HCEO WORKING PAPER SERIES

The Gender Gap in Self-Promotion

Christine L. Exley
Judd B. Kessler

Working Paper

2019-058

10/2019

HUMAN CAPITAL AND
ECONOMIC OPPORTUNITY
GLOBAL WORKING GROUP

The University of Chicago
1126 E. 59th Street Box 107
Chicago IL 60637

www.hceconomics.org

The Gender Gap in Self-Promotion

Christine L. Exley and Judd B. Kessler*

September 27, 2019

Abstract

In job applications, job interviews, performance reviews, and a wide range of other environments, individuals are explicitly asked or implicitly invited to assess their own performance. In a series of experiments, we find that women rate their performance less favorably than equally performing men. This gender gap in self-promotion is notably persistent. It stays just as strong when we eliminate gender differences in confidence about performance and when we eliminate strategic incentives to engage in self-promotion. Because of the prevalence of self-promotion opportunities, this self-promotion gap may contribute to the persistent gender gap in education and labor market outcomes.

*Exley: clexley@hbs.edu, Harvard Business School; Kessler: judd.kessler@wharton.upenn.edu, The Wharton School, University of Pennsylvania.

1 Introduction

At various points in their educational and professional lives — in college and professional school applications, in job applications and interviews, in performance reviews — individuals are explicitly asked to report on their ability and performance. In myriad other contexts, individuals face implicit invitations or opportunities to talk about their ability and performance. In response to these explicit and implicit opportunities, individuals convey a level of competence and success, which we call their level of "self-promotion."[1]

Focusing on a field that readers of this article may know well — the economics profession — it is clear that the level of self-promotion may play a role at a variety of points in the career of an economist. Even before the academic job market, self-promotion can affect how one writes a graduate school application, how one talks about his or her research ideas and technical skills, whether one gets the attention of desired advisors, and how one is perceived in internal seminars. During the academic job market, self-promotion can have direct implications for how a job market paper and a job talk are received and how much of joint work is attributed to a candidate rather than coauthors. Self-promotion can continue to have impacts throughout an academic career, as the way economists talk about their research — and their contribution to the research — can affect citations, prominence, as well as tenure and promotion decisions.

Given the pervasiveness of self-promotion in educational and labor market environments, one might be worried about the potential for a gender gap in self-promotion. If women describe their performance and ability less favorably than equally capable men do, a gender gap in self-promotion might have implications for observed gender gaps in educational and labor market environments. However, there is scant research on how individuals describe their own performance and ability in strategic settings — in which self-assessments are communicated to others and can influence labor market outcomes — or about how such self-promotion varies with gender.[2]

In this paper, we examine self-promotion by having experimental participants complete an analytical task: answering 20 questions from the Armed Services Vocational Aptitude Battery (ASVAB). Participants report beliefs about the number of questions they answered correctly on the ASVAB test (our measure of "confidence") and then respond to subjective, quantitative self-assessment questions about their performance (our measures of "self-promotion"). In our main study version, the *Public* version, participants are aware that one of their responses to a self-assessment question will be reported to a potential employer who will use that response — and only that re-

[1] In standard parlance, self-promotion often has a negative connotation. We do not imply one here and instead follow literature from psychology that refers to more positive self-assessments as engaging in more self-promotion. For instance, Rudman (1998) defines self-promotion as follows: "Designed to augment one's status and attractiveness, self-promotion includes pointing with pride to one's accomplishments, speaking directly about one's strengths and talents, and making internal rather than external attributions for achievements." There is a literature that suggests that extensive self-promotion can lead to backlash, which we discuss below.

[2] The lack of research is understandable given that self-promotion may be hard to measure and may not be easily observable. In contrast, there is a large literature that elicits beliefs without strategic motives (see, e.g., the discussion in Charness, Rustichini and Van de Ven (2018), which we discuss in more detail below).

sponse — to decide whether to hire them and how much to pay them. Data from employers confirm that self-promotion pays: more positive responses to these self-assessment questions increase the chance that participants are hired and the wages they are paid when hired.

We document a large and persistent gender gap in self-promotion. Despite no gender difference in performance on the analytical task, as documented in Section 3.1, women's responses to the self-assessment questions suggest they performed less well than the men. The effects are large. For example, when asked to indicate their agreement on a scale from 0 to 100 with a statement that reads "I performed well on the test," the average woman reports her performance as being 15 points lower out of 100 than the average man. That is, the average man rates himself a 61 out of 100 and the average women rates herself 25% lower, a 46 out of 100, despite the fact that the average performer in both groups answered 10 out of 20 ASVAB questions correctly. After documenting the gender gap in self-promotion in Section 3.2 — including the robustness of these results to several different types of self-assessment questions — the remainder of the paper examines underlying mechanisms of this gap and attempts to close it.

In Section 3.3, we show that the gender gap in self-promotion is *not* driven by a gender gap in confidence. Consistent with related literature discussed towards the end of our introduction, we observe a gender gap in confidence, as measured by participants' beliefs in the number of ASVAB questions they answered correctly. We show, however, that the self-promotion gap persists when we provide participants with *perfect information* about their absolute and relative performance.[3] We find that women still engage in less self-promotion than men, even when both are told that they answered the same number of questions correctly and both are told their exact place in the relative distribution (e.g., when both are told that they answered 15 out of 20 questions correctly and that their score was better than 80%, and worse than 12%, of prior participants).

In Section 3.4, we show that the gender gap in self-promotion is *not* driven by the strategic incentives associated with making statements to a potential employer. In the *Private* version of our study, participants answer the same self-assessment questions privately: there are no employers that participants have an incentive to impress, and how participants answer questions cannot influence their payments. While responses are less favorable in this version — evidence that participants are engaging in more self-promotion when incentivized to do so — we find a similarly sized gender gap. As in the *Public* version, the gender gap persists in the *Private* version even after we inform participants of their absolute and relative performance. Thus, the underlying gender gap in self-promotion is not caused by a gender gap in willingness to inflate self-assessments for personal gain.

In Section 3.5, we demonstrate the robustness of the gender gap in self-promotion to an additional study version, the *Ambiguous* version, which introduces some chance of "being caught" if participants inflate their self-assessments too much. This version is identical to our *Public* version except that participants are told that information about their actual performance (i.e., how many

[3] We are able to provide perfect information by focusing on self-promotion about *past* performance. In particular, we fully inform participants of their past performance and ask them about their past performance.

questions they answered correctly on the ASVAB) *could be* communicated to employers along with their response to a self-assessment question about their performance. The chance of true performance being communicated neither increases nor decreases the gender gap in self-promotion. We observe similar results both when participants are informed and are not informed of their absolute and relative performance.

Taken together, our results document a large gender gap in self-promotion that arises in unambiguous and ambiguous labor market environments. In addition, we rule out potential explanations for this gap. It is neither driven by a gender difference in confidence nor a gender difference in willingness to inflate self-assessments in the face of incentives to do so.[4] This leaves open many avenues for future research on what drives a gender difference in self-promotion, several of which we provide direct evidence for or against in Section 4.

Our results make contributions to three related literatures. First, motivated by the well-documented empirical evidence on the gender gaps in pay and in occupational and industry representation (Goldin, 2014; Blau and Kahn, 2017), there is a large literature in experimental economics that documents gender gaps in important labor market decisions. Relative to men, women are less willing to select a challenging task (Gneezy and Rustichini, 2003), to negotiate (Small et al., 2007), to enter a competition (Niederle and Vesterlund, 2007; Croson and Gneezy, 2009; Niederle and Vesterlund, 2011; Niederle, 2016), to contribute their ideas or assume a "leadership" position in a group (Coffman, 2014; Born, Ranehill and Sandberg, 2018; Coffman, Flikkema and Shurchkov, 2019), to claim credit for their contributions (Isaksson, 2018), and to apply for a job (Coffman, Collis and Kulkarni, 2019*b*).[5] We add self-promotion to the list of behaviors with gender gaps that might contribute to disparities in labor market outcomes. In addition to its direct effects, self-promotion may also interact with other labor market decisions, including those listed above. That is, self-promotion can affect how well one performs conditional on entering a negotiation, whether one is hired conditional on applying for a job, and whether one is selected to represent their group conditional on speaking up.[6]

Second, our results contribute to the literature on the gender gap in confidence in male-typed domains (Lundeberg, Fox and Punćcohaŕ, 1994; Niederle and Vesterlund, 2007; Coffman, 2014;

[4]In addition, in Sections 4.1 through 4.5, we provide evidence against other potential causes of the gender gap in self-promotion. We provide evidence against men and women having: different standards for good work (see Section 4.1), different beliefs about how employers respond to self-promotion (see Section 4.4), and different beliefs about how much they deserve to earn from the study (see Section 4.5). We also provide evidence against our results being driven by differences in aversion to lying (see Section 4.2) or differences in expected backlash (see Section 4.3).

[5]Two literatures that can help to explain some of these gender gaps include the literature on gender differences in confidence, which we discuss below, and the literature on the gender differences in risk, as reviewed in Eckel and Grossman (2008) and Croson and Gneezy (2009).

[6]Some prior work has examined gender differences in how women communicate. Such communication may involve discussions of performance, such as via the evaluation of chat conversations in negotiations (Exley, Niederle and Vesterlund, Forthcoming) or in group decision-making contexts (Coffman, Flikkema and Shurchkov, 2019). Unlike our paper, this prior work generally finds no substantial differences in the communication strategies of men versus women, but these chats are neither limited to a description of own's performance nor easily quantifiable. Differences in communication style have been observed when a partner's gender is known, as in Huang and Low (2017).

Bordalo et al., 2018; Isaksson, 2018; Coffman, Collis and Kulkarni, 2019*b*) as well as the literature on potential reasons for its persistence, such as gender differences in how individuals update their beliefs (Ertac, 2011; Mobius et al., 2014; Buser, Gerhards and Van der Weele, 2018; Coutts, 2018).[7] Given the importance of gender stereotypes in explaining these gender differences (Coffman, 2014; Bordalo et al., 2018; Coffman, Flikkema and Shurchkov, 2019), we follow much of this literature by focusing on a male-typed domain to align with gender gaps in labor market outcomes that motivate our study. Our paper replicates the gender gap in confidence in a male-typed domain and then deviates from the extant literature in two important ways.

We differ from almost all of the prior literature by focusing on a *strategic setting*, in which beliefs are communicated to others who can affect participants' payoffs.[8] As in many environments of interest outside the lab, participants in the *Public* version of our study may want to inflate their self-assessments about how well they performed, since potential employers will use these self-assessments to determine whether to hire them and how much to pay them. We are aware of only two other papers that elicit beliefs about performance in a strategic setting.[9] Reuben, Sapienza and Zingales (2014) shows that, in a setting where gender is known, men more than women inflate estimates about how well they performed on a task when these estimates are provided to "employers" who may hire them for a future task. Charness, Rustichini and Van de Ven (2018) finds that men, but not women, inflate the likelihood that they scored in the top half of their group on a cognitive ability task when doing so might deter other participants from entering a competition with them.[10]

We also differ from the prior literature — including both Reuben, Sapienza and Zingales (2014) and Charness, Rustichini and Van de Ven (2018) — by focusing on *subjective beliefs* rather than the objective beliefs that are typically elicited to measure confidence. Unlike objective beliefs about how many questions a participant answered correctly (as in Reuben, Sapienza and Zingales (2014)) or whether the participant is in the top half of performers (as in Charness, Rustichini and Van de Ven (2018)), subjective beliefs cannot be right or wrong (i.e., there is no correct response when a participant is asked to indicate agreement with the statement: "I performed well on the test"). When individuals communicate about their performance and ability in practice, they regularly report subjective beliefs.[11] A better understanding of subjective beliefs is important not only because they are often conveyed (in lieu of, or in conjunction with, objective beliefs) to potential employers, colleagues, supervisors, and other evaluators, but also because they operate differently from objective beliefs. We observe a gender gap in self-promotion (i.e., subjective beliefs)

[7]For example, Coffman, Collis and Kulkarni (2019*b*) shows that increasing the informativeness of a signal about performance does not reduce the size of observed gender gaps in performance beliefs.

[8]Prior literature has typically elicited beliefs that are incentivized for accuracy but are not shared with others.

[9]Schwardmann and van der Weele (Forthcoming) examine a different form of strategic beliefs: they show that individuals are more likely to inflate their *privately* held beliefs about their own ability when they know their future payment will depend on convincing others of their high ability.

[10]In addition to this work, there is a related literature on deception. For papers that investigate gender differences in cheating, see, for examples, Dreber and Johannesson (2008) and Erat and Gneezy (2012). Adams, Kuhn and Waddell (2019) finds a stronger correlation between cheating and stated confidence among men relative to women.

[11]This distinction is related to work about verifiable versus unverifiable signals of support as in Kessler (2017).

that persists even when we fully eliminate the gender gap in confidence (i.e., objective beliefs) by informing participants of both their absolute and relative performance.

Third, our results contribute to psychological literature related to self-promotion. This literature largely focuses on gender-specific backlash in response to self-promotion — for an example in the context of negotiations, see Bowles, Babcock and Lai (2007); for a review paper, see Rudman and Phelan (2008) — and there is noticeably little work on gender differences in self-promotion itself.[12] Not only does our paper provide evidence of a gender gap in self-promotion per se, we show that this gap arises in a strategic environment *absent gender-specific backlash concerns*. Employers in our *Public* version do not observe participant gender and only learn how a participant answers one self-assessment question. Similarly, the potential additional information that employers learn in our *Ambiguous* version is restricted to a participant's actual performance. While gender-specific backlash is not relevant in our study, we speculate that the existence of gender-specific backlash outside of our studies could cause our gender gap in self-promotion if it has led women to internalize the risks of engaging in self-promotion. One related question for future work is how a gender difference in self-promotion might respond to the potential for gender-specific backlash, for example if information about gender were communicated to employers along with self-assessments.[13]

The rest of the paper proceeds as follows. Section 2 describes the design of our studies. Section 3 presents evidence of the gender gap in self-promotion and our main results on mechanisms and robustness. Section 4 discusses additional potential mechanisms and other directions for future research. Section 5 concludes.

2 Design and Data

In October 2018, we recruited 900 workers on Amazon Mechanical Turk (MTurk) to complete one of the three main versions of our study.[14] Each participant was randomized into one of our study versions and received a guaranteed $2 completion fee for the 20-minute study. In addition, one out of four parts in the study was randomly selected to determine a possible bonus payment for each participant. After participants completed all four parts of the study, they took a short

[12]Indeed, Moss-Racusin and Rudman (2010), which documents gender differences in self-promotion in a simulated job interview, writes: "Despite its relevance to career success, to our knowledge, researchers have yet to examine gender differences in self-promotion."

[13]See Alston (2019) for evidence on how women, when given the opportunity to hide their gender on a job application for a male-typed job, are keen to do so. If future work makes the gender of workers known, self-promotion could also be examined in contexts where gender discrimination has been documented. For reviews of the literature on discrimination, see Riach and Rich (2002) and Blau and Kahn (2017). For recent evidence on gender discrimination, see also Reuben, Sapienza and Zingales (2014); Milkman, Akinola and Chugh (2015); Baert, De Pauw and Deschacht (2016); Bohnet and Bazerman (2016); Sarsons (2017*a*,*b*); Bohren, Imas and Rosenberg (2019); Bohren et al. (2019); Coffman, Exley and Niederle (2019); Kessler, Low and Sullivan (2019). For evidence about how others respond differently to men versus women in the language they use, see Bohren, Imas and Rosenberg (2018).

[14]To be eligible, workers must have previously completed at least 100 tasks on MTurk, with a 95% or better approval rating from prior employers, and workers must be working from an United States IP address. In each of the three study versions, the median age is 34 years old, and the median educational attainment is a Bachelor's Degree. The percentage of male participants varies from 46% – 52% across the three versions of our study, but these differences are not statistically significant.

follow-up survey that collected demographic information.

Each of the three study versions are detailed in the following Sections 2.1 through 2.3. In Section 2.1, we describe the *Public* version of our study, upon which the other two versions are based. In Sections 2.2 and 2.3, we describe the *Private* and *Ambiguous* versions of our study, highlighting how they differ from the *Public* version. In Section 2.4, we describe — and present results from — an additional study version, involving 300 participants in the role of "employers," which are relevant for the *Public* and *Ambiguous* versions of our study.[15]

2.1 The *Public* Version

The *Public* version of our study has four parts, described in sequence as follows (and see Appendix B.1 for screenshots).

Part 1: ASVAB and Confidence

In Part 1 of the study, participants are asked to take a test comprising of 20 multiple choice questions from the Armed Services Vocational Aptitude Battery (ASVAB). They have up to 30 seconds to answer each question, and there are four questions each from the following five categories: General Science, Arithmetic Reasoning, Math Knowledge, Mechanical Comprehension, and Assembling Objects. Participants are informed that "In addition to being used by the military to determine which jobs armed service members are qualified for, performance on the ASVAB is often used as a measure of cognitive ability by academic researchers." If Part 1 is randomly selected for payment, a participant's bonus payment is equal to 5 cents times the number of ASVB questions answered correctly.

After participants answer the 20 ASVAB questions, and before they continue to Part 2, they are asked: "Out of the 20 questions on the test you took in part 1, how many questions do you think you answered correctly?" This question is not incentivized, and participants can select any number from 0 to 20.

Part 2: Uninformed Self-Promotion

In Part 2, participants are asked five questions about their performance on the ASVAB. Participants are told that if Part 2 is randomly selected for payment, one of the responses to one of the questions will be shared with another MTurk participant called their "employer." The employer will see the response to the randomly selected question — and only that question (i.e., not any of the other responses or any information about actual performance) — and will determine whether to hire them and how much to pay them if hired.

More specifically, if an employer chooses not to hire a participant, the participant will earn a bonus of 25 cents, and the employer will earn a bonus of 100 cents. If an employer chooses to hire a participant, the employer will choose a wage between 25 and 100 cents, which will be the bonus

[15] In addition, we use data from 100 participants from a prior study who completed the same ASVAB test described below, in order to provide information to study participants on their relative performance. We also analyze data from 399 MTurk workers who evaluated free-response comments generated by study participants (as described below). Including these 499 participants, this paper involves a total of 1,699 study participants.

for the participant. The employer's bonus payment will then equal: 100 cents minus the wage paid to the participant plus 5 cents times the number of questions the participant answered correctly on the ASVAB.[16]

To encourage participants to reflect on their performance, the first question in Part 2 is a free-response question that states: "Please describe how well you think you performed on the test that you took in part 1 and why." The remaining four questions are the subjective, quantitative self-assessment questions that we analyze for the remainder of the paper.[17]

The next two questions elicit self-assessments about how well the participants performed on the ASVAB test in more quantifiable ways. The first states: "Please indicate how well you think you performed on the test you took in part 1." It allows participants to select from one of the following six answers: terrible, very poor, neutral, good, very good, and excellent. The second elicits a more continuous measure, asking participants to indicate the extent to which they agree, on a scale from 0 (entirely disagree) to 100 (entirely agree), with the following statement: "I performed well on the test I took in part 1."

The final two questions elicit self-assessments when there is also room for participants to hold preferences and beliefs about a related, hypothetical job. Using the same 0 to 100 scale described above, these questions ask participants to indicate the extent to which they agree with the following statements: "I would apply for a job that required me to perform well on the test I took in part 1" and "I would succeed in a job that required me to perform well on the test I took in part 1."

Part 3: Informed Self-Promotion

In Part 3, participants are asked precisely the same five questions about their performance on the ASVAB and participants are told that if Part 3 is randomly selected for payment, one of the answers to one of the questions will be shared with their employer.

Before answering these questions, however, participants learn precise information about their *absolute* and *relative* performance on the ASVAB test. In particular, participants are told exactly how many of the 20 questions they answered correctly (i.e., their absolute performance) and they are compared to 100 other participants who were asked the same questions as part of a prior study and told how many of those participants answered more questions correctly versus fewer questions correctly (i.e., their relative performance). As an attention check, participants must also correctly report how many of the 20 ASVAB questions they answered correctly before proceeding to answer the questions in Part 3. Note that providing relative performance information as well as absolute performance information serves to eliminate potential gender differences in beliefs about where a

[16]Note that employer earnings are based on the number of correct answers that the participant completed previously. This means that participants do not have to complete additional tasks and the decision environment avoids any potential uncertainty that might arise about future performance.

[17]The free-response question can also theoretically be interpreted as providing an opportunity for self-promotion. Analyzing this free-response data is fraught, however, as the text is hard to evaluate and can convey information such as gender and competence that makes measuring self-promotion per se difficult. Nevertheless, we attempt to learn what we can from this data by having 399 MTurk participants evaluate the responses, and we summarize those findings in Appendix A.4.

score falls in the performance distribution.

Part 4: Deservingness Question

In Part 4, participants are asked one question that measures deservingness for earnings from our experiment: "Out of a maximum amount of 100 cents, what amount of bonus payment, in cents, do you think you deserve for your performance on the test you took in part 1?" If this part is randomly selected as the part that counts, their bonus payment equals whatever amount they indicate from 0 to 100 cents. This question allows us to control for a potential gender difference in deservingness (i.e., how much participants believe they deserve to earn from the study) or in the desire to earn money from the experiment. As highlighted in Appendix Section 4.5, we find no statistically significant (nor economically meaningful) gender differences on reports of deservingness, which means gender differences observed in response to the self-assessment questions cannot be attributed to differential deservingness or desire to earn money from the study.

2.2 The *Private* Version

The *Private* version of our study proceeds exactly as the *Public* version except that participants provide their Part 2 and Part 3 self-assessments in a non-strategic, non-incentivized setting. In particular, there is no mention of any "employer," and participants are told that if Part 2 or Part 3 is randomly selected as the part-that-counts, their bonus payment will equal 25 cents regardless of how they answer the self-assessment questions.

2.3 The *Ambiguous* Version

The *Ambiguous* version of our study proceeds exactly as the *Public* version except that participants are told that there is *some chance* that their employers will be informed of how many questions they answered correctly on the ASVAB test. This chance is left ambiguous and meant to create a form of deterrence from overstating performance on the self-assessment questions. Since employers are free to choose any wage they want, participants may fear retribution through a low wage if they overstate how well they did in response to the self-assessment questions. Broadly, the *Ambiguous* version was designed to capture some of the ambiguity present in labor market settings, where applicants or employees are aware that signals about true performance may be available to employers and where individuals may fear retribution if an employer learns that an applicant or employee has misrepresented his or her performance.

2.4 The *Employers* Version

In November 2018, we recruited 300 workers on MTurk to complete the *Employers* version of our study using the same criteria as in the main study versions (see footnote 14). Each employer received a guaranteed $1.50 completion fee for the 15-minute study. In addition, two of their decisions, out of the 21 decisions in the study, was randomly selected to determine a possible bonus payment for them and for associated "workers," participants in the *Public* and *Ambiguous* study versions. After employers completed all decisions of the study, they took a short follow-up survey that collected demographic information.

For each decision, employers are informed that they must decide whether to hire a worker, and, if so, how much to pay that worker. Payoffs from hiring a worker are as described above. If an employer chooses not to hire a worker, the employer earns a bonus of 100 cents and the worker earns a bonus of 25 cents. If an employer chooses to hire a worker, the employer must also choose a wage between 25 and 100 cents. The worker will receive that wage, and that employer's bonus payment will equal 100 cents minus the wage paid to the worker plus 5 cents times the number of questions the worker answered correctly on the ASVAB test. The only information employers receive about a worker before hiring them is how the worker answered one of the four self-assessment questions.

Employers make hiring and wage decisions via the strategy method. That is, they respond to each of the six possible answers to the Likert-scale question and five randomly selected answers (i.e., numbers from 0 to 100) to the each of the three other questions.[18] Employers respond to each of these 21 answers to the self-assessment questions by deciding whether to hire a worker who gave this answer and how much to pay the worker if hired. Employers are told that two of the 21 answers will be selected for payment and that they will be paired with a worker who gave that answer. For these two answers, employers and paid based on their hiring and wage decisions and — if they chose to hire the worker who gave the answer — the worker's performance on the ASVAB test.

Table 1: *Employer Version*, Wage Regressions

Subjective Question:	Performance	Performance Bucket	Apply	Succeed
Answer	0.21***	4.26***	0.22***	0.21***
	(0.02)	(0.27)	(0.02)	(0.02)
Constant	22.70***	18.95***	21.94***	22.76***
	(0.75)	(0.70)	(0.61)	(0.78)
N	1490	1788	1490	1490

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. SEs are clustered by employer. Results are from OLS regressions of the wage received by the participant (25 cents if not hired and the chosen wage if hired). *Answer* is the answer provided by the participant in response to the Performance question in Column 1, the Performance-Bucket question in Column 2, the Apply question in Column 3, and the Succeed question in Column 4. In the Performance question, participants indicate the extent of their agreement (from 0 – 100) with the following statement: "I performed well on the test I took." In the Performance-Bucket question, participants select a bucket (coded from 1 for the lowest to 6 for the highest) when asked to "indicate how well you think you performed on the test." In the Apply question, participants indicate the extent their agreement (from 0 – 100) with the following statement: "I would apply for a job that required me to perform well on the test I took." In the Success question, participants indicate the extent of their agreement (from 0 – 100) with the following statement: "I would succeed in a job that required me to perform well on the test I took." Data are from the hiring decisions in the *Employer* version.

As expected, self-promotion pays. Employers respond to more positive self-assessments by being more likely to hire workers and by paying them more when hired. Table 1 shows how higher answers

[18] As noted above, these questions ask participants to state their agreement with the following statements: "I performed well on the test I took in part 1," "I would apply for a job that required me to perform well on the test I took in part 1," and "I would succeed in a job that required me to perform well on the test I took in part 1." Employers face all hiring decisions related to one self-assessment question before moving on to the next self-assessment question, but the order in which they face answers to each self-assessment question is randomized.

on each of the four self-assessment questions affects wages paid to workers. In all specifications, the coefficient on *Answer* is positive and significant. Columns (1), (3), and (4) show that employers pay on average 0.21 or 0.22 cents more for every point participants add to their self-assessment on the 100-point scale. Column (2) shows that employers increase wage paid by 4.26 cents for each increase on the six-point Likert scale. These results highlight that participants have an incentive to inflate their self-assessments to increase their expected study earnings.

3 Results

In this section, we present results from our three main study versions, the *Public*, *Private*, and *Ambiguous* versions. In Section 3.1, we show that there is no gender difference in performance on the ASVAB test and show that, nevertheless, we replicate a large gender gap in confidence (i.e., women report that they answered fewer questions correctly than equally performing men). In Section 3.2, we document a gender gap in the responses to the self-assessment questions elicited in Part 2 of the *Public* version of our experiment — which we call the gender gap in self-promotion — and show that it is present for all four of our self-assessment questions. In Section 3.3, we show that the self-promotion gap persists in Part 3 of the *Public* version of our study — after participants have been informed of their absolute and relative performance — demonstrating that the gender gap in self-promotion is not driven by the gender gap in confidence. In Section 3.4, we show that the gender gap in self-promotion persists in the *Private* version of our study, demonstrating that it is not driven by differences in willingness to inflate self assessments in the presence of incentives to do so. In Section 3.5, we show that the gender gap in self-promotion persists in the *Ambiguous* version of our study, suggesting it is robust to environments where performance information may become available to employers.

3.1 Performance and Confidence

In this section, we show that there is no gender difference in performance on the ASVAB test but that we nevertheless replicate a gender gap in confidence.[19] Since participants answer the 20 ASVAB questions and report their objective performance beliefs before they know which of the three study versions to which they have been randomized, we pool the results from the three study versions in presenting the results.[20]

Panel A of Figure 1 shows CDFs of the number of ASVAB questions answered correctly by male participants and by female participants. There is no gender gap in performance. On average, women answer 9.96 questions correctly and men answer 9.82 questions correctly. The difference is not statistically significant, and our data allow us to rule out small differences (the 95% confidence interval is a difference of -0.3 to 0.61 questions).[21]

[19] That we find a gender gap in confidence reassures us that our environment is not particularly different from the settings that researchers have explored in related work.

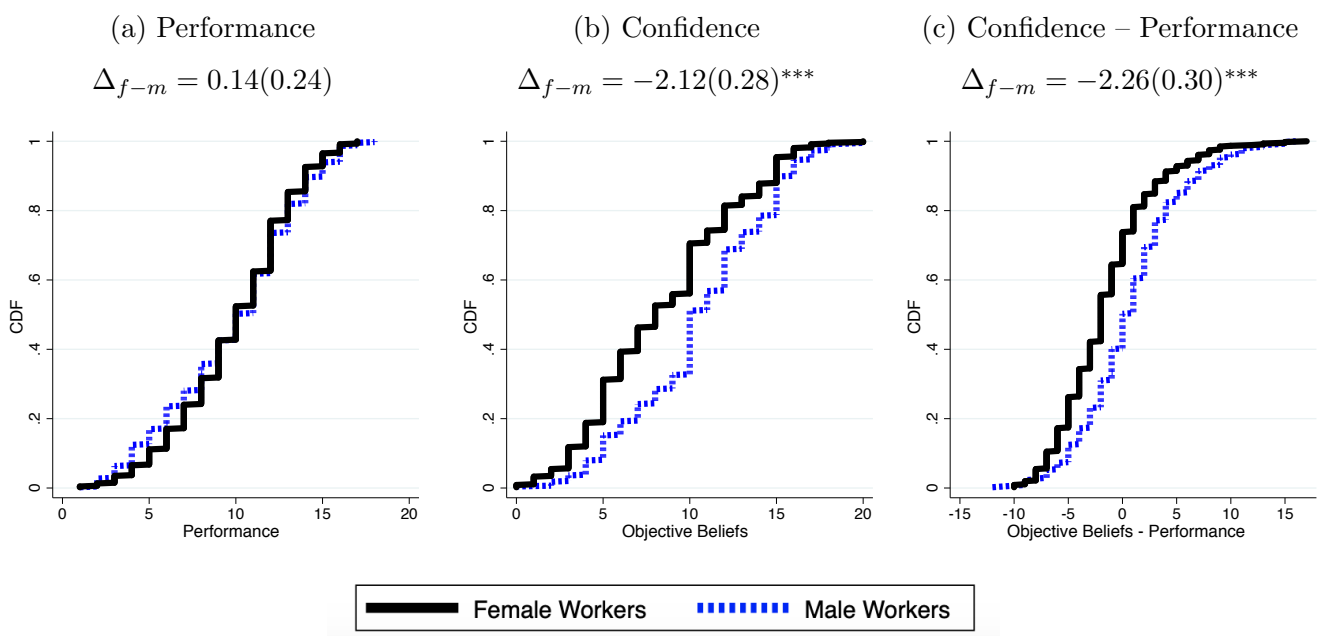
[20] As expected, Appendix A.1 shows that these results do not differ by study version.

[21] Looking beyond the means to the full distributions, a Kolmogorov–Smirnov test fails to reject the null of no difference between the CDFs ($p > 0.29$).

Despite the lack of a gender difference in performance, Panels B and C of Figure 1 replicate a common finding in the literature: women, relative to men, believe they performed worse on the test. Panel B shows raw performance beliefs (i.e., confidence). On average, men believe they answered 10.57 questions correctly while women believe they answered only 8.45 questions correctly. The mean difference is statistically significant ($p < 0.01$), and the distributions are statistically significantly different (a Kolmogorov–Smirnov test yields $p < 0.01$). Panel C shows the difference between confidence and performance and shows that the gender gap in confidence persists. Again, the mean difference is statistically significant ($p < 0.01$), and the distributions are statistically significantly different (a Kolmogorov–Smirnov test yields $p < 0.01$). Looking at where the CDFs cross 0, we see that the gender gap in confidence is driven both by more women than men underestimating their actual performance and more men than women overestimating their actual performance.

Table 2 presents the corresponding regression results. Column 1 confirms that there is no statistically significant gender difference in performance (the coefficient on *Female* is not different from 0), and the remaining columns confirm the statistically significant gender gaps in confidence, including when considering the raw data only (Column 2), when a linear term controls for performance (Column 3), when controlling for performance with dummies for each possible score (Column 4), and when the outcome variable directly captures the difference between confidence and performance (Column 5). In all cases, the coefficient on *Female* is negative, large, and statistically significant.

Figure 1: Performance and Confidence



Δ_{f-m} indicates the average gender difference of the associated outcome (i.e., the average outcome among women minus that among men). The associated robust standard error is shown in parentheses and the statistical significance is indicated as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table note applies to other figures that include Δ_{f-m} estimates except that standard errors are clustered at the participant level whenever possible.

Table 2: Performance and Confidence Regressions

DV:	*Performance*	*Confidence*			*Confidence − Performance*
	(1)	(2)	(3)	(4)	(5)
Female	0.14	-2.12***	-2.17***	-1.98***	-2.26***
	(0.24)	(0.28)	(0.26)	(0.26)	(0.30)
Performance			0.37***		
			(0.04)		
Constant	9.82***	10.57***	6.93***		0.75***
	(0.18)	(0.19)	(0.46)		(0.22)
N	900	900	900	900	900
Performance FEs	No	No	No	Yes	No

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The SEs are robust. Results are from OLS regressions of the noted dependent variable (DV). *Performance* equals the number of questions a participant correctly answered out of the 20 ASVAB questions. *Confidence* equals the number of questions a participant believes he or she correctly answered. *Female* is an indicator for the participant being female. Performance FEs are dummies for each possible performance on the ASVAB. Data are from all three study versions.

3.2 Is there a gender gap in self-promotion?

How do participants respond to the self-assessment questions? We turn first to the *Public* version of the experiment in which participants are told that one of their responses will be shared with a potential employer, and that this is all the employer will know when making a hiring and wage decision. Figure 2 shows responses to the four quantitative self-assessment questions in Part 2 of the *Public* version (i.e., before participants learn their absolute and relative performance in Part 3).

All four panels show large gender gaps in responses to the self-assessment questions. Women engage in less self-promotion than men. Panel A shows results from the question that asks participants to respond to the statement "I performed well on the test I took in part 1" on a scale from 0 (entirely disagree) to 100 (entirely agree). Women provide statistically significantly lower responses ($p < 0.01$) and the distributions are statistically significantly different (a Kolmogorov–Smirnov test yields $p < 0.01$). We obtain similar results in Panel B for the six-point Likert scale question: "Please indicate how well you think you performed on the test you took in part 1" ($p < 0.01$ for the t-test; $p < 0.01$ for the Kolmogorov–Smirnov test). Panels C and D show results from the self-assessment questions that allow participants to hold preferences and beliefs about a related, hypothetical job. Participants respond to the statements "I would apply for a job that required me to perform well on the test I took in part 1" (Panel C) and "I would succeed in a job that required me to perform well on the test I took in part 1" (Panel D) on a scale from 0 (entirely disagree) to 100 (entirely agree). We again see statistically significant differences in response to both questions ($p < 0.01$ for both t-tests; $p < 0.01$ for both Kolmogorov–Smirnov tests).

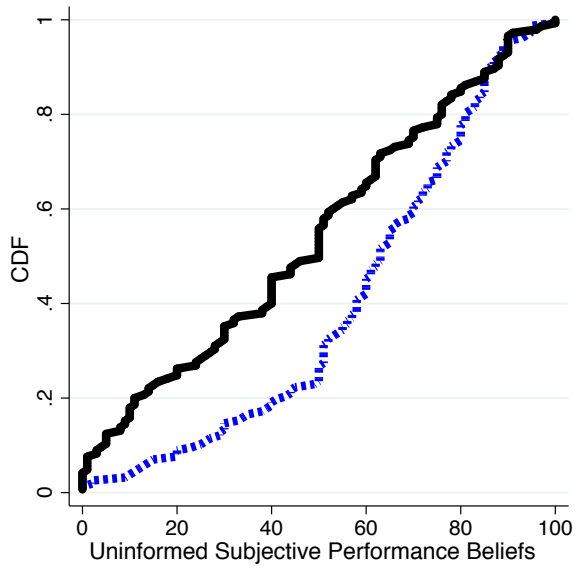
The first three columns of Table 3 confirm the statistical significance of these gender gaps, including when controlling for performance on the ASVAB. Column (1) presents results when considering the raw data only, column (2) adds a linear control for performance, and column (3) controls for

performance with dummies for each possible test score 0 to 20. As expected, the coefficient on *Female* remains negative, large, and statistically significant for all four self-assessment questions.

Figure 2: *Public Version*, Self-Promotion

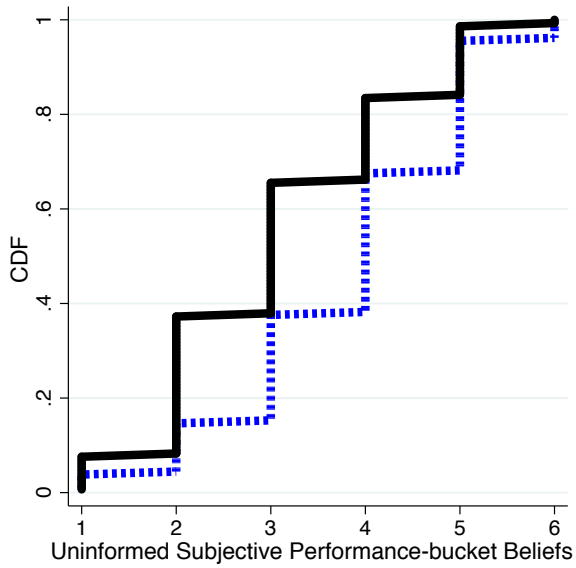
(a) Performance Beliefs

$\Delta_{f-m} = -14.77(3.07)^{***}$



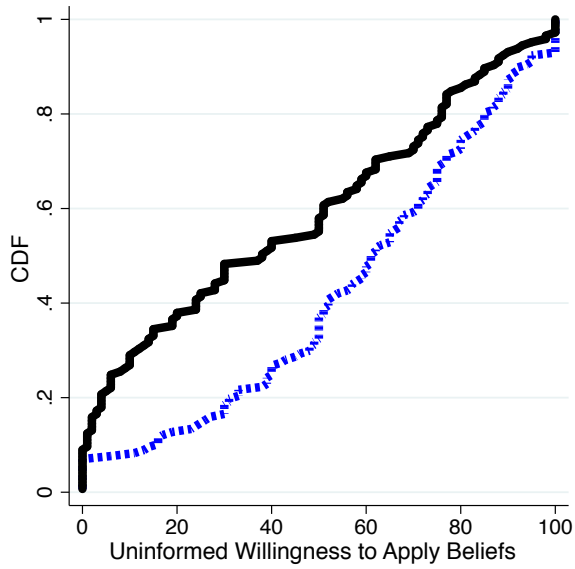
(b) Performance-Bucket Beliefs

$\Delta_{f-m} = -0.73(0.14)^{***}$



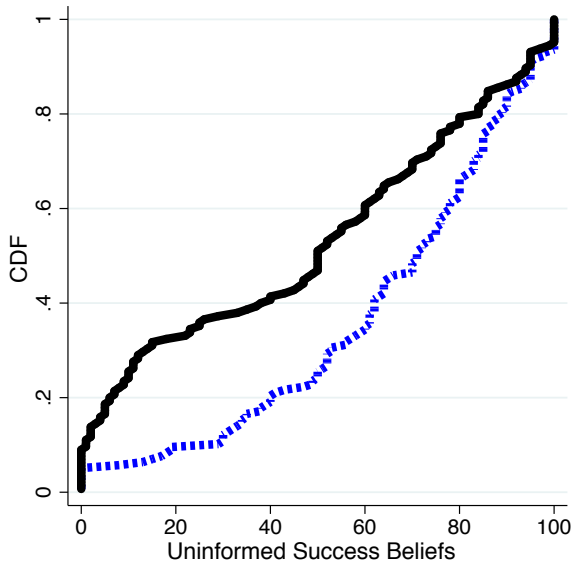
(c) Willingness to Apply Beliefs

$\Delta_{f-m} = -17.70(3.55)^{***}$



(d) Success Beliefs

$\Delta_{f-m} = -17.95(3.57)^{***}$



Female Workers Male Workers

Table 3: *Public Version*, Self-Promotion Regressions

Data:	Uninformed Self-Promotion			Informed Self-Promotion		
	(1)	(2)	(3)	(4)	(5)	(6)
DV = Performance Beliefs						
Female	-14.77***	-14.53***	-12.68***	-10.97***	-10.34***	-7.01**
	(3.07)	(3.03)	(2.96)	(3.27)	(3.05)	(2.90)
Performance		1.05**			2.80***	
		(0.42)			(0.43)	
Constant	60.52***	50.23***		59.43***	31.92***	
	(1.88)	(4.71)		(2.11)	(5.25)	
DV = Performance-Bucket Beliefs						
Female	-0.73***	-0.73***	-0.59***	-0.62***	-0.59***	-0.40***
	(0.14)	(0.14)	(0.13)	(0.15)	(0.14)	(0.13)
Performance		0.03			0.12***	
		(0.02)			(0.02)	
Constant	3.81***	3.55***		3.85***	2.67***	
	(0.10)	(0.24)		(0.11)	(0.27)	
DV = Willingness to Apply Beliefs						
Female	-17.70***	-17.34***	-15.31***	-15.55***	-14.92***	-10.73***
	(3.55)	(3.50)	(3.46)	(3.72)	(3.55)	(3.40)
Performance		1.59***			2.75***	
		(0.48)			(0.50)	
Constant	58.63***	43.00***		60.30***	33.29***	
	(2.27)	(5.38)		(2.44)	(5.93)	
DV = Success Beliefs						
Female	-17.95***	-17.58***	-15.09***	-16.19***	-15.44***	-11.73***
	(3.57)	(3.51)	(3.46)	(3.73)	(3.48)	(3.30)
Performance		1.63***			3.29***	
		(0.47)			(0.47)	
Constant	65.04***	48.97***		65.00***	32.60***	
	(2.15)	(5.29)		(2.33)	(5.69)	
N	302	302	302	302	302	302
Performance FEs	No	No	Yes	No	No	Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. SEs are clustered by participant. Results are from OLS regressions of the noted dependent variable (DV). *Performance Beliefs* indicates the extent of a participant's agreement (from 0 – 100) with the following statement: "I performed well on the test I took." *Performance-Bucket Beliefs* indicates which Likert-scale response (coded from 1 for the lowest to 6 for the highest) a participant selects when asked to "indicate how well you think you performed on the test." *Willingness to Apply Beliefs* indicates the extent of a participant's agreement (from 0 – 100) with the following statement: "I would apply for a job that required me to perform well on the test I took." *Success Beliefs* indicates the extent of a participant's agreement (from 0 – 100) with the following statement: "I would succeed in a job that required me to perform well on the test I took." *Female* is an indicator that the participant is female. *Performance* equals the number of questions a participant correctly answered out of the 20 ASVAB questions. Performance FEs are dummies for each possible performance on the ASVAB. Data in columns in (1) through (3) are from uninformed self-assessment questions elicited in Part 2 of the *Public* version and data in columns (4) through (6) are from informed self-assessment questions elicited in Part 3 of the *Public* version.

3.3 Is the gender gap in self-promotion driven by confidence?

In this section, we investigate whether the gender gap in self-promotion identified in Section 3.2 is driven by the gender gap in confidence (i.e., beliefs about performance) identified in Section 3.1.

We do this by analyzing the responses from Part 3 of our *Public* version, in which participants are informed about their absolute and relative performance on the ASVAB test and then answer the same self-assessment questions about their performance. Unlike settings in which men and women might respond differently to partially informative signals (e.g., because the signal is about past performance when the outcome of interest is about future performance), this information fully closes the gender gap in confidence as participants are perfectly informed about their past performance and are asked about their past performance. Moreover, to ensure participants attend to the performance information, they must accurately report back the number of questions they answered correctly on the ASVAB test before providing answers in Part 3. Any gender gap in self-promotion that persists in Part 3 cannot be driven by a gender gap in confidence. For a more detail discussion of the advantages to our approach of controlling for confidence by design, see Section 4.8.

Columns (4) through (6) of Table 3 report on data from Part 3 of the *Public* version. The first thing to note is that self assessments respond to the information. The coefficient on *Performance* in column (5) is always larger than the corresponding coefficient in column (2), demonstrating that subjective assessments are more responsive to performance when performance is known.[22]

Nevertheless, in columns (4) through (6), the coefficient on *Female* remains negative, large, and statistically significant. This finding shows that the gender gap in self-promotion is not driven by the gender gap in confidence.[23] Comparing the responses from Part 2 to Part 3, we see that the coefficients on *Female* are *directionally* smaller in Part 3, suggesting a potential role for confidence in explaining the magnitude of the initial gender gap in self-promotion. That said, fully informing participants of their performance does not statistically significantly alter the self-promotion gap: the coefficients on *Female* in Part 2 and Part 3 are never statistically significantly different.

3.4 Is the gender gap in self-promotion driven by incentives to inflate self-assessments?

In this section, we investigate data from the *Private* version of our study to test whether the gender gap in self-promotion persists in the absence of an incentive to make favorable subjective assessments to a potential employer. If the self-promotion gap were due to men inflating their assessments more than women in response to such incentives, the gender gap would shrink — or disappear — in the *Private* version.

Table 4 reports results from the *Private* version and shows that the gender gap in self-promotion remains strong and significant in the absence of incentives to inflate subjective assessments. Columns

[22] When we test for the statistical significance of this difference in regressions interacting *Performance* with participants being informed, we find interaction p-values < 0.01 for the two performance questions, an interaction p-value < 0.1 for the apply question, and an interaction p-value < 0.05 for the success question.

[23] Corresponding CDFs for each question, following the structure of Figure 2, are shown in Appendix Figure A.4.

Table 4: *Private Version*, Self-Promotion Regressions

Sample:	Uninformed Self-Promotion			Informed Self-Promotion		
	(1)	(2)	(3)	(4)	(5)	(6)
DV = Performance Beliefs						
Female	-15.80***	-15.72***	-13.46***	-10.96***	-10.79***	-8.01***
	(2.88)	(2.84)	(2.93)	(3.14)	(2.91)	(2.88)
Performance		1.31***			3.04***	
		(0.46)			(0.48)	
Constant	53.95***	41.06***		50.82***	20.91***	
	(1.99)	(5.15)		(2.26)	(5.66)	
DV = Performance-Bucket Beliefs						
Female	-0.68***	-0.68***	-0.56***	-0.44***	-0.44***	-0.33**
	(0.13)	(0.13)	(0.13)	(0.15)	(0.14)	(0.14)
Performance		0.04*			0.11***	
		(0.02)			(0.02)	
Constant	3.52***	3.15***		3.48***	2.38***	
	(0.09)	(0.25)		(0.11)	(0.28)	
DV = Willingness to Apply Beliefs						
Female	-19.26***	-19.19***	-17.57***	-16.33***	-16.20***	-13.25***
	(3.33)	(3.30)	(3.51)	(3.56)	(3.45)	(3.53)
Performance		1.18**			2.23***	
		(0.50)			(0.54)	
Constant	53.74***	42.14***		52.48***	30.51***	
	(2.41)	(5.57)		(2.56)	(6.12)	
DV = Success Beliefs						
Female	-18.60***	-18.52***	-16.46***	-16.13***	-15.98***	-13.15***
	(3.46)	(3.43)	(3.61)	(3.61)	(3.48)	(3.53)
Performance		1.30**			2.49***	
		(0.53)			(0.55)	
Constant	57.60***	44.75***		54.98***	30.44***	
	(2.44)	(5.96)		(2.55)	(6.24)	
N	304	304	304	304	304	304
Performance FEs	No	No	Yes	No	No	Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. SEs are clustered by participant. Results are from OLS regressions of the noted dependent variable (DV). *Performance Beliefs* indicates the extent of a participant's agreement (from 0 – 100) with the following statement: "I performed well on the test I took." *Performance-Bucket Beliefs* indicates which Likert-scale response (coded from 1 for the lowest to 6 for the highest) a participant selects when asked to "indicate how well you think you performed on the test." *Willingness to Apply Beliefs* indicates the extent of a participant's agreement (from 0 – 100) with the following statement: "I would apply for a job that required me to perform well on the test I took." *Success Beliefs* indicates the extent of a participant's agreement (from 0 – 100) with the following statement: "I would succeed in a job that required me to perform well on the test I took." *Female* is an indicator that the participant is female. *Performance* equals the number of questions a participant correctly answered out of the 20 ASVAB questions. Performance FEs are dummies for each possible performance on the ASVAB. Data in columns in (1) through (3) are from uninformed self-assessment questions elicited in Part 2 of the *Private* version and data in columns (4) through (6) are from informed self-assessment questions elicited in Part 3 of the *Private* version.

(1) through (3) report on data from Part 2 of the *Private* version, and columns (4) through (6) report on data from Part 3 of the *Private* version. First, we note that the constant terms in Table 4 are smaller then the corresponding constant terms in Table 3.[24] This pattern suggests that participants do inflate their subjective assessments in response to incentives (i.e., they engage in more self-promotion in the *Public* version than the *Private* version). However, the coefficients on *Female* in Table 4 are similar in size to the coefficients in Table 3, revealing that the gender gap in self-promotion is not driven by a differential response to incentives.

3.5 Does the gender gap in self-promotion persist in a more ambiguous environment?

In this section, we investigate data from the *Ambiguous* version of our study to explore whether the gender gap in self-promotion persists when participants are aware that information about their actual performance could be communicated to employers along with a response to a subjective question. Such ambiguity may be present in environments of interest outside the lab, where signals about performance and ability may be available when individuals engage in self-promotion.

Table 5 reports results from the *Ambiguous* version and shows that the gender gap in self-promotion remains strong and significant in the presence of such ambiguity. Following the structure of the other tables, columns (1) through (3) report on data from Part 2 of the *Ambiguous* version and columns (4) through (6) report on data from Part 3 of the *Ambiguous* version. The coefficients on *Female* in Table 5, while directionally smaller than the coefficients observed in Tables 3 and 4, is still negative, large, and statistically significant, suggesting that the gender gap in self-promotion is robust to the presence of ambiguity.

[24] In regressions when we include data from both the *Private* and *Public* versions (absent performance controls), the *Public* coefficient yields a p-value < 0.05 for all questions when participants are uninformed or informed (except for the uninformed apply question, which has $p = 0.14$).

Table 5: *Ambiguous Version*, Self-Promotion Regressions

Sample:	Uninformed Self-Promotion			Informed Self-Promotion		
	(1)	(2)	(3)	(4)	(5)	(6)
DV = Performance Beliefs						
Female	-9.79***	-10.67***	-9.15***	-7.76**	-9.41***	-7.24**
	(2.90)	(2.89)	(2.93)	(3.09)	(2.93)	(2.83)
Performance		1.24***			2.34***	
		(0.39)			(0.40)	
Constant	59.35***	47.23***		58.26***	35.40***	
	(1.96)	(4.30)		(2.18)	(4.94)	
DV = Performance-Bucket Beliefs						
Female	-0.52***	-0.54***	-0.47***	-0.42***	-0.47***	-0.36***
	(0.13)	(0.13)	(0.13)	(0.14)	(0.14)	(0.14)
Performance		0.02			0.07***	
		(0.02)			(0.02)	
Constant	3.76***	3.61***		3.77***	3.13***	
	(0.09)	(0.23)		(0.11)	(0.26)	
DV = Willingness to Apply Beliefs						
Female	-14.43***	-14.82***	-12.82***	-10.46***	-11.76***	-9.11***
	(3.28)	(3.29)	(3.29)	(3.44)	(3.35)	(3.38)
Performance		0.54			1.84***	
		(0.46)			(0.46)	
Constant	60.29***	54.96***		58.03***	40.07***	
	(2.15)	(5.10)		(2.33)	(5.34)	
DV = Success Beliefs						
Female	-10.70***	-11.32***	-9.24***	-9.60***	-10.85***	-8.07**
	(3.30)	(3.29)	(3.32)	(3.40)	(3.31)	(3.29)
Performance		0.88*			1.77***	
		(0.45)			(0.44)	
Constant	63.45***	54.85***		62.78***	45.47***	
	(2.18)	(5.11)		(2.30)	(5.29)	
N	294	294	294	294	294	294
Performance FEs	No	No	Yes	No	No	Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. SEs are clustered by participant. Results are from OLS regressions of the noted dependent variable (DV). *Performance Beliefs* indicates the extent of a participant's agreement (from 0 – 100) with the following statement: "I performed well on the test I took." *Performance-Bucket Beliefs* indicates which Likert-scale response (coded from 1 for the lowest to 6 for the highest) a participant selects when asked to "indicate how well you think you performed on the test." *Willingness to Apply Beliefs* indicates the extent of a participant's agreement (from 0 – 100) with the following statement: "I would apply for a job that required me to perform well on the test I took." *Success Beliefs* indicates the extent of a participant's agreement (from 0 – 100) with the following statement: "I would succeed in a job that required me to perform well on the test I took." *Female* is an indicator that the participant is female. *Performance* equals the number of questions a participant correctly answered out of the 20 ASVAB questions. Performance FEs are dummies for each possible performance on the ASVAB. Data in columns in (1) through (3) are from uninformed self-assessment questions elicited in Part 2 of the *Ambiguous* version and data in columns (4) through (6) are from informed self-assessment questions elicited in Part 3 of the *Ambiguous* version.

4 Discussion

In the prior sections, we ruled out two potential drivers of the gender gap in self-promotion — confidence and responses to strategic incentives — and demonstrated the robustness of the gap. This leaves open opportunities for future work to identify the underlying causes of this self-promotion gap. To help guide that future work, we provide additional results from our data. Sections 4.1 through 4.7 highlight potential mechanisms that might be driving the gender gap in self-promotion and provide evidence from our data suggesting why they are unlikely to be relevant or why they might be worth exploring in future work. Section 4.8 demonstrates why it was important that we controlled for confidence through our experimental design and considers the potential role of self-promotion in assessments of confidence, both of which may be relevant for future work.

4.1 Do women and men have different beliefs about what constitutes "good" work?

One potential reason a gender gap in self-promotion might arise is that men and women might have different standards for good work. Such a difference might manifest as women believing that scoring a 15 out of 20 (and being in the 88th percentile) is only "good" and worth a 70/100 on the 100-point scale while men believe such a score is "very good" and worth an 85/100. While we cannot assess differential views of what constitutes good work directly, we can compare how male and female *employers* respond to reports of self-promotion in their hiring and wage decisions. If men and women systematically differ on their views of self-assessments (e.g., women thinking that 70/100 maps to a score of 15, worthy of a wage of 50 cents; and men thinking that a 70/100 maps to a score of 10, worthy of a wage of 25 cents), we would expect such differences to affect how male and female employers use self-assessments in their hiring and wage decisions. Appendix Table A.7 shows that male and female employers do not respond differently to participants' responses to any of the four self-assessment questions in determining wages paid to workers (in either level or slope). Consequently, we think it is unlikely that the self-promotion gap observed in our study is due to gender differences in views about what constitutes good work.[25]

4.2 Can differences in aversion to lying explain our results?

Building off of the seminal paper on deception (Gneezy, 2005), Dreber and Johannesson (2008) shows that men are more likely than women to lie in order to achieve a small financial gain.[26] A natural question, then, is whether the gender gap in self-promotion relates to gender differences in

[25] It is worth noting, however, that using employer data to make such an assessment is potentially confounded by differences in prosocial preferences by gender, as male and female employers might differ in how much they wish to reward other experimental participants (conditional on the same expected performance). In this context, however, such gender differences in preferences would need to perfectly offset another gender difference in order for there to be no gender difference in the mapping of self-promotion to wages (in either level or slope).

[26] This gender difference has been replicated in other studies (see, e.g., Erat and Gneezy (2012) and Houser, Vetter and Winter (2012)) and has been shown to depend on: whether the lie is harmful or prosocial towards others (Erat and Gneezy, 2012), the size of the stakes (Childs, 2012), and culture (Gylfason, Arnardottir and Kristinsson, 2013).

lying. On one hand, we note that participants cannot explicitly lie in our study. There is no correct or true answer to the self-assessment questions because the questions are inherently subjective. On the other hand, it could be that purposefully inflating self-assessments in order to increase the expected financial gain from employers is similar to lying. To the extent that lying plays a role in self-assessments, we find evidence against this driving our results. As detailed in Section 3.4, both men and women inflate their self-assessments in the *Public* version relative to self-assessments in the *Private* version. However, men do *not* inflate their self-assessments more than women, and the gender gap in self-promotion is just as large in the *Private* version as in the *Public* version.

4.3 Can differential concerns about backlash explain our results?

In line with the literature on backlash mentioned in our Introduction, a gender gap in self-promotion could reflect women, relative to men, expecting more backlash from employers if they engage in self-promotion. This could arise due to men and women having different beliefs about backlash in general. It could also arise because of fears about gender-specific backlash: women facing more backlash *because* they are women. In our study, concerns about gender-specific backlash are absent because employers neither learn the gender — nor any other identifiable characteristics — of workers. Data from our *Ambiguous* version helps to speak about concerns about backlash more generally. In that version, employers might learn about actual performance and could then engage in backlash by choosing to "punish" workers who seem to have inflated their self-assessments. That the gender gap in self-promotion is similarly sized in the *Ambiguous* version as in the *Public* version suggests that differential beliefs about backlash are not driving the results of our study.

While our study eliminated the possibility of gender-specific backlash by design, as we note in the Introduction, gender-specific backlash could contribute to our results if it has led women — more so than men — to internalize the costs of engaging in self-promotion. Such internalization could potentially even affect self-promotion in the *Private* version of our study where no employers are present. We consequently view exploring how self-promotion responds to environment that makes backlash concerns more relevant as an interesting and important avenue for future work.

4.4 Do women expect self-promotion to be less financially beneficial?

Putting aside backlash concerns per se, a gender gap in self-promotion might still reflect men and women holding different beliefs about how employers will use their responses to determine pay. For example, Coffman, Collis and Kulkarni (2019*a*) shows that — even when controlling for beliefs about future performance on a task — women, relative to men, believe that an employer is less likely to promote them. Beliefs about the relationship between self-promotion and the resulting pay from employers, however, cannot explain why the gender gap in self-promotion persists in our *Private* version. In addition, that male and female *employers* do not respond differently to self-promotion, as detailed in Section 4.1, further casts doubt on the possibility for such differences in beliefs in our context more generally.

4.5 Do women believe they deserve (or want to earn) less money from the study?

One reason that women might engage in less self-promotion is that they do not believe they deserve to earn as much money from the study as the men do. First, we note that this channel could not explain the persistence of the gender gap in the *Private* version of our study, in which self-promotion cannot affect study earnings. Further evidence against this channel comes from Part 4 of the study, which directly asks participants how much money they believe they deserve for participating in the study in an incentivized way (i.e., if Part 4 is randomly selected for payment, participants earn the amount they state). Appendix Figure A.9 and Appendix Table A.5 show that there is no difference in how men and women respond to this question about how much they believe they deserve to earn from the study. Consequently, we think it is unlikely that the self-promotion gap is due to differences in deservingness or a desire for study earnings.

4.6 Are the results controlling for confidence driven by consistency?

When considering gender gaps in self-promotion that may arise in the labor market, consistency could play a role. Individuals may assess their performance initially and then have repeated opportunities to engage in self-promotion about it. Initial self-promotion — which could take place before participants get information about their performance — could then affect subsequent self-promotion. More generally, self-promotion at one point in time may influence self-promotion at a later point in time.

Participants in our study always answer the set of self-assessment questions first without information about their performance and then again after they were informed of their absolute and relative performance. Consequently, one could imagine that consistency plays a role in the persistence of the gender gap in self-promotion that we observe in our experiment after we have eliminated the role of the confidence gap. We designed our experiment in this way to allow for the potentially important role of consistency, which is unavoidable in labor market environments. Nevertheless, evidence for such consistency effects appears muted in our setting.

Results from Section 3.3 show that participants are statistically significantly more responsive to their performance after being informed of that performance, indicating that they are changing their self-promotion (in sensible ways) when informed. We can also assess the potential role for consistency in our data by looking at the fraction of participants who change their answers to the self-assessment questions between Part 2 and Part 3. As shown in Appendix Table A.6, when participants learn that their performance was different than the belief they reported in Part 1, 81% – 86% of participants change their answers to each of the three self-assessment questions on the 0 to 100 scale, and 49% of participants change their Likert-scale selection. There also appears to be substantial switching between Part 2 and Part 3 by the approximately 10% of participants who correctly guessed their true performance in Part 1. This data suggests a limited role for a consistency motive — at least a motive to stick to a specific report — in driving the self-promotion

gap after participants are informed of their performance. Future work could further explore the role for consistency in driving a self-promotion gap.

4.7 Does the gap differ when it may reflect more than beliefs about past performance?

We asked participants four self-assessment questions to measure self-promotion. As described in Section 2, we asked two questions solely about past performance and two questions about applying for and succeeding in a hypothetical job that was related to the ASVAB task. While the first two questions only pertain to beliefs about past performance, the latter two questions also pertain to preferences and beliefs about the hypothetical job. The latter two questions are therefore more similar in spirit to the approach in Coffman, Collis and Kulkarni (2019*a*), which investigates gender gaps in decisions about applying to jobs and thus gaps that make relevant preferences and beliefs over those jobs. A reasonable hypothesis is that the gender gap in self-promotion may differ — and perhaps be bigger — if women have a distaste for a job or lack confidence in their (not yet realized) future performance in the job.

As can be seen by comparing the size of the gender gap on the three questions we asked on the 0 to 100 scale, the gender gaps in the apply and succeed questions are directionally larger than than the gender gap in the performance question. It is never the case, however, that the size of the gender gap is statistically significantly different between the performance question and either of the other two questions. Thus, these results provide only suggestive evidence that the self-promotion gap may be larger when more than beliefs about past performance are relevant.

We see exploring the self-promotion gap in a variety of richer settings — that make relevant more than just beliefs about past performance — as a fruitful avenue for future work. However, we view it as a strength of our current paper to focus on beliefs about past performance, since this feature allows us to entirely eliminate the role of confidence (by providing perfect information on past performance).

4.8 Is there a relationship between self-promotion and confidence?

We show that confidence does not drive the self-promotion gap. We achieve this by fully informing individuals about their past performance and only asking them about their past performance, thereby eliminating the role of confidence through our experimental design. This contrasts with prior work that has been unable to (fully) eliminate the role of confidence on other behaviors because it provides only a noisy signal about performance, because information about past performance is provided when beliefs about future performance are relevant, or both. In addition, however, this prior work has typically controlled for confidence statistically — running regressions that use confidence as a linear or non-parametric control. We can conduct a similar exercise using only data from Part 1 and Part 2 of our study. The results of this exercise are instructive.

Appendix Table A.4 reports on the regressions described above for the *Public* and *Private* versions of the study, controlling linearly for confidence (i.e., objective performance beliefs) in

columns (2) and (5) and with dummies for each possible belief in columns (3) and (6). Controlling for confidence in this way decreases the coefficient on *Female* relative to the uncontrolled coefficients shown in columns (1) and (4). In fact, for the self-promotion measures of past performance, the gender gap is no longer significant. This finding highlights two issues that are worth considering in future work.

First, using confidence measures as a control in this way can lead to misleading estimates. If we had run these regressions, we would have misinterpreted the role of confidence in self-promotion by concluding that confidence substantially explains the self-promotion gap. Meanwhile, by controlling for confidence in our design, we observe that confidence does not explain the self-promotion gap. Such conflicting results should serve as a guide (and a warning) for future researchers.

Second, the results in Appendix Table A.4 highlight a correlation between confidence and self-promotion, even though our results in Section 3.3 show that the confidence gap is not causing the self-promotion gap. Such a correlation could arise due to improper measurement of confidence, due to the two variables being correlated with some omitted variable, or because the confidence gap is — at least in part — driven by the self-promotion gap (i.e., the causal relationship may go in the opposite direction). This relationship is worth exploring in future work.

5 Conclusion

We have documented a gender gap in self-promotion. When communicating to potential employers, women systematically provide less favorable assessments of their own past performance and potential future ability than equally performing men. The gap is not a function of the gender gap in confidence; we find that it persists when participants are perfectly informed of their absolute and relative performance on the relevant task. The gap is not driven by differential responses to an incentive to inflate self-assessments; we find that it is present, and that it is just as large, in a study version absent employers in which participants do not have an incentive to make favorable reports. In addition, we find that the gap is robust to an ambiguous environment where information about true performance — along with the self reports — might become known to employers.

We focus our work on self-promotion because we view it as an understudied behavior that could have important implications for labor market outcomes. Among other contexts, individuals are often explicitly invited to engage in self-promotion: in applications to educational institutions, in job applications, in job interviews, and in performance reviews. Many additional environments provide implicit opportunities to engage in self-promotion (e.g., when casually discussing work or work related issues with colleagues or superiors, when discussing private contribution to group work, and when advocating for oneself in the workplace).

The frequency of opportunities to engage in self-promotion means that it is has the potential to interact with other gender gaps that have been observed in the literature. For example, prior work has demonstrated a gender gap in a willingness to negotiate. Self-promotion can have direct consequences for whether negotiation is successful, so a gender gap in self-promotion might contribute

to differential outcomes for men and women who choose to negotiate. Similar arguments can be made for job applications (a gender gap in self-promotion might affect the likelihood a candidate will be hired), and whether one is selected to represent a group (a gender gap in self-promotion might affect how successful one is when they choose to speak up). Thus, we view future research on how self-promotion interacts with other labor market decisions as particularly promising.

In considering future work, we also note that two other avenues would be interesting to pursue. First, given the prevalence of self-promotion opportunities, how can the gender gap in self-promotion be mitigated? While our results cast doubt on approaches that focus on providing more information on performance or emphasizing the strategic incentives at play, they leave open the possibility that norms and views about what is "appropriate" may be able to mitigate the gender gap (He, Kang and Lacetera, 2019; Coffman, Collis and Kulkarni, 2019*a*).[27] A related issue is that the gender gap documented in our study may reflect women having internalized — more so than men — that self-promotion is inappropriate or that it is risky due to backlash concerns.

Second, do employers respond differently to the self-promotion of men and women in settings, unlike those examined in our studies, in which an employee's gender is known? On one hand, if employers expect that men, more than women, inflate their self-assessments, this may mitigate any gender gap in labor market outcomes that arises from a gender gap in self-promotion. However, Reuben, Sapienza and Zingales (2014) provides evidence against the empirical relevance of this possibility. That paper finds that men, more than women, tend to inflate their performance estimates, but it finds that employers do not (fully) account for this. On the other hand, the literature on backlash (discussed in the Introduction) makes clear the potential for greater backlash for women relative to men, which could exacerbate gender gaps in self-promotion and their associated impact on labor market outcomes.

[27] He, Kang and Lacetera (2019) shows that the gender gap in willingness to enter competition is eliminated when individuals must "opt-out" of a competition rather than "opt-in" to a competition. Coffman, Collis and Kulkarni (2019*a*) shows that the gender gap in willingness to apply to an advanced job is eliminated when individuals are provided with clear guidance as to the conditions under which they should apply (e.g., if they have scored above some threshold on a skills-assessment test). Both the "opt-out" frame of the first paper and the additional information in the second paper may succeed, in part, by clarifying the norms of the decision environment. That a reduction in ambiguity may mitigate gender gaps is also reminiscent of findings from the negotiations literature (Bowles and McGinn, 2008; Mazei et al., 2015; Leibbrandt and List, 2015).

References

Adams, Nathan R., Michael A. Kuhn, and Glen R. Waddell. 2019. "Confidence and Contrition: Is Cheating Internalized in Performance Assessments?" *Working Paper*.

Alston, Mackenzie Alston. 2019. "The (Perceived) Cost of Being Female: An Experimental Investigation of Strategic Responses to Discrimination." *Working Paper*.

Baert, Stijn, Ann-Sophie De Pauw, and Nick Deschacht. 2016. "Do employer preferences contribute to sticky floors?" *ILR Review*, 69(3): 714–736.

Blau, Francine D., and Lawrence M. Kahn. 2017. "The Gender Wage Gap: Extent, Trends. and Explanations." *Journal of Economic Literature*, 55(3).

Bohnet, Iris, Alexandra van Geen, and Max Bazerman. 2016. "When Performance Trumps Gender Bias: Joint Versus Separate Evaluation." *Management Science*, 62(5): 1225–1234.

Bohren, Aislinn, Alex Imas, and Michael Rosenberg. 2018. "The Language of Discrimination: Using Experimental versus Observational Data." *AEA Papers and Proceedings*, 108(169–74).

Bohren, J Aislinn, Alex Imas, and Michael Rosenberg. 2019. "The Dynamics of Discrimination: Theory and Evidence." *American Economic Review*.

Bohren, J. Aislinn, Kareem Haggag, Alex Imas, and Devin G. Pope. 2019. "Inaccurate Statistical Discrimination." *NBER Working Paper*.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer1. 2018. "Belief about Gender." *American Economic Review*.

Born, Andreas, Eva Ranehill, and Anna Sandberg. 2018. "A man's world? – The impact of a male dominated environment on female leadership." *Working Paper*.

Bowles, Hannah Riley, and Kathleen L. McGinn. 2008. "Gender in job negotiations: a two-level game." *Negotiation Journal*, 24(4): 393–410.

Bowles, Hannah Riley, Linda Babcock, and Lei Lai. 2007. "Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask." *Organizational Behavior and Human Decision Processes*, 103(1): 84–103.

Buser, Thomas, Leonie Gerhards, and Joël J Van der Weele. 2018. "Measuring responsiveness to feedback as a personal trait." *Journal of Risk and Uncertainty*, 56(2): 165–192.

Charness, Gary, Aldo Rustichini, and Jeroen Van de Ven. 2018. "Self-confidence and strategic behavior." *Experimental Economics*, 21(1): 72–98.

Childs, Jason. 2012. "Gender differences in lying." *Economics Letters*, 114(2): 147–149.

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

Coffman, Katherine B., Christine L. Exley, and Muriel Niederle. 2019. "When gender discrimination isn't about gender." *Working Paper.*

Coffman, Katherine B., Manuela R. Collis, and Leena Kulkarni. 2019*a*. "Do Clearly Stated Qualifications Help Attract Talented Female Applicants?" *Working Paper.*

Coffman, Katherine, Clio Bryant Flikkema, and Olga Shurchkov. 2019. "Gender Stereotypes in Deliberation and team decisions." *Working Paper.*

Coffman, Katherine, Manuela Collis, and Leena Kulkarni. 2019*b*. "Stereotypes and Belief Updating." *Working Paper.*

Coutts, Alexander. 2018. "Good news and bad news are still news: Experimental evidence on belief updating." *Experimental Economics*, 1–27.

Croson, Rachel, and Uri Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature*, 47(2): 448–474.

Dreber, Anna, and Magnus Johannesson. 2008. "Gender differences in deception." *Economics Letter*, 99(1).

Eckel, Catherine C., and Philip J. Grossman. 2008. "Men, Women and Risk Aversion: Experimental Evidence." In *Handbook of Experimental Economics Results*. 1061–1073.

Erat, Sanjiv, and Uri Gneezy. 2012. "White Lies." *Management Science*, 50(4): 723–733.

Ertac, Seda. 2011. "Does self-relevance affect information processing? Experimental evidence on the response to performance and non-performance feedback." *Journal of Economic Behavior & Organization*, 80(3): 532–545.

Exley, Christine L., Muriel Niederle, and Lise Vesterlund. Forthcoming. "Knowing When to Ask: The Cost of Leaning-in." *Journal of Political Economy.*

Gneezy, Uri. 2005. "Deception: The Role of Consequences." *The American Economic Review*, 95(1): pp. 384–394.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. 2003. "Performance in competitive environments: Gender differences." *The Quarterly Journal of Economics*, 118(3): 1049–1074.

Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review*, 104(4): 1091–1119.

Gylfason, Haukur Freyr, Audur Arna Arnardottir, and Kari Kristinsson. 2013. "More on gender differences in lying." *Economic Letters*, 119(1): 94–95.

He, Joyce, Sonia Kang, and Nicola Lacetera. 2019. "Leaning in or not leaning out? Opt-out choice framing attenuates gender differences in the decision to compete." *Working Paper.*

Houser, Daniel, Stefan Vetter, and Joachim Winter. 2012. "Fairness and cheating." *European Economic Review*, 56: 1645–1655.

Huang, Jennie, and Corinne Low. 2017. "Trumping Norms: Lab evidence on aggressive communication before and after the 2016 US presidential election." *American Economic Review, Papers & Proceedings.*

Isaksson, Siri. 2018. "It Takes Two: Gender Differences in Group Work." *Working Paper.*

Kessler, Judd B. 2017. "Announcements of support and public good provision." *American Economic Review*, 107(12): 3760–87.

Kessler, Judd B, Corinne Low, and Colin D Sullivan. 2019. "Incentivized Resume Rating: Eliciting Employer Preferences without Deception." *American Economic Review.*

Leibbrandt, Andreas, and John A. List. 2015. "Do women avoid salary negotiations? Evidence from a large-scale natural field experiment." *Management Science*, 61(9): 2016–2024.

Lundeberg, Mary A, Paul W Fox, and Judith Punćcohaŕ. 1994. "Highly confident but wrong: Gender differences and similarities in confidence judgments." *Journal of educational psychology*, 86(1).

Mazei, Jens, Joachim Hüffmeier, Philipp Alexander Freund, Alice F. Stuhlmacher, Lena Bilke, and Guido Hertel. 2015. "A meta-analysis on gender differences in negotiation outcomes and their moderators." *Psychological Bulletin*, 141(1): 85–104.

Milkman, Katherine L, Modupe Akinola, and Dolly Chugh. 2015. "What happens before? A field experiment exploring how pay and representation differentially shape bias on the pathway into organizations." *Journal of Applied Psychology*, 100(6).

Mobius, Markus M., Muriel Niederle, Paul Niehaus, and Tanya S. Rosenblat. 2014. "Managing Self-Confidence: Theory and Experimental Evidence." *Working Paper.*

Moss-Racusin, Corinne A, and Laurie A Rudman. 2010. "Disruptions in women's self-promotion: The backlash avoidance model." *Psychology of women quarterly*, 34(2): 186–202.

Niederle, Muriel. 2016. "Gender." In *Handbook of Experimental Economics.* Vol. 2, , ed. John Kagel and Alvin E. Roth, 481–553. Princeton University Press.

Niederle, Muriel, and Lise Vesterlund. 2007. "Do Women shy away from competition? Do men compete too much?" *Quarterly Journal of Economics*, 122(3): 1067–1101.

Niederle, Muriel, and Lise Vesterlund. 2011. "Gender and Competition." *Annual Review of Economics*, 3: 601–630.

Reuben, Ernesto, Paola Sapienza, and Luigi Zingales. 2014. "How stereotypes impair women's careers in science." *Proceedings of the National Academy of Sciences*, 111(12): 4403–4408.

Riach, P. A., and J. Rich. 2002. "Field Experiments of Discrimination in the Market Place." *The Economic Journal*, 112(483).

Rudman, Laurie A. 1998. "Self-promotion as a risk factor for women: the costs and benefits of counter-stereotypical impression management." *Journal of personality and social psychology*, 74(3).

Rudman, Laurie A, and Julie E Phelan. 2008. "Backlash effects for disconfirming gender stereotypes in organizations." *Research in organizational behavior*, 28(6-79).

Sarsons, Heather. 2017*a*. "Interpreting Signals in the Labor Market: Evidence from Medical Referrals." *Working Paper*.

Sarsons, Heather. 2017*b*. "Recognition for Group Work: Gender Differences in Academia†." *American Economic Review: Papers & Proceedings*, 107(5): 141–145.

Schwardmann, Peter, and Joël van der Weele. Forthcoming. "Deception and Self-Deception." *Nature Human Behavior*.

Small, Deborah A., Michele Gelfand, Linda Babcock, and Hilary Gettman. 2007. "Who goes to the bargaining table? The influence of gender and framing on the initiation of negotiation." *Journal of personality and social psychology*, 93(4): 583–599.

A Appendix

A.1 Performance and Confidence by Study Version

This section reports on performance and confidence data by study version. In all versions, there is no gender gap in performance but there is a gender gap in confidence.

Figure A.1: *Public Version*, Performance and Confidence

(a) Performance $\Delta_{f-m} = -0.23(0.42)$

(b) Confidence $\Delta_{f-m} = -2.67(0.48)^{***}$

(c) Confidence – Performance $\Delta_{f-m} = -2.44(0.54)^{***}$

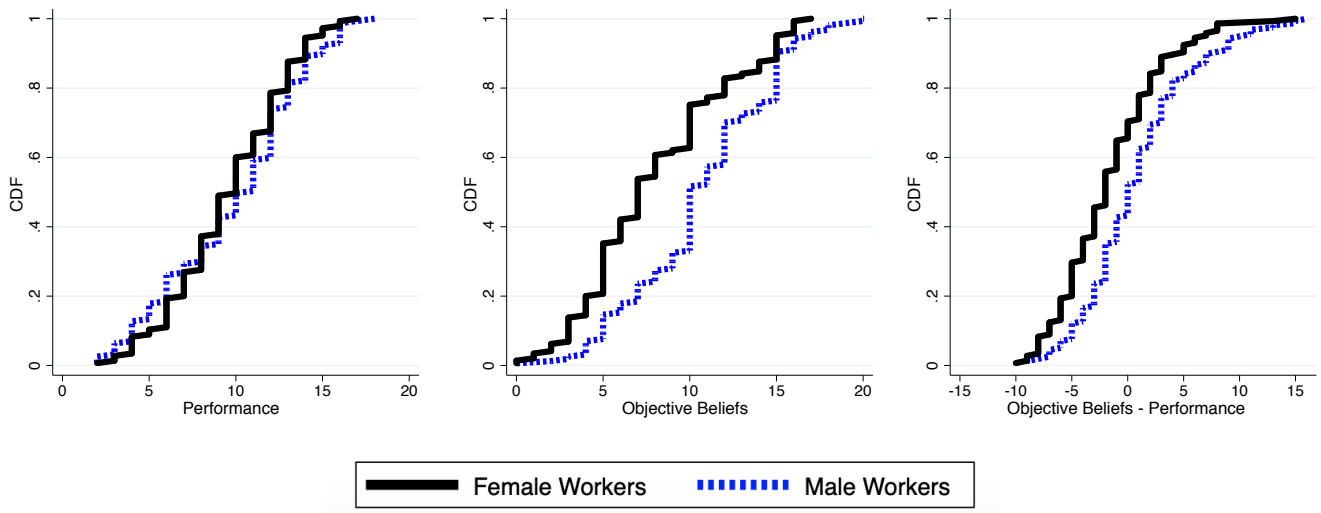


Table A.1: In the *Public Version*, Performance and Confidence Regressions

DV:	*Performance*	*Confidence*			*Confidence − Performance*
	(1)	(2)	(3)	(4)	(5)
Female	-0.23	-2.67***	-2.60***	-2.20***	-2.44***
	(0.42)	(0.48)	(0.46)	(0.45)	(0.54)
Performance			0.30***		
			(0.07)		
Constant	9.84***	10.67***	7.73***		0.83**
	(0.31)	(0.32)	(0.82)		
N	302	302	302	302	302
Performance FEs	No	No	No	Yes	No

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. SEs are clustered by participant. Results are from OLS regressions of the noted dependent variable (DV). *Performance* equals the number of questions a participant correctly answered out of the 20 ASVAB questions. *Confidence* equals the number of questions a participant believes he or she correctly answered. *Female* is an indicator for the participant being female. Performance FEs are dummies for each possible performance on the ASVAB. Data are from the *Public* version.

Figure A.2: *Private Version*, Performance and Confidence

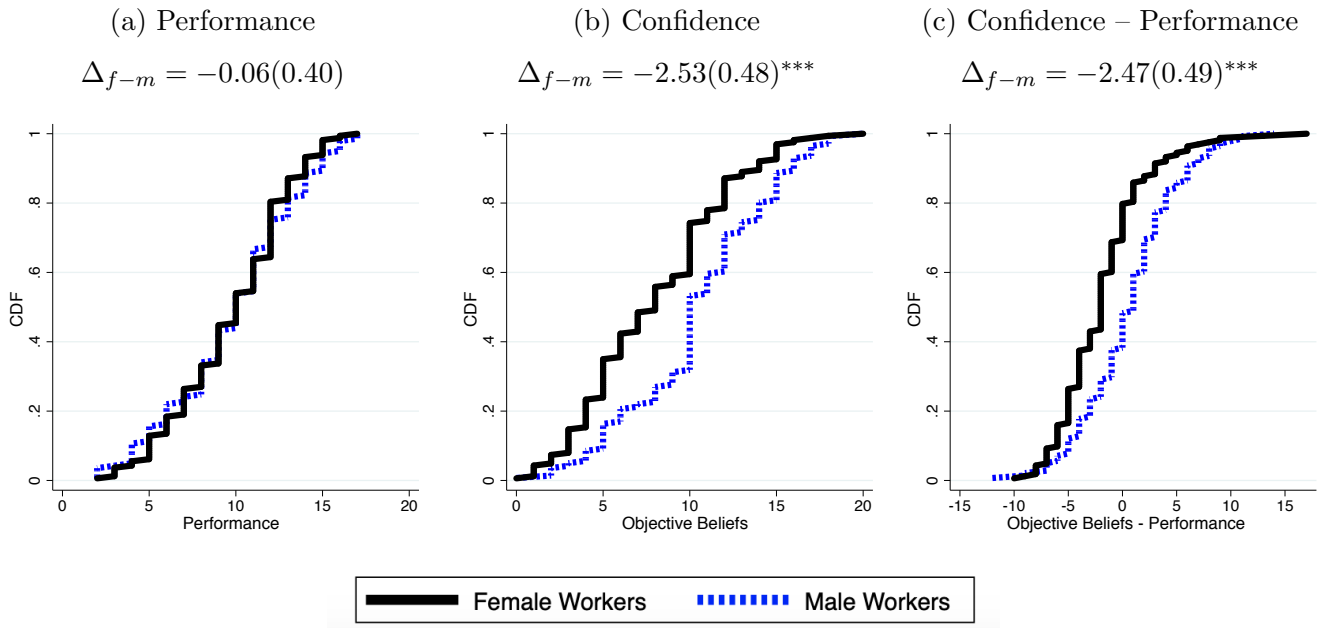


Table A.2: In the *Private Version*, Performance and Confidence Regressions

DV:	*Performance*	*Confidence*			*Confidence − Performance*
	(1)	(2)	(3)	(4)	(5)
Female	-0.06	-2.53***	-2.51***	-2.36***	-2.47***
	(0.40)	(0.48)	(0.44)	(0.46)	(0.49)
Performance			0.47***		
			(0.07)		
Constant	9.84***	10.50***	5.87***		0.65*
	(0.31)	(0.35)	(0.77)	(0.73)	
N	304	304	304	304	304
Performance FEs	No	No	No	Yes	No

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. SEs are clustered by participant. Results are from OLS regressions of the noted dependent variable (DV). *Performance* equals the number of questions a participant correctly answered out of the 20 ASVAB questions. *Confidence* equals the number of questions a participant believes he or she correctly answered. *Female* is an indicator for the participant being female. Performance FEs are dummies for each possible performance on the ASVAB. Data are from the *Private* version.

Figure A.3: *Ambiguous Version*, Performance and Confidence

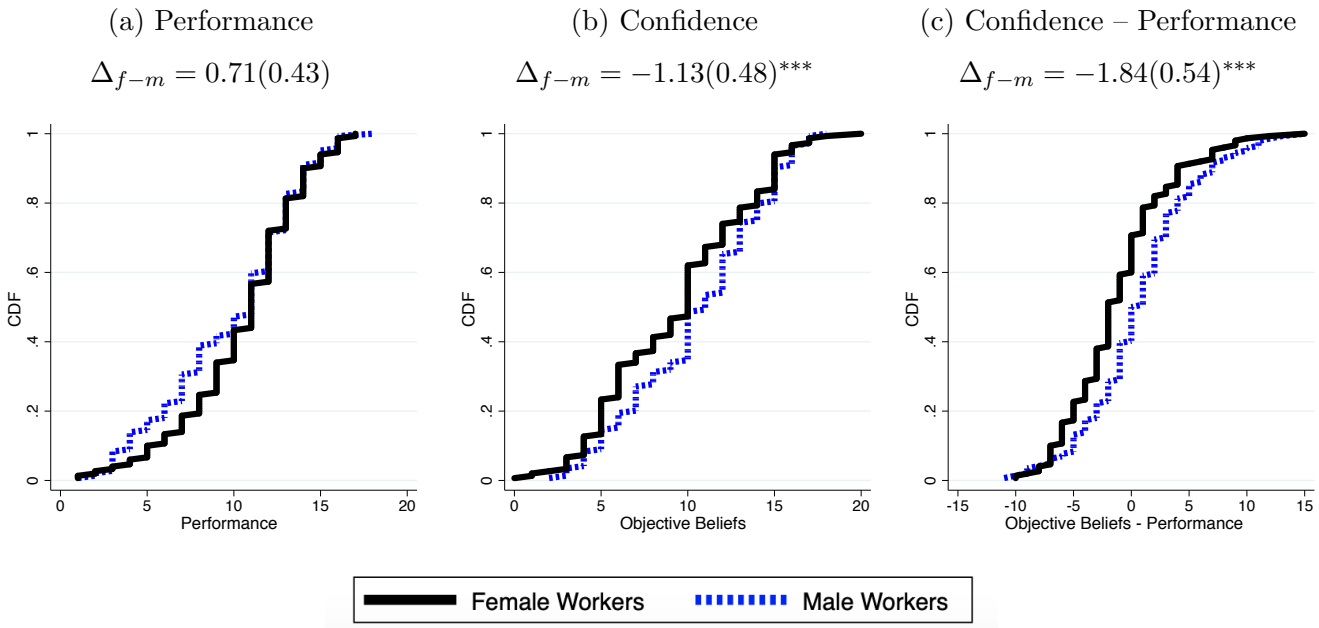


Table A.3: In the *Ambiguous Version*, Performance and Confidence Regressions

DV:	*Performance*	*Confidence*			*Confidence − Performance*
	(1)	(2)	(3)	(4)	(5)
Female	0.71	-1.13**	-1.36***	-1.06**	-1.84***
	(0.43)	(0.48)	(0.46)	(0.46)	(0.54)
Performance			0.33***		
			(0.07)		
Constant	9.78***	10.53***	7.32***		0.75*
	(0.33)	(0.33)	(0.78)		(0.40)
N	294	294	294	294	294
Performance FEs	No	No	No	Yes	No

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. SEs are clustered by participant. Results are from OLS regressions of the noted dependent variable (DV). *Performance* equals the number of questions a participant correctly answered out of the 20 ASVAB questions. *Confidence* equals the number of questions a participant believes he or she correctly answered. *Female* is an indicator for the participant being female. Performance FEs are dummies for each possible performance on the ASVAB. Data are from the *Ambiguous* version.

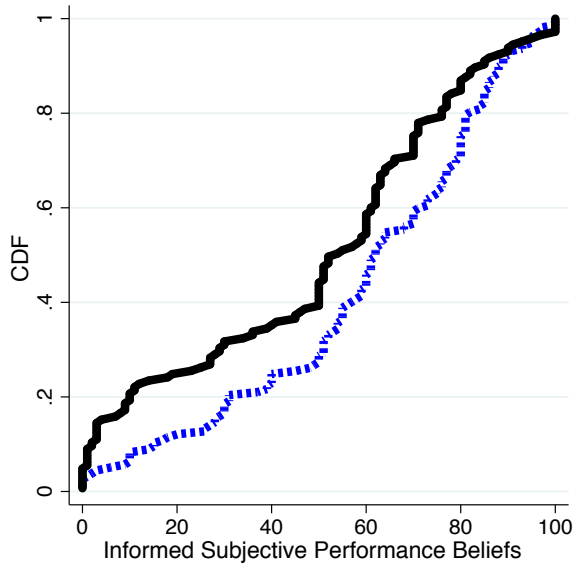
A.2 Additional Figures and Tables on Self-Promotion

A.2.1 Public Version

Figure A.4: *Public Version*, Informed Self-Promotion

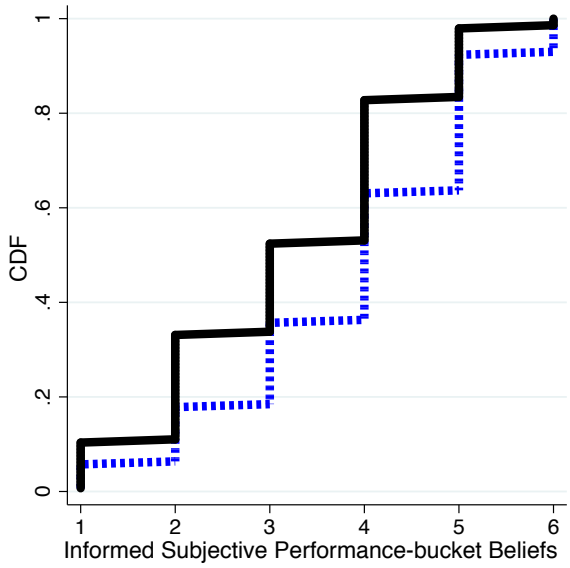
(a) Performance Beliefs

$\Delta_{f-m} = -10.97(3.27)^{***}$



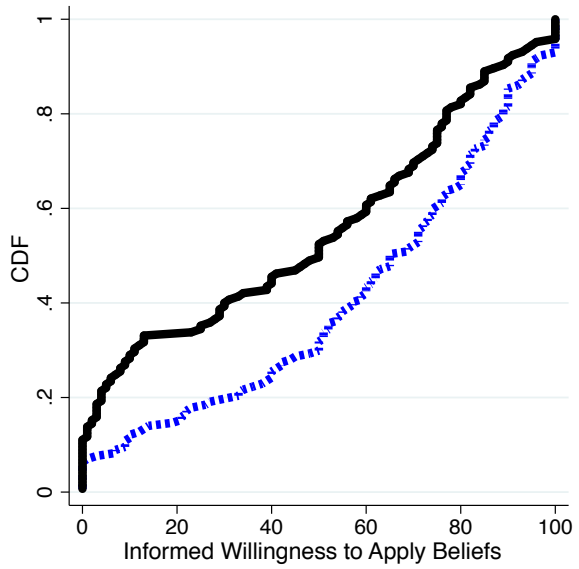
(b) Performance-Bucket Beliefs

$\Delta_{f-m} = -0.62(0.15)^{***}$



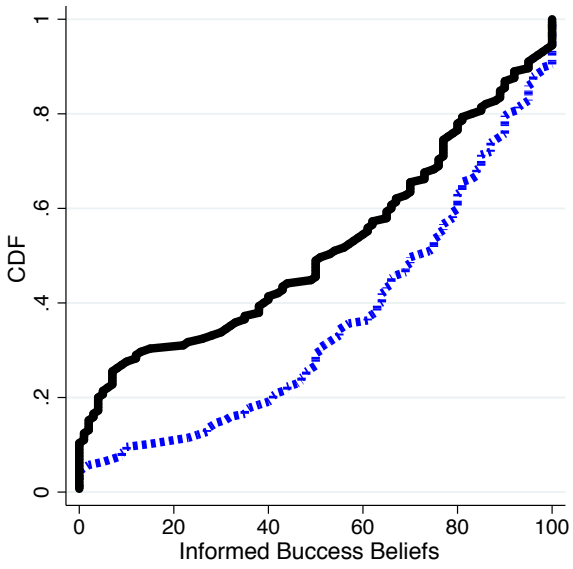
(c) Willingness to Apply Beliefs

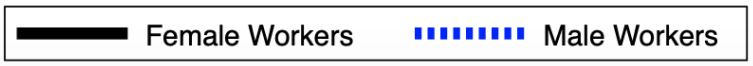
$\Delta_{f-m} = -15.55(3.72)^{***}$



(d) Success Beliefs

$\Delta_{f-m} = -16.19(3.73)^{***}$



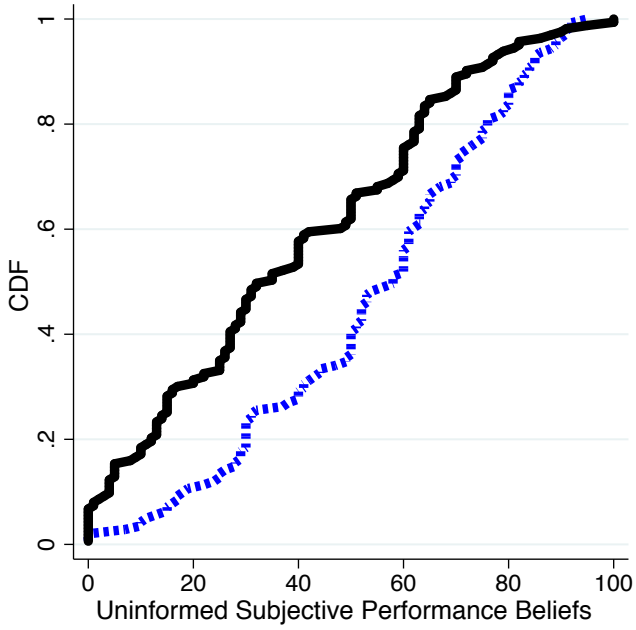


A.2.2 Private Version

Figure A.5: *Private Version*, Uninformed Self-Promotion

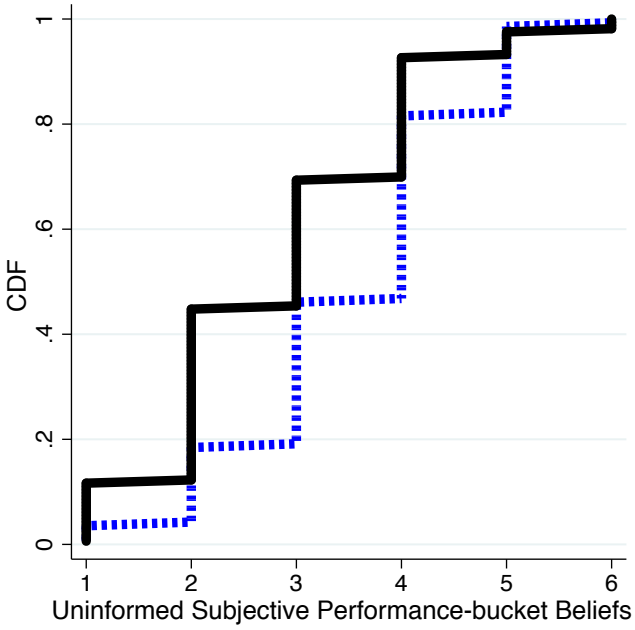
(a) Performance Beliefs

$\Delta_{f-m} = -15.80(2.88)^{***}$



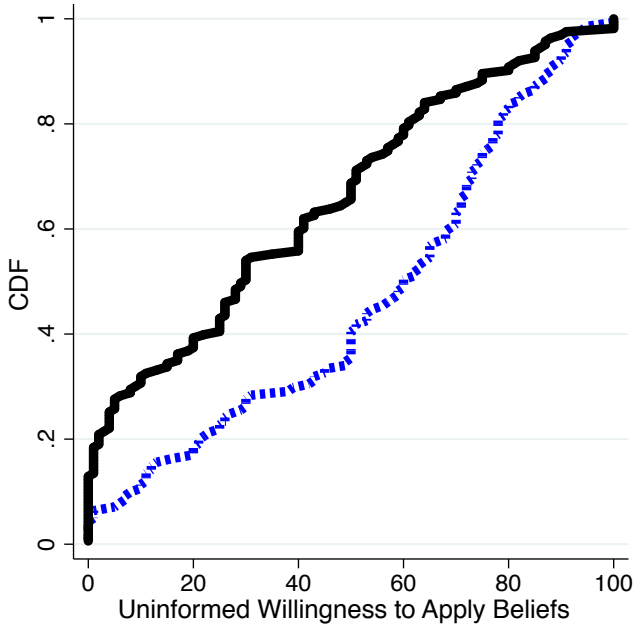
(b) Performance-Bucket Beliefs

$\Delta_{f-m} = -0.68(0.13)^{***}$



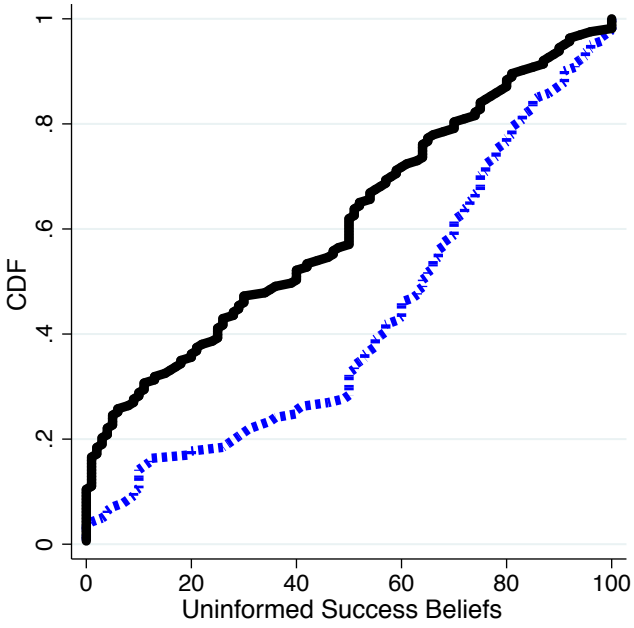
(c) Willingness to Apply Beliefs

$\Delta_{f-m} = -19.26(3.33)^{***}$



(d) Success Beliefs

$\Delta_{f-m} = -18.60(3.46)^{***}$

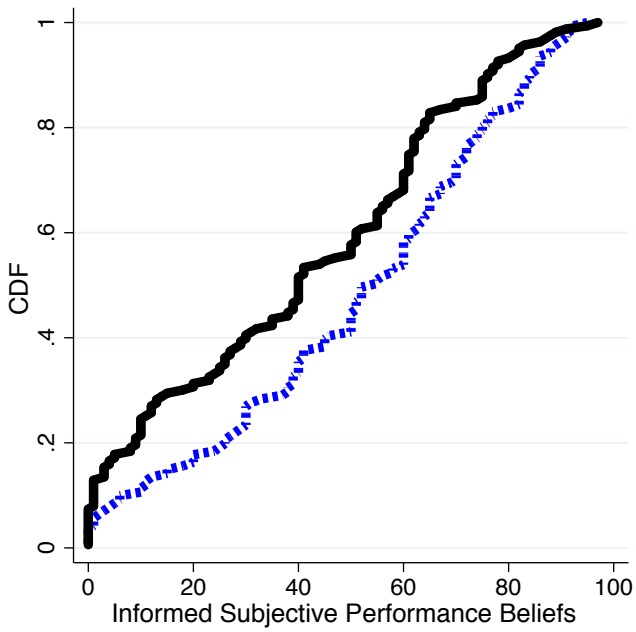


Female Workers Male Workers

Figure A.6: *Private Version*, Informed Self-Promotion

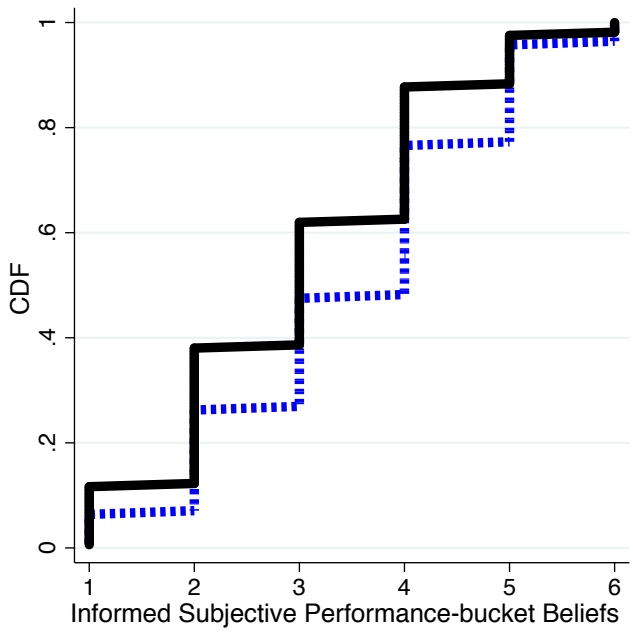
(a) Performance Beliefs

$\Delta_{f-m} = -10.96(3.14)^{***}$



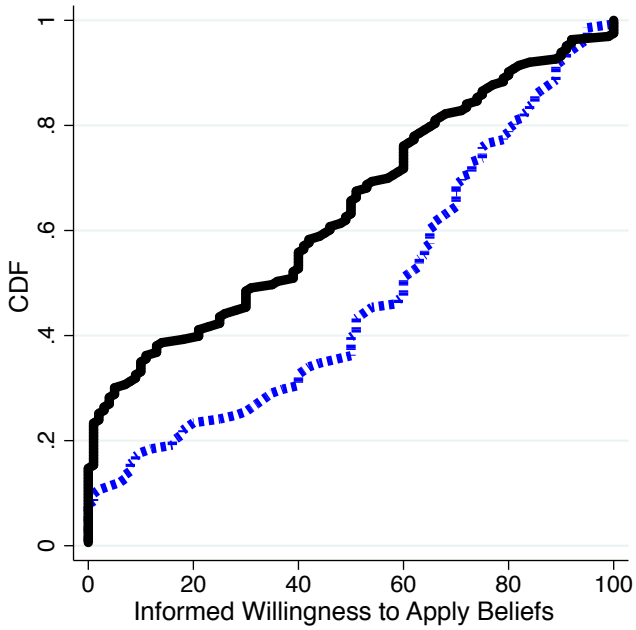
(b) Performance-Bucket Beliefs

$\Delta_{f-m} = -0.44(0.15)^{***}$



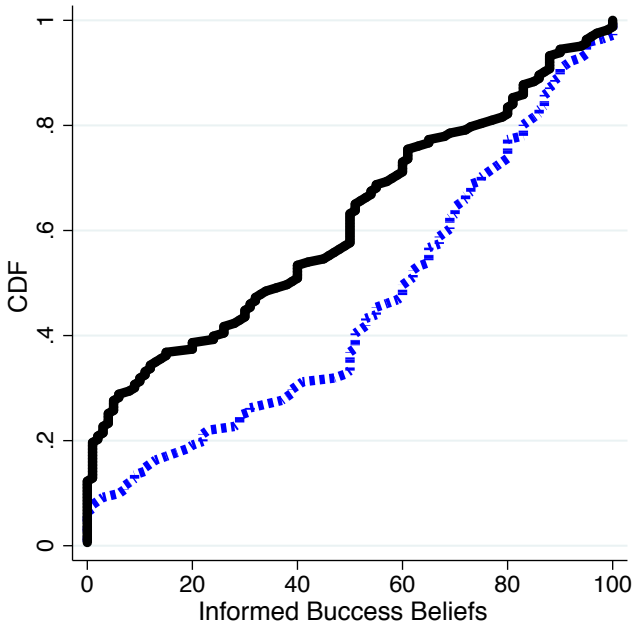
(c) Willingness to Apply Beliefs

$\Delta_{f-m} = -16.33(3.56)^{***}$



(d) Success Beliefs

$\Delta_{f-m} = -16.13(3.61)^{***}$



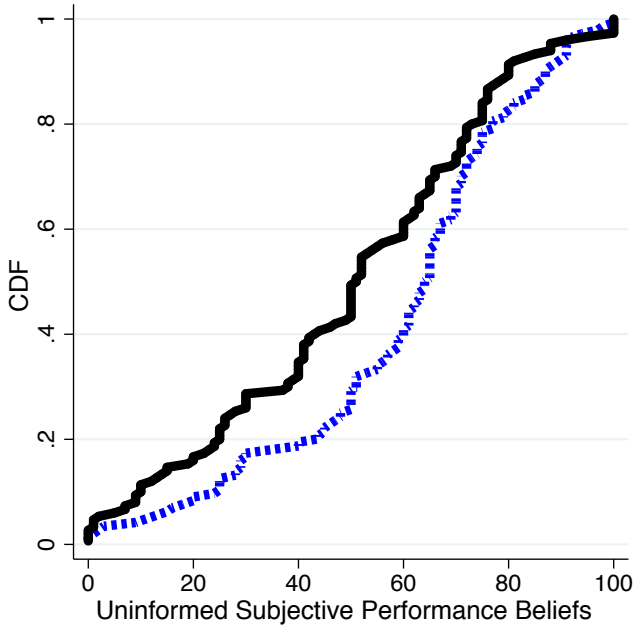
Female Workers Male Workers

A.2.3 Ambiguous Version

Figure A.7: *Ambiguous Version*, Uninformed Self-Promotion

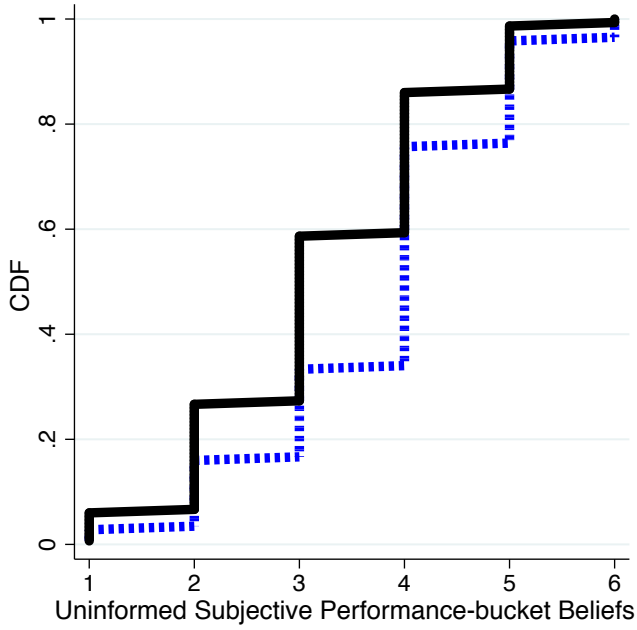
(a) Performance Beliefs

$\Delta_{f-m} = -9.79(2.90)^{***}$



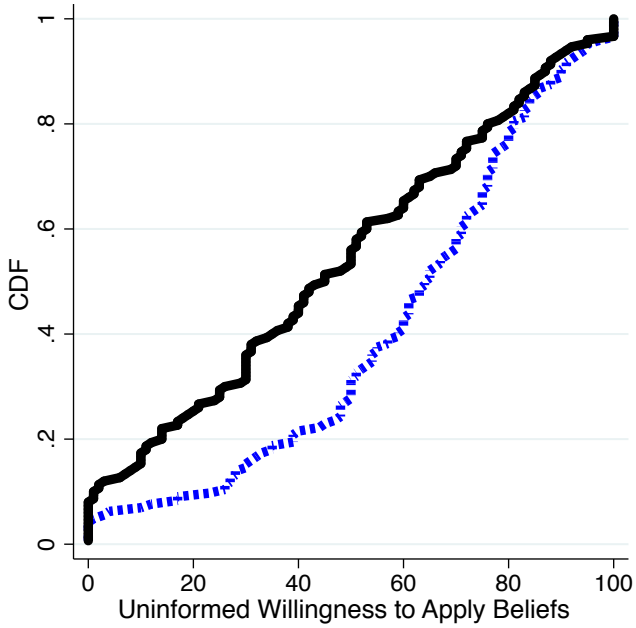
(b) Performance-Bucket Beliefs

$\Delta_{f-m} = -0.52(0.13)^{***}$



(c) Willingness to Apply Beliefs

$\Delta_{f-m} = -14.43(3.28)^{***}$



(d) Success Beliefs

$\Delta_{f-m} = -10.70(3.30)^{***}$

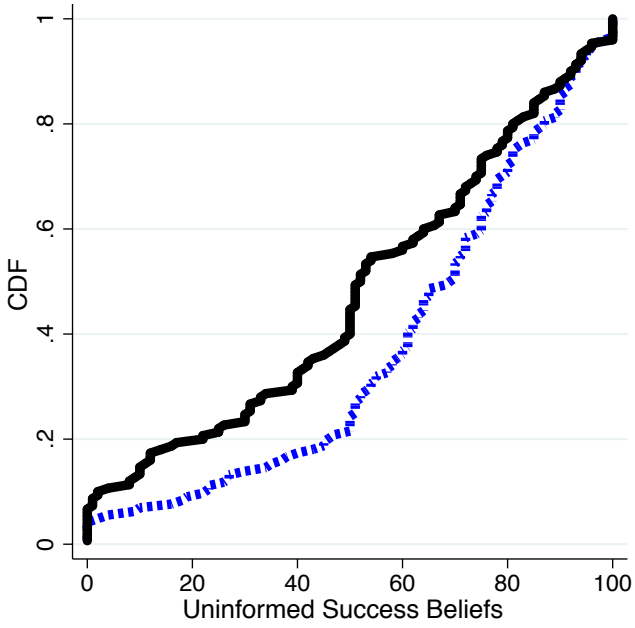
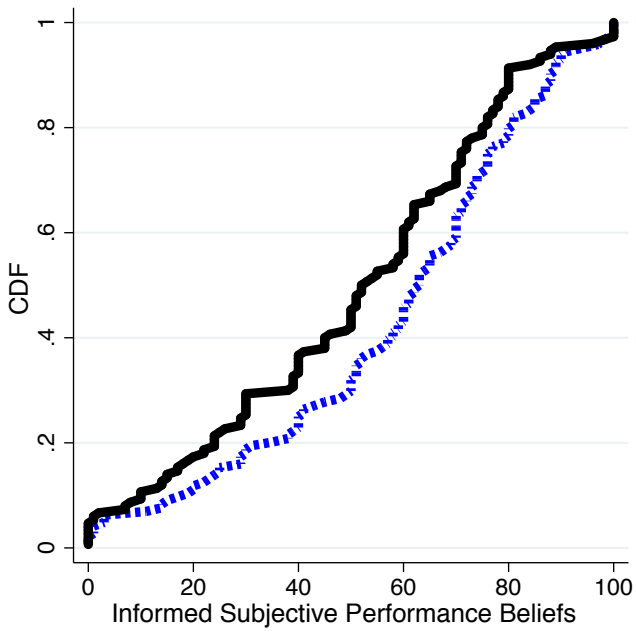


Figure A.8: *Ambiguous Version*, Informed Self-Promotion

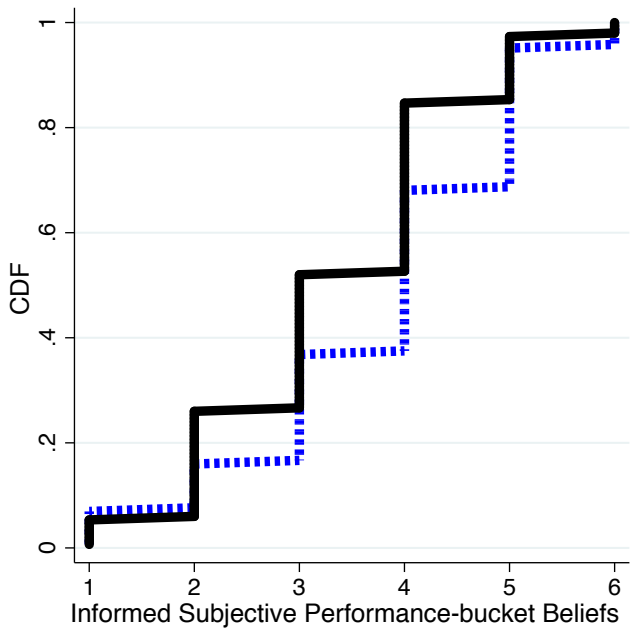
(a) Performance Beliefs

$\Delta_{f-m} = -7.76(3.09)^{**}$



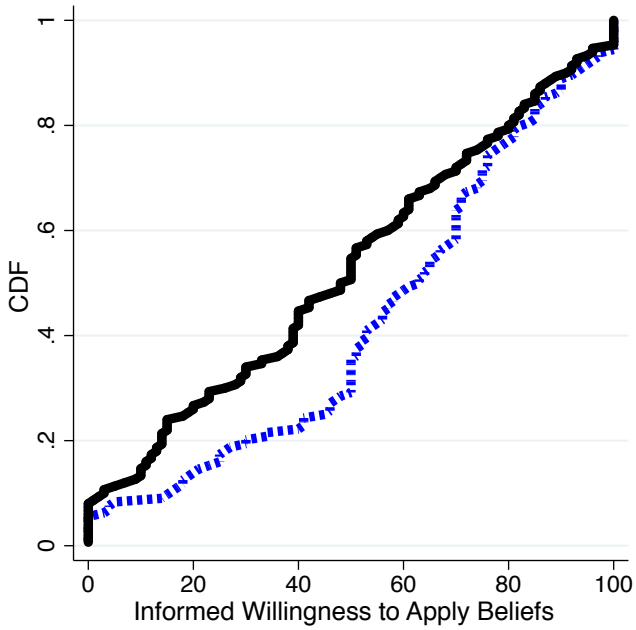
(b) Performance-Bucket Beliefs

$\Delta_{f-m} = -0.42(0.14)^{***}$



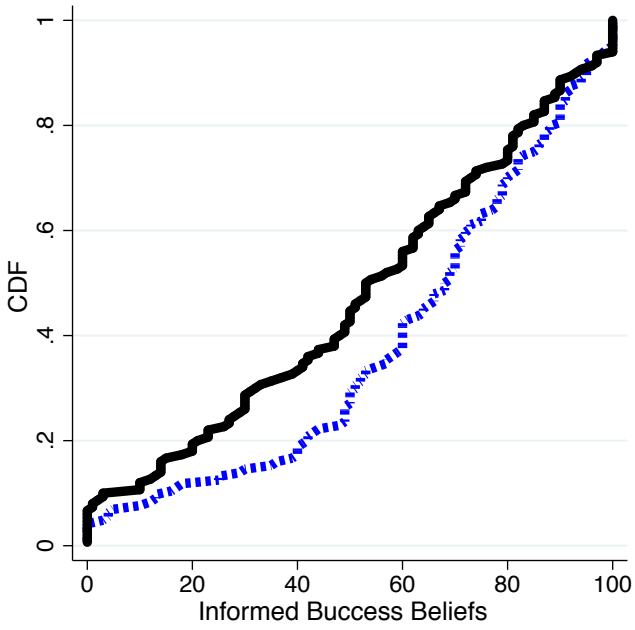
(c) Willingness to Apply Beliefs

$\Delta_{f-m} = -10.46(3.44)^{***}$



(d) Success Beliefs

$\Delta_{f-m} = -9.60(3.40)^{***}$



Female Workers Male Workers

A.2.4 Controlling For Confidence

Table A.4: *Public* and *Private Versions*, Self-Promotion Regressions with Confidence Controls

Sample:	Public Version			Private Version		
	(1)	(2)	(3)	(4)	(5)	(6)
DV = Performance Beliefs						
Female	-12.68***	-3.72	-3.81	-13.46***	-3.58	-2.70
	(2.96)	(2.37)	(2.44)	(2.93)	(2.32)	(2.37)
Confidence		4.08***			4.18***	
		(0.29)			(0.27)	
DV = Performance-Bucket Beliefs						
Female	-0.59***	-0.19*	-0.19	-0.56***	-0.10	-0.07
	(0.13)	(0.11)	(0.12)	(0.13)	(0.10)	(0.11)
Confidence		0.18***			0.19***	
		(0.01)			(0.01)	
DV = Willingness to Apply Beliefs						
Female	-15.31***	-7.14**	-6.71**	-17.57***	-8.31***	-7.12**
	(3.46)	(3.16)	(3.33)	(3.51)	(3.16)	(3.24)
Confidence		3.72***			3.92***	
		(0.38)			(0.35)	
DV = Success Beliefs						
Female	-15.09***	-6.70**	-6.48**	-16.46***	-7.54**	-6.51*
	(3.46)	(3.07)	(3.11)	(3.61)	(3.36)	(3.54)
Confidence		3.82***			3.77***	
		(0.38)			(0.37)	
Performance FEs	Yes	Yes	Yes	Yes	Yes	Yes
Confidence FEs	No	No	Yes	No	No	Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. SEs are clustered by participant. Results are from OLS regressions of the noted dependent variable (DV). *Performance Beliefs* indicates the extent of a participant's agreement (from 0 – 100) with the following statement: "I performed well on the test I took." *Performance-Bucket Beliefs* indicates which Likert-scale response (coded from 1 for the lowest to 6 for the highest) a participant selects when asked to "indicate how well you think you performed on the test." *Willingness to Apply Beliefs* indicates the extent of a participant's agreement (from 0 – 100) with the following statement: "I would apply for a job that required me to perform well on the test I took." *Success Beliefs* indicates the extent of a participant's agreement (from 0 – 100) with the following statement: "I would succeed in a job that required me to perform well on the test I took." *Female* is an indicator that the participant is female. *Performance* equals the number of questions a participant correctly answered out of the 20 ASVAB questions. Performance FEs are dummies for each possible performance on the ASVAB. *Confidence* equals the number of questions a participant believes he or she correctly answered. Confidence FEs are dummies for each possible response to the question about how many ASVAB questions the participant correctly answered. Data in columns in (1) through (3) are from uninformed self-assessment questions elicited in Part 2 of the *Public* version and data in columns (4) through (6) are from uninformed self-assessment questions elicited in Part 3 of the *Private* version.

A.2.5 Controlling for Deservingness

Figure A.9: *All Versions*, Deservingness Measure

$$\Delta_{f-m} = -1.55(2.24)$$

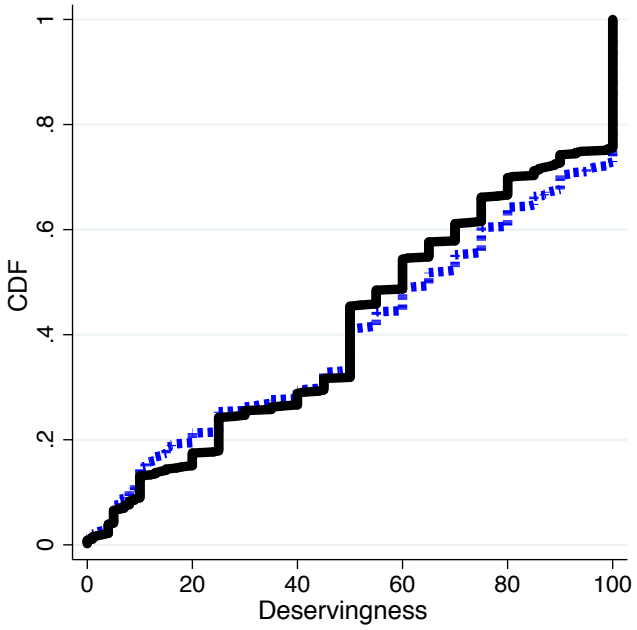


Table A.5: *All Versions*, Deservingness Regressions

	(1)	(2)	(3)
Female	-1.55	-2.24	-2.49
	(2.24)	(1.91)	(1.94)
Performance		4.90***	
		(0.25)	
Constant	60.70***	12.53***	
	(1.65)	(3.13)	
Performance FEs	No	No	Yes
N	900	900	900

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. SEs are clustered at the employer-level. Results are from OLS regressions of the deservingness measure in Part 4 and thus in response to the following question: "Out of a maximum amount of 100 cents, what amount of bonus payment, in cents, do you think you deserve for your performance on the test you took in part 1." *Female* is an indicator for the a female participant. *Performance*s equals the number of questions a participant correctly answered out of the 20 ASVAB questions. Performance FEs are dummies for each possible performance on the ASVAB. Data are from all versions.

A.2.6 Examining Consistency

Table A.6: Consistency of Self-Promotion

	All Versions		Public Version		Private Version		Ambiguous Version	
Confidence − Performance	$\neq 0$	$= 0$	$\neq 0$	$= 0$	$\neq 0$	$= 0$	$\neq 0$	$= 0$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Performance	0.86	0.68	0.89	0.61	0.86	0.76	0.83	0.66
Performance-Bucket	0.49	0.39	0.47	0.30	0.52	0.42	0.47	0.41
Apply	0.81	0.66	0.82	0.43	0.78	0.64	0.82	0.84
Succeed	0.81	0.73	0.81	0.57	0.80	0.73	0.80	0.84
N	812	88	279	23	271	33	262	32

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. These results show the fraction of participants who changed their answers to each self-assessment question after learning their performance information. The odd columns restrict to the set of participants who learn that their performance beliefs were initially wrong after learning their performance information while the even columns restrict to the set of participants who learn that their performance beliefs were initially correct (i.e., their performance equals their believed performance). Data are from all versions in Columns 1–2, the *Public* version in Columns 3–4, the *Private* version in Columns 5–6, and the *Ambiguous* version in Columns 7–8.

A.3 The Employers Version

Table A.7: *Employer Version*, Wage Regressions

Self-Assessment Question:	Performance	Performance Bucket	Apply	Succeed
	(1)	(2)	(3)	(4)
Answer	0.21***	4.23***	0.23***	0.22***
	(0.03)	(0.41)	(0.02)	(0.02)
*Female*Answer*	-0.01	0.06	-0.01	-0.02
	(0.03)	(0.55)	(0.03)	(0.03)
Female	-1.30	-2.23	-1.37	-1.21
	(1.51)	(1.40)	(1.24)	(1.56)
Constant	23.37***	20.11***	22.66***	23.39***
	(1.28)	(1.08)	(1.03)	(1.22)
N	1490	1788	1490	1490

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. SEs are clustered by employer. Results are from OLS regressions of the wage received by the participant (25 cents if not hired and the chosen wage if hired). *Answer* is the answer provided by the participant in response to the Performance question in Column 1, the Performance-Bucket question in Column 2, the Apply question in Column 3, and the Succeed question in Column 4. In the Performance question, participants indicate the extent of their agreement (from 0 – 100) with the following statement: "I performed well on the test I took." In the Performance-Bucket question, participants select a bucket (coded from 1 for the lowest to 6 for the highest) when asked to "indicate how well you think you performed on the test." In the Apply question, participants indicate the extent their agreement (from 0 – 100) with the following statement: "I would apply for a job that required me to perform well on the test I took." In the Success question, participants indicate the extent of their agreement (from 0 – 100) with the following statement: "I would succeed in a job that required me to perform well on the test I took." *Female* is an indicator for a female employer. Data are from the hiring decisions in the *Employer* version.

A.4 The Free-Response Evaluators Versions

In February 2019, we recruited 400 workers on MTurk to complete the *Free-Response Evaluators* versions of our study using the same criteria as in the main study versions (see footnote 14). We collected data from 399 workers.[28] Each participant received a guaranteed $1.50 completion fee for the 15-minute study. In addition, one of their decisions, out of the 21 decisions in the study, was selected to determine a possible bonus payment for them, and if relevant, for an associated "worker."[29] After evaluators completed all decisions of the study, they took a short follow-up survey that collected demographic information.

The evaluators were randomly assigned either to make 21 hiring decisions (n=198) or to make 21 sets of predictions (n=201). Before making each decision or set of predictions, the evaluator was provided with the text entered by a participant to the free-response question: "Please describe how well you think you performed on the test that you took in part 1 and why." The free response either came from Part 2 or Part 3. Evaluators were randomly assigned these 21 free responses from the set of eligible free responses written by the participants from the three main versions of the study.[30]

Evaluators assigned to make hiring decisions were asked whether they would like to hire the participant who provided that free response and, if so, how much to pay them. The payoffs for the evaluator and associated participant are the same as described in the *Employers* version.[31] While similar to the *Employers* version, there are many more possible free responses than answers to the quantitative self-promotion questions, which means our analysis on hiring decisions is underpowered relative to the *Employers* version, since we only have at most a few evaluators reacting to each free response.

Evaluators assigned to make predictions were instead asked to predict whether the participant who wrote the free response was male or female and how many questions, out of 20, that participant answered correctly on the ASVAB. The payoffs for evaluators are determined as follows. One of the two predictions from one of the 21 sets was randomly selected. If the prediction was correct, the evaluator received a bonus payment of 50 cents.[32]

Relative to the *Employers* version, there are three important differences when considering the results in the *Free-Response Evaluators* versions. First, since there is no objective way to rank free-response answers, we cannot examine how hiring decisions or predictions vary as the responses improve (as we did when examining the impact of a one unit increase on the 0–100 scales in the *Employers* version). Second,

[28] One worker was excluded from participation for having previously participated in the study but was counted as being recruited.

[29] Each participant who completed the *Public* or *Ambiguous* versions of our study was matched with an employer from the *Employer* version of our study and received corresponding payoffs from their employers' hiring decisions. By contrast, in the *Free-Response Evaluators* versions, only select workers from the *Public* and *Ambiguous* versions were matched with an evaluator and received corresponding payoffs, rather than everyone. Since we also wanted evaluators to provide data on the free responses from the *Private* version, evaluators were (accurately) told that one of their decisions would be selected to count but *not* that one of their decisions would be randomly selected to count (as this would have required putting 0% weight on free responses from the *Private* version in the randomization).

[30] Each of our 900 participants provided a response in Part 2 and a response in Part 3. Not all of the 1800 possible free responses were evaluated, however. A research assistant — blinded to sex and study version — deemed 130 free responses "ineligible" due to the answer not relating to the question asked or due to severe grammar and/or spelling issues that made an answer incomprehensible. Consequently, the evaluators were each randomly shown 21 free-responses from the set of 1670 eligible free responses.

[31] As explained in footnote 29, however, free responses from the *Private* version were never selected for payment.

[32] Unlike hiring decisions, the randomly selected prediction can come from a participant from any of our three main study versions.

while evaluators are not informed of the gender of the associated worker, they may be able to infer gender — to some degree — given how the free responses are written. Below, we test this hypothesis using data from the predictions. Third, as noted above, given the large number of possible free responses, we are underpowered to consider the effect of specific free responses.

For these reasons, we favor the analysis of our quantitative self-assessment questions presented in the main text to examine self-promotion. Here, however, we investigate the hiring decisions and predictions from the *Free-Response Evaluators* versions, to present several interesting (but inherently secondary) results. Given our power issues, we combine free responses from the three main study versions (i.e., the *Public*, *Private*, and *Ambiguous* versions).[33] In cases where multiple evaluators faced a decision about the same free response, we use the average decision provided by the evaluators (e.g., if a free response is predicted to be written by a female participant by one evaluator but a male participant by another evaluator, that participant is recorded as being predicted to be female with a 0.50 probability).

Appendix Table A.8 presents results from regressions testing whether the gender of the free response author affects the hiring decisions and predictions of evaluators. The structure of the table mirrors the structure of tables in the main text (e.g., columns (1) and (4) have no controls, (2) and (5) have linear performance controls, and (3) and (6) have dummies for each level of performance). Panel A shows that evaluators predict that free responses provided by female participants come from lower-performing workers. This evidence is relatively consistent with our findings from the quantitative self-assessment questions since women appear to provide less favorable subjective descriptions of their performance. Panel B shows that, although these evaluators are not informed of the gender of the participant associated with the free response, evaluators can infer gender — to some degree — when viewing the responses. Evaluators are significantly more likely to predict that a response was written by a female participant when it was indeed written by a female participant. Panel C shows that the relationship between the gender of the worker and evaluators' hiring decisions is inconclusive. Based on the free response answers, evaluators pay directionally, but not significantly, less to female worker. We note that there are several possible explanations for this last finding. For instance, a preference to hire workers believed to be higher performing (who are more likely to be male, per our first finding) may counteract a preference to hire workers believed to be female (who are more likely to be female, per our second finding). In other words, hiring decisions based off of the free responses may conflate performance beliefs and other preferences. As mentioned in footnote 17 in the main text of the paper, this difficulty with the free-response data contributes to our decision to focus our self-promotion analysis on the self-assessment questions we explore in the main text of the paper.

[33]Note that an additional limitation is that approximately 10% of free responses were not evaluated either because they were deemed ineligible, as explained in footnote 30, or because they were never randomly selected to be shown to an evaluator. The results are qualitatively similar when restricting to the data from each of these three version, with one possible exception: the gender difference in the wage data is largely statistically insignificant but is sometimes directionally negative and sometimes directionally positive, depending on the study version.

Table A.8: *All Versions*, Free Response Regressions

Data:	Uninformed Free Responses			Informed Free Responses		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: DV = Predicted Performance						
Female	-0.82***	-0.78***	-0.67***	-0.51**	-0.41*	-0.35
	(0.23)	(0.22)	(0.22)	(0.24)	(0.22)	(0.23)
Performance		0.20***			0.35***	
		(0.03)			(0.04)	
Constant	12.16***	10.02***		12.36***	8.67***	
	(0.17)	(0.41)		(0.18)	(0.45)	
N	749	749	749	773	773	773
Panel B: DV = Predicted Probability Female						
Female	0.08***	0.09***	0.08***	0.09***	0.09***	0.09***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Performance		0.01***			0.01*	
		(0.00)			(0.00)	
Constant	0.37***	0.23***		0.34***	0.27***	
	(0.02)	(0.04)		(0.02)	(0.04)	
N	749	749	749	773	773	773
Panel C: DV = Wage						
Female	-1.28	-1.26	-1.44*	-0.96	-0.84	-0.66
	(0.82)	(0.81)	(0.81)	(0.99)	(0.98)	(1.04)
Performance		0.54***			0.68***	
		(0.11)			(0.13)	
Constant	33.58***	28.03***		35.45***	28.43***	
	(0.60)	(1.15)		(0.76)	(1.59)	
N	743	743	743	755	755	755
Performance FEs	No	No	Yes	No	No	Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. SEs are clustered by participant. Results are from OLS regressions of the noted dependent variable (DV). *Predicted Performance* equals the number of questions that an evaluator predicts a participant correctly answered out of the 20 ASVAB questions. *Predicted Probability Female* equals the probability with which an evaluator predicted a participant to have been female. *Wage* equals the wage given to the participant by an evaluator. *Female* is an indicator that the participant is a female. *Performance* equals the number of questions a participant correctly answered out of the 20 ASVAB questions. Performance FEs are dummies for each possible performance on the ASVAB. Data in columns in (1) through (3) are from uninformed free responses elicited in Part 2 and data in columns (4) through (6) are from informed free responses elicited in Part 3 of all three study versions.

B Experimental instructions

B.1 Instructions for *Public* version

Prior to participating in the study, participants must correctly answer a captcha and consent to participate in the study. At the end of the study, participants must complete a short follow-up survey to gather demographic and control information.

The study begins by informing each participant of the $2 study completion fee and of the opportunity to earn additional payment for themselves. Figure B.1 shows how this payment information is explained along with the corresponding understanding question that the participant must answer correctly to proceed.

Figure B.1: Payment Information

Overview: This study will consist of 4 parts and a short follow-up survey. Part 1 is the longest, so you should expect to spend more time completing part 1 and less time completing each of the subsequent parts 2 - 4. Following certain instructions, you will be asked understanding questions. You must answer these understanding questions correctly in order to proceed to complete the study.

Your Payment: For completing this study, you are guaranteed to receive $2 within 24 hours. In addition, one part out of the 4 parts will be randomly selected as the part-that-counts. Any amount you earn in the part-that-counts will be distributed to you as a bonus payment.

Understanding Question: Which of the following statements is true?

For completing this study, I will receive $2 within 24 hours, but I do NOT have a chance of receiving any additional bonus payment.

For completing this study, I will receive $2 within 24 hours, and I will also receive the amount I earn in the part-that-counts as additional bonus payment.

For completing this study, I will receive $2 within 24 hours, and I will also receive the total amount I earn across all parts as additional bonus payment.

The instructions for Part 1 and an example of a corresponding ASVAB question are displayed in Figures B.2 and B.3, respectively.

Figure B.2: Instructions for Part 1

Overview: This study will consist of 4 parts and a short follow-up survey. Part 1 is the longest, so you should expect to spend more time completing part 1 and less time completing each of the subsequent parts 2 - 4. Following certain instructions, you will be asked understanding questions. You must answer these understanding questions correctly in order to proceed to complete the study.

Your Payment: For completing this study, you are guaranteed to receive $2 within 24 hours. In addition, one part out of the 4 parts will be randomly selected as the part-that-counts. Any amount you earn in the part-that-counts will be distributed to you as a bonus payment.

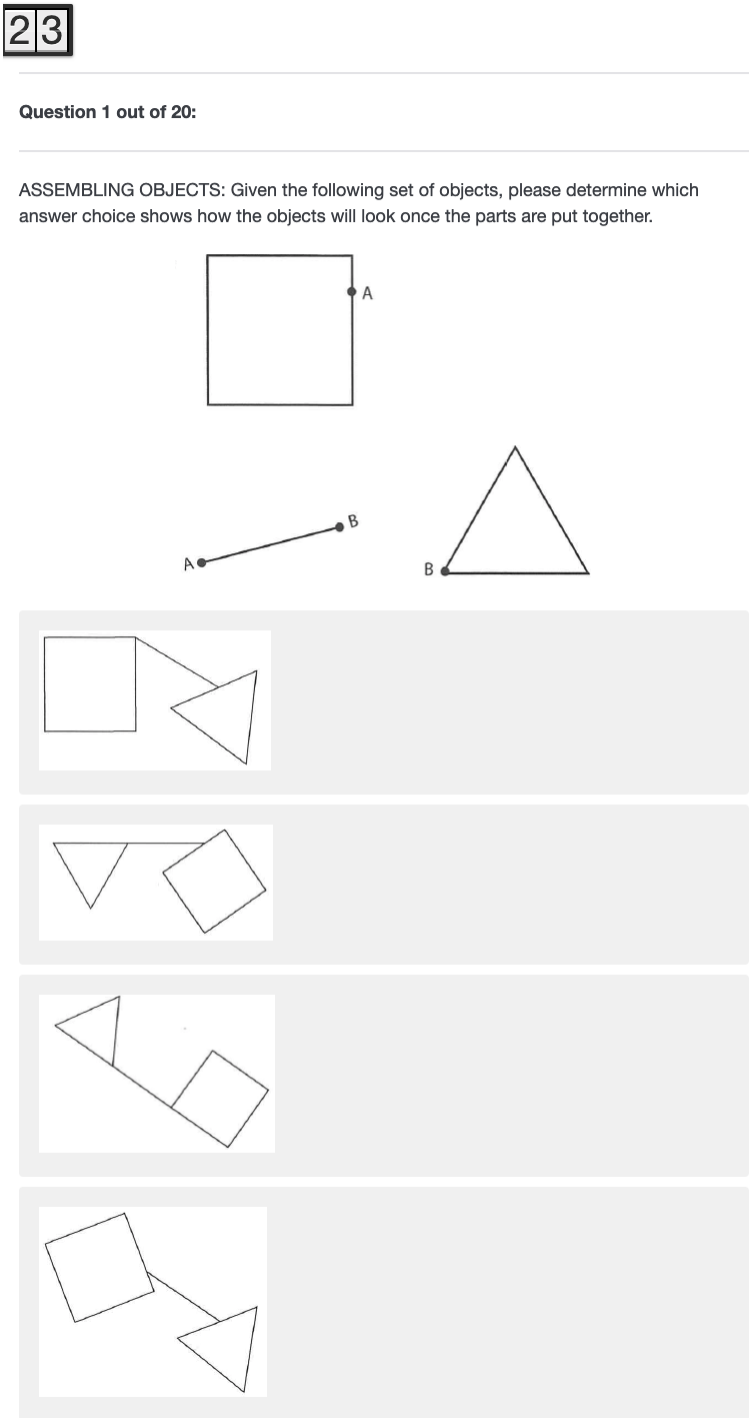
Understanding Question: Which of the following statements is true?

For completing this study, I will receive $2 within 24 hours, but I do NOT have a chance of receiving any additional bonus payment.

For completing this study, I will receive $2 within 24 hours, and I will also receive the amount I earn in the part-that-counts as additional bonus payment.

For completing this study, I will receive $2 within 24 hours, and I will also receive the total amount I earn across all parts as additional bonus payment.

Figure B.3: Part 1: Example of ASVAB question



After completing the ASVAB questions in Part 1 but before proceeding to Part 2, participants are asked to about their confidence as shown in Figure B.4.

Figure B.4: Confidence Question

Congrats! You have now completed part 1 out of 4.

Before pushing the arrow to proceed onto the next part in this study, please answer the following question.

Out of the 20 questions on the test you took in part 1, how many questions do you think you answered correctly?

Participants then receive instructions for Part 2 (see Figure B.5), must correctly answer understanding questions about those instructions (see Figure B.6), and then are asked the self-assessment questions about their performance (see Figure B.7).

Figure B.5: Part 2 Instructions

Instructions for Part 2 out of 4:

In part 2, you will be asked several questions -- on the next page -- related to your performance on the test you completed in part 1.

One of your answers to these questions will be shown to "your part 2 employer," who will be another MTurk worker who completes a different version of this study. Your part 2 employer can decide whether to hire you and, if so, how much to pay you.

Prior to deciding whether to hire you and, if so, how much to pay you, your part 2 employer will NOT be informed of how many questions you answered correctly on the test in part 1.

If this part is randomly selected as the part-that-counts, the additional payment given to your part 2 employer and to you will be determined as follows:

- If your part 2 employer chooses NOT to hire you, your additional payment will equal 25 cents and your part 2 employer's additional payment will equal 100 cents.

- If your part 2 employer chooses to hire you, your additional payment will equal how much they choose to pay you, and your part 2 employer's additional payment will equal (i) 100 cents minus how much they choose to pay you, plus (ii) 5 cents times the number of questions you answered correctly on the test in part 1. Your part 2 employer can choose to pay you any amount between 25 cents and 100 cents.

Figure B.6: Part 2 Understanding Questions

Understanding Question: If this part is randomly selected as the part-that-counts, your additional payment...

will equal 25 cents for sure.

will equal 5 cents times the number of questions you answered correctly on the test in part 1.

will equal 25 cents if you are not hired and the amount your part 2 employer chooses to pay you if you are hired.

Understanding Question: When deciding how much to pay you, your part 2 employer will only know...

how many questions you answered correctly on the test you took in part 1.

how you answer one of the questions -- on the next page -- about your performance on the test you took in part 1.

how you answer all of the questions -- on the next page -- about your performance on the test you took in part 1.

Figure B.7: Part 2 Self-Assessment Questions about Performance

Now, please answer the five questions below to complete part 2. Note that, although the final three questions appear in the same block, they are three separate questions.

Please describe how well you think you performed on the test that you took in part 1 and why.

Please indicate how well you think you performed on the test you took in part 1.

Terrible	Very Poor	Neutral	Good	Very Good	Exceptional

On a scale from 0 (entirely disagree) to 100 (entirely agree), please indicate the extent to which you agree with each of the following statements:

Entirely Disagree		Strongly Disagree		Disagree		Somewhat Disagree		Neither Disagree Nor Agree		Somewhat Agree		Agree		Strongly Agree		Entirely Agree
0		10		20		30		40		50		60		70		80		90		100

I performed well on the test I took in part 1.

I would apply for a job that required me to perform well on the test I took in part 1.

I would succeed in a job that required me to perform well on the test I took in part 1.

After completing Part 2, participants are provided with perfect information on their absolute and relative performance as well as required to correctly report back their absolute performance as shown in Figure B.8.

Figure B.8: Absolute and Relative Performance Information

Congrats! You have now completed part 2 out of 4.

Before pushing the arrow to proceed onto the next part in this study, please read the information below on how well you performed on the test in part 1 and answer the corresponding understanding question.

You answered **0 questions correctly out of the 20 questions**. As a result, compared to 100 other participants who were asked the exact same questions as you were, you answered more questions correctly than 0 of them and fewer questions correctly than 100 of them.

Understanding Question: Out of the 20 questions on the test you took in part 1, how many questions did you answer correctly?

In Part 3, participants are then provided with the same instructions (see Figure B.9), understanding questions (see Figure B.10), and self-assessment questions about their performance (see Figure B.11) as they were in Part 2.

Figure B.9: Part 3 Instructions

Instructions for Part 3 out of 4:

In part 3, you will be asked several questions -- on the next page -- related to your performance on the test you completed in part 1.

One of your answers to these questions will be shown to "your part 3 employer," who will be another MTurk worker who completes a different version of this study. Your part 3 employer can decide whether to hire you and, if so, how much to pay you.

Prior to deciding whether to hire you and, if so, how much to pay you, your part 3 employer will NOT be informed of how many questions you answered correctly on the test in part 1 (even though you were informed of this information on the previous page).

If this part is randomly selected as the part-that-counts, the additional payment given to your part 3 employer and to you will be determined as follows:

- If your part 3 employer chooses NOT to hire you, your additional payment will equal 25 cents and your part 3 employer's additional payment will equal 100 cents.

- If your part 3 employer chooses to hire you, your additional payment will equal how much they choose to pay you, and your part 3 employer's additional payment will equal (i) 100 cents minus how much they choose to pay you, plus (ii) 5 cents times the number of questions you answered correctly on the test in part 1. Your part 3 employer can choose to pay you any amount between 25 cents and 100 cents.

Figure B.10: Part 3 Understanding Questions

Understanding Question: If this part is randomly selected as the part-that-counts, your additional payment...

will equal 25 cents for sure.

will equal 5 cents times the number of questions you answered correctly on the test in part 1.

will equal 25 cents if you are not hired and the amount your part 3 employer chooses to pay you if you are hired.

Understanding Question: When deciding how much to pay you, your part 3 employer will only know...

how many questions you answered correctly on the test you took in part 1.

how you answer one of the questions -- on the next page -- about your performance on the test you took in part 1.

how you answer all of the questions -- on the next page -- about your performance on the test you took in part 1.

Figure B.11: Part 3 Self-Assessment Questions about Performance

Now, please answer the five questions below to complete part 3. Note that, although the final three questions appear in the same block, they are three separate questions.

Please describe how well you think you performed on the test that you took in part 1 and why.

Please indicate how well you think you performed on the test you took in part 1.

Terrible	Very Poor	Neutral	Good	Very Good	Exceptional

On a scale from 0 (entirely disagree) to 100 (entirely agree), please indicate the extent to which you agree with the following statement:

Entirely Disagree		Strongly Disagree		Disagree		Somewhat Disagree		Neither Disagree Nor Agree		Somewhat Agree		Agree		Strongly Agree		Entirely Agree
0		10		20		30		40		50		60		70		80		90		100

I performed well on the test I took in part 1.

I would apply for a job that required me to perform well on the test I took in part 1.

I would succeed in a job that required me to perform well on the test I took in part 1.

Finally, participants receive instructions about and are asked to answer the deservingness question in Part 4 (see Figure B.12)

Figure B.12: Part 4 Instructions and Deservingness Question

Instructions for Part 4 out of 4:

To complete part 4, please answer the one question below. If this part is randomly selected as the part-that-counts, your additional payment will equal whatever amount you answer in this question.

Out of a maximum amount of 100 cents, what amount of bonus payment, in cents, do you think you deserve for your performance on the test you took in part 1?

B.2 Instructions for the *Private* version

The *Private* version of the study proceeds in the same manner as the *Public* version of the study except for the instructions about Part 2 and Part 3, as participants are simply informed that they will receive 25 cents regardless of how they answer the self-assessment questions about their performance. See Figure B.13 for these instructions and the corresponding understanding question.

Figure B.13: The *Private* version: Part 2 Instructions and Understanding Question

Instructions for Part 2 out of 4:

In part 2, you will be asked several questions -- on the next page -- related to your performance on the test you completed in part 1.

If this part is randomly selected as the part-that-counts, your additional payment will equal 25 cents regardless of how you answer these questions. Thus, we ask that you please answer these questions carefully and honestly.

Understanding Question: If this part is randomly selected as the part-that-counts, your additional payment...

- will equal 25 cents for sure.
- will equal 5 cents times the number of questions you answered correctly on the test in part 1.
- will depend on how you answer the questions -- on the next page -- about your performance on the test you took in part 1.

B.3 Instructions for the *Ambiguous* version

The *Ambiguous* version of the study proceeds in the same manner as the *Public* version of the study except for the instructions about Part 2 and Part 3, as participants are informed that there is some chance that their employer will learn their actual performance. See Figures B.14 and B.15 for these instructions and the corresponding understanding question, respectively.

Figure B.14: The *Ambiguous* version: Part 2 Instructions

Instructions for Part 2 out of 4:

In part 2, you will be asked several questions -- on the next page -- related to your performance on the test you completed in part 1.

There is some chance that one of your answers to these questions will be shown to "your part 2 employer," who will be another MTurk worker who completes a different version of this study. Your part 2 employer can decide whether to hire you and, if so, how much to pay you.

Prior to deciding whether to hire you and, if so, how much to pay you, there is also some chance that your part 2 employer will be informed of how many questions you answered correctly on the test in part 1.

However, while your part 2 employer may learn one of your answers to the questions -- on the next page -- related to your performance on the test in part 1 and/or how many questions you answered correctly on the test in part 1, it is also possible that your part 2 employer will not learn any information related to your performance prior to deciding whether to hire you and, if so, how much to pay you.

If this part is randomly selected as the part-that-counts, the additional payment given to your part 2 employer and to you will be determined as follows:

- If your part 2 employer chooses NOT to hire you, your additional payment will equal 25 cents and your part 2 employer's additional payment will equal 100 cents.

- If your part 2 employer chooses to hire you, your additional payment will equal how much they choose to pay you, and your part 2 employer's additional payment will equal (i) 100 cents minus how much they choose to pay you, plus (ii) 5 cents times the number of questions you answered correctly on the test in part 1. Your part 2 employer can choose to pay you any amount between 25 cents and 100 cents.

Figure B.15: The *Ambiguous* version: Part 2 Self-Assessment Questions about Performance

Understanding Question: If this part is randomly selected as the part-that-counts, your additional payment...

will equal 25 cents for sure.

will equal 5 cents times the number of questions you answered correctly on the test in part 1.

will equal 25 cents if you are not hired and the amount your part 2 employer chooses to pay you if you are hired.

Understanding Question: When deciding how much to pay you, your part 2 employer will...

definitely know how many questions you answered correctly on the test you took in part 1.

definitely know how you answer all of the questions -- on the next page -- about your performance on the test you took in part 1.

will know nothing about your performance on the test in part 1, or instead will know one of your answers to the questions – on the next page -- related to your performance on the test in part 1 and/or how many questions you answered correctly on the test in part 1.