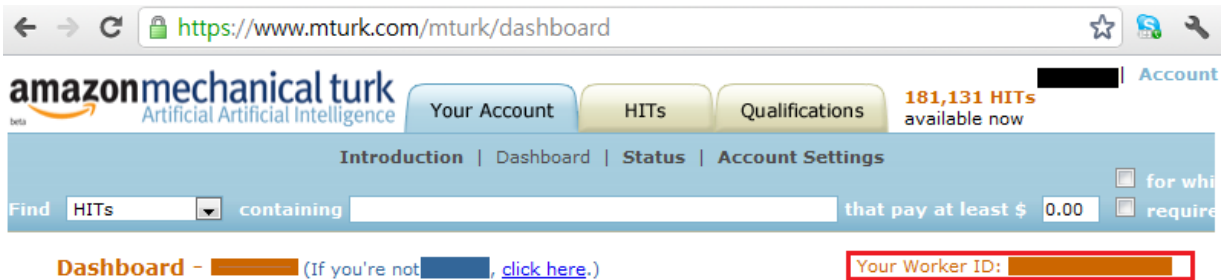
Timing

These page timer metrics will not be displayed to the recipient.

First Click	0 seconds
Last Click	0 seconds
Page Submit	0 seconds
Click Count	0 clicks

To see if you are eligible to participate, please enter your Mechanical Turk Worker ID into the box below and then click NEXT.

Please see below for where you can find your Worker ID. Your Worker ID starts with the letter A and has 12-14 letters or numbers. It must be all CAPITAL letters and no spaces. It is NOT your email address.



Study Overview

Thank you for choosing to participate. From this point forward, we ask you to be focused on the tasks of the study.

In this study, you will be completing two sessions. The first session is the one you are currently in. The second session will be an MTurk HIT that will open on **December 7th**, and will remain open until **December 9th**. On December 7th, you will receive a HIT invitation through email and the MTurk platform to participate in the second session. You will be given until **December 9th** to participate in the second session. We will send you reminders on December 6th and December 9th to help you remember to participate in the second session.

You will only be allowed to participate in the second session if you have completed the first session.

Please take a minute to check that you are available between December 7th and December 9th to participate the second session, and add this to your calendar.

You will receive a participation fee of \$1.00 for your participation in today's session, and a participation fee of \$5.00 for your participation in the second session. In addition, you may earn additional compensation based upon your decisions in the study. We will explain clearly how your decisions influence your additional payment as you move through the study.

The participation fee for today's session will be paid within 48 hours of completing today's session. The participation fee for the second session will be paid within 48 hours of completing the second session. Any additional payment you may have earned in the study will be paid within two weeks of completing the second session.

Overview of Today's Session

Today's session consists of answering questions about yourself that are commonly asked in job interviews to evaluate candidates.

We will explain the questions clearly as you move through the session. Note that in this

session, other MTurkers are also participating and answering the same questions you are answering.

Understanding and attention questions

This survey contains understanding and attention questions. If you answer an understanding question incorrectly, or if you fail to answer an attention question within the given time frame, you may be dismissed from the HIT and will not receive any completion payment or any additional payment. For that reason, you should read the instructions carefully and should not navigate away from the survey page during your participation.

Will you be available for the second session of this study?

- I will be available to take the second part of this study sometime on December 7th - December 9th
- I will not be available

→

Powered by Qualtrics [↗](#)

Here is your first attention check.

In the space below, please spell the word "human" backwards. Please use all lowercase letters and insert no spaces between the letters.



Powered by Qualtrics 

Before continuing, please tell us a bit about yourself.

What is your sex?

- Male
- Female

What is your age bracket?

- 18-30
- 31-50
- 51 or older

What region best describes where you live?

- Northeast
- South
- Midwest
- West

→

Powered by Qualtrics 

Interview questions

In this part, you will be asked 3 questions about yourself that are commonly asked in job interviews to evaluate candidates. You will have **5 minutes to answer each question**.

Please answer these questions as if you were being interviewed for a job opening. Note that your answers may be viewed by other participants in a future session.

Each answer **must be at least 60 words**. If your answer to any of the three questions is less than 60 words, or if you submit an answer too quickly (in 2 minutes or less), you may be dismissed from the study and will not be eligible for additional payment. Additionally, as we would like to make sure to get your honest responses, copy-paste has been disabled.

When you click next, the first question will appear. The computer will time 5 minutes for you and will count the number of words you have typed so far. After the time expires, the next question will appear and the timer and word-counter will restart.

→

Powered by Qualtrics [↗](#)

Timing

These page timer metrics will not be displayed to the recipient.

First Click	0 seconds
Last Click	0 seconds
Page Submit	0 seconds
Click Count	0 clicks

0447

Question 1. What is something you have achieved that you are proud of, and why?

Your answer must have at least 60 more words to be valid.

→

Timing

These page timer metrics will not be displayed to the recipient.

First Click	0 seconds
Last Click	0 seconds
Page Submit	0 seconds
Click Count	0 clicks

0448

Question 2. Describe a difficult task you were faced with and how you addressed it.

Your answer must have at least 60 more words to be valid.

→

Timing

These page timer metrics will not be displayed to the recipient.

First Click	0 seconds
Last Click	0 seconds
Page Submit	0 seconds
Click Count	0 clicks

0446

Question 3. What are you passionate about, and why?

Your answer must have at least 60 more words to be valid.

→

Congratulations! You completed the interview questions.

You have now completed today's session of the study. Thank you so much for your time.

Remember there is a second session of this study in 3 weeks from today on **December 7th - December 9th**. The second session will last approximately 20 minutes. You will earn \$5.00 plus a chance of earning additional payment by completing the second session.

→

Powered by Qualtrics 

Your survey code is: X_48Dkf50

To receive your \$1 payment for participating today, click "Accept HIT" in the Mechanical Turk window (if you haven't already), enter this completion code, and the click "Submit".

Please do so now.

You will receive your payment of \$1.00 for completing this HIT within 48 hours of submitting your survey code. If you have any questions, please email John-Henry at labor.experiment83@gmail.com.

In 20 days, you will receive an email with the access link for the second session of the study.

You may close this window once you have entered your completion code into Amazon's Mechanical Turk platform. Thank you for your participation.

Powered by Qualtrics 

This is Session 2 of a two-session study. You will need to have completed Session 1 to participate in today's session.

What happens if I say yes, but I change my mind later?

You can leave the research at any time. It will not be held against you.

If I take part in this research, how will my privacy be protected? What happens to the information you collect?

Efforts will be made to limit the use and disclosure of your Personal Information, including research study participation, to people who have a need to review this information. We cannot promise complete secrecy. Organizations that may inspect and copy your information include the IRB and other representatives of this organization.

We may share the answers you provide to any questions in this study with other participants of the study. However, we will not share any of your identifying information with other participants, nor will we post any of your identifying information online. The results of this study may be used in reports, presentations, or publications but your name will not be used. All results will be shared only in aggregate form. If identifiers are removed from your identifiable private information or identifiable samples that are collected during this research, that information or those samples could be used for future research studies or distributed to another investigator for future research studies without your additional informed consent.

Compensation

You will receive \$5 for completing the second session today. Additional payments will be earned depending on decisions and performance during the study. These additional payments will be made within two weeks of your completion of the second session.

Who can I talk to?

If you have questions, concerns, or complaints, or think the research has hurt you, talk to the research team at labor.economics83@gmail.com. The primary investigator of this study is **Katherine Coffman, Harvard Business School**.

This research has been reviewed and approved by the Harvard University Area Institutional Review Board ("IRB"). You may talk to them at (617) 496-2847 or cuhs@harvard.edu if:

- Your questions, concerns, or complaints are not being answered by the research team.
- You cannot reach the research team.
- You want to talk to someone besides the research team.
- You have questions about your rights as a research subject.
- You want to get information or provide input about this research.

You may print a copy of this information sheet for your own records.

By clicking "Accept", you acknowledge that you are 18 years or older, have read the information stated and consent to participate in this study.

Accept

Powered by Qualtrics 

Before you proceed to the survey, please complete the captcha below.

QaH6!-V

Please type in the characters that appear in the above picture.

→

Powered by Qualtrics A

Timing

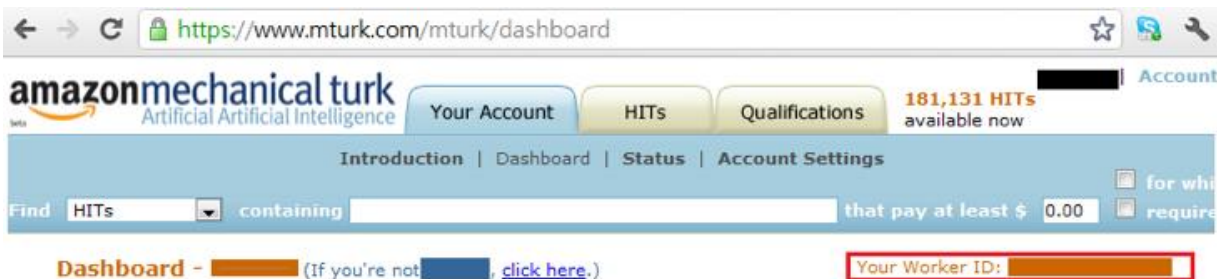
These page timer metrics will not be displayed to the recipient.

First Click	0 seconds
Last Click	0 seconds
Page Submit	0 seconds
Click Count	0 clicks

Only individuals who completed their participation in the first session of this study 3 weeks ago are eligible to participate in today's session.

To see if you are eligible to participate, please enter your Mechanical Turk Worker ID into the box below and then click NEXT.

Please see below for where you can find your Worker ID. Your Worker ID starts with the letter A and has 12-14 letters or numbers. It must be all CAPITAL letters and no spaces. It is NOT your email address.



Study Overview

Thank you for choosing to participate. From this point forward, we ask you to be focused on the tasks of the study.

This is the second and final session of a two-session study. You will receive a participation fee of \$5.00 for your participation in today's session. In addition, you may earn additional compensation based upon your answers in the two sessions. We will explain clearly how your answers influence your additional payment as you move through the study.

The participation fee for today's session will be paid within 48 hours of completing today's session. Any additional payment you may have earned in the study will be paid within two weeks of completing today's session.

Overview of Today's Session

During the first session of this study 3 weeks ago, you answered questions about yourself that are commonly asked in job interviews to evaluate candidates.

In today's session, you will be asked a number of questions related to your participation in the first session. We will explain the questions clearly as you move through the session.

Note that in this session, other MTurkers are also participating and answering the same questions you are answering.

Understanding and attention questions

This survey contains understanding and attention questions. If you answer an understanding question incorrectly, or if you fail to answer an attention question within the given time frame, you may be dismissed from the HIT and will not receive any completion payment or any additional payment. For that reason, you should read the instructions carefully and should not navigate away from the survey page during your participation.

→

Here is your first attention check.

In the space below, please spell the word "human" backwards. Please use all lowercase letters and insert no spaces between the letters.

→

Powered by Qualtrics [↗](#)

Your participation in Session 1

Recall that during the first session of this study 3 weeks ago, you answered the following questions about yourself:

- What is your sex? (Male / Female)
- What is your age bracket? (18-30 / 31-50 / 51 or older)
- What region best describes where you live? (Northeast / South / Midwest / West)

And you answered the following questions that are commonly asked in job interviews:

Question 1. What is something you have achieved that you are proud of, and why?

Question 2. Describe a difficult task you were faced with and how you addressed it.

Question 3. What are you passionate about, and why?

Time between Session 1 and today

After all participants completed the first session of this study, we randomly selected 10 participants, and showed their information (sex, age bracket, and region of residence) and their answers to the three job interview questions to a Human Resources professional with years of experience evaluating job interviews. At the end of today's session, we will inform you of whether you are one of the 10 randomly-chosen participants whose information and answers to the job interview questions we showed to the Human Resources professional.

If you are one of the 10 randomly-chosen participants whose information and answers to the job interview questions we showed to the Human Resources professional, then in the time between the first session of the study 3 weeks ago and today's session, the Human Resources professional saw the information about yourself that you provided at the beginning of the first session (sex, age bracket, and region of residence) and your answers to the three interview questions (and NO other information about yourself), and assigned you a score on each of the following four traits that employers tend to value in general:

- **intellectual curiosity:** an interest in learning about the world and understanding the experiences around you as well as your own experiences
- **a tendency to strive for achievement:** an ability and tendency to accomplish your goals, including professional and personal goals
- **assertiveness:** an ability and tendency to take charge and direct the activities of others
- **tolerance to stress:** an ability and tendency to cope with stressful situations without

being negatively affected by them

We then computed the average of your scores on the four traits, and set this average as your overall Interview Score.

Similarly, the Human Resources professional saw the information (sex, age bracket, and region of residence) and the answers to the three interview questions given by the other 9 randomly-chosen participants of the first session, and assigned them scores on the four traits. We then computed the average of the scores on the four traits for each of the 9 participants, and set this average as their Interview Score.

Before you advance to the next page, we would like to check your understanding of the instructions. Please answer whether each of the following statements is true or false.

1. 10 randomly-chosen participants were evaluated on the following four traits: intellectual curiosity, a tendency to strive for achievement, assertiveness, and tolerance to stress

- True
- False

2. The evaluation was provided by a Human Resources professional

- True
- False

3. The evaluator saw the sex, age bracket, and region of residence, and the answers to the three interview questions given by each of the 10 randomly-chosen participants who were evaluated.

- True
- False

We will now ask you some questions regarding the evaluation by the Human Resources professional.

→

Powered by Qualtrics 

Suppose you are one of the 10 randomly-chosen participants whose information (sex, age bracket, and region of residence) and answers to the job interview questions we showed to the Human Resources professional.

If we compared you to the 9 other randomly-chosen participants whose information (sex, age bracket, and region of residence) and answers we also showed to the Human Resources professional, where would you rank in terms of your Interview Score (the average of your scores on the four traits of intellectual curiosity, a tendency to strive for achievement, assertiveness, and tolerance to stress, as determined by the Human Resources professional)?

Please provide a guess from 1 to 10 of your rank relative to the other 9 participants on the Interview Score. **If you guess correctly and you are one of the 10 randomly selected people who were evaluated, you will earn an extra \$10 in bonus payment.**

Selecting 1 means that you think you scored better than all 9 participants.
Selecting 10 means that you think you scored worse than all 9 participants.

- Rank 1 - I scored better than all 9 participants
- Rank 2
- Rank 3
- Rank 4
- Rank 5
- Rank 6
- Rank 7
- Rank 8
- Rank 9
- Rank 10 - I scored worse than all 9 participants

→

How sure do you feel about your guess that you ranked 5 out of 10?

- 1. Not at all sure
- 2. A little sure
- 3. Moderately sure
- 4. Very sure
- 5. Completely sure

→

Powered by Qualtrics 

Suppose you are one of the 10 randomly-chosen participants whose information (sex, age bracket, and region of residence) and answers to the interview questions we showed to the Human Resources professional.

How interested are you in learning your rank on the Interview Score?

- 1. Not at all interested
- 2. Slightly interested
- 3. Moderately interested
- 4. Very interested
- 5. Extremely interested

The arrow button will take you to additional, related questions.



As you will see, some questions in the following pages will involve the possibility of "completing sliders". Before you move to the next page, we'd like to explain now what it means to complete a slider.

To complete a slider, you must scroll the slider bar to a specific position. The specific position is a randomly-chosen value between 0 and 100, and is indicated with the slider. The slider is not completed successfully until its bar is scrolled exactly to the indicated position. If you are asked to complete sliders, you will have as much time as you would like to complete the sliders. But, you must complete all sliders successfully in order to advance to the next page.

To get a feel for what it means to complete sliders, below there are 2 sliders. Please scroll each of them to the indicated position. You will be able to advance to the next page only after you complete both sliders successfully.

Please complete the following sliders.

0 10 20 30 40 50 60 70 80 90 100

Slide to 60

Slide to 28



We wish to know more about your preferences for finding out how you ranked on the Interview Score.

In this part you will see 11 questions. In each question you must choose between option A and option B.

Option A means that you decide to be told your rank on the Interview Score at the end of today's session.

Option B means that you decide NOT to be told your rank on the Interview Score at the end of today's session.

In addition, either option may also require you to complete a specific number of sliders at the end of today's session.

Your responses to the 11 questions will determine whether you find out how you ranked on the Interview Score, if you are one of the 10 randomly-chosen participants whose information (sex, age bracket, and region of residence) and answers to the interview questions we showed to the Human Resources professional. Your responses to the 11 questions will also determine whether you must complete sliders (and how many) to complete today's session.

After you've responded to all 11 questions, the computer will randomly select one question as the decision-that-counts. The computer will use your decision (A or B) in only that question to determine what happens. Therefore, you should decide honestly in each question which option you prefer, because each question has an equal chance of being selected, and whatever you choose in that question will happen.

Note that if you choose not to learn your rank on the Interview Score in the decision-that-counts, you will not be able to change your decision later in today's session.

If you choose to learn your rank on the Interview Score, and you are one of the 10 randomly-chosen participants whose information (sex, age bracket, and region of residence) and answers to the job interview questions we showed to the Human Resources professional, you will see your rank at the end of today's session and will be

asked to type in your rank back to us to confirm that you received the information.

Please continue to question 1.

→

Powered by Qualtrics 

Attention check

Instructions: For this question, we want to make sure you are reading everything. Please select the first option which indicates that you are unsure. Do not choose the option that claims to be the right answer.

What four traits will the Human Resources professional assign you a score on?

- I am not sure
- Intellectual curiosity, a tendency to strive for achievement, assertiveness, and tolerance to stress
- Imagination, clarity, experience, and originality

→

Powered by Qualtrics [↗](#)

Question 1. Which option do you prefer?

- Option A: You are told your rank on the Interview Score.
- Option B: You are NOT told your rank on the Interview Score.

→

Powered by Qualtrics [↗](#)

Question 2. Which option do you prefer?

- Option A: You are told your rank on the Interview Score, AND you must complete 2 sliders at the end of today's session.
- Option B: You are NOT told your rank on the Interview Score.

→

Powered by Qualtrics [↗](#)

Question 3. Which option do you prefer?

- Option A: You are told your rank on the Interview Score, AND you must complete 5 sliders at the end of today's session.
- Option B: You are NOT told your rank on the Interview Score.

→

Powered by Qualtrics [↗](#)

Question 4. Which option do you prefer?

- Option A: You are told your rank on the Interview Score, AND you must complete 10 sliders at the end of today's session.
- Option B: You are NOT told your rank on the Interview Score.

→

Powered by Qualtrics [↗](#)

Question 5. Which option do you prefer?

- Option A: You are told your rank on the Interview Score, AND you must complete 50 sliders at the end of today's session.
- Option B: You are NOT told your rank on the Interview Score.

→

Powered by Qualtrics [↗](#)

Question 6. Which option do you prefer?

- Option A: You are told your rank on the Interview Score, AND you must complete 100 sliders at the end of today's session.
- Option B: You are NOT told your rank on the Interview Score.

→

Powered by Qualtrics [↗](#)

Question 7. Which option do you prefer?

- Option A: You are told your rank on the Interview Score.
- Option B: You are NOT told your rank on the Interview Score, AND you must complete 2 sliders at the end of today's session.

→

Powered by Qualtrics [↗](#)

Question 8. Which option do you prefer?

- Option A: You are told your rank on the Interview Score.
- Option B: You are NOT told your rank on the Interview Score, AND you must complete 5 sliders at the end of today's session.

→

Powered by Qualtrics [↗](#)

Question 9. Which option do you prefer?

- Option A: You are told your rank on the Interview Score.
- Option B: You are NOT told your rank on the Interview Score, AND you must complete 10 sliders at the end of today's session.

→

Powered by Qualtrics [↗](#)

Question 10. Which option do you prefer?

- Option A: You are told your rank on the Interview Score.
- Option B: You are NOT told your rank on the Interview Score, AND you must complete 50 sliders at the end of today's session.

→

Powered by Qualtrics [↗](#)

Question 11. Which option do you prefer?

- Option A: You are told your rank on the Interview Score.
- Option B: You are NOT told your rank on the Interview Score, AND you must complete 100 sliders at the end of today's session.

→

Powered by Qualtrics [↗](#)

Please answer a few short questions before we tell you which question was chosen as the decision-that-counts.

→

Powered by Qualtrics 

Suppose you learn that you ranked **better** than you believed on the Interview Score. On a scale 1-10:

How much do you think this feedback would influence your own evaluation of your abilities in terms of the traits of intellectual curiosity, a tendency to strive for achievement, assertiveness, and tolerance to stress?

1. Not at all 2. 3. 4. 5. 6. 7. 8. 9. 10. Extremely

To what extent would this feedback that you performed better than you believed lead you to change your beliefs about your ability to perform well on a job interview?

1. Not at all 2. 3. 4. 5. 6. 7. 8. 9. 10. Extremely

To what extent would this feedback that you performed better than you believed lead you to change your beliefs about your capabilities in other aspects of your life?

1. Not at all 2. 3. 4. 5. 6. 7. 8. 9. 10. Extremely

→

Suppose you learn that you ranked **worse** than you believed on the Interview Score. On a scale 1-10:

How much do you think this feedback would influence your own evaluation of your abilities in terms of the traits of intellectual curiosity, a tendency to strive for achievement, assertiveness, and tolerance to stress?

1. Not at all 2. 3. 4. 5. 6. 7. 8. 9. 10. Extremely

To what extent would this feedback that you performed worse than you believed lead you to change your beliefs about your ability to perform well on a job interview?

1. Not at all 2. 3. 4. 5. 6. 7. 8. 9. 10. Extremely

To what extent would this feedback that you performed worse than you believed lead you to change your beliefs about your capabilities in other aspects of your life?

1. Not at all 2. 3. 4. 5. 6. 7. 8. 9. 10. Extremely

→

Timing

These page timer metrics will not be displayed to the recipient.

First Click	0 seconds
Last Click	0 seconds
Page Submit	0 seconds
Click Count	0 clicks

Choose one or more races that you consider yourself to be.

- White
- Asian
- Black or African American
- Latinx

What is the highest education degree you have received?

- Less than high school degree
- High school degree or equivalent (for example, GED)
- Some college but no degree
- Associate degree
- Bachelor degree
- Graduate degree

Choose ALL the categories that describe your current employment situation

- An employee working less than 40 hours per week
- An employee working 40 or more hours per week
- Self-employed
- Unemployed
- Looking for a job
- Retired

Not able to work

Did you attend high school in the United States?

Yes

No

How difficult did you find the instructions for this study today?

Extremely easy

Moderately easy

Slightly easy

Neither easy nor difficult

Slightly difficult

Moderately difficult

Extremely difficult

→

Powered by Qualtrics [↗](#)

Timing

These page timer metrics will not be displayed to the recipient.

First Click	0 seconds
Last Click	0 seconds
Page Submit	0 seconds
Click Count	0 clicks

Please indicate your agreement with the following two statements as they relate to you personally.

In the past, I have worried whether I have been treated or evaluated unfairly because of my sex.

- Strongly disagree
- Disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Agree
- Strongly agree

In the future, if I were trying to find or keep a job, I think employers will treat or evaluate me differently because of my sex.

- I think employers will treat or evaluate me SUBSTANTIALLY LESS FAVORABLY than others because of my sex
- I think employers will treat or evaluate me SLIGHTLY LESS FAVORABLY than others because of my sex
- I think employers will treat or evaluate me EQUALLY to others because of my sex
- I think employers will treat or evaluate me SLIGHTLY MORE FAVORABLY than others because of my sex
- I think employers will treat or evaluate me SUBSTANTIALLY MORE FAVORABLY than others because of my sex

Timing

These page timer metrics will not be displayed to the recipient.

First Click	0 seconds
Last Click	0 seconds
Page Submit	0 seconds
Click Count	0 clicks

Consider the 10 randomly-chosen participants of this study whose information (sex, age bracket, and region of residence) and answers to the job interview questions were shown to the Human Resources professional. Tell us whether you think men or women, on average, obtained a better Interview Score (the average of the scores on the four traits of intellectual curiosity, a tendency to strive for achievement, assertiveness, and tolerance to stress, as determined by the Human Resources professional).

- Women obtained a much better Interview Score than men on average
- Women obtained a slightly better Interview Score than men on average
- No difference between women and men on average
- Men obtained a slightly better Interview Score than women on average
- Men obtained a much better Interview Score than women on average

→

Congratulations, you finished the session!

You are NOT one of the 10 participants randomly selected to have their answers to the job interview questions evaluated. Therefore, you will not receive information about your Interview Score.

You now completed the study. Thank you so much for your time. Please press the arrow to receive your survey completion code which lets you claim your earnings.



Powered by Qualtrics 

Appendix F. Experimental Instructions, Forecast Study (Cognitive Test and Interview)

Eliciting Beliefs of Demand for Performance Feedback

Start of Block: Consent

consent What is the purpose of this research? In this study, we are interested in your beliefs about how other people behaved in a previous study. In particular, in a previous study that we conducted, we asked participants how willing they were to find out how they performed on a task. In this study, we will ask you to guess how willing these past participants were to find out how they performed.

What can I expect if I take part in this research?

You will complete an online survey. It should take you no more than 15 minutes. You will not interact with any other participants.

You will earn \$4 for completing the study. This payment will be made to you through the Prolific platform within 48 hours of your participation. In addition, you may earn \$0.25 in additional bonus pay if your guesses about past participants were more accurate. This additional payment will be made through Prolific within two weeks of your participation.

Your answers will be linked to your Prolific ID at the time of the study. However, we will delete your Prolific ID from the stored dataset after completing study payments. Only data without your Prolific ID will be analyzed or shared. No other identifying information will be collected.

This survey contains understanding and attention questions. If you answer an understanding question incorrectly, or if you fail to answer an attention question within the given time frame, you may be dismissed from the study and may not receive payment. For that reason, you should read the instructions carefully and should not navigate away from the survey page during your participation.

consent2 What should I know about a research study? Whether or not you take part is up to you. Your participation is completely voluntary. You can choose not to take part. You can agree to take part and later change your mind. Your decision will not be held against you. Your refusal to participate will not result in any consequences or loss of benefits to which you are otherwise entitled to receive. If you wish to stop participating at any time, simply close the web browser. You can ask all the questions you want before you decide.

Risks and benefits

There are no foreseeable risks associated with this study, nor are there direct benefits to you.

You may not be told everything

As part of this research design, you may not be told everything about the purpose of this study. In addition, while you will have complete and truthful information about the procedures of the version of the study that you

are participating in, there may be other versions of the study with different procedures. You will not be told about these other versions. These other versions have no impact on your payments.

Who can I talk to?

This study is being conducted by researchers at Harvard Business School and Katz School of Business. If you have questions, concerns, or complaints, or think the research has hurt you, talk to the research team at Katherine Coffman, Harvard Business School, 617 495 6538 or kcoffman@hbs.edu, or David Klinowski, Katz School of Business, dklinowski@katz.pitt.edu. If you wish to contact someone independent of the research team, you may contact the Human Subjects Protection Advocate of the Human Research Protection Office, University of Pittsburgh, at 1-866-212-2668.

End of Block: Consent

Start of Block: Overview

Overview

Study Overview Thank you for choosing to participate. From this point forward, we ask you to be focused on the tasks of the study. In this study, we are interested in your beliefs about how other people behaved in a previous study.

In particular, in a previous study that we conducted, we asked participants how willing they were to find out how they performed on a task relative to others. In this study, we will ask you to guess how willing these past participants were to find out how their performance compared to others.

In the next pages, you will receive more information about exactly what task these previous participants performed and what questions they were asked. You will then provide your guesses.

You will receive \$4 for completing the study.

Also, at the end of this study, we will randomly choose one of the guesses you make. If that guess is within 5 percentage points of the correct answer (plus or minus 5 percentage points), we will pay you an additional \$0.25 as a bonus payment.

Understanding and Attention Questions This survey contains understanding and attention questions.

If you answer an understanding question incorrectly, or if you fail to answer an attention question within the given time frame, you may be dismissed from the study and will not receive any completion payment or any additional payment.

For that reason, you should read the instructions carefully and should not navigate away from the survey page during your participation.

Page Break

ac_human Here is your first attention check.

In the space below, please spell the word "human" backwards. Please use all lowercase letters and insert no spaces between the letters.

End of Block: Overview

Start of Block: Initial Demographics

sex Before continuing, please tell us a bit about yourself.

What is your gender?

- Man (5)
 - Woman (6)
 - Non-binary / third gender (7)
 - Transgender (8)
 - Prefer not to say (9)
-

age What is your age bracket?

- 18-30 (1)
 - 31-50 (2)
 - 51 or older (3)
 - Prefer not to say (4)
-

region What U.S. region best describes where you live?

- Northeast (1)
- South (2)
- Midwest (3)
- West (4)
- Other (5)
- Prefer not to say (6)

Page Break

Bridge Great! Thank you. We'll now start explaining the task that the previous participants completed.

End of Block: Initial Demographics

Start of Block: Interview Study Intro

Q41 The task that the previous study participants completed was to answer 3 interview questions.

These were questions that are commonly asked in job interviews to evaluate candidates. The specific questions were:

1. *What is something you have achieved that you are proud of, and why?*
2. *Describe a difficult task you were faced with and how you addressed it.*
3. *What are you passionate about, and why?*

Participants had to answer each question as if they were being interviewed for a job opening. They had 5 minutes to answer each question. Answers had to be typed on their computer or electronic device, and each answer had to be at least 60 words.

After answering the interview questions, the first session of the study concluded. Participants were asked to return to a second session, to be held three weeks later. In the three weeks between the first and second session, a Human Resources (HR) professional evaluated some of the answers to the interview questions.

In the second session of the study, we measured the previous study participants' willingness to receive information on how the HR professional evaluated their answers.

Next, we tell you more details of the evaluation by the HR professional and of how we measured the previous study participants' willingness to find out how they performed on the evaluation.

Display This Question:

`if blind_pool = non-blind`

Nonblind **Details of the evaluation by the HR professional** *After all participants completed the first session of the study, we randomly selected 10 participants, and showed their information (sex, age bracket, and region of residence) and their answers to the three job interview questions to a HR professional with years of experience evaluating job interviews. The HR professional assigned each participant a score from 1 - 10 (with 10 being the best score) on the following four traits that employers tend to value in general: - **intellectual curiosity**: an interest in learning about the world and understanding the experiences around you as well as your own experiences - **a tendency to strive for achievement**: an ability and tendency to accomplish your goals, including professional and personal goals - **assertiveness**: an ability and tendency to take charge and direct the activities of others - **tolerance to stress**: an ability and tendency to cope with stressful situations without being negatively affected by them* We then computed the average of the participant's scores on the four traits, and set this average as the participant's overall Interview Score. Finally, we calculated the participant's rank relative to the 9 other randomly selected participants whose answers were evaluated by the HR professional.

The rank is a number between 1 and 10. Rank 1 means that the participant obtained a better Interview Score than all other nine participants, and Rank 10 means that the participant obtained a worse Interview Score than all other nine participants.

Display This Question:

`if blind_pool = blind`

Blind **Details of the evaluation by the HR professional** *After all participants completed the first session of the study, we randomly selected 10 participants, and showed only their answers to the three job interview questions to a HR professional with years of experience evaluating job interviews. The HR professional assigned each participant a score from 1 - 10 (with 10 being the best score) on the following four traits that employers tend to value in general: - **intellectual curiosity**: an interest in learning about the world and understanding the experiences around you as well as your own experiences - **a tendency to strive for achievement**: an ability and tendency to accomplish your goals, including professional and personal goals - **assertiveness**: an ability and tendency to take charge and direct the activities of others - **tolerance to stress**: an ability and tendency to cope with stressful situations without being negatively affected by them* We then computed the average of the participant's scores on the four traits, and set this average as the participant's overall Interview Score. Finally, we calculated the participant's rank relative to the 9 other randomly selected participants whose answers were evaluated by the HR professional. The rank is a number between 1 and 10. Rank 1 means that the participant obtained a better Interview Score than all other nine participants, and Rank 10 means that the participant obtained a worse Interview Score than all other nine participants.

Q65 In the second session of the study, we described to all participants the procedure followed by the HR professional to evaluate the interview answers and how we computed their Interview Score rank.

Then, we asked participants about their preferences for finding out how they performed. We will now explain to you exactly how we asked them about their preferences.





Test sliders As you will see in the following pages, some of the questions we asked the previous study participants involved the possibility of "completing sliders." Before you move to those pages, we'd like to explain now what it means to complete a slider.

To complete a slider, you must scroll the slider bar to a specific position. The specific position is a randomly chosen number between 0 and 100, and is indicated with a slider. The slider is not completed successfully until its bar is scrolled exactly to the indicated position.

When completing sliders, previous study participants had as much time as they wanted to complete the sliders. But, they had to complete all sliders successfully in order to advance to the next page.

To get a feel for what it means to complete sliders, below are 2 sliders. Please scroll each of them to the indicated position. You will be able to advance to the next page only after you complete both sliders successfully.

0 10 20 30 40 50 60 70 80 90 100

Slide to 60 ()	
Slide to 28 ()	

Options **To measure how willing participants were to find out how they performed, we asked them a series of 11 questions.**

In each question, they had to choose between two options:

- **Option A:** Yes, tell me how my interview score ranked compared to the other participants
- **Option B:** No, do NOT tell me how my interview score ranked compared to the other participants

As you will see, the "cost" (in terms of required sliders) of choosing Option A or Option B varied across the 11 questions. Sometimes, participants had to complete sliders in order to choose one of the options.

End of Block: Interview Study Intro

Start of Block: Eliciting Beliefs Interview Men

Q341 Your job now is to guess the answers given by MEN in the previous study.

In particular, we randomly selected 100 MEN from the previous study. You will guess how many of those 100 MEN chose to find out how they performed for each question below.

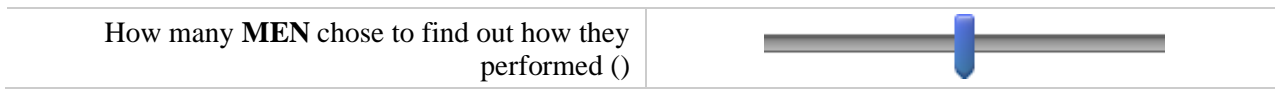
Recall that we will randomly select one of the guesses you make in this study and pay you an additional \$0.25 if you are within 5 percentage points of the correct answer.

Q342 First, participants were asked to choose between these options:

- **Option A:** Yes, tell me how my interview score compared to the other participants
 - **Option B:** No, do NOT tell how me my interview score compared to the other participants
-

Q343 **YOUR GUESS:** How many of the **100 MEN** chose Option A (to find out how they performed) in the question above?

0 10 20 30 40 50 60 70 80 90 100



Page Break






Q344 In the next five questions, we asked whether the previous study participants would be willing to complete X sliders in order to find out how they performed. We varied the number X across the five different questions.

For each X , participants were asked to choose between these options:

- **Option A:** Yes, tell me how my interview score compared to the other participants AND I will complete X sliders
- **Option B:** No, do NOT tell me how my interview score compared to the other participants

YOUR GUESSES: For each option below, how many of the **100 MEN** chose Option A (to find out how they performed) when...

0 10 20 30 40 50 60 70 80 90 100

They had to complete 2 sliders to find out their rank ()	
They had to complete 5 sliders to find out their rank ()	
They had to complete 10 sliders to find out their rank ()	
They had to complete 50 sliders to find out their rank ()	
They had to complete 100 sliders to find out their rank ()	

Page Break

Q345 In the final five questions, we asked whether the previous study participants would be willing to complete X sliders in order to NOT FIND OUT how they performed. We varied the number X across the five different questions.

For each X , participants were asked to choose between these options:

- **Option A:** Yes, tell me how my interview score compared to the other participants
- **Option B:** No, do NOT tell me how my interview score compared to the other participants AND I will complete X sliders

YOUR GUESSES: For each option below, how many of the **100 MEN** chose Option A (to find out how they performed) when...



End of Block: Eliciting Beliefs Interview Men

Start of Block: TransitionMW

TransitionMW Thank you for providing your guesses! Now that you have provided your answers about men, we would like to ask you about your guesses for women.

End of Block: TransitionMW

Start of Block: Eliciting Beliefs Interview Women

Q346 **Your job now is to guess the answers given by WOMEN in the previous study.**

In particular, we randomly selected 100 WOMEN from the previous study. You will guess how many of those 100 WOMEN chose to find out how they performed for each question below.

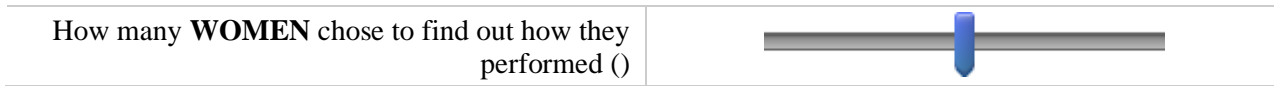
Recall that we will randomly select one of the guesses you make in this study and pay you an additional \$0.25 if you are within 5 percentage points of the correct answer.

Q347 First, participants were asked to choose between these options:

- **Option A:** Yes, tell me how my interview score compared to the other participants
 - **Option B:** No, do NOT tell me how my interview score compared to the other participants
-

Q348 **YOUR GUESS:** How many of the **100 WOMEN** chose Option A (to find out how they performed) in the question above?

0 10 20 30 40 50 60 70 80 90 100



Page Break






Q349 In the next five questions, we asked whether the previous study participants would be willing to complete X sliders in order to find out how they performed. We varied the number X across the five different questions.

For each X , participants were asked to choose between these options:

- **Option A:** Yes, tell me how my interview score compared to the other participants AND I will complete X sliders
- **Option B:** No, do NOT tell me how my interview score compared to the other participants

YOUR GUESSES: For each option below, how many of the **100 WOMEN** chose Option A (to find out how they performed) when...

0 10 20 30 40 50 60 70 80 90 100

They had to complete 2 sliders to find out their rank ()	
They had to complete 5 sliders to find out their rank ()	
They had to complete 10 sliders to find out their rank ()	
They had to complete 50 sliders to find out their rank ()	
They had to complete 100 sliders to find out their rank ()	

Page Break

Q350 In the final five questions, we asked whether the previous study participants would be willing to complete X sliders in order to NOT FIND OUT how they performed. We varied the number X across the five different questions.

For each X , participants were asked to choose between these options:

- **Option A:** Yes, tell me how my interview score compared to the other participants
- **Option B:** No, do NOT tell me how my interview score compared to the other participants AND I will complete X sliders

YOUR GUESSES: For each option below, how many of the **100 WOMEN** chose Option A (to find out how they performed) when...



End of Block: Eliciting Beliefs Interview Women

Start of Block: TransitionWM

TransitionWM Thank you for providing your guesses! Now that you have provided your answers about women, we would like to ask you about your guesses for men.

End of Block: TransitionWM

Start of Block: Cognitive Skills Intro

int6 The task that the previous study participants completed was a **cognitive skills test**.

The test contained 30 questions. Participants had 5 minutes to work on the test. They received additional payment for each correct answer, and no penalty was assessed for skipped questions or incorrect answers.

Page Break



Comprehension check What was the subject matter of the test that the previous participants took?

- Literature (1)
- Cognitive Skills (2)
- Biology (3)

Display This Question:

If What was the subject matter of the test that the previous participants took? = Cognitive Skills

Q127 That's correct! Please proceed.

Display This Question:

If What was the subject matter of the test that the previous participants took? != Cognitive Skills

Q129 Your answer is **incorrect**. The previous participants completed a **cognitive skills test**.

Please proceed.

Page Break

Q130 When you click next, you will be able to review the test questions that the previous participants worked on. Note that you do not have to answer these test questions yourself, but you must spend at least two minutes reviewing the questions. Understanding the task that the previous participants completed may help you to provide better guesses in the next section.

End of Block: Cognitive Skills Intro

Start of Block: ASVAB Easy

Q131 Timing

First Click (1)

Last Click (2)

Page Submit (3)

Click Count (4)



Q106 GENERAL SCIENCE: Which planet orbits closest to Earth?

- Mercury (1)
- Venus (2)
- Jupiter (3)
- Saturn (4)



Q107 GENERAL SCIENCE: Which substance is a good thermal conductor?

- plastic (1)
 - rubber (2)
 - porcelain (3)
 - aluminum (4)
-



Q108 GENERAL SCIENCE: Which of the following is NOT a nucleotide of DNA?

- adenine (1)
 - guanine (2)
 - thymine (3)
 - uracil (4)
-



Q109 GENERAL SCIENCE: What position do most fungi occupy in a food web?

- producers (1)
 - primary consumers (2)
 - secondary consumers (3)
 - decomposers (4)
-



Q110 ARITHMETIC REASONING: If three burgers and two orders of fries cost \$26.50 and a burger costs \$6.50, how much does one order of fries cost?

- \$1.75 (1)
 - \$3.50 (2)
 - \$6.75 (3)
 - \$7.00 (4)
-



Q111 ARITHMETIC REASONING: A worker was paid \$15,036 for 7 months of work. If he received the same amount each month, how much was he paid for the first 2 months?

- \$2,148 (1)
 - \$4,296 (2)
 - \$5,137 (3)
 - \$6,444 (4)
-



Q112 ARITHMETIC REASONING: A high school football team played 12 games in a season. If they won 75 percent of their games, how many games did they lose?

- 3 (1)
 - 4 (2)
 - 6 (3)
 - 9 (4)
-



Q113 ARITHMETIC REASONING: A fruit stand sells apples, bananas, and oranges at a ratio of 3:2:1. If the fruit stand sells 20 bananas, how many total pieces of fruit does the fruit stand sell?

- 10 (1)
 - 30 (2)
 - 40 (3)
 - 60 (4)
-

Q114 ARITHMETIC REASONING: Aprille has \$50 to buy the items on her shopping list. Assuming there is no sales tax, about how much change will Aprille receive after buying all items on her list?

Aprille's List

	ITEM	PRICE	
Hammer	\$13.24		
Screwdriver	\$11.99	Wrench	Nails
\$4.27			\$5.60

\$12 (1)

\$13 (2)

\$14 (3)

\$15 (4)



Q115 ARITHMETIC REASONING: At the grocery store, apples cost \$1.89 per pound and oranges cost \$2.19 per pound. How much would it cost to purchase 2 pounds of apples and 1.5 pounds of oranges?

- \$6.62 (1)
- \$7.07 (2)
- \$7.14 (3)
- \$7.22 (4)
-



Q116 ARITHMETIC REASONING: Kim and Chris are writing a book together. Kim wrote twice as many pages as Chris, and together they wrote 240 pages. How many pages did Chris write?

- 80 (1)
- 100 (2)
- 120 (3)
- 160 (4)
-

Q117 ARITHMETIC REASONING: Out of 1,560 students at Ward Middle School, 15% want to take French. Which expression represents how many students want to take French?

- 1560/15 (1)
 - 1560 x 15 (2)
 - 1560 x 0.15 (3)
 - 1560/0.15 (4)
-



Q118 ARITHMETIC REASONING: A group of 20 friends is planning a road trip. They have 3 cars that seat 4 people, 3 cars that seat 5 people, and 1 car that seats 6 people. What is the fewest number of cars they can take on the trip if each person needs his or her own car seat?

- 3 cars (1)
 - 4 cars (2)
 - 5 cars (3)
 - 6 cars (4)
-



Q119 MATH KNOWLEDGE: What is the value of $(\frac{1}{2})^3$?

- 1/8 (1)
 - 1/6 (2)
 - 1/4 (3)
 - 3/8 (4)
-



Q120 MATH KNOWLEDGE: What is the remainder when 397 is divided by 4?

- 0 (1)
 - 1 (2)
 - 2 (3)
 - 3 (4)
-



Q121 MATH KNOWLEDGE: What is $\frac{5}{8}$ as a percent?

- 1.6% (1)
 - 16% (2)
 - 0.625% (3)
 - 62.5% (4)
-



Q122 MATH KNOWLEDGE: Rectangular water tank A is 5 feet long, 10 feet wide, and 4 feet tall. Rectangular tank B is 5 feet long, 5 feet wide, and 4 feet tall. If the same amount of water is poured into both tanks and the height of the water in Tank A is 1 foot, how high will the water be in Tank B?

- 1 foot (1)
 - 2 feet (2)
 - 3 feet (3)
 - 4 feet (4)
-



Q123 MATH KNOWLEDGE: What is 498,235 rounded to the nearest thousands?

- 498,000 (1)
 - 498,200 (2)
 - 499,000 (3)
 - 499,200 (4)
-

Q124 MATH KNOWLEDGE: What is the value of $3x + 7y - 4$ if $x = 8$ and $y = 2$?

- 34 (1)
 - 38 (2)
 - 42 (3)
 - 58 (4)
-



Q125 MECHANICAL COMPREHENSION: What is the unit for power?

- watt (W) (1)
 - joule (J) (2)
 - newton (N) (3)
 - coulomb (C) (4)
-



Q126 MECHANICAL COMPREHENSION: In the gear train shown below, Gear B will move _____ and Gear C will move _____.

- clockwise; clockwise (1)
 - clockwise; counterclockwise (2)
 - counterclockwise; counterclockwise (3)
 - counterclockwise; clockwise (4)
-



Q127 MECHANICAL COMPREHENSION: Which of the following terms defines friction?

- a force that opposes motion (1)
 - a force that pushes down onto a surface (2)
 - a force that rotates an object (3)
 - a force that increases the force of gravity (4)
-



Q128 MECHANICAL COMPREHENSION: What is the gear ratio for an input gear with 1,000 teeth and an output gear with 3,500 teeth?

- 3.5 (1)
 - 35 (2)
 - 350 (3)
 - 3,500 (4)
-



Q129 ASSEMBLING OBJECTS: Given the following set of objects, please determine which answer choice shows how the objects will look once the parts are put together.

- (1)
 - (2)
 - (3)
 - (4)
-



Q130 ASSEMBLING OBJECTS: Given the following set of objects, please determine which answer choice shows how the objects will look once the parts are put together.

- (1)
 - (2)
 - (3)
 - (4)
-



Q131 ASSEMBLING OBJECTS: Given the following set of objects, please determine which answer choice shows how the objects will look once the parts are put together.

- (1)
 - (2)
 - (3)
 - (4)
-



Q132 ASSEMBLING OBJECTS: Given the following set of objects, please determine which answer choice shows how the objects will look once the parts are put together.

- (1)
 - (2)
 - (3)
 - (4)
-



Q133 ASSEMBLING OBJECTS: Given the following set of objects, please determine which answer choice shows how the objects will look once the parts are put together.

- (1)
 - (2)
 - (3)
 - (4)
-



Q134 ASSEMBLING OBJECTS: Given the following set of objects, please determine which answer choice shows how the objects will look once the parts are put together.

- (1)
- (2)
- (3)
- (4)



Q135 ASSEMBLING OBJECTS: Given the following set of objects, please determine which answer choice shows how the objects will look once the parts are put together.

- (1)
- (2)
- (3)
- (4)

End of Block: ASVAB Easy

Start of Block: ASVAB Hard

Q132 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)



Q137 GENERAL SCIENCE: What organism has cells that contain mitochondria?

- whale (1)
 - mushroom (2)
 - tulip (3)
 - all of the above (4)
-



Q138 GENERAL SCIENCE: A box sliding down a ramp experiences all of the following forces EXCEPT

- tension. (1)
 - friction. (2)
 - gravity. (3)
 - normal. (4)
-



Q139 GENERAL SCIENCE: Which pH level is classified as a base?

- 1 (1)
 - 4 (2)
 - 6 (3)
 - 8 (4)
-



Q140 GENERAL SCIENCE: What factor is an abiotic part of an ecosystem?

- producers (1)
 - consumers (2)
 - water (3)
 - decomposers (4)
-



Q141 GENERAL SCIENCE: Which organelle makes proteins?

- mitochondria (1)
 - cytoplasm (2)
 - vacuoles (3)
 - ribosomes (4)
-



Q142 GENERAL SCIENCE: The exchange of gases happens in which parts of the respiratory system?

- alveoli (1)
 - brochi (2)
 - lobes (3)
 - trachea (4)
-



Q143 GENERAL SCIENCE: Enzymes are an example of which type of macromolecule?

- lipids (1)
 - DNA (2)
 - RNA (3)
 - proteins (4)
-



Q144 GENERAL SCIENCE: What is the part of a neuron that sends electrical signals away from the neuron cell?

- axon (1)
 - dendrite (2)
 - stimulus (3)
 - myelin (4)
-



Q145 ARITHMETIC REASONING: The population of a town was 7,250 in 2014 and 7,374 in 2015. what was the percent increase from 2014 to 2015 to the nearest tenth of a percent?

- 1.5% (1)
 - 1.6% (2)
 - 1.7% (3)
 - 1.8% (4)
-



Q146 ARITHMETIC REASONING: Convert 8 pounds, 8 ounces to kilograms to the nearest tenth of a kilogram.

- 3.9 kilograms (1)
 - 4.1 kilograms (2)
 - 17.6 kilograms (3)
 - 18.7 kilograms (4)
-



Q147 ARITHMETIC REASONING: Miguel works at a car dealership and is paid a 2 percent commission on every car he sells. If he sells one car for \$15,000 and two cars for \$12,900 each, how much will he be paid in commissions?

- \$300 (1)
 - \$558 (2)
 - \$816 (3)
 - \$5,580 (4)
-



Q148 ARITHMETIC REASONING: If the smallest angle in a non-right triangle is 20° and the shortest side is 14, what is the length of the longest side if the largest angle is 100° ?

- 12.78 (1)
 - 34.31 (2)
 - 40.31 (3)
 - 70.02 (4)
-



Q149 MATH KNOWLEDGE: Simplify: $[(3x^2y^2)^2]/[33x-2y^3]$

- $3x6y$ (1)
 - $(x6y)/3$ (2)
 - $x^4/(3y)$ (3)
 - $(3x^4)/y$ (4)
-



Q150 MATH KNOWLEDGE: Which expression is equivalent to $(x + 3)(x - 2)(x + 4)$?

- $x^3 - 2x + 24$ (1)
 - $x^3 + 5x - 24$ (2)
 - $x^3 + 9x^2 - 24$ (3)
 - $x^3 + 5x^2 - 2x - 24$ (4)
-



Q151 MATH KNOWLEDGE: Which of the following is a solution of the given equation?

$$4(m + 4)^2 - 4m^2 + 20 = 276$$

- 3 (1)
 - 4 (2)
 - 6 (3)
 - 12 (4)
-



Q152 MATH KNOWLEDGE: Which of the following could be the perimeter of a triangle with two sides that measure 13 and 5?

- 24.5 (1)
 - 26.5 (2)
 - 36 (3)
 - 37 (4)
-



Q153 MATH KNOWLEDGE: Which of the following is an equation of the line that passes through the points (4, -3) and (-2, 9) in the xy-plane?

- $y = -2x + 5$ (1)
 - $y = -\frac{1}{2}x - 1$ (2)
 - $y = \frac{1}{2}x - 5$ (3)
 - $y = 2x - 11$ (4)
-



Q154 MATH KNOWLEDGE: If $j = 4$, what is the value of $2(j - 4)4 - j + \frac{1}{2}j$?

- 0 (1)
 - 2 (2)
 - 2 (3)
 - 4 (4)
-



Q155 MATH KNOWLEDGE: Solve for y: $3y + 2x = 15z$

- $y = (3/15z) - 2x$ (1)
 - $y = (-2x + 15z)/3$ (2)
 - $y = -(2/3)x + 15z$ (3)
 - $y = -2x + 5z$ (4)
-



Q156 MATH KNOWLEDGE: What is the area of the shape?

- 6 mm² (1)
 - 16 mm² (2)
 - 64 mm² (3)
 - 128 mm² (4)
-



Q157 MATH KNOWLEDGE: Solve for x: $4x + 12 = x - 3$

- $x = -5$ (1)
 - $x = -3$ (2)
 - $x = 1.8$ (3)
 - $x = 5$ (4)
-



Q158 MATH KNOWLEDGE: Two spheres are tangent to each other. One has a volume of 36π , and the other has a volume of 288π . What is the greatest distance between a point on one of the spheres and a point on the other sphere?

- 6 (1)
 - 9 (2)
 - 18 (3)
 - 36 (4)
-



Q159 MATH KNOWLEDGE: A wedge from a cylindrical piece of cheese was cut as shown. If the entire wheel of cheese weighed 73 pounds before the wedge was removed, what is the approximate remaining weight of the cheese?

- 12.17 pounds (1)
 - 37.00 pounds (2)
 - 55.00 pounds (3)
 - 60.83 pounds (4)
-



Q160 MATH KNOWLEDGE: Which of the following is a solution to the inequality $2x + y \leq -10$?

- (0, 0) (1)
 - (10, 2) (2)
 - (10, 10) (3)
 - (-10, -10) (4)
-

Q161 MECHANICAL COMPREHENSION: What is the net force acting on the block below? (Assume positive is to your right)

- 100 N (1)
 - 0 N (2)
 - 700 N (3)
 - 120,000 N (4)
-



Q162 MECHANICAL COMPREHENSION: What is the potential energy of a person who weighs 150 N when she is on a stool 1 m in the air?

- 15 J (1)
 - 30 J (2)
 - 150 J (3)
 - 300 J (4)
-



Q163 MECHANICAL COMPREHENSION: The arrow labeled F in the diagram of a box on an include below represents

- normal force (1)
- gravity (2)
- friction (3)
- tension (4)



Q164 MECHANICAL COMPREHENSION: A machine performs 100 J of work in 10 s. What is the power used by the machine?

- 0 W (1)
- 10 W (2)
- 100 W (3)
- 1,000 W (4)



Q165 MECHANICAL COMPREHENSION: What is the mechanical advantage of a wheel and axle if the axle radius (input) is 0.1 m and the wheel radius (output) is 0.5 m?

- 1/10 (1)
- $\frac{1}{5}$ (2)
- $\frac{1}{4}$ (3)
- $\frac{1}{2}$ (4)



Q166 ASSEMBLING OBJECTS: Given the following set of objects, please determine which answer choice shows how the objects will look once the parts are put together.

- (1)
- (2)
- (3)
- (4)

End of Block: ASVAB Hard

Start of Block: CogTest Preferences for Feedback

Cogprefint After participants completed the test, we calculated two pieces of information for each participant:

1. The number of questions they answered correctly.
2. Their rank relative to 9 other randomly chosen participants who already completed the same test. That is, we calculated how their performance compared to 9 other randomly chosen participants (first, second, third, ..., last).

After participants completed the test, we asked them about their preferences for receiving these two pieces of information. We will now explain to you exactly how we asked them about their preferences.

Page Break

Sliderint As you will see in the following pages, some of the questions we asked the previous study participants involved the possibility of "completing sliders." Before you move to those pages, we'd like to explain now what it means to complete a slider.

To complete a slider, you must scroll the slider bar to a specific position. The specific position is a randomly chosen number between 0 and 100, and is indicated with a slider. The slider is not completed successfully until its bar is scrolled exactly to the indicated position.

When completing sliders, previous study participants had as much time as they wanted to complete the sliders. But, they had to complete all sliders successfully in order to advance to the next page.

To get a feel for what it means to complete sliders, below are 2 sliders. Please scroll each of them to the indicated position. You will be able to advance to the next page only after you complete both sliders successfully.



Slider practice Please complete the following sliders.

0 10 20 30 40 50 60 70 80 90 100

Scroll the slider to the number 29 ()	
Scroll the slider to the number 82 ()	

Page Break

Q136 To measure how willing participants were to find out how they performed, we asked them a series of 3 questions.

In each question, they had to choose between two options:

- **Option A:** Yes, tell me my test score and how I ranked compared to the 9 other participants
- **Option B:** No, do NOT tell me my test score nor how I ranked compared to the 9 other participants

As you will see, the "cost" (in terms of required sliders) of choosing Option A or Option B varied across the three questions. Sometimes, participants had to complete sliders in order to choose one of the options.

End of Block: CogTest Preferences for Feedback

Start of Block: Eliciting Beliefs Cogtest Men

Q298 Your job now is to guess the answers given by MEN in the previous study.

In particular, we randomly selected 100 MEN from the previous study. You will guess how many of those 100 MEN chose to find out how they performed for each question below.

Recall that we will randomly select one of the guesses you make in this study and pay you an additional \$0.25 if you are within 5 percentage points of the correct answer.

Q299 First, participants were asked to choose between these options:

- **Option A:** Yes, tell me my test score and how I ranked compared to the other participants
- **Option B:** No, do NOT tell me my test score nor how I ranked compared to the other participants

Q300 YOUR GUESS: How many of the **100 MEN** chose Option A (to find out how they performed) in the question above?

0 10 20 30 40 50 60 70 80 90 100

How many **MEN** chose to find out how they performed? ()



Page Break

Q301 Next, we asked whether the previous study participants would be willing to complete 10 sliders in order to find out how they performed.

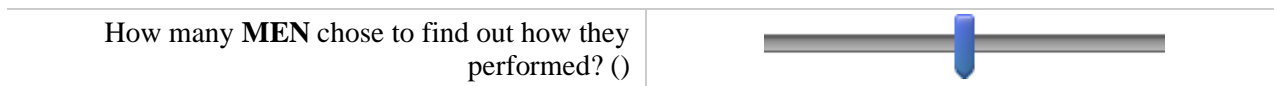
Participants were asked to choose between these options:

-- **Option A:** Yes, tell me my test score and how I ranked compared to the other participants AND I will complete 10 sliders

-- **Option B:** No, do NOT tell me my test score and how I ranked compared to the other participants

YOUR GUESS: How many of the **100 MEN** chose Option A (to find out how they performed) for this question.

0 10 20 30 40 50 60 70 80 90 100



Page Break

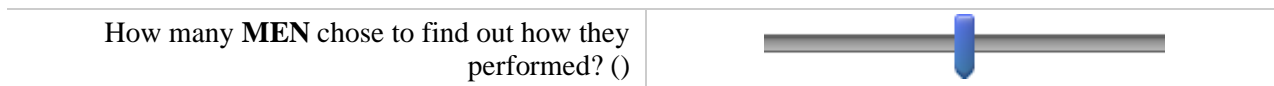
Q302 Finally, we asked whether the previous study participants would be willing to complete 10 sliders in order to **NOT** find out how they performed.

Participants were asked to choose between these options:

- **Option A:** Yes, tell me my test score and how I ranked compared to the other participants
- **Option B:** No, do NOT tell me my test score nor how I ranked compared to the other participants, AND I will complete 10 sliders

YOUR GUESS: How many of the **100 MEN** chose Option A (to find out how they performed) for this question.

0 10 20 30 40 50 60 70 80 90 100



End of Block: Eliciting Beliefs Cogtest Men

Start of Block: Eliciting Beliefs Cogtest Women

Q336 **Your job now is to guess the answers given by WOMEN in the previous study.**

In particular, we randomly selected 100 WOMEN from the previous study. You will guess how many of those 100 WOMEN chose to find out how they performed for each question below.

Recall that we will randomly select one of the guesses you make in this study and pay you an additional \$0.25 if you are within 5 percentage points of the correct answer.

Q337 First, participants were asked to choose between these options:

- **Option A:** Yes, tell me my test score and how I ranked compared to the other participants
- **Option B:** No, do NOT tell me my test score nor how I ranked compared to the other participants

Q338 **YOUR GUESS:** How many of the **100 WOMEN** chose Option A (to find out how they performed) in the question above?

0 10 20 30 40 50 60 70 80 90 100



Page Break

Q339 Next, we asked whether the previous study participants would be willing to complete 10 sliders in order to find out how they performed.

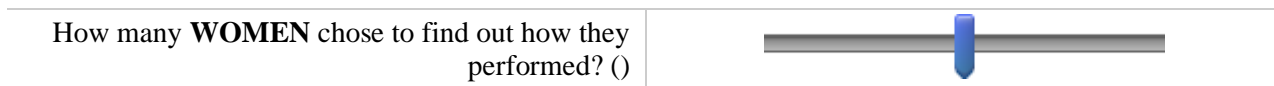
Participants were asked to choose between these options:

-- **Option A:** Yes, tell me my test score and how I ranked compared to the other participants AND I will complete 10 sliders

-- **Option B:** No, do NOT tell me my test score and how I ranked compared to the other participants

YOUR GUESS: How many of the **100 WOMEN** chose Option A (to find out how they performed) for this question.

0 10 20 30 40 50 60 70 80 90 100



Page Break

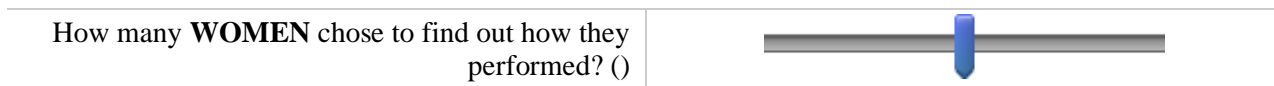
Q340 Finally, we asked whether the previous study participants would be willing to complete 10 sliders in order to **NOT** find out how they performed.

Participants were asked to choose between these options:

- **Option A:** Yes, tell me my test score and how I ranked compared to the other participants
- **Option B:** No, do NOT tell me my test score nor how I ranked compared to the other participants, AND I will complete 10 sliders

YOUR GUESS: How many of the **100 WOMEN** chose Option A (to find out how they performed) for this question.

0 10 20 30 40 50 60 70 80 90 100



End of Block: Eliciting Beliefs Cogtest Women

Start of Block: General questions

Task_qual Overall, for the task participants completed in the previous study, how would you describe differences in men's and women's preferences for finding out how they performed?

- Men want to find out how they performed much more than women (2)
 - Men want to find out how they performed somewhat more than women (3)
 - There is no gender difference in interest in finding out how they performed (4)
 - Women want to find out how they performed somewhat more than men (5)
 - Women want to find out how they performed much more than men (11)
-

General_qual Thinking more generally -- not just for the task participants completed in the previous study -- how would you describe gender differences in preferences for finding out how they performed in educational and professional settings, such as tasks in school and work?

- Men want to find out how they performed much more than women (2)
- Men want to find out how they performed somewhat more than women (3)
- There is no gender difference in interest in finding out how they performed (4)
- Women want to find out how they performed somewhat more than men (5)
- Women want to find out how they performed much more than men (11)

Page Break



prolific_id What is your Prolific ID?

Please note that this response should auto-fill with the correct id.

Thanks You've finished the study! Thank you!

End of Block: General questions
