APPENDICES (FOR ONLINE PUBLICATION ONLY)

Experimental Instructions

Table Of Contents

This paper involved four main study waves. Section G presents the full instructions for the main *Evaluator Study* and its treatments. Section F presents the full instructions for the *Worker Study* and its treatments. Section I presents the full instructions for additional *Evaluator* studies. Section H presents the full instructions for additional *Worker* studies.

F   Full Instructions for the *Worker Study*

F.1   Instructions for the *Baseline* Treatment of *Worker Study*

After consenting to participate in the study, each participant is informed of the $3 study completion fee and of the opportunity to earn additional payment. Figure F.1.1 shows the overview participants are given and the corresponding comprehension question they must answer correctly in order to proceed. Then, participants proceed to Part 1, which involves a 10-item Math and Science Test. Figure F.1.2 shows the Part 1 instructions and the corresponding comprehension question they must answer correctly in order to proceed.

Participants are then asked two questions about what would characterize poor performance and poor math and science skills (Classifier Question 1 and Classifier Question 2), as shown in Figure F.1.3.

Participants then proceed to the Part 2 instructions, which are related to predicting their own performance on the test via a series of self-evaluation questions. Figure F.1.4 shows the Part 2 instructions and the corresponding comprehension questions that participants need to answer correctly in order to proceed. Participants answer 17 self-evaluation questions (see Appendix Table A1 for corresponding labels of these self-evaluation questions), which are presented in randomized order (Figures F.1.5-F.1.13).

After completing Part 2, participants complete a short follow-up survey that collects additional control and demographic information.

Figure F.1.1: Study Overview, the *Baseline* Treatment of *Worker Study*

Overview: This study will consist of 2 parts and a short follow-up survey. 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 will receive $3 as a completion payment. In addition, one part out of the 2 parts will be randomly selected as the part-that-counts. Any amount you earn in the part-that-counts will be given to you as a bonus payment.

Understanding Question: Which of the following statements is true?

For completing this study, I will receive nothing.

For completing this study, I will receive $3 for sure, and I will have no chance of a bonus payment.

For completing this study, I will receive $3 for sure. In addition, I will receive any amount I earn in the part-that-counts as a bonus payment.

Figure F.1.2: Part 1 Instructions, *Baseline* Treatment of the *Worker Study*

Instructions for Part 1 out of 2:

In Part 1, you will complete a math and science test. On the test, you will be asked to answer up to 10 questions. Each question will test your math and science skills. Specifically, you will be asked about general science, arithmetic reasoning, math knowledge, mechanical comprehension, and assembling objects. Performance on this test is often used as a measure of cognitive ability by academic researchers.

You will be presented with each of the 10 questions on separate pages. You will be given up to 20 seconds to answer each question, although you may push the arrow at the bottom of the page to answer a question before the 20 seconds are up.

If Part 1 is randomly selected as the part-that-counts, your additional payment will equal 10 cents times the number of questions you answer correctly on this test.

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

> will not depend on how many questions you answer correctly on the test.

> will be lower if you answer more questions correctly on the test.

> will be higher if you answer more questions correctly on the test.

Figure F.1.3: Classifier Questions, *Baseline* Treatment of the *Worker Study*

Before proceeding to Part 2, please answer the following two questions:

An individual's performance on the math and science test was **poor** if the number of questions the individual answered correctly was **less than or equal to...**

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

An individual's performance on the math and science test was **indicative of poor math and science skills** if the number of questions the individual answered correctly was **less than or equal to...**

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

Instructions for Part 2 out of 2:

In part 2, you will be asked to make 17 predictions related to **your performance on the math and science test** you took in part 1.

In some of these predictions, you will be asked to guess the right answer to a multiple-choice question. In each of those predictions, you will earn $1 if your guess is right.

In the other predictions, you will be asked to guess the percent chance of some outcome being true on a slider that ranges from 0% to 100%. In each of those predictions, to secure the largest chance of earning $1 from the prediction, you should report your most-accurate guess. To learn the precise rule that determines how much you earn from these predictions click here.

If part 2 is randomly selected as the part-that-counts, one of your 17 predictions will be randomly selected as the prediction-that-counts, and your additional payment will equal the amount you earn in the prediction-that-counts.

Thus, to maximize your chance of earning an additional payment of $1, you should provide your most-accurate guess when making each prediction.

Understanding Question: To maximize your chance of additional bonus payment, how should you make predictions in this part?

> It doesn't matter

> As accurately as possible

> Randomly

Understanding Question: If this part is randomly selected as the part-that-counts, how much additional payment will you receive?

> I will receive what I earn from all predictions in this part.

> I will receive what I earn from the prediction-that-counts in this part.

> Nothing

Figure F.1.5: Self-Evaluation Question 0, *Baseline* Treatment of the *Worker Study*

Prediction X out of 17: Out of the 10 questions on the math and science test, what do you think is **the number you answered correctly**?

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

Figure F.1.6: Self-Evaluation Questions 1B and 1C, *Baseline* Treatment of the *Worker Study*

Prediction X out of 17: Did you get **3 or more questions right** out of the 10 questions on the math and science test?

| No | Yes |

Prediction X out of 17: What is the **percent chance** that you got **3 or more questions right** out of the 10 questions on the math and science test?

| Extremely unlikely | | Somewhat unlikely | | Neither likely nor unlikely | | | Somewhat likely | | | Extremely likely |
| 0 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |

% chance that you got 3 or more questions right

Figure F.1.7: Self-Evaluation Questions 2B and 2C, *Baseline* Treatment of the *Worker Study*

Prediction X out of 17: Did you get **5 or more questions right** out of the 10 questions on the math and science test?

| No | Yes |

Prediction X out of 17: What is the **percent chance** that you got **5 or more questions right** out of the 10 questions on the math and science test?

| Extremely unlikely | | Somewhat unlikely | | Neither likely nor unlikely | | | Somewhat likely | | | Extremely likely |
| 0 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |

% chance that you got 5 or more questions right

Figure F.1.8: Self-Evaluation Questions 3B and 3C, *Baseline* Treatment of the *Worker Study*

Prediction X out of 17: Did you get **7 or more questions right** out of the 10 questions on the math and science test?

No	Yes

Prediction X out of 17: What is the **percent chance** that you got **7 or more questions right** out of the 10 questions on the math and science test?

Extremely unlikely	Somewhat unlikely	Neither likely nor unlikely	Somewhat likely	Extremely likely
0 10	20 30	40 50 60	70 80	90 100

% chance that you got 7 or more questions right

Figure F.1.9: Self-Evaluation Questions 4B and 4C, *Baseline* Treatment of the *Worker Study*

Please review the below information to make the next two predictions:
You scored in the top half if your score (i.e., the number of questions you got right) is greater than or equal to the scores of at least 50% of other participants. These other participants involve 50 randomly-selected men and 50 randomly-selected women from the set of all other participants who take this study.

Prediction X out of 17: Did you **score in the top half** when compared to **other participants** who took the study?

No	Yes

Prediction X out of 17: What is the **percent chance** that you **scored in the top half** when compared to **other participants**?

Extremely unlikely	Somewhat unlikely	Neither likely nor unlikely	Somewhat likely	Extremely likely
0 10	20 30	40 50 60	70 80	90 100

% chance that you scored in top half when compared to other participants

Figure F.1.10: Self-Evaluation Questions 5B and 5C, *Baseline* Treatment of the *Worker Study*

Please review the below information to make the next two predictions:
When compared to women, you scored in the top half if your score (i.e., the number of questions you got right) is greater than or equal to the scores of at least 50% of 100 women. These women involve 100 randomly-selected women from the set of all other participants who are women and take this study.

Prediction X out of 17: Did you score in the top half when compared to women?

No	Yes

Prediction X out of 17: What is the percent chance that you scored in the top half when compared to women?

Extremely unlikely		Somewhat unlikely		Neither likely nor unlikely		Somewhat likely			Extremely likely
0	10	20	30	40	50	60	70	80	90	100

% chance that you scored in top half when compared to women

Figure F.1.11: Self-Evaluation Questions 6B and 6C, *Baseline* Treatment of the *Worker Study*

Please review the below information to make the next two predictions:
When compared to men, you scored in the top half if your score (i.e., the number of questions you got right) is greater than or equal to the scores of at least 50% of 100 men. These men involve 100 randomly-selected men from the set of all other participants who are men and take this study.

Prediction X out of 17: Did you score in the top half when compared to men?

No	Yes

Prediction X out of 17: What is the percent chance you scored in the top half when compared to men?

Extremely unlikely		Somewhat unlikely		Neither likely nor unlikely		Somewhat likely			Extremely likely
0	10	20	30	40	50	60	70	80	90	100

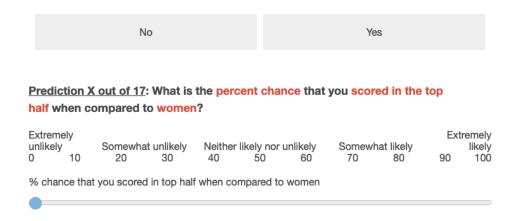% chance that you scored in top half when compared to men

Figure F.1.12: Self-Evaluation Questions 7B and 7C, *Baseline* Treatment of the *Worker Study*

Please review the below information to make the next two predictions:

Recall that, prior to beginning Part 2, you were asked a question about which scores you believed were poor. You will be matched with an "evaluator" who was also asked this question, and your evaluator is said to have described your performance as poor if they indicated that your score was poor. Your evaluator will be randomly selected from the set of all other workers who also complete the study and is equally likely to be a man or a woman.

Prediction X out of 17: Did your evaluator describe your performance on the math and science test as poor?

No	Yes

Prediction X out of 17: What is the percent chance that your evaluator described your performance on the math and science test as poor?

Extremely unlikely		Somewhat unlikely		Neither likely nor unlikely			Somewhat likely			Extremely likely
0	10	20	30	40	50	60	70	80	90	100

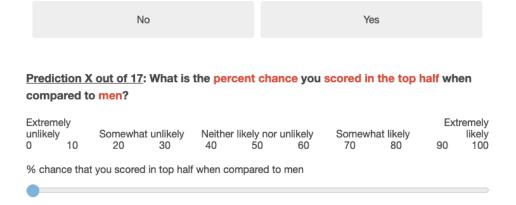% chance that your performance was described as poor

Figure F.1.13: Self-Evaluation Questions 8B and 8C, *Baseline* Treatment of the *Worker Study*

Please review the below information to make the next two predictions:
Recall that, prior to beginning Part 2, you were asked a question about which scores you believed were indicative of poor math and science skills. You will be matched with an "evaluator" who was also asked this question, and your evaluator is said to have described your performance as indicative of poor math and science skills if they indicated that your score was indicative of poor math and science skills. Your evaluator will be randomly selected from the set of all other workers who also complete the study and is equally likely to be a man or a woman.

Prediction X out of 17: Did your evaluator describe your performance on the math and science test as indicative of poor math and science skills?

No	Yes

Prediction X out of 17: What is the percent chance that your evaluator described your performance on the math and science test as indicative of poor math and science skills?

Extremely unlikely		Somewhat unlikely		Neither likely nor unlikely			Somewhat likely			Extremely likely
0	10	20	30	40	50	60	70	80	90	100

% chance that your performance was described as indicative of poor math and science skills

F.2 Instructions for the *Strategic Incentives* Treatment of the *Worker Study*

Relative to the *Baseline* treatment of the *Worker Study* (Section F.1), all that differs in the *Strategic Incentives* treatment of the *Worker Study* is the Part 2 instructions. In this condition, workers are informed that one of their answers may be shown to their employer who will determine how much they earn if Part 2 is randomly selected as the part-that-counts. New Figures F.2.1 and F.2.2 below show the Part 2 instructions and the corresponding comprehension questions that participants need to answer correctly in order to proceed. All other screens look identical to the *Baseline* treatment of the *Worker Study*, shown above.

Figure F.2.1: Part 2 Instructions, *Strategic Incentives* Treatment of the *Worker Study*

Instructions for Part 2 out of 2:

In part 2, you will be asked to make 17 predictions related to **your performance on the math and science test** you took in part 1.

In some of these predictions, you will be asked to guess the right answer to a multiple-choice question. In the other predictions, you will be asked to guess the percent chance of some outcome being true on a slider that ranges from 0% to 100%.

One of your predictions will be randomly selected as the prediction-that-counts.

Your answer to the prediction-that-counts will be shown to "your employer," who will be another Prolific worker who completes a different version of this study. Your employer will decide whether to hire you.

Aside from your answer to the prediction-that-counts, your employer will not be provided with any information on you or on your performance. For instance, your employer will NOT be informed of any demographic information about you, and your employer will NOT be informed of how many questions you answered correctly on the math and science test.

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

- If your employer chooses NOT to hire you, your additional payment will equal 50 cents and your employer's additional payment will equal 50 cents.

- If your employer chooses to hire you, your additional payment will equal 100 cents and your employer's additional payment will equal 10 cents times the number of questions you answered correctly on the test.

Understanding Question: If Part 2 is randomly selected as the part-that-counts, your additional payment...

will be higher if your predictions are more accurate

will NOT depend on how accurate your predictions are

Understanding Question: If Part 2 is randomly selected as the part-that-counts, will your employer learn how many questions you answered correctly on the math and science test?

No - they will only be provided with my answer to one of my predictions

Yes

12

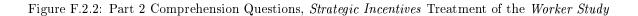Figure F.2.2: Part 2 Comprehension Questions, *Strategic Incentives* Treatment of the *Worker Study*

Understanding Question: If Part 2 is randomly selected as the part-that-counts, your additional payment...

will be higher if you are hired by your employer

will NOT depend on whether you are hired by your employer

Understanding Question: If Part 2 is randomly selected as the part-that-counts and your employer hires you, your employer's additional payment...

will be higher if you have a good performance and lower if you have a bad performance on the math and science test

will NOT depend on your performance on the math and science test

G   Full Instructions for the *Evaluator Study*

All participants in this study are randomized to be asked about male or female workers (or "group-1" or "group-2" workers in some conditions) and to be in one of six treatments described below.

G.1   Instructions for the *Baseline* Treatment of the *Evaluator Study*

After consenting to participate in the study, each participant is informed of the $2 study completion fee and of the opportunity to earn additional payment. Figures G.1.1, G.1.2, and G.1.3 show the overview and comprehension questions we give to participants who are randomized to evaluate **female workers**. They must answer comprehension questions correctly in order to proceed. Then, participants provide their prior beliefs (Figure G.1.4). Subsequently, they are provided with information on female workers' self-evaluations and asked to provide their posterior beliefs (Figure G.1.5). After this, they are asked to provide their overconfidence and underconfidence beliefs (Figure G.1.6). Finally, all participants take a short survey of five randomized bonus questions, as shown in Figures G.1.7-G.1.12, and a follow-up survey that collects additional control and demographic information.

For participants who are randomized to be asked about **male workers**, "female" is replaced by "male" everywhere, and the self-evaluation information provided in Figures G.1.5 and G.1.6 changes from 80% to **56%**.

Figure G.1.1: Study Overview, *Baseline* Treatment of the *Evaluator Study*

<u>**Main Instructions (Page 1 out of 2)**</u>

<u>**Overview:**</u>
This study will consist of 3 predictions and a short follow-up survey. For completing this study, you are guaranteed to receive $2 within 24 hours. In addition, any additional payment you earn will be distributed to you as a bonus payment.

<u>**The Workers**</u>:
In a prior study, an approximately equal number of men and women (called "workers") completed a math and science test with 10 questions. Each question tested their math and science skills by asking them about general science, arithmetic reasoning, math knowledge, mechanical comprehension, and assembling objects. Performance on questions like these is often used as a measure of cognitive ability by academic researchers. A worker's score on the test equals the number of questions they answer correctly, and a worker earns 10 cents times their score.

<u>**Your Predictions**</u>:
You will be asked to make 3 predictions related to the **performance of workers on the math and science test**.

To maximize your chance of earning an additional payment of $1, you should provide your most-accurate guess when making each prediction. This is because each prediction will ask you to guess the percent chance of some outcome being true. In each of those predictions, to secure the largest chance of earning $1 from the prediction, you should report your most-accurate guess. To learn the precise payment rule that determines how much you earn from these predictions click here.

One of your 3 predictions will be randomly selected as the prediction-that-counts, and your additional payment will equal the amount you earn in the prediction-that-counts.

Understanding Question: To maximize your chance of additional bonus payment, how should you make predictions in this part?

> It doesn't matter

> As accurately as possible

> Randomly

Understanding Question: How much additional payment will you receive?

> I will receive what I earn from all predictions in this part

> I will receive what I earn from the prediction-that-counts

> Nothing

15

Figure G.1.2: Instructions about Female Workers, *Baseline* Treatment of the *Evaluator Study*

Main Instructions (Page 2 out of 2)

In each prediction, we will ask you to make a prediction about the performance of your worker. Below please learn more about your workers and the types of predictions we will ask you to make about your workers.

Your Workers
In each prediction, your worker will be randomly selected from the following group: all of the female workers who had performances in the "middle" (when compared to all female and male workers) on the math and science test. Your worker in one prediction will never be the same as your worker in another prediction. **Thus, you will never be asked about the same worker twice.**

Workers who had performances in the middle neither performed the best nor performed the worst. According to the number of questions they got right on the math and science test, workers who had performances in the middle performed better than or equal to at least one-quarter of all workers, and they performed worse than or equal to at least one-quarter of all workers.

You will sometimes be provided with information on beliefs held by the relevant workers when they were asked to make predictions about their own performance.

Types of Predictions
In each prediction, you will be asked to predict the **percent chance that some outcome is true.** Sometimes, you will be asked to predict the percent chance that your worker in that prediction had an evaluator who described the worker's performance as indicative of poor math and science skills. Other times, you will be asked to predict the percent chance that your worker is overconfident or underconfident when asked to make predictions about their own performance. Thus, please note the following:

- Each worker had an evaluator who was randomly selected from the set of all other workers who completed the study and who was equally likely to be a man or a woman. The evaluators answered a question about which scores they believed were indicative of poor math and science skills.
- An evaluator is said to have described a worker's performance as indicative of poor math and science skills if they indicated the worker's score was indicative of poor math and science skills.
- Thus, when an evaluator choose how to describe a worker's performance, the evaluator effectively knew how many questions the worker got right but did not know anything else about the worker, such as the worker's gender.

Worker Predictions:
In some predictions, you will be informed of the average prediction made by the female workers who could be randomly selected to be your worker in that prediction

Thus, please note that: after completing the math and science test, workers made predictions about their performance on the math and science test. When making these predictions, they were not given any information on their own performance and knew that they that should report their most-accurate guess to maximize their chance of earning an additional bonus payment of $1.

16

Figure G.1.3: Comprehension Questions about Female Workers, *Baseline* Treatment of the *Evaluator Study*

Understanding Question: When an evaluator describes a worker's performance, do they know the gender of the worker?

Yes

No

Understanding Question: A worker should expect to earn more...

if they provided more accurate predictions

if they got hired (regardless of how accurate their predictions were)

Understanding Question: In each prediction, my worker will be randomly selected from the following group:

all of the female workers

all of the female workers who had performances in the middle

all of the male workers who had performances in the middle

Figure G.1.4: Prior Belief about Female Workers, *Baseline* Treatment of the *Evaluator Study*

Prediction 1 out of 3:
Please provide an integer answer (0, 1, 2,...100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your female worker will be randomly selected from the following group: all of the female workers who had performances in the middle.

What do you think is the percent chance that your female worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills?

Figure G.1.5: Posterior Belief about Female Workers, *Baseline* Treatment of the *Evaluator Study*

Prediction 2 out of 3:
Please provide an integer answer (0, 1, 2,...100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your female worker will again be randomly selected from the following group: all of the female workers who had performances in the middle.

After completing the math and science test, 80% of workers in that group predicted that they had an evaluator who described their performance as indicative of poor math and science skills.

What do you think is the percent chance that your female worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills?

<div style="border:1px solid black; width:200px; height:40px;"></div>

Figure G.1.6: Over/Underconfidence Beliefs about Female Workers, *Baseline* Treatment of the *Evaluator Study*

<u>Prediction 3 out of 3:</u>
Please provide an integer answer (0, 1, 2,...100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your female worker will again be randomly selected from the following group: all of the female workers who had performances in the middle.

After completing the math and science test, 80% of workers in that group predicted that they had an evaluator who described their performance as indicative of poor math and science skills.

Below, we will ask you to make predictions about the percent chance that your worker is overconfident or underconfident, depending on that worker's performance. Specifically, we will ask you to make one guess for each performance that your worker could have had. **For determining how much money you earn from this prediction, we will then only consider your guess that corresponds to the prediction that your worker actually made.**

<u>Overconfidence Prediction</u>: If your female worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills, what do you think is the percent chance that your female worker is overconfident because they predicted that they had an evaluator who did NOT describe their performance as indicative of poor math and science skills?

<u>Underconfidence Prediction</u>: If your female worker in this prediction had an evaluator who did NOT describe their performance as indicative of poor math and science skills, what do you think is the percent chance that your female worker is underconfident because they predicted that they had an evaluator who described their performance as indicative of poor math and science skills?

Figure G.1.7: Bonus Questions Instructions, *Baseline* Treatment of the *Evaluator Study*

Bonus Questions

Recall that you are guaranteed to receive $2 within 24 hours. In addition, you will receive $1 as a bonus payment if you provided the correct answer to one randomly selected prediction out of the predictions you have already made.

Now, you have the chance to earn an additional $1 as bonus payment (for up to a total of $2 as bonus payment).

In particular, you will now be asked to answer 5 bonus questions. One of the bonus questions will be randomly selected as the bonus-question-that-counts. If you provide the correct answer in the bonus-question-that-counts, you will earn an additional $1 as bonus payment.

Figure G.1.8: Bonus Question 1: Bayesian Updating, *Baseline* Treatment of the *Evaluator Study*

Bonus Question X out of 5:
There are two doctors at a hospital: Doctor Bailey and Doctor Grey.

- Doctor Bailey has 100 patients and 10% are female.
- Doctor Grey has 100 different patients and 70% are female.
- We put Doctor Bailey's and Doctor Grey's patient files together and randomly shuffle all 200 of them.
- We draw one file from the stack at random, and the patient from this file is male.

What is the percent chance that this patient is Doctor Bailey's patient?
Please provide an integer answer (0, 1, 2,...100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

Figure G.1.9: Bonus Question 2: CRT1, *Baseline* Treatment of the *Evaluator Study*

Bonus Question X out of 5:
A cookie and a peppermint cost $1.10 in total. The cookie costs a dollar more than the peppermint.

How much does the peppermint cost (in cents)?
Please omit the "cents" symbol and only write in the corresponding number (e.g., 0, 1, 2,...)

Figure G.1.10: Bonus Question 3: CRT2, *Baseline* Treatment of the *Evaluator Study*

Bonus Question X out of 5:
If it takes 5 machines 5 minutes to make 5 microwaves, how many minutes would it take 100 machines to make 100 microwaves?
Please omit "minutes" from your answer and only write in the corresponding number (e.g., 0, 1, 2,...)

Figure G.1.11: Bonus Question 4: CRT3, *Baseline* Treatment of the *Evaluator Study*

Bonus Question X out of 5:
A virus spreads through a population. Every day, the number of infected people doubles.

If it takes 48 days for the entire population to catch the virus, how many days would it take for half of the population to catch the virus?
Please omit "days" from your answer and only write in the corresponding number (e.g., 0, 1, 2,...)

Figure G.1.12: Bonus Question 5: Base Rate Neglect, *Baseline* Treatment of the *Evaluator Study*

Bonus Question X out of 5:
A cab was involved in a hit and run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:

- 85% of the cabs in the city are Green and 15% are Blue.
- A witness identified the cab as Blue. The court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time.

What is the percent chance (rounded to the nearest whole number) that the cab involved in the accident was Blue rather than Green?
Please provide an integer answer (0, 1, 2,...100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

G.2 Instructions for the *Attention* Treatment of the *Evaluator Study*

The *Attention* treatment of the *Evaluator Study* differs from the *Baseline* treatment of the *Evaluator Study* (Section G.1) only in the order of the predictions made by the participant.

After consenting to participate in the study, each participant is informed of the $2 study completion fee and of the opportunity to earn additional payment. The screenshots for this study are identical to those above in the *Baseline* Treatment but are shown in a different order. Figures G.1.1, G.1.2, and G.1.3 show the overview and comprehension questions we give to participants who are randomized to evaluate **female workers**. They must answer comprehension questions correctly in order to proceed. Then, participants provide their prior beliefs (Figure G.1.4). Next, they are provided with information on workers' self-evaluations and asked to provide their over/underconfidence beliefs (Figure G.1.6). After this, they are asked to provide their posterior beliefs (Figure G.1.5). Finally, all participants take a short survey of five randomized bonus questions, as previously shown in Figures G.1.7-G.1.12, and a follow-up survey that collects additional control and demographic information.

For participants who are randomized to be asked about **male workers**, 'female" is replaced by "male" everywhere, and the self-evaluation information provided in Figures G.1.5 and G.1.6 changes from 80% to **56%**.

G.3 Instructions for the *Calculation* Treatment of the *Evaluator Study*

The *Calculation* treatment of the *Evaluator Study* differs from the *Attention* treatment of the *Evaluator Study* (Section G.2) only in the decision screen that elicits their posterior beliefs, highlighted via the new Figure shown below.

After consenting to participate in the study, each participant is informed of the $2 study completion fee and of the opportunity to earn additional payment. Figures G.1.1, G.1.2, and G.1.3 show the overview and comprehension questions we give to participants who are randomized to evaluate **female workers**. They must answer comprehension questions correctly in order to proceed. Then, participants provide their prior beliefs (Figure G.1.4). Next, they are provided with information on workers' self-evaluations and asked to provide their over/underconfidence beliefs (Figure G.1.6). After this, they are asked to provide their posterior beliefs (new Figure G.3.1 below). Finally, all participants take a short survey of five randomized bonus questions, as previously shown in Figures G.1.7-G.1.12, and a follow-up survey that collects additional control and demographic information.

For participants who are randomized to be asked about **male workers**, 'female" is replaced by "male" everywhere, and the self-evaluation information provided in Figures G.1.6 and G.3.1 changes from 80% to **56%**.

23

Figure G.3.1: Posterior Belief about Female Workers, *Calculation* Treatment of the *Evaluator Study*

<u>**Prediction 3 out of 3**</u>:
Please provide an integer answer (0, 1, 2,...100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your female worker will again be randomly selected from the following group: all of the female workers who had performances in the middle.

After completing the math and science test, 80% of workers in that group predicted that they had an evaluator who described their performance as indicative of poor math and science skills.

There is a very well-known theory in probability and statistics (called <u>Bayes' Rule</u>) that gives a mathematical way to update your guess after receiving some new information. Given the information above on what female workers thought about their own performance, and given how likely you thought female workers are to be overconfident or underconfident, Bayes' Rule would say that your updated guess (from Prediction 1) would be X%.

We are telling you this just in case it is helpful for you. You do NOT have to use Bayes' Rule to update your guess.

What do you think is the percent chance that your female worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills?

G.4 Instructions for the *Baseline, Unknown Gender* Treatment of the *Evaluator Study*

The *Baseline, Unknown Gender* treatment differs from the *Baseline* treatment of the *Evaluator Study* (Section G.1) only in that participants are not told the gender of their worker and "group-1 workers" and "group-2 workers" replace "male workers" and "female workers," respectively.

After consenting to participate in the study, each participant is informed of the $2 study completion fee and of the opportunity to earn additional payment. Figures G.1.1 (above), G.4.1, and G.4.2 show the overview participants who are randomized to evaluate **group-2 workers** are given and the corresponding comprehension questions they must answer correctly in order to proceed. Then, participants provide their prior beliefs (new Figure G.4.3), are provided with information on group-2 workers' self-evaluations and asked to provide their posterior beliefs (new Figure G.4.4), and are asked to provide their overconfidence and underconfidence beliefs (new Figure G.4.5). Finally, all participants take a short survey of five randomized bonus questions, as previously shown in Figures G.1.7-G.1.12, and a follow-up survey that collects additional control and demographic information.

For participants who are randomized to be asked about **group-1 workers** (considered "male workers" in the *Baseline* Treatment of the *Evaluator Study* (Section G.1)), "group-2" is replaced by "group-1" everywhere, and the self-evaluation information provided in Figures G.4.4 and G.4.5 changes from 80% to **56%**.
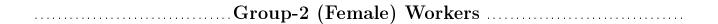
25

Figure G.4.1: Instructions about Group-2 Workers, *Baseline, Unknown Gender* Treatment of the *Evaluator Study*

Main Instructions (Page 2 out of 2)

In each prediction, we will ask you to make a prediction about the performance of your worker. Below please learn more about your workers and the types of predictions we will ask you to make about your workers.

Your Workers
In each prediction, your worker will be randomly selected from the following group: all of the group-2 workers who had performances in the "middle" (when compared to all group-1 and group-2 workers) on the math and science test. Your worker in one prediction will never be the same as your worker in another prediction. **Thus, you will never be asked about the same worker twice.**

Workers who had performances in middle neither performed the best nor performed the worst. According to the number of questions they got right on the math and science test, workers who had performances in the middle performed better than or equal to at least one-quarter of all workers, and they performed worse than or equal to at least one-quarter of all workers.

We assigned each worker to group-1 or group-2 based on an answer they provided to a question in the follow-up survey. While you will not be informed of their answer to this follow-up survey question, you will sometimes be provided with information on beliefs held by the relevant workers when they were asked to make predictions about their own performance.

Types of Predictions
In each prediction, you will be asked to predict the **percent chance that some outcome is true.** Sometimes, you will be asked to predict the percent chance that your worker in that prediction had an evaluator who described the worker's performance as indicative of poor math and science skills. Other times, you will be asked to predict the percent chance that your worker is overconfident or underconfident when asked to make predictions about their own performance. Thus, please note the following:

- Each worker had an evaluator who was randomly selected from the set of all other workers who completed the study and who was equally likely to be a man or a woman. The evaluators answered a question about which scores they believed were indicative of poor math and science skills.
- An evaluator is said to have described a worker's performance as indicative of poor math and science skills if they indicated the worker's score was indicative of poor math and science skills.
- Thus, when an evaluator choose how to describe a worker's performance, the evaluator effectively knew how many questions the worker got right but did not know anything else about the worker, such as the worker's gender.

Figure G.4.2: Comprehension Questions about Group-2 Workers, *Baseline, Unknown* Treatment of the *Evaluator Study*

Understanding Question: When an evaluator describes a worker's performance, do they know the gender of the worker?

> Yes

> No

Understanding Question: A worker should expect to earn more...

> if they provided more accurate predictions

> if they got hired (regardless of how accurate their predictions were)

Understanding Question: In each prediction, my worker will be randomly selected from the following group:

> all of the group-2 workers

> all of the group-2 workers who had performances in the middle

> all of the group-1 workers who had performances in the middle

Figure G.4.3: Prior Belief about Group-2 Workers, *Baseline, Unknown* Treatment of the *Evaluator Study*

Prediction 1 out of 3:
Please provide an integer answer (0, 1, 2,...100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your group-2 worker will be randomly selected from the following group: all of the group-2 workers who had performances in the middle.

What do you think is the percent chance that your group-2 worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills?

> [ ]

Figure G.4.4: Posterior Belief about Group-2 Workers, *Baseline, Unknown* Treatment of the *Evaluator Study*

Prediction 2 out of 3:

Please provide an integer answer (0, 1, 2,...100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your group-2 worker will again be randomly selected from the following group: all of the group-2 workers who had performances in the middle.

After completing the math and science test, 80% of workers in that group predicted that they had an evaluator who described their performance as indicative of poor math and science skills.

What do you think is the percent chance that your group-2 worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills?

[   ]

Figure G.4.5: Over/Underconfidence Beliefs about Group-2 Workers, *Baseline, Unknown* Treatment of the *Evaluator Study*

<u>**Prediction 3 out of 3:**</u>
Please provide an integer answer (0, 1, 2,...100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your group-2 worker will again be randomly selected from the following group: all of the group-2 workers who had performances in the middle.

After completing the math and science test, 80% of workers in that group predicted that they had an evaluator who described their performance as indicative of poor math and science skills.

Below, we will ask you to make predictions about the percent chance that your worker is overconfident or underconfident, depending on that worker's performance. Specifically, we will ask you to make one guess for each performance that your worker could have had. **For determining how much money you earn from this prediction, we will then only consider your guess that corresponds to the prediction that your worker actually made.**

<u>**Overconfidence Prediction**</u>**: If your group-2 worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills, what do you think is the percent chance that your group-2 worker is overconfident because they predicted that they had an evaluator who did NOT describe their performance as indicative of poor math and science skills?**

```
┌─────────────────┐
│                 │
└─────────────────┘
```

<u>**Underconfidence Prediction**</u>**: If your group-2 worker in this prediction had an evaluator who did NOT describe their performance as indicative of poor math and science skills, what do you think is the percent chance that your group-2 worker is underconfident because they predicted that they had an evaluator who described their performance as indicative of poor math and science skills?**

```
┌─────────────────┐
│                 │
└─────────────────┘
```

G.5 Instructions for the *Attention, Unknown Gender* Treatment of the *Evaluator Study*

The *Attention, Unknown Gender* treatment differs from the *Attention* treatment of the *Evaluator Study* (Section G.2) in the same way that the *Baseline, Unknown Gender* treatment (Section G.4) differs from the *Baseline* treatment of the *Evaluator Study* (Section G.1). Participants are not told the gender of their worker and "group-1 workers" and "group-2 workers" replace "male workers" and "female workers," respectively.

G.6 Instructions for the *Calculation, Unknown Gender* Treatment of the *Evaluator Study*

The *Calculation, Unknown Gender* treatment differs from the *Calculation* treatment of the *Evaluator Study* (G.3) in the same way that the *Baseline, Unknown Gender* treatment (Section G.4) differs from the *Baseline* treatment of the *Evaluator Study* (Section G.1). Participants are not told the gender of their worker and "group-1 workers" and "group-2 workers" replace "male workers" and "female workers," respectively.
;

H   Full Instructions for Additional Worker Studies

H.1   Instructions for the *Worker (Undergraduate Students) Study*

The *Worker (Undergraduate Students) Study* surveys undergraduate students of a university.

After consenting to participate in the study, each participant is informed of the $10 study completion fee and of the opportunity to earn additional payment. Figure H.1.1 shows the overview participants are given and the corresponding comprehension question they must answer correctly in order to proceed. Participants then proceed to Part 1. Figure H.1.2 shows the Part 1 instructions and the corresponding comprehension question they must answer correctly in order to proceed. After completing Part 1, participants are asked two questions about what would characterize poor test performance and poor math and science skills (Classifier Question 1 and Classifier Question 2), as previously shown in Figure F.1.3.

Participants then proceed to Part 2. Figure H.1.3 shows the Part 2 instructions and the corresponding comprehension questions that participants need to answer correctly in order to proceed. Participants then answer 13 self-evaluation questions (see Appendix Table A1 for corresponding labels of these self-evaluation questions). In addition to 7 self-evaluation questions of the *Baseline* treatment of the *Worker Study* (Figures F.1.5-F.1.8 above), participants were asked 6 more self-evaluation questions (additional Figures H.1.4-H.1.7 below; Figure H.1.5 shows the additional instructions and comprehension question for Figures H.1.6 and H.1.7). These self-evaluation questions are presented in a randomized order (with the constraint that Figure H.1.6 and H.1.7 are consecutive).

After completing Part 2, participants complete a short follow-up survey that collects additional control and demographic information.

Figure H.1.1: Study Overview, *Worker (Undergraduate Students) Study*

Overview: This study will consist of 2 parts and a short follow-up survey. 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 will receive an Amazon gift card that will be emailed to you. The amount of your gift card is guaranteed to be at least $10. In addition, one part out of the 2 parts will be randomly selected as the part-that-counts. Any amount you earn in the part-that-counts will be added to the $10 to determine the total amount on your gift card.

Understanding Question: Which of the following statements is true?

> For completing this study, I will receive an Amazon gift card that is worth no more than $10.

> For completing this study, I will receive an Amazon gift card that is worth the sum of $10 and any amount I earn in the part-that-counts.

> For completing this study, I will receive an Amazon gift card that is worth the amount I earn in the part-that-counts.

Figure H.1.2: Part 1 Instructions, *Worker (Undergraduate Students) Study*

<u>Instructions for Part 1 out of 2:</u>

In Part 1, you will complete a math and science test. On the test, you will be asked to answer up to 10 questions. Each question will test your math and science skills. Specifically, you will be asked about general science, arithmetic reasoning, math knowledge, mechanical comprehension, and assembling objects. Performance on this test is often used as a measure of cognitive ability by academic researchers.

You will be presented with each of the 10 questions on separate pages. You will be given up to 15 seconds to answer each question, although you may push the arrow at the bottom of the page to answer a question before the 15 seconds are up.

If Part 1 is randomly selected as the part-that-counts, your additional payment will equal $1 times the number of questions you answer correctly on this test.

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

will not depend on how many questions you answer correctly on the test.

will be lower if you answer more questions correctly on the test.

will be higher if you answer more questions correctly on the test.

Figure H.1.3: Part 2 Instructions, *Worker (Undergraduate Students) Study*

Instructions for Part 2 out of 2:

In part 2, you will be asked to make 13 predictions related to **your performance on the math and science test** you took in part 1.

In some of these predictions, you will be asked to guess the right answer to a multiple-choice question. If each of those predictions, you will earn $5 if your guess is right.

In the other predictions, you will be asked to guess the percent chance of some outcome being true on a slider that ranges from 0% to 100%. In each of those predictions, to secure the largest chance of earning $5 from the prediction, you should report your most-accurate guess. To learn the precise rule that determines how much you earn from these predictions click here.

If part 2 is randomly selected as the part-that-counts, one of your 13 predictions will be randomly selected as the prediction-that-counts, and your additional payment will equal the amount you earn in the prediction-that-counts.

Thus, to maximize your chance of earning an additional payment of $5, you should provide your most-accurate guess when making each prediction.

Understanding Question: To maximize your chance of additional bonus payment, how should you make predictions in this part?

It doesn't matter

As accurately as possible

Randomly

Understanding Question: If this part is randomly selected as the part-that-counts, how much additional payment will you receive?

I will receive what I earn from all predictions in this part.

I will receive what I earn from the prediction-that-counts in this part.

Nothing

Figure H.1.4: Self-Evaluation Questions New-1B and New-1C, *Worker (Undergraduate Students) Study*

Prediction X out of 13: Did you get **9 or more questions right** out of the 10 questions on the math and science test?

No	Yes

Prediction X out of 13: What is the **percent chance** that you got **9 or more questions right** out of the 10 questions on the math and science test?

Extremely unlikely		Somewhat unlikely		Neither likely nor unlikely			Somewhat likely			Extremely likely
0	10	20	30	40	50	60	70	80	90	100

% chance that you got 9 or more questions right

Figure H.1.5: Self-Evaluation Questions New-2B, 2C, 3B and C Instructions, *Worker (Undergraduate Students) Study*

Additional Instructions

Recall that, prior to beginning Part 2, you were asked the following two questoins:

- **Question 1:** An individual's performance on the math and science test was **poor** if the number of questions the individual answered correctly was **equal to or less than...**
- **Question 2:** An individual's performance on the math and science test was **indicative of poor math and science skills** if the number of questions the individual answered correctly was **equal to or less than...**

In the next 4 predictions, we will pair you with an evaluator and will ask you to predict how your evaluator described your performance on the math and science test given their answers to the above two questions. Your evaluator will be randomly selected from the set of other participants who completed this study and will be equally likely to be a man or a woman. Then, how your evaluator described your perfromance will be determined as follows:

- If the evaluator indicated that an individual who answered the same number of questions correctly on the test as you did had a performance that was poor in Question 1, then your evaluator has described your performance as poor.
- If the evaluator indicated that an individual who answered the same number of questions correctly on the test as you did had a performance that was indicative of poor math and sceience skills in Question 2, then your evaluator has described your performance as being indicative of poor math and science skills.

For example, if your evaluator indicated that a score of less than 5 was indicative of poor math and science skills, and if you scored a 4, then your evaluator has described your performance as being indicative of poor math and science skills.

Understanding Question: How is the evaluator selected?

> The evaluator is randomly selected from the set of other participants who completed this study.

> The evaluator is selected to have the same performance as you did on the math and science test.

> The evaluator is selected in some other way.

Figure H.1.6: Self-Evaluation Questions New-2B and 2C, *Worker (Undergraduate Students) Study*

Prediction X out of 13: Did your evaluator describe your performance on the math and science test as **poor**?

No	Yes

Prediction X out of 13: What is the **percent chance** that your evaluator described your performance on the math and science test as **poor**?

Extremely unlikely		Somewhat unlikely		Neither likely nor unlikely			Somewhat likely			Extremely likely
0	10	20	30	40	50	60	70	80	90	100

% chance that your performance was described as poor

Figure H.1.7: Self-Evaluation Questions New-3B and 3C, *Worker (Undergraduate Students) Study*

Prediction X out of 13: Did your evaluator describe your performance on the math and science test as **indicative of poor math and science skills**?

No	Yes

Prediction X out of 13: What is the **percent chance** that your evaluator described your performance on the math and science test as **indicative of poor math and science skills**?

Extremely unlikely		Somewhat unlikely		Neither likely nor unlikely			Somewhat likely			Extremely likely
0	10	20	30	40	50	60	70	80	90	100

% chance that your performance was described as indicative of poor math and science skills

I Full Instructions for Additional Evaluator Studies

I.1 Full Instructions for the *Evaluator (Alternative Questions) Study*

All participants in this study are randomized to be asked about male or female workers.

After consenting to participate in the study, each participant is informed of the \$3 study completion fee and of the opportunity to earn additional payment. Figures I.1.1-I.1.4 show the overview participants who are randomized to evaluate **female workers** are given and the corresponding comprehension questions they must answer correctly in order to proceed.

Then, participants are provided with additional instructions about their prior beliefs (Figure I.1.5), are asked to provide their prior beliefs relating to six different outcomes that are presented in a random order (Figures I.1.6-I.1.11), are provided with additional instructions about their posterior beliefs (Figure I.1.12), are asked to provide their posterior beliefs relating to six different outcomes that are presented in a random order (Figures I.1.13-I.1.18), are provided with additional instructions about their overconfidence and underconfidence beliefs (Figure I.1.19), and are asked to provide their overconfidence and underconfidence beliefs relating to six different outcomes that are presented in a random order (Figures I.1.20-I.1.25). Finally, all participants complete a follow-up survey that collects additional control and demographic information.

For evaluators who are instead asked to evaluate **male workers**, "female" is replaced by "male" everywhere. In addition to this, see Figures I.1.26-I.1.31 for posterior belief questions about **male workers** and the corresponding self-evaluation information provided for each question.

Figure I.1.1: Study Overview, *Evaluator (Alternative Questions) Study*

<u>**Main Instructions (Page 1 out of 2)**</u>

<u>**Overview:**</u>
This study will consist of 18 predictions and a short follow-up survey. For completing this study, you are guaranteed to receive $3 within 24 hours. In addition, any additional payment you earn will be distributed to you as a bonus payment.

<u>**The Workers**</u>:
In a prior study, an approximately equal number of men and women (called "workers") completed a math and science test with 10 questions. Each question tested their math and science skills by asking them about general science, arithmetic reasoning, math knowledge, mechanical comprehension, and assembling objects. Performance on questions like these is often used as a measure of cognitive ability by academic researchers. A worker's score on the test equals the number of questions they answer correctly, and a worker earns 10 cents times their score.

<u>**Your Predictions**</u>:
You will be asked to make 18 predictions related to the **performance of workers on the math and science test**.

To maximize your chance of earning an additional payment of $1, you should provide your most-accurate guess when making each prediction. This is because each prediction will ask you to guess the percent chance of of some outcome being true. In each of those predictions, to secure the largest chance of earning $1 from the prediction, you should report your most-accurate guess. To learn the precise payment rule that determines how much you earn from these predictions <u>click here.</u>

One of your 18 predictions will be randomly selected as the prediction-that-counts, and your additional payment will equal the amount you earn in the prediction-that-counts.

Understanding Question: To maximize your chance of additional bonus payment, how should you make predictions in this part?

| It doesn't matter |
| As accurately as possible |
| Randomly |

Understanding Question: How much additional payment will you receive?

| I will receive what I earn from all predictions in this part |
| I will receive what I earn from the prediction-that-counts |
| Nothing |

39

Figure I.1.2: Instructions about Female Workers, *Evaluator (Alternative Questions) Study*

Main Instructions (Page 2 out of 2)

In each prediction, we will ask you to make a prediction about the performance of your worker. Below please learn more about your workers and the types of predictions we will ask you to make about your workers.

Your workers

In each prediction, your worker will be randomly selected from the group of female workers who had performances in the "middle" (when compared to all male and female workers) on the math and science test. Specifically, your female worker will be randomly selected from the group of all female workers who had performances in the middle (when compared to all male and female workers). Your worker in one prediction will never be the same as your worker in another prediction. **Thus, you will never be asked about the same worker twice.**

Workers who had performances in middle neither performed the best nor performed the worst. According to the number of questions they got right on the math and science test, workers who had performances in the middle performed better than or equal to at least one-quarter of all workers, and they performed worse than or equal to at least one-quarter of all workers.

Figure I.1.3: Instructions about Female Workers, cont., *Evaluator (Alternative Questions) Study*

Types of Predictions
You will be asked to make four different types of predictions about the **percent chance that some outcome is true.**

In one set of predictions, you will be asked to predict the **percent chance that your worker in that prediction got at least some number of questions right** on the math and science test.

In a second set of predictions, you will be asked to predict the **percent chance that your worker in that prediction had an evaluator who described their performance as poor or as indicative of poor math and science skills**. Thus, please note the following:

- Each worker has an evaluator who is randomly selected from the set of all other workers who complete the study and who is equally likely to be a man or a woman. The evaluators answered a question about which scores they believed were poor and a question about which scores they believed were indicative of poor math and science skills.
- An evaluator is said to have described a worker's performance as poor or as indicative of poor math and science skills if they indicated the worker's score was poor or was indicative of poor math and science skills, respectively.
- Thus, when an evaluator chooses how to describe a worker's performance, the evaluator effectively knows how many questions the worker got right but does not know anything else about the worker, such as the worker's gender.

In a third set of predictions, you will be asked to predict the **percent chance that your worker in that prediction scored in the top half.** Thus, please note the following:

- A worker scored in the top half if their score (i.e., the number of questions they got right) was greater than or equal to the scores of 50% of other participants.
- These other participants involve 50 randomly-selected men and 50 randomly-selected women from the set of all other participants who took this study.

In a fourth of predictions, you will be asked to predict the **percent chance that your worker in that prediction is overconfident or underconfident** when asked to make predictions about their own performance.

Figure I.1.4: Comprehension Questions about Female Workers, *Evaluator (Alternative Questions) Study*

Understanding Question: In each prediction, my worker will be randomly selected from...

the entire group of female workers

the group of female workers who had performances in the middle

the group of male workers who had performances in the middle

Understanding Question: When an evaluator describes a worker's performance, do they know the gender of the worker?

Yes

No

Figure I.1.5: Prior Belief Instructions about Female Workers, *Evaluator (Alternative Questions) Study*

Additional Instructions

The next set of predictions will ask you to make predictions about the percent chance that your female worker had some performance on the math and science test How we classify workers according to their performance will be defined on each decision screen.

Figure I.1.6: Prior Belief (3+) about Female Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

What do you think is the percent chance that your female worker in this prediction got 3 or more questions right on the test?

Figure I.1.7: Prior Belief (5+) about Female Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

What do you think is the percent chance that your female worker in this prediction got 5 or more questions right on the test?

Figure I.1.8: Prior Belief (7+) about Female Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

What do you think is the percent chance that your female worker in this prediction got 7 or more questions right on the test?

Figure I.1.9: Prior Belief (poor-2) about Female Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

What do you think is the percent chance that your female worker in this prediction had an evaluator who described her performance as poor on the test?

Figure I.1.10: Prior Belief (main self-evaluation) about Female Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

What do you think is the percent chance that your female worker in this prediction had an evaluator who described her performance as indicative of poor math and science skills on the test?

Figure I.1.11: Prior Belief (top-half) about Female Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

What do you think is the percent chance that your female worker in this prediction scored in the top half on the test?

Figure I.1.12: Posterior Belief Instructions about Female Workers, *Evaluator (Alternative Questions) Study*

Additional Instructions

In each of next predictions, you will be informed of the average prediction made by all of the female workers who could be randomly selected to be your worker in that prediction.

Worker Predictions:
After completing the math and science test, workers made predictions about their performance on the math and science test. When making these predictions, they were not given any information on their own performance and knew that they that should report their most-accurate guess to maximize their chance of earning an additional bonus payment of $1.

Understanding Question: A worker should expect to earn more...

if they provided more accurate predictions

if they got hired (regardless of how accurate their predictions were)

Figure I.1.13: Posterior Belief (3+) about Female Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test, **78% of female workers predicted** that they got 3 or more questions right.

What do you think is the percent chance that your female worker in this prediction **got 3 or more questions right**?

Figure I.1.14: Posterior Belief (5+) about Female Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test, **26% of female workers predicted** that they got 5 or more questions right.

What do you think is the percent chance that your female worker in this prediction **got 5 or more questions right**?

[ ]

Figure I.1.15: Posterior Belief (7+) about Female Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test, **4% of female workers predicted** that they got 7 or more questions right.

What do you think is the percent chance that your female worker in this prediction **got 7 or more questions right**?

[ ]

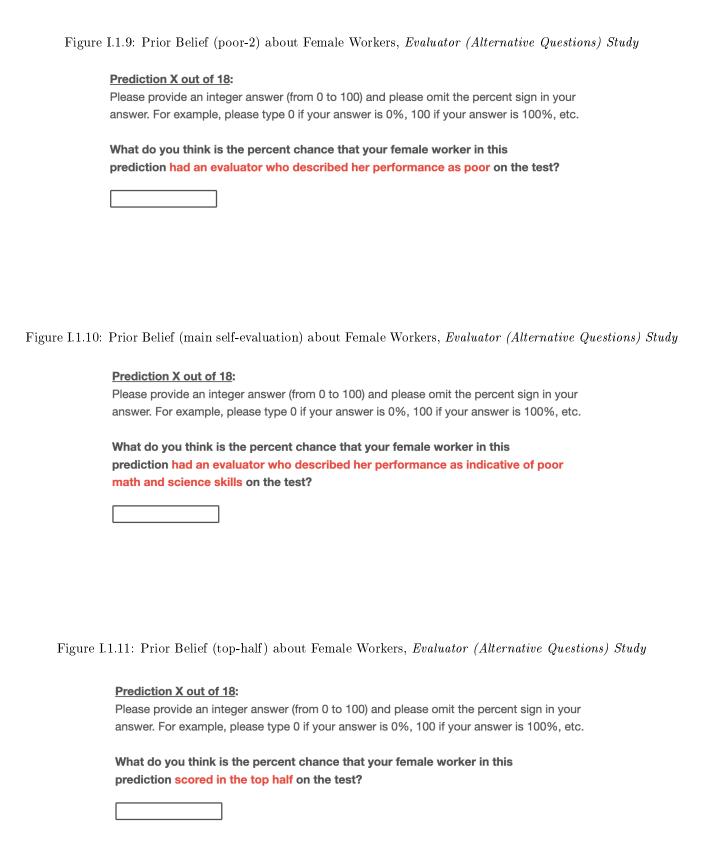Figure I.1.16: Posterior Belief (poor-2) about Female Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test, **81% of female workers predicted** that they had an evaluator who described her performance as poor.

What do you think is the percent chance that your female worker in this prediction **had an evaluator who described her performance as poor**?

[ ]

Figure I.1.17: Posterior Belief (main self-evaluation) about Female Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test, 80% of female workers predicted that they had an evaluator who described her performance as indicative of poor math and science skills.

What do you think is the percent chance that your female worker in this prediction had an evaluator who described her performance as indicative of poor math and science skills?
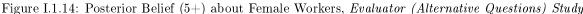
Figure I.1.18: Posterior Belief (top-half) about Female Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test, 26% of female workers predicted that they scored in the top half.

What do you think is the percent chance that your female worker in this prediction scored in the top half?

Figure I.1.19: Over/Underconfidence Beliefs Instructions about Female Workers, *Evaluator (Alternative Questions) Study*

Additional Instructions

In the next prediction, we will ask you to make predictions about the percent chance that your worker is overconfident or underconfident, depending on that worker's performance. Specifically, we will ask you to make one guess for each performance that your worker could have had. **For determining how much money you earn from the next prediction, we will then only consider your guess that corresponds to the prediction that your worker actually made.**
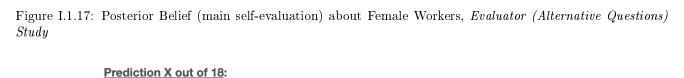
Figure I.1.20: Over/Underconfidence Beliefs (3+) about Female Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

Overconfidence Prediction: If your female worker in this prediction got fewer than 3 questions right, what do you think is the percent chance that she is overconfident because she predicted that she got 3 or more questions right?

```
┌─────────────────┐
│                 │
└─────────────────┘
```

Underconfidence Prediction: If your female worker in this prediction got 3 or more questions right, what do you think is the percent chance that she is underrconfident because she predicted that she got fewer than 3 questions right?

```
┌─────────────────┐
│                 │
└─────────────────┘
```

Figure I.1.21: Over/Underconfidence Beliefs (5+) about Female Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

Overconfidence Prediction: If your female worker in this prediction got fewer than 5 questions right, what do you think is the percent chance that she is overconfident because she predicted that she got 5 or more questions right?

Underconfidence Prediction: If your female worker in this prediction got 5 or more questions right, what do you think is the percent chance that she is underrconfident because she predicted that she got fewer than 5 questions right?

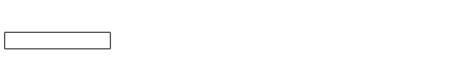Figure I.1.22: Over/Underconfidence Beliefs (7+) about Female Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

Overconfidence Prediction: If your female worker in this prediction got fewer than 7 questions right, what do you think is the percent chance that she is overconfident because she predicted that she got 7 or more questions right?

Underconfidence Prediction: If your female worker in this prediction got 7 or more questions right, what do you think is the percent chance that she is underrconfident because she predicted that she got fewer than 7 questions right?

Figure I.1.23: Over/Underconfidence Beliefs (poor-2) about Female Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

Overconfidence Prediction: If your female worker in this prediction had an evaluator who described her performance as poor, what do you think is the percent chance that she is overconfident because she predicted that she had an evaluator who did NOT describe her performance as poor?

```
[                    ]
```

Underconfidence Prediction: If your female worker in this prediction had an evaluator who did NOT describe her performance as poor, what do you think is the percent chance that she is underrconfident because she predicted that she had an evaluator who described her performance as poor?

```
[                    ]
```

Figure I.1.24: Over/Underconfidence Beliefs (main self-evaluation) about Female Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

Overconfidence Prediction: If your female worker in this prediction had an evaluator who described her performance as indicative of poor math and science skills, what do you think is the percent chance that she is overconfident because she predicted that she had an evaluator who did NOT describe her performance as indicative of poor math and science skills?

```
[                    ]
```

Underconfidence Prediction: If your female worker in this prediction had an evaluator who did NOT describe her performance as indicative of poor math and science skills, what do you think is the percent chance that she is underrconfident because she predicted that she had an evaluator who described her performance as indicative of poor math and science skills?

```
[                    ]
```

Figure I.1.25: Over/Underconfidence Beliefs (top-half) about Female Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

Overconfidence Prediction: If your female worker in this prediction did NOT score in the top half, what do you think is the percent chance that she is overconfident because she predicted that she scored in the top half?

```
[                    ]
```

Underconfidence Prediction: If your female worker in this prediction scored in the top half, what do you think is the percent chance that she is underrconfident because she predicted that she did NOT score in the top half?

```
[                    ]
```

Figure I.1.26: Posterior Belief (3+) about Male Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test, 83% of male workers predicted that they got 3 or more questions right.

What do you think is the percent chance that your male worker in this prediction got 3 or more questions right?

Figure I.1.27: Posterior Belief (5+) about Male Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test, 47% of male workers predicted that they got 5 or more questions right.

What do you think is the percent chance that your male worker in this prediction got 5 or more questions right?

Figure I.1.28: Posterior Belief (7+) about Male Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test, 10% of male workers predicted that they got 7 or more questions right.

What do you think is the percent chance that your male worker in this prediction got 7 or more questions right?

Figure I.1.29: Posterior Belief (poor-2) about Male Workers, *Evaluator (Alternative Questions) Study*
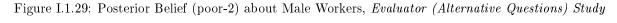
Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test, 60% of male workers predicted that they had an evaluator who described his performance as poor.

What do you think is the percent chance that your male worker in this prediction had an evaluator who described his performance as poor?

> [   ]

Figure I.1.30: Posterior Belief (main self-evaluation) about Male Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test, 56% of male workers predicted that they had an evaluator who described his performance as indicative of poor math and science skills.

What do you think is the percent chance that your male worker in this prediction had an evaluator who described his performance as indicative of poor math and science skills?

> [   ]

Figure I.1.31: Posterior Belief (top-half) about Male Workers, *Evaluator (Alternative Questions) Study*

Prediction X out of 18:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test, 46% of male workers predicted that they scored in the top half.

What do you think is the percent chance that your male worker in this prediction scored in the top half?

> [   ]

I.2 Full Instructions for the *Evaluator (Attention, Top Half) Study*

All participants in this study are randomized to be asked about male or female workers and are asked to evaluate male or female workers based on whether they scored in the top half of 50 randomly selected male workers and 50 randomly selected female workers.

After consenting to participate in the study, each participant is informed of the $2 study completion fee and of the opportunity to earn additional payment. Figures G.1.1 (above) and I.2.1 (below) show the overview participants who are randomized to evaluate **female workers** are given and the corresponding comprehension questions they must answer correctly in order to proceed.

Participants provide their prior beliefs (Figure I.2.2), are provided with information on female workers' self-evaluations and asked to provide their overconfidence and underconfidence beliefs (Figure I.2.3), and then are asked to provide their posterior beliefs (Figure I.2.4). Finally, all participants take a follow-up survey that collects additional control and demographic information.

For participants who are randomized to be asked about **male workers**, "female" is replaced by "male" everywhere, and the self-evaluation information provided in Figures I.2.3 and I.2.4 changes from 26% to **46%**.

Figure I.2.1: Study Overview, *Evaluator (Attention, Top Half) Study*

In each prediction, we will ask you to make a prediction about the performance of your worker. Below please learn more about your workers and the types of predictions we will ask you to make about your workers.

Your Workers
In each prediction, your worker will be randomly selected from the following group: all of the female workers who had performances in the "middle" (when compared to all female and male workers) on the math and science test. Your worker in one prediction will never be the same as your worker in another prediction. **Thus, you will never be asked about the same worker twice.**

Workers who had performances in the middle neither performed the best nor performed the worst. According to the number of questions they got right on the math and science test, workers who had performances in the middle performed better than or equal to at least one-quarter of all workers, and they performed worse than or equal to at least one-quarter of all workers.

You will sometimes be provided with information on beliefs held by the relevant workers when they were asked to make predictions about their own performance.

Types of Predictions
In each prediction, you will be asked to predict the **percent chance that some outcome is true**. Sometimes, you will be asked to predict the percent chance that your worker in that prediction scored in the top or bottom half. Other times, you will be asked to predict the percent chance that your worker is overconfident or underconfident when asked to make predictions about their own performance. Thus, please note the following:

- Other "comparison" participants involve 50 randomly selected men and 50 randomly selected women from the set of all other participants who took this study.
- A worker scored in the top half if their score (i.e., the number of questions they got right) was greater than or equal to the scores of 50% of the other comparison participants.
- A worker scored in the bottom half if their score (i.e., the number of questions they got right) was less than the scores of 50% of the other comparison participants.

Worker Predictions:
In some predictions, you will be informed of the average prediction made by the female workers who could be randomly selected to be your worker in that prediction.

Thus, please note: after completing the math and science test, workers made predictions about their performance on the math and science test. When making these predictions, they were not given any information on their own performance and knew that they that should report their most accurate guess to maximize their chance of earning an additional bonus payment of $1.

Understanding Question: A worker should expect to earn more...

if they provided more accurate predictions

if they got hired (regardless of how accurate their predictions were)

Understanding Question: In each prediction, my worker will be randomly selected from the following group:

all of the female workers

all of the female workers who had performances in the middle

all of the male workers who had performances in the middle

Figure I.2.2: Prior Belief about Female Workers, *Evaluator (Attention, Top Half) Study*

<u>**Prediction 1 out of 3**</u>:
Please provide an integer answer (0, 1, 2, ..., 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your female worker will be randomly selected from the following group: all of the female workers who had performances in the middle.

What do you think is the percent chance that your female worker in this prediction scored in the top half?

<div style="border:1px solid #000; width:200px; height:30px;"></div>

Figure I.2.3: Over/Underconfidence Beliefs about Female Workers, *Evaluator (Attention, Top Half) Study*

<u>**Prediction 2 out of 3**</u>:
Please provide an integer answer (0, 1, 2, ..., 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your female worker will again be randomly selected from the following group: all of the female workers who had performances in the middle.

After completing the math and science test, 26% of workers in that group predicted that they scored in the top half.

Below, we will ask you to make predictions about the percent chance that your worker is overconfident or underconfident, depending on that worker's performance. Specifically, we will ask you to make one guess for each performance that your worker could have had. **For determining how much money you earn from this prediction, we will then only consider your guess that corresponds to the prediction that your worker actually made.**

<u>Overconfidence Prediction</u>**: If your female worker in this prediction scored in the bottom half, what do you think is the percent chance that your female worker is overconfident because they predicted that they scored in the top half?**

<div style="border:1px solid #000; width:180px; height:30px;"></div>

<u>Underconfidence Prediction</u>**: If your female worker in this prediction scored in the top half, what do you think is the percent chance that your female worker is underconfident because they predicted that they scored in the bottom half?**

<div style="border:1px solid #000; width:180px; height:30px;"></div>

Figure I.2.4: Posterior Belief about Female Workers, *Evaluator (Attention, Top Half) Study*

<u>Prediction 3 out of 3</u>:
Please provide an integer answer (0, 1, 2, ..., 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your female worker will again be randomly selected from the following group: all of the female workers who had performances in the middle.

After completing the math and science test, 26% of workers in that group predicted that they scored in the top half.

What do you think is the percent chance that your female worker in this prediction scored in the top half?

I.3 Full Instructions for *Evaluator (Full Distribution) Study*

In the *Evaluator (Full Distribution) Study*, all participants in this study are randomized to be asked about male or female workers and are asked to consider all male or female workers rather than only those with performances in the "middle."

After consenting to participate in the study, each participant is informed of the $2 study completion fee and of the opportunity to earn additional payment. Figures G.1.1 (above), I.3.1, and I.3.2 show the overview participants who are randomized to evaluate **female workers** are given and the corresponding comprehension questions they must answer correctly in order to proceed.

Participants provide their prior beliefs (Figure I.3.3), are provided with information on female workers' self-evaluations and asked to provide their posterior beliefs (Figure I.3.4), and then are asked to provide their overconfidence and underconfidence beliefs (Figure I.3.5). Finally, all participants take a follow-up survey that collects additional control and demographic information.

For participants who are randomized to be asked about **male workers**, "female" is replaced by "male" everywhere, and the self-evaluation information provided in Figures I.3.4 and I.3.5 changes from 76% to **57%**.

Figure I.3.1: Study Overview, *Evaluator (Full Distribution) Study*

Main Instructions (Page 2 out of 2)

In each prediction, we will ask you to make a prediction about the performance of your worker. Below please learn more about your workers and the types of predictions we will ask you to make about your workers.

Your Workers:
In each prediction, your worker will be randomly selected from the following group: **all of the female workers** who took the math and science test.

Types of Predictions
In each prediction, you will be asked to predict the percent chance that some outcome is true. Sometimes, you will be asked to predict the percent chance that your worker in that prediction had an evaluator who described the worker's performance as indicative of poor math and science skills. Other times, you will be asked to predict the percent chance that your worker is overconfident or underconfident when asked to make predictions about their own performance. Thus, please note the following:

- Each worker had an evaluator who was randomly selected from the set of all other workers who completed the study and who was equally likely to be a man or a woman. The evaluators answered a question about which scores they believed were indicative of poor math and science skills.
- An evaluator is said to have described a worker's performance as indicative of poor math and science skills if they indicated the worker's score was indicative of poor math and science skills.
- Thus, when an evaluator chose how to describe a worker's performance, the evaluator effectively knew how many questions the worker got right but did not know anything else about the worker, such as the worker's gender.

Worker Predictions:
In some predictions, you will be informed of the average prediction made by the female workers who could be randomly selected to be your worker in that prediction.

Thus, please note: after completing the math and science test, workers made predictions about their performance on the math and science test. When making these predictions, they were not given any information on their own performance and knew that they that should report their most accurate guess to maximize their chance of earning an additional bonus payment of $1.

Figure I.3.2: Comprehension Questions about Female Workers, *Evaluator (Full Distribution) Study*

Understanding Question: When an evaluator describes a worker's performance, do they know the gender of the worker?

Yes

No

Understanding Question: A worker should expect to earn more...

if they provided more accurate predictions

if they got hired (regardless of how accurate their predictions were)

Understanding Question: In each prediction, my worker will be randomly selected from the following group:

all of the female workers

all of the male workers

Figure I.3.3: Prior Belief about Female Workers, *Evaluator (Full Distribution) Study*

Prediction 1 out of 3:
Please provide an integer answer (0, 1, 2, ..., 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your female worker will be randomly selected from the group of all female workers.

What do you think is the percent chance that your female worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills?

[   ]

Figure I.3.4: Posterior Belief about Female Workers, *Evaluator (Full Distribution) Study*

Prediction 2 out of 3:
Please provide an integer answer (0, 1, 2, ..., 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your female worker will again be randomly selected from the group of all female workers.

After completing the math and science test, 76% of workers in that group predicted that they had an evaluator who described their performance as indicative of poor math and science skills.

What do you think is the percent chance that your female worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills?

[   ]

Figure I.3.5: Over/Underconfidence Beliefs about Female Workers, *Evaluator (Full Distribution) Study*

<u>**Prediction 3 out of 3:**</u>
Please provide an integer answer (0, 1, 2, ..., 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your female worker will again be randomly selected from the group of all female workers.

After completing the math and science test, 76% of workers in that group predicted that they had an evaluator who described their performance as indicative of poor math and science skills.

Below, we will ask you to make predictions about the percent chance that your worker is overconfident or underconfident, depending on that worker's performance. Specifically, we will ask you to make one guess for each performance that your worker could have had. **For determining how much money you earn from this prediction, we will then only consider your guess that corresponds to the prediction that your worker actually made.**

<u>**Overconfidence Prediction**</u>**: If your female worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills, what do you think is the percent chance that your female worker is overconfident because they predicted that they had an evaluator who did NOT describe their performance as indicative of poor math and science skills?**

[ ]

<u>**Underconfidence Prediction**</u>**: If your female worker in this prediction had an evaluator who did NOT describe their performance as indicative of poor math and science skills, what do you think is the percent chance that your female worker is underconfident because they predicted that they had an evaluator who described their performance as indicative of poor math and science skills?**

[ ]

I.4   Full Instructions for the *Evaluator (Professional Evaluators) Study*

All participants in this study are randomized to be asked about male or female workers or about "group-1" or "group-2" workers.

I.4.1   Instructions for the *Baseline* Treatment of the *Evaluator (Professional Evaluators) Study*

The *Baseline* treatment of the *Evaluator (Professional Evaluators) Study* is similar to the *Baseline* treatment of the *Evaluator Study* (Section G.1) with the major difference being that, in the *Evaluator (Professional Evaluators) Study*, the participants are asked about workers from the *Worker (Undergraduate Students) Study* (Section H.1) rather than other Prolific workers. In addition, participants in this study—according to self-reported data collected via Prolific's internal screening questions—met the following two criteria: (1) they have experience in making hiring decisions (i.e. have been responsible for hiring job candidates) and (2) they have experience in a management position.

After consenting to participate in the study, each participant is informed of the $2 study completion fee and of the opportunity to earn additional payment. Figures I.4.1-I.4.3 below show the overview participants who are randomized to evaluate **female workers** are given and the corresponding comprehension questions they must answer correctly in order to proceed. Then, participants provide their prior beliefs (Figure I.4.4), are provided with information on female workers' self-evaluations and asked to provide their posterior beliefs (Figure I.4.5), and are asked to provide their overconfidence and underconfidence beliefs (Figure I.4.6). Finally, participants take a short survey of five randomized bonus questions, as previously shown in Figures G.1.7-G.1.12, and a follow-up survey that collects additional control and demographic information.

For participants who are randomized to be asked about **male workers**, "female" is replaced by "male" everywhere, and the self-evaluation information provided in Figures I.4.5 and I.4.6 changes from 59% to **32%**.

Figure I.4.1: Study Overview, *Baseline* Treatment of the *Evaluator (Professional Evaluators) Study*

<u>**Main Instructions (Page 1 out of 2)**</u>

<u>**Overview:**</u>
This study will consist of 3 predictions and a short follow-up survey. For completing this study, you are guaranteed to receive $2 within 24 hours. In addition, any additional payment you earn will be distributed to you as a bonus payment.

<u>**The Workers**</u>:
In a prior study, we recruited an approximately equal number of male and female undergraduate students from a large midwestern university. These students were assigned the role of "workers" and completed a math and science test with 10 questions. Each question tested their math and science skills by asking them about general science, arithmetic reasoning, math knowledge, mechanical comprehension, and assembling objects. Performance on questions like these is often used as a measure of cognitive ability by academic researchers. A worker's score on the test equals the number of questions they answer correctly, and a worker earns 10 cents times their score.

<u>**Your Predictions**</u>:
You will be asked to make 3 predictions related to the **performance of workers on the math and science test**.

To maximize your chance of earning an additional payment of $1, you should provide your most-accurate guess when making each prediction. This is because each prediction will ask you to guess the percent chance of some outcome being true. In each of those predictions, to secure the largest chance of earning $1 from the prediction, you should report your most-accurate guess. To learn the precise payment rule that determines how much you earn from these predictions <u>click here.</u>

One of your 3 predictions will be randomly selected as the prediction-that-counts, and your additional payment will equal the amount you earn in the prediction-that-counts.

Understanding Question: To maximize your chance of additional bonus payment, how should you make predictions in this part?

It doesn't matter
As accurately as possible
Randomly

Understanding Question: How much additional payment will you receive?

I will receive what I earn from all predictions in this part
I will receive what I earn from the prediction-that-counts
Nothing

Figure I.4.2: Instructions about Female Workers, *Baseline* Treatment of the *Evaluator (Professional Evaluators) Study*

Your Workers

In each prediction, your worker will be randomly selected from the following group: all of the male undergraduate students who completed the prior study and expect to graduate in Spring 2023. Your worker in one prediction will never be the same as your worker in another prediction. **Thus, you will never be asked about the same worker twice.**

You will sometimes be provided with information on beliefs held by the relevant workers when they were asked to make predictions about their own performance.

Types of Predictions

In each prediction, you will be asked to predict the **percent chance that some outcome is true.** Sometimes, you will be asked to predict the percent chance that your worker in that prediction had an evaluator who described the worker's performance as indicative of poor math and science skills. Other times, you will be asked to predict the percent chance that your worker is overconfident or underconfident when asked to make predictions about their own performance. Thus, please note the following:

- Each worker had an evaluator who was randomly selected from the set of all other workers who completed the prior study and who was equally likely to be a man or a woman. The evaluators answered a question about which scores they believed were indicative of poor math and science skills.
- An evaluator is said to have described a worker's performance as indicative of poor math and science skills if they indicated the worker's score was indicative of poor math and science skills.
- Thus, when an evaluator choose how to describe a worker's performance, the evaluator effectively knew how many questions the worker got right but did not know anything else about the worker, such as the worker's gender.

Worker Predictions:

In some predictions, you will be informed of the average prediction made by the male workers who could be randomly selected to be your worker in that prediction.

Thus, please note that: after completing the math and science test, workers made predictions about their performance on the math and science test. When making these predictions, they were not given any information on their own performance and knew that they that should report their most accurate guess to maximize their chance of earning an additional bonus payment of $1.

Figure I.4.3: Comprehension Questions about Female Workers, *Baseline* Treatment of the *Evaluator (Professional Evaluators) Study*

Understanding Question: When an evaluator describes a worker's performance, do they know the gender of the worker?

Yes

No

Understanding Question: A worker should expect to earn more...

if they provided more accurate predictions

if they got hired (regardless of how accurate their predictions were)

Understanding Question: In each prediction, my worker will be randomly selected from the following group:

all of the male undergraduate students who completed the prior study

all of the male undergraduate students who completed the prior study and expect to graduate in Spring 2023

all of the female undergraduate students who completed the prior study and expect to graduate in Spring 2023

Figure I.4.4: Prior Belief about Female Workers, *Baseline* Treatment of the *Evaluator (Professional Evaluators) Study*

<u>**Prediction 1 out of 3**</u>:
Please provide an integer answer (0, 1, 2,...100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your female worker will be randomly selected from the following group: all of the female undergraduate students who completed the prior study and expected to graduate in Spring 2023.

What do you think is the percent chance that your female worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills?

Figure I.4.5: Posterior Belief about Female Workers, *Baseline* Treatment of the *Evaluator (Professional Evaluators) Study*

Prediction 2 out of 3:

Please provide an integer answer (0, 1, 2,...100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your female worker will again be randomly selected from the following group: all of the female undergraduate students who completed the prior study and expected to graduate in Spring 2023.

After completing the math and science test, 59% of workers in that group predicted that they had an evaluator who described their performance as indicative of poor math and science skills.

What do you think is the percent chance that you female worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills?

Figure I.4.6: Over/Underconfidence Beliefs about Female Workers, *Baseline* Treatment of the *Evaluator (Professional Evaluators) Study*

Prediction 3 out of 3:
Please provide an integer answer (0, 1, 2,...100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your female worker will again be randomly selected from the following group: all of the female undergraduate students who completed the prior study and expected to graduate in Spring 2023.

After completing the math and science test, 59% of workers in that group predicted that they had an evaluator who described their performance as indicative of poor math and science skills.

Below, we will ask you to make predictions about the percent chance that your worker is overconfident or underconfident, depending on that worker's performance. Specifically, we will ask you to make one guess for each performance that your worker could have had. **For determining how much money you earn from this prediction, we will then only consider your guess that corresponds to the prediction that your worker actually made.**

Overconfidence Prediction: If your female worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills, what do you think is the percent chance that your female worker is overconfident because they predicted that they had an evaluator who did NOT describe their performance as indicative of poor math and science skills?

[_____]

Underconfidence Prediction: If your female worker in this prediction had an evaluator who did NOT describe their performance as indicative of poor math and science skills, what do you think is the percent chance that your female worker is underconfident because they predicted that they had an evaluator who described their performance as indicative of poor math and science skills?

[_____]

I.4.2 Instructions for the *Baseline, Unknown Gender* Treatment of the *Evaluator (Professional Evaluators) Study*

The *Baseline, Unknown Gender* treatment of the *Evaluator (Professional Evaluators) Study* is the same as the *Baseline* treatment of the *Evaluator (Professional Evaluators) Study* (Section I.4.1) except "male" and "female" is replaced with "group-1" and "group-2," respectively, and worker gender is unknown to participants.

After consenting to participate in the study, each participant is informed of the $2 study completion fee and of the opportunity to earn additional payment. Figures I.4.1 (shown above), I.4.7, and I.4.8 show the overview participants who are randomized to evaluate **group-2 workers** are given and the corresponding comprehension questions they must answer correctly in order to proceed. Then, participants provide their prior beliefs (Figure I.4.9), are provided with information on group-2 workers' self-evaluations and asked to provide their posterior beliefs (Figure I.4.10), and are asked to provide their overconfidence and underconfidence beliefs (Figure I.4.11). Finally, participants take a short survey of five randomized bonus questions, as previously shown in Figures G.1.7-G.1.12, and a follow-up survey that collects additional control and demographic information.

For participants who are randomized to be asked about **group-1 workers** (considered "male workers" in the Treatment of the *Evaluator (Professional Evaluators) Study* (Section I.4.1)), "group-2" is replaced by "group-1" everywhere, and the self-evaluation information provided in Figures I.4.10 and I.4.11 changes from 59% to **32%**.

Figure I.4.7: Instructions about Group-2 Workers, *Baseline* Treatment of the *Evaluator (Professional Evaluators) Study*

Main Instructions (Page 2 out of 2)

In each prediction, we will ask you to make a prediction about the performance of your worker. Below please learn more about your workers and the types of predictions we will ask you to make about your workers.

Your Workers
In each prediction, your worker will be randomly selected from the following group: all of the group-2 undergraduate students who completed the prior study and expect to graduate in Spring 2023. Your worker in one prediction will never be the same as your worker in another prediction. **Thus, you will never be asked about the same worker twice.**

We assigned each worker to group-1 or group-2 based on an answer they provided to a question in the follow-up survey. While you will not be informed of their answer to this follow-up survey question, you will sometimes be provided with information on beliefs held by the relevant workers when they were asked to make predictions about their own performance.

Types of Predictions
In each prediction, you will be asked to predict the **percent chance that some outcome is true.** Sometimes, you will be asked to predict the percent chance that your worker in that prediction had an evaluator who described the worker's performance as indicative of poor math and science skills. Other times, you will be asked to predict the percent chance that your worker is overconfident or underconfident when asked to make predictions about their own performance. Thus, please note the following:

- Each worker had an evaluator who was randomly selected from the set of all other workers who completed the prior study and who was equally likely to be a man or a woman. The evaluators answered a question about which scores they believed were indicative of poor math and science skills.
- An evaluator is said to have described a worker's performance as indicative of poor math and science skills if they indicated the worker's score was indicative of poor math and science skills.
- Thus, when an evaluator choose how to describe a worker's performance, the evaluator effectively knew how many questions the worker got right but did not know anything else about the worker, such as the worker's gender.

Worker Predictions:
In some predictions, you will be informed of the average prediction made by the group-2 workers who could be randomly selected to be your worker in that prediction.

Thus, please note that: after completing the math and science test, workers made predictions about their performance on the math and science test. When making these predictions, they were not given any information on their own performance and knew that they that should report their most accurate guess to maximize their chance of earning an additional bonus payment of $1.

Figure I.4.8: Comprehension Questions about Group-2 Workers, *Baseline* Treatment of the *Evaluator (Professional Evaluators) Study*

Understanding Question: When an evaluator describes a worker's performance, do they know the gender of the worker?

Yes

No

Understanding Question: A worker should expect to earn more...

if they provided more accurate predictions

if they got hired (regardless of how accurate their predictions were)

Understanding Question: In each prediction, my worker will be randomly selected from the following group:

all of the group-2 undergraduate students who completed the prior study

all of the group-2 undergraduate students who completed the prior study and expect to graduate in Spring 2023

all of the group-1 undergraduate students who completed the prior study and expect to graduate in Spring 2023

Figure I.4.9: Prior Belief about Group-2 Workers, *Baseline* Treatment of the *Evaluator (Professional Evaluators) Study*

Prediction 1 out of 3:
Please provide an integer answer (0, 1, 2,...100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your group-2 worker will be randomly selected from the following group: all of the group-2 undergraduate students who completed the prior study and expected to graduate in Spring 2023.

What do you think is the percent chance that your group-2 worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills?

Figure I.4.10: Posterior Belief about Group-2 Workers, *Baseline* Treatment of the *Evaluator (Professional Evaluators) Study*

<u>Prediction 2 out of 3</u>:

Please provide an integer answer (0, 1, 2,...100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your group-2 worker will again be randomly selected from the following group: all of the group-2 undergraduate students who completed the prior study and expected to graduate in Spring 2023.

After completing the math and science test, 59% of workers in that group predicted that they had an evaluator who described their performance as indicative of poor math and science skills.

What do you think is the percent chance that you group-2 worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills?

```
[                    ]
```

Figure I.4.11: Over/Underconfidence Beliefs about Group-2 Workers, *Baseline* Treatment of the *Evaluator (Professional Evaluators) Study*

<u>**Prediction 3 out of 3**</u>:

Please provide an integer answer (0, 1, 2,...100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your group-2 worker will again be randomly selected from the following group: all of the group-2 undergraduate students who completed the prior study and expected to graduate in Spring 2023.

After completing the math and science test, 59% of workers in that group predicted that they had an evaluator who described their performance as indicative of poor math and science skills.

Below, we will ask you to make predictions about the percent chance that your worker is overconfident or underconfident, depending on that worker's performance. Specifically, we will ask you to make one guess for each performance that your worker could have had. **For determining how much money you earn from this prediction, we will then only consider your guess that corresponds to the prediction that your worker actually made.**

<u>**Overconfidence Prediction**</u>**: If your group-2 worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills, what do you think is the percent chance that your group-2 worker is overconfident because they predicted that they had an evaluator who did NOT describe their performance as indicative of poor math and science skills?**

```

```

<u>**Underconfidence Prediction**</u>**: If your group-2 worker in this prediction had an evaluator who did NOT describe their performance as indicative of poor math and science skills, what do you think is the percent chance that your group-2 worker is underconfident because they predicted that they had an evaluator who described their performance as indicative of poor math and science skills?**

```

```

I.5 Full Instructions for the *Evaluator (Extended) Study*

All participants in this study are randomized to be asked about male or female workers and to be in one of four treatments described below.

I.5.1 Instructions for the *Baseline* Treatment of the *Evaluator (Extended) Study*

After consenting to participate in the study, each participant is informed of the $3 study completion fee and of the opportunity to earn additional payment. Figures I.5.1-I.5.3 below show the overview participants who are randomized to evaluate **female workers** are given and the corresponding comprehension questions they must answer correctly in order to proceed. Then, participants provide their prior beliefs (Figure I.5.4), are provided with information on 20 female workers' self-evaluations and asked to provide their posterior beliefs (see Figures I.5.5 and I.5.6 for additional instructions and an example). Participants then are provided with additional instructions and asked to provide their posterior belief about the average self-evaluation of female workers (Figures I.5.7 and I.5.8) and are asked to provide their overconfidence and underconfidence beliefs (Figures I.5.9 and I.5.10). Finally, all participants complete a follow-up survey that collects additional control and demographic information.

For evaluators who are randomized to be asked about **male workers**, 'female" is replaced by "male" everywhere, and the self-evaluation information provided in Figure I.5.8 changes from 80% to **56%**.

Figure I.5.1: Study Overview, *Baseline* Treatment of the *Evaluator (Extended) Study*

<u>**Main Instructions (Page 1 out of 2)**</u>

<u>**Overview:**</u>
This study will consist of a series of predictions and a short follow-up survey. For completing this study, you are guaranteed to receive $3 within 24 hours. In addition, any additional payment you earn will be distributed to you as a bonus payment.

<u>**The Workers**</u>:
In a prior study, an approximately equal number of men and women (called "workers") completed a math and science test with 10 questions. Each question tested their math and science skills by asking them about general science, arithmetic reasoning, math knowledge, mechanical comprehension, and assembling objects. Performance on questions like these is often used as a measure of cognitive ability by academic researchers. A worker's score on the test equals the number of questions they answer correctly, and a worker earns 10 cents times their score.

<u>**Your Predictions**</u>:
You will be asked to make 23 predictions related to the **performance of workers on the math and science test**.

To maximize your chance of earning an additional payment of $1, you should provide your most-accurate guess when making each prediction. This is because each prediction will ask you to guess the percent chance of of some outcome being true. In each of those predictions, to secure the largest chance of earning $1 from the prediction, you should report your most-accurate guess. To learn the precise payment rule that determines how much you earn from these predictions <u>click here.</u>

One of your 23 predictions will be randomly selected as the prediction-that-counts, and your additional payment will equal the amount you earn in the prediction-that-counts.

Understanding Question: To maximize your chance of additional bonus payment, how should you make predictions in this part?

> It doesn't matter

> As accurately as possible

> Randomly

Understanding Question: How much additional payment will you receive?

> I will receive what I earn from all predictions in this part

> I will receive what I earn from the prediction-that-counts

> Nothing

Figure I.5.2: Instructions about Female Workers, *Baseline* Treatment of the *Evaluator (Extended) Study*

Main Instructions (Page 2 out of 2)

In each prediction, we will ask you to make a prediction about the performance of your worker. Below please learn more about your workers and the types of predictions we will ask you to make about your workers.

Your workers
In each prediction, your worker will be randomly selected from the group of female workers who had performances in the "middle" (when compared to all male and female workers) on the math and science test. Specifically, your female worker will be randomly selected from the group of all female workers who had performances in the middle (when compared to all male and female workers). Your worker in one prediction will never be the same as your worker in another prediction. **Thus, you will never be asked about the same worker twice.**

Workers who had performances in middle neither performed the best nor performed the worst. According to the number of questions they got right on the math and science test, workers who had performances in the middle performed better than or equal to at least one-quarter of all workers, and they performed worse than or equal to at least one-quarter of all workers.

Types of Predictions
In each prediction, you will be asked to predict the **percent chance that some outcome is true.** Sometimes, you will be asked to predict the percent chance that your worker in that prediction had an evaluator who described their performance as indicative of poor math and science skills. Other times, you will be asked to predict the percent chance that your worker is overconfident or underconfident when asked to make predictions about their own performance. Thus, please note the following:

- Each worker has an evaluator who is randomly selected from the set of all other workers who complete the study and who is equally likely to be a man or a woman. The evaluators answered a question about which scores they believed were indicative of poor math and science skills.
- An evaluator is said to have described a worker's performance as indicative of poor math and science skills if they indicated the worker's score was indicative of poor math and science skills.
- Thus, when an evaluator chooses how to describe a worker's performance, the evaluator effectively knows how many questions the worker got right but does not know anything else about the worker, such as the worker's gender.

Figure I.5.3: Comprehension Questions about Female Workers, *Baseline* Treatment of the *Evaluator (Extended) Study*

Understanding Question: In each prediction, my worker will be randomly selected from...

the entire group of workers

the group of female workers with performances in the middle

the group of male workers with performances in the middle

Understanding Question: When an evaluator describes a worker's performance, do they know the gender of the worker?

Yes

No

Figure I.5.4: Prior Belief about Female Workers, *Baseline* Treatment of the *Evaluator (Extended) Study*

Prediction X out of 23:
In each prediction, please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

What do you think is the percent chance that your female worker in this prediction had an evaluator who described her performance as indicative of poor math and science skills?

Figure I.5.5: Additional Instructions about Worker-Specific Posterior Belief about Female Workers, *Baseline* Treatment of the *Evaluator (Extended) Study*

Additional Instructions

In each of the next predictions, you will be informed of the prediction made by your worker when that worker was asked to make a prediction about their own performance.

Worker Predictions:
After completing the math and science test, workers made predictions about their performance on the math and science test. When making these predictions, they were not given any information on their own performance and knew that they that should report their most-accurate guess to maximize their chance of earning an additional bonus payment of $1.

Understanding Question: A worker should expect to earn more...

> if they provided more accurate predictions

> if they got hired (regardless of how accurate their predictions were)

Figure I.5.6: Example of Worker-Specific Posterior Belief about Female Workers, *Baseline* Treatment of the *Evaluator (Extended) Study*

Prediction X out of 23:
In each prediction, please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test, your female worker in this prediction predicted that there is a 50% chance that her evaluator described her performance as indicative of poor math and science skills.

What do you think is the percent chance that your female worker in this prediction had an evaluator who described her performance as indicative of poor math and science skills?

[ ]

Figure I.5.7: Additional Instructions about Posterior Belief about Female Workers, *Baseline* Treatment of the *Evaluator (Extended) Study*

Additional Instructions

In the next prediction, rather than being informed of the prediction made by your worker in that prediction, you will be informed of the average prediction made by all of the female workers who could be randomly selected to be your worker in that prediction.

Figure I.5.8: Posterior Belief about Female Workers, *Baseline* Treatment of the *Evaluator (Extended) Study*

Prediction X out of 23:
In each prediction, please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test, 80% of female workers predicted that their evaluator described their performance as indicative of poor math and science skills.

What do you think is the percent chance that your female worker in this prediction had an evaluator who described her performance as indicative of poor math and science skills?

Figure I.5.9: Additional Instructions about Over/Underconfidence Beliefs about Female Workers, *Baseline* Treatment of the *Evaluator (Extended) Study*

Additional Instructions

In the next prediction, we will ask you to make predictions about the percent chance that your worker is overconfident or underconfident, depending on that worker's performance. Specifically, we will ask you to make one guess for each performance that your worker could have had. **For determining how much money you earn from the next prediction, we will then only consider your guess that corresponds to the prediction that your worker actually made.**

Figure I.5.10: Over/Underconfidence Beliefs about Female Workers, *Baseline* Treatment of the *Evaluator (Extended) Study*

<u>**Prediction X out of 23**</u>**:**

In each prediction, please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.
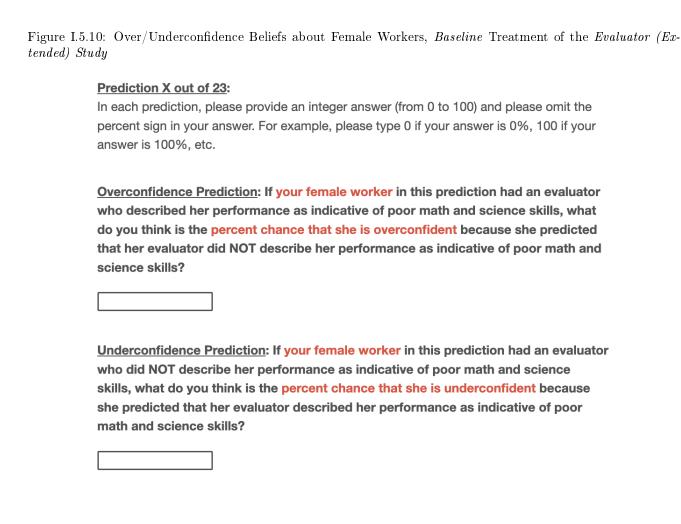
<u>**Overconfidence Prediction**</u>**: If your female worker in this prediction had an evaluator who described her performance as indicative of poor math and science skills, what do you think is the percent chance that she is overconfident because she predicted that her evaluator did NOT describe her performance as indicative of poor math and science skills?**

```
┌─────────────────┐
│                 │
└─────────────────┘
```

<u>**Underconfidence Prediction**</u>**: If your female worker in this prediction had an evaluator who did NOT describe her performance as indicative of poor math and science skills, what do you think is the percent chance that she is underconfident because she predicted that her evaluator described her performance as indicative of poor math and science skills?**

```
┌─────────────────┐
│                 │
└─────────────────┘
```

I.5.2 Instructions for the *Joint Evaluations* Treatment of the *Evaluator (Extended) Study*

For the *Evaluator (Extended) Study*, the *Joint Evaluations* treatment differs from the *Baseline* treatment (Section I.5.1) by asking about both a male worker and a female worker on each decision screen.

After consenting to participate in the study, each participant is informed of the $3 study completion fee and of the opportunity to earn additional payment. Figures I.5.1 (shown above) and I.5.11 (below) show the overview participants who are randomized to evaluate **female and male workers** are given and the corresponding comprehension questions they must answer correctly in order to proceed. Then, participants provide their prior beliefs (Figure I.5.12), are provided with the self-evaluations of 20 female and 20 male workers and asked to provide their posterior beliefs (see Figure I.5.13 for an example), are asked for their posterior belief about male and female workers' average self-evaluations (Figure I.5.14), and are asked to provide their overconfidence and underconfidence beliefs (Figure I.5.15). Finally, all participants complete a follow-up survey that collects additional control and demographic information.

Figure I.5.11: Study Overview, *Joint Evaluations* Treatment of the *Evaluator (Extended) Study*

<u>**Main Instructions (Page 2 out of 2)**</u>

In each prediction set, we will ask you to make make two predictions: one about a female worker and one about a male worker. Below please learn more about your workers and the types of predictions we will ask you to make about your workers.

<u>**Your workers**</u>
In each prediction set, your two workers will be randomly selected from the group of male and female workers who had performances in the "middle" (when compared to all male and female workers) on the math and science test. Specifically, your male worker will be randomly selected from the group of all male workers who had performances in the middle (when compared to all male and female workers), and your female worker will be randomly selected from the group of all female workers who had performances in the middle (when compared to all male and female workers). Your worker in one prediction will never be the same as your worker in another prediction. **Thus, you will never be asked about the same worker twice.**

Workers who had performances in middle neither performed the best nor performed the worst. According to the number of questions they got right on the math and science test, workers who had performances in the middle performed better than or equal to at least one-quarter of all workers, and they performed worse than or equal to at least one-quarter of all workers.

<u>**Types of Predictions**</u>
In each prediction, you will be asked to predict the **percent chance that some outcome is true.** Sometimes, you will be asked to predict the percent chance that your worker in that prediction had an evaluator who described their performance as indicative of poor math and science skills. Other times, you will be asked to predict the percent chance that your worker is overconfident or underconfident when asked to make predictions about their own performance. Thus, please note the following:

- Each worker has an evaluator who is randomly selected from the set of all other workers who complete the study and who is equally likely to be a man or a woman. The evaluators answered a question about which scores they believed were indicative of poor math and science skills.
- An evaluator is said to have described a worker's performance as indicative of poor math and science skills if they indicated the worker's score was indicative of poor math and science skills.
- Thus, when an evaluator chooses how to describe a worker's performance, the evaluator effectively knows how many questions the worker got right but does not know anything else about the worker, such as the worker's gender.

Understanding Question: In each prediction, my female/male worker will be randomly selected from...

the entire group of female/male workers

the group of female/male workers with performances in the middle

Understanding Question: When an evaluator describes a worker's performance, do they know the gender of the worker?

Yes

No

Figure I.5.12: Prior Belief about Female and Male Workers, *Joint Evaluations* Treatment of the *Evaluator (Extended) Study*

Prediction Set X out of 13:

In each prediction, please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

What do you think is the percent chance that your male worker in this prediction set had an evaluator who described his performance as indicative of poor math and science skills?

```
┌──────────────┐
│              │
└──────────────┘
```

What do you think is the percent chance that your female worker in this prediction set had an evaluator who described her performance as indicative of poor math and science skills?

```
┌──────────────┐
│              │
└──────────────┘
```

Figure I.5.13: Example of Worker-Specific Posterior Belief about Female and Male Workers, *Joint Evaluations* Treatment of the *Evaluator (Extended) Study*

<u>**Prediction Set X out of 13**</u>:

In each prediction, please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test:

- **your female worker in this prediction set predicted that there is a 85% chance that her evaluator described her performance as indicative of poor math and science skills, and**
- **your male worker in this prediction set predicted that there is a 92% chance that his evaluator described his performance as indicative of poor math and science skills.**

For your workers in this prediction set, what do you think is the percent chance that their evaluator described their performance as indicative of poor math and science skills?

Percent chance your female worker in this prediction set had an evaluator who described her performance as indicative of poor math and science skills:

Percent chance your male worker in this prediction set had an evaluator who described his performance as indicative of poor math and science skills:

Figure I.5.14: Posterior Belief about Female and Male Workers, *Joint Evaluations* Treatment of the *Evaluator (Extended) Study*
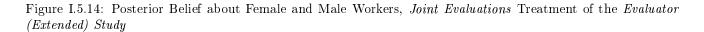
<u>**Prediction Set X out of 13**</u>:
In each prediction, please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test:

- **56% of male workers predicted** that their evaluator described their performance as indicative of poor math and science skills, and
- **80% of female workers predicted** that their evaluator described their performance as indicative of poor math and science skills.

For your workers in prediction set, what do you think is the percent chance that their evaluator described their performance as indicative of poor math and science skills?

Percent chance your male worker in this prediction set had an evaluator who described his performance as indicative of poor math and science skills:

[   ]

Percent chance your female worker in this prediction set had an evaluator who described her performance as indicative of poor math and science skills:

[   ]

Figure I.5.15: Over/Underconfidence Beliefs about Female and Male Workers, *Joint Evaluations* Treatment of the *Evaluator (Extended) Study*

<u>**Prediction Set X out of 13**</u>:
In each prediction, please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

<u>**Overconfidence Prediction**</u>**: If your male/female worker in prediction set had an evaluator who described their performance as indicative of poor math and science skills, what do you think is the percent chance that your male/female worker is overconfident because they predicted that their evaluator did NOT describe their performance as indicative of poor math and science skills?**

Percent chance your male worker in this prediction set is overconfident:

[   ]

Percent chance your female worker in this prediction set is overconfident:

[   ]

<u>**Underconfidence Prediction**</u>**: If your male/female worker in this prediction set had an evaluator who did NOT describe their performance as indicative of poor math and science skills, what do you think is the percent chance that your male/female worker is underconfident because they predicted that their evaluator described their performance as indicative of poor math and science skills?**

Percent chance your male worker in this prediction set is underconfident:

[   ]

Percent chance your female worker in this prediction set is underconfident:

[   ]

I.5.3   Instructions for the *Strategic Incentives* Treatment of the *Evaluator (Extended) Study*

For the *Evaluator (Extended) Study*, the *Strategic Incentives* treatment differs from the *Baseline* treatment (Section I.5.1) only in that participants are instead asked about workers who face strategic incentives.

After consenting to participate in the study, each participant is informed of the $3 study completion fee and of the opportunity to earn additional payment. Figures I.5.1 (shown above) and I.5.16 (below) show the overview participants who are randomized to evaluate **female workers** are given and the corresponding comprehension questions they must answer correctly in order to proceed. Then, participants provide their prior beliefs (Figure I.5.17), are provided with information on 20 female workers' self-evaluations and asked to provide their posterior beliefs (see new Figures I.5.18 and I.5.19 below for additional instructions and an example). Participants then are provided with additional instructions and asked to provide their posterior belief about the average self-evaluation of female workers (Figures I.5.20 and I.5.21) and are asked to provide their overconfidence and underconfidence beliefs (FiguresI.5.22 and I.5.23). Finally, all participants complete a follow-up survey that collects additional control and demographic information.

For evaluators who are randomized to be asked about **male workers**, "female" is replaced by "male" everywhere, and the self-evaluation information provided in Figure I.5.21 changes from 74% to **57%**.

Additional Instructions

In each of the next predictions, you will be informed of the prediction made by your worker when that worker was asked to make a prediction about their own performance.

Worker Predictions:
After completing the math and science test, workers made predictions about their performance on the math and science test. When making these predictions, they were not given any information on their own performance and knew that an employer would decide whether to hire them after the employer learned one of their predictions. They knew that the employer would only learn this prediction before deciding to hire them or not—the employer would not learn any demographic information about the worker and would not learn the worker's true score on the math and science test.

Specifically, a worker knew that:
- If their employer chooses NOT to hire them, then the worker would earn 50 cents and their employer would earn 50 cents.
- If their employer chooses to hire them, then the worker would earn 100 cents and their employer would earn 10 cents times the number of questions that the worker answered correctly on the test.

Understanding Question: A worker should expect to earn more...

if they provided more accurate predictions

if they got hired (regardless of how accurate their predictions were)

Figure I.5.17: Prior Belief about Female Workers, *Strategic Incentives* Treatment of the *Evaluator (Extended) Study*

Prediction X out of 23:

In each prediction, please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

What do you think is the percent chance that your female worker in this prediction had an evaluator who described her performance as indicative of poor math and science skills?

Figure I.5.18: Additional Instructions about Worker-Specific Posterior Belief about Female Workers, *Strategic Incentives* Treatment of the *Evaluator (Extended) Study*

Additional Instructions

In each of the next predictions, you will be informed of the prediction made by your worker when that worker was asked to make a prediction about their own performance.

Worker Predictions:
After completing the math and science test, workers made predictions about their performance on the math and science test. When making these predictions, they were not given any information on their own performance and knew that an employer would decide whether to hire them after the employer learned one of their predictions. They knew that the employer would only learn this prediction before deciding to hire them or not—the employer would not learn any demographic information about the worker and would not learn the worker's true score on the math and science test.

Specifically, a worker knew that:
    - If their employer chooses NOT to hire them, then the worker would earn 50 cents and their employer would earn 50 cents.
    - If their employer chooses to hire them, then the worker would earn 100 cents and their employer would earn 10 cents times the number of questions that the worker answered correctly on the test.

Understanding Question: A worker should expect to earn more...

if they provided more accurate predictions

if they got hired (regardless of how accurate their predictions were)

90

Figure I.5.19: Worker-Specific Posterior Belief about Female Workers, *Strategic Incentives* Treatment of the *Evaluator (Extended) Study*

Prediction X out of 23:
In each prediction, please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test, your female worker in this prediction predicted that there is a 50% chance that her evaluator described her performance as indicative of poor math and science skills.

What do you think is the percent chance that your female worker in this prediction had an evaluator who described her performance as indicative of poor math and science skills?

```
┌─────────────────┐
│                 │
└─────────────────┘
```

Figure I.5.20: Additional Instructions about Posterior Belief about Female Workers, *Strategic Incentives* Treatment of the *Evaluator (Extended) Study*

Additional Instructions

In the next prediction, rather than being informed of the prediction made by your worker in that prediction, you will be informed of the average prediction made by all of the female workers who could be randomly selected to be your worker in that prediction.

Figure I.5.21: Posterior Belief about Female Workers, *Strategic Incentives* Treatment of the *Evaluator (Extended) Study*

Prediction X out of 23:
In each prediction, please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test, 74% of female workers predicted that their evaluator described their performance as indicative of poor math and science skills.

What do you think is the percent chance that your female worker in this prediction had an evaluator who described her performance as indicative of poor math and science skills?

```
┌─────────────────┐
│                 │
└─────────────────┘
```

Figure I.5.22: Additional Information about Over/Underconfidence Beliefs about Female Workers, *Strategic Incentives* Treatment of the *Evaluator (Extended) Study*

Additional Instructions

In the next prediction, we will ask you to make predictions about the percent chance that your worker is overconfident or underconfident, depending on that worker's performance. Specifically, we will ask you to make one guess for each performance that your worker could have had. **For determining how much money you earn from the next prediction, we will then only consider your guess that corresponds to the prediction that your worker actually made.**

Figure I.5.23: Over/Underconfidence Beliefs about Female Workers, *Strategic Incentives* Treatment of the *Evaluator (Extended) Study*
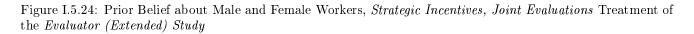
Prediction X out of 23:

In each prediction, please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

Overconfidence Prediction: If your female worker in this prediction had an evaluator who described her performance as indicative of poor math and science skills, what do you think is the percent chance that she is overconfident because she predicted that her evaluator did NOT describe her performance as indicative of poor math and science skills?

[   ]

Underconfidence Prediction: If your female worker in this prediction had an evaluator who did NOT describe her performance as indicative of poor math and science skills, what do you think is the percent chance that she is underconfident because she predicted that her evaluator described her performance as indicative of poor math and science skills?

[   ]

I.5.4 Instructions for the *Joint Evaluations, Strategic Incentives* Treatment of the *Evaluator (Extended) Study*

The *Joint Evaluations, Strategic Incentives* treatment differs from the *Joint Evaluations* treatment (Section I.5.2) in the same way as the *Strategic Incentives* treatment (Section I.5.3) differs from the *Baseline* treatment (Section I.5.1). Participants are asked about workers who face strategic incentives and are asked about both a male worker and a female worker on each decision screen. See Figures I.5.24, I.5.25, I.5.26, and I.5.27 for the prior belief, worker-specific posterior belief, posterior belief about average self-evaluations, and overconfidence and underconfidence beliefs questions about **female workers** and **male workers**, respectively.

Figure I.5.24: Prior Belief about Male and Female Workers, *Strategic Incentives, Joint Evaluations* Treatment of the *Evaluator (Extended) Study*

Prediction X out of 23:

In each prediction, please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

What do you think is the percent chance that your female worker in this prediction had an evaluator who described her performance as indicative of poor math and science skills?

[   ]

What do you think is the percent chance that your male worker in this prediction had an evaluator who described his performance as indicative of poor math and science skills?

[   ]

Figure I.5.25: Worker-Specific Posterior Belief about Male and Female Workers, *Strategic Incentives, Joint Evaluations* Treatment of the *Evaluator (Extended) Study*

<u>**Prediction X out of 23**</u>:

In each prediction, please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test:

- **your female worker** in this prediction set predicted that there is a **90%** chance that her evaluator described her performance as indicative of poor math and science skills, and
- **your male worker** in this prediction set predicted that there is a **80%** chance that his evaluator described his performance as indicative of poor math and science skills.

For your workers in this prediction set, what do you think is the percent chance that their evaluator described their performance as indicative of poor math and science skills?

Percent chance your female worker in this prediction set had an evaluator who described her performance as indicative of poor math and science skills:

[   ]

Percent chance your male worker in this prediction set had an evaluator who described his performance as indicative of poor math and science skills:

[   ]

Figure I.5.26: Posterior Belief about Male and Female Workers, *Strategic Incentives, Joint Evaluations* Treatment of the *Evaluator (Extended) Study*

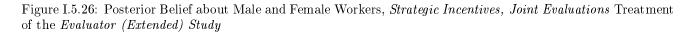<u>**Prediction Set X out of 13**</u>:

In each prediction, please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

After completing the math and science test:

- **57% of male workers predicted** that their evaluator described their performance as indicative of poor math and science skills, and
- **74% of female workers predicted** that their evaluator described their performance as indicative of poor math and science skills.

For your workers in prediction set, what do you think is the percent chance that their evaluator described their performance as indicative of poor math and science skills?

Percent chance your male worker in this prediction set had an evaluator who described his performance as indicative of poor math and science skills:

┌─────────────────┐
│   │
└─────────────────┘

Percent chance your female worker in this prediction set had an evaluator who described her performance as indicative of poor math and science skills:

┌─────────────────┐
│   │
└─────────────────┘

Figure I.5.27: Over/Underconfidence Beliefs about Male and Female Workers, *Strategic Incentives, Joint Evaluations* Treatment of the *Evaluator (Extended) Study*

Prediction X out of 23:
In each prediction, please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

Overconfidence Prediction: If your female/male worker in prediction set had an evaluator who described their performance as indicative of poor math and science skills, what do you think is the percent chance that your female/male worker is overconfident because they predicted that their evaluator did NOT describe their performance as indicative of poor math and science skills?

Percent chance your female worker in this prediction set is overconfident:

[ ]

Percent chance your male worker in this prediction set is overconfident:

[ ]

Underconfidence Prediction: If your female/male worker in this prediction set had an evaluator who did NOT describe their performance as indicative of poor math and science skills, what do you think is the percent chance that your female/male worker is underconfident because they predicted that their evaluator described their performance as indicative of poor math and science skills?

Percent chance your female worker in this prediction set is underconfident:

[ ]

Percent chance your male worker in this prediction set is underconfident:

[ ]

97

I.6 Full Instructions for the *Evaluator (Additional Demographics) Study*

All participants in this study are randomized to be asked about male or female workers.

After consenting to participate in the study, each participant is informed of the $1.50 study completion fee and of the opportunity to earn additional payment. Figures I.6.1-I.6.3 show the overview participants who are randomized to evaluate **female workers** are given and the corresponding comprehension questions they must answer correctly in order to proceed. Then, participants provide their prior beliefs (Figure I.6.4), posterior beliefs (Figure I.6.5), and their overconfidence and underconfidence beliefs (Figure I.6.6). Finally, all participants take a short survey of five randomized bonus questions, as shown above in Figures G.1.7-G.1.12, and a follow-up survey that collects additional control and demographic information.

For participants who are randomized to be asked about **male workers**, "female" is replaced by "male" everywhere, and the self-evaluation information provided in Figure I.6.5 changes from 68% to **38%**.

Main Instructions (Page 1 out of 2)

Overview:
This study will consist of 3 predictions and a short follow-up survey. For completing this study, you are guaranteed to receive $1.50 within 24 hours. In addition, any additional payment you earn will be distributed to you as a bonus payment.

The Workers:
In a prior study, an approximately equal number of men and women (called "workers") completed a math and science test with 10 questions. Each question tested their math and science skills by asking them about general science, arithmetic reasoning, math knowledge, mechanical comprehension, and assembling objects. Performance on questions like these is often used as a measure of cognitive ability by academic researchers. A worker's score on the test equals the number of questions they answer correctly, and a worker earns 10 cents times their score.

Your Predictions:
You will be asked to make 3 predictions related to the **performance of workers on the math and science test**.

To maximize your chance of earning an additional payment of $1, you should provide your most-accurate guess when making each prediction. This is because each prediction will ask you to guess the percent chance of of some outcome being true. In each of those predictions, to secure the largest chance of earning $1 from the prediction, you should report your most-accurate guess. To learn the precise payment rule that determines how much you earn from these predictions click here.

One of your 3 predictions will be randomly selected as the prediction-that-counts, and your additional payment will equal the amount you earn in the prediction-that-counts.

Understanding Question: To maximize your chance of additional bonus payment, how should you make predictions in this part?

It doesn't matter

As accurately as possible

Randomly

Understanding Question: How much additional payment will you receive?

I will receive what I earn from all predictions in this part

I will receive what I earn from the prediction-that-counts

Nothing

Figure I.6.2: Instructions about Female Workers, *Evaluator (Additional Demographics) Study*

<u>**Main Instructions (Page 2 out of 2)**</u>

In each prediction, we will ask you to make a prediction about the performance of your worker. Below please learn more about your workers and the types of predictions we will ask you to make about your workers.

<u>**Your workers**</u>
In each prediction, your worker will be randomly selected from the group of workers who:

- **work full time,**
- **are between 26 and 40 years old**
- **live in the Southern region of the United States,**
- **have at least completed some college education, and**
- **are female.**

Your worker in one prediction will never be the same as your worker in another prediction. **Thus, you will never be asked about the same worker twice.**

Figure I.6.3: Instructions about Female Workers cont., *Evaluator (Additional Demographics) Study*

Types of Predictions

In each prediction, you will be asked to predict the **percent chance that some outcome is true.** Sometimes, you will be asked to predict the percent chance that your worker in that prediction had an evaluator who described their performance as indicative of poor math and science skills. Other times, you will be asked to predict the percent chance that your worker is overconfident or underconfident when asked to make predictions about their own performance. Thus, please note the following:

- Each worker has an evaluator who is randomly selected from the set of all other workers who complete the study and who is equally likely to be a man or a woman. The evaluators answered a question about which scores they believed were indicative of poor math and science skills.
- An evaluator is said to have described a worker's performance as indicative of poor math and science skills if they indicated the worker's score was indicative of poor math and science skills.
- Thus, when an evaluator chooses how to describe a worker's performance, the evaluator effectively knows how many questions the worker got right but does not know anything else about the worker, such as the worker's gender.

Worker Predictions:

In some predictions, you will be informed of the average prediction made by the group of workers who could be randomly selected to be your worker in that prediction.

Thus please note that: after completing the math and science test, workers made predictions about their performance on the math and science test. When making these predictions, they were not given any information on their own performance and knew that they that should report their most-accurate guess to maximize their chance of earning an additional bonus payment of $1.

Figure I.6.4: Prior Belief about Female Workers, *Evaluator (Additional Demographics) Study*

Prediction 1 out of 3:

Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your worker will be a randomly selected from the group of female workers who got 5 questions right on the math and science test.

What do you think is the percent chance that your worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills?

☐

Figure I.6.5: Posterior Belief about Female Workers *Evaluator (Additional Demographics) Study*
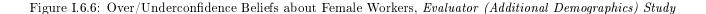
Prediction 2 out of 3:

Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your worker will again be a randomly selected from the group of female workers who got 5 questions right on the math and science test.

After completing the math and science test, **68% of female workers who got 5 questions right predicted** that they had an evaluator who described their performance as indicative of poor math and science skills.

What do you think is the percent chance that your worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills?

> [   ]

Figure I.6.6: Over/Underconfidence Beliefs about Female Workers, *Evaluator (Additional Demographics) Study*

Prediction 3 out of 3:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your worker will again be a randomly selected from the group of female workers who got 5 questions right on the math and science test.

Below, we will ask you to make predictions about the percent chance that your worker is overconfident or underconfident, depending on that worker's performance. Specifically, we will ask you to make one guess for each performance that your worker could have had. **For determining how much money you earn from this prediction, we will then only consider your guess that corresponds to the prediction that your worker actually made.**

Overconfidence Prediction: If your worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills, what do you think is the percent chance that your worker is overconfident because they predicted that they had an evaluator who did NOT describe their performance as indicative of poor math and science skills?

☐

Underconfidence Prediction: If your worker in this prediction had an evaluator who did NOT describe their performance as indicative of poor math and science skills, what do you think is the percent chance that your worker is underconfident because they predicted that they had an evaluator who described their performance as indicative of poor math and science skills?

☐

I.7 Full Instructions for the *Evaluator (Known Performance) Study*

All participants in this study are randomized to be asked about male or female workers.

After consenting to participate in the study, each participant is informed of the $1.50 study completion fee and of the opportunity to earn additional payment. Previous Figure I.6.1 and Figures I.7.1-I.7.3 show the overview participants who are randomized to evaluate **female workers** are given and the corresponding comprehension questions they must answer correctly in order to proceed. Then, participants provide their prior beliefs (Figure I.7.4), posterior beliefs (Figure I.7.5), and their overconfidence and underconfidence beliefs (see Figure I.7.6). Finally, all participants take a short survey of five randomized bonus questions, as previously shown in Figures G.1.7-G.1.12, and a follow-up survey that collects additional control and demographic information.

For participants who are randomized to be asked about **male workers**, "female" is replaced by "male" everywhere, and the self-evaluation information provided in Figure I.7.5 changes from 68% to **41%**.

Figure I.7.1: Study Overview, *Evaluator (Known Performance) Study*

In each prediction, we will ask you to make a prediction about the performance of your worker. Below please learn more about your workers and the types of predictions we will ask you to make about your workers.

Your workers
In each prediction, your worker will be randomly selected from the group of all female workers who got 5 questions right on the math and science test.

Your worker in one prediction will never be the same as your worker in another prediction. **Thus, you will never be asked about the same worker twice.**

Figure I.7.2: Instructions about Female Workers, *Evaluator (Known Performance) Study*

Types of Predictions

In each prediction, you will be asked to predict the **percent chance that some outcome is true.** Sometimes, you will be asked to predict the percent chance that your worker in that prediction had an evaluator who described their performance as indicative of poor math and science skills. Other times, you will be asked to predict the percent chance that your worker is overconfident or underconfident when asked to make predictions about their own performance. Thus, please note the following:

- Each worker has an evaluator who is randomly selected from the set of all other workers who complete the study and who is equally likely to be a man or a woman. The evaluators answered a question about which scores they believed were indicative of poor math and science skills.
- An evaluator is said to have described a worker's performance as indicative of poor math and science skills if they indicated the worker's score was indicative of poor math and science skills.
- Thus, when an evaluator chooses how to describe a worker's performance, the evaluator effectively knows how many questions the worker got right but does not know anything else about the worker, such as the worker's gender.

Worker Predictions:

In some predictions, you will be informed of the average prediction made by the group of workers who could be randomly selected to be your worker in that prediction.

Thus please note that: after completing the math and science test, workers made predictions about their performance on the math and science test. When making these predictions, they were not given any information on their own performance and knew that they that should report their most-accurate guess to maximize their chance of earning an additional bonus payment of $1.

Figure I.7.3: Comprehension Questions about Female Workers, *Evaluator (Known Performance) Study*

Understanding Question: When an evaluator describes a worker's performance, do they know the gender of the worker?

Yes

No

Understanding Question: A worker should expect to earn more...

if they provided more accurate predictions

if they got hired (regardless of how accurate their predictions were)

Figure I.7.4: Prior Belief about Female Workers, *Evaluator (Known Performance) Study*

Prediction 1 out of 3:
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your worker will be a randomly selected from the group of female workers who got 5 questions right on the math and science test.

What do you think is the percent chance that your worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills?

Figure I.7.5: Posterior Belief about Female Workers, *Evaluator (Known Performance) Study*

Prediction 2 out of 3:

Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your worker will again be a randomly selected from the group of female workers who got 5 questions right on the math and science test.

After completing the math and science test, **68% of female workers who got 5 questions right predicted** that they had an evaluator who described their performance as indicative of poor math and science skills.

What do you think is the percent chance that your worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills?

Figure I.7.6: Over/Underconfidence Beliefs about Female Workers, *Evaluator (Known Performance) Study*

<u>**Prediction 3 out of 3:**</u>
Please provide an integer answer (from 0 to 100) and please omit the percent sign in your answer. For example, please type 0 if your answer is 0%, 100 if your answer is 100%, etc.

In this prediction, your worker will again be a randomly selected from the group of female workers who got 5 questions right on the math and science test.

Below, we will ask you to make predictions about the percent chance that your worker is overconfident or underconfident, depending on that worker's performance. Specifically, we will ask you to make one guess for each performance that your worker could have had. **For determining how much money you earn from this prediction, we will then only consider your guess that corresponds to the prediction that your worker actually made.**

<u>Overconfidence Prediction</u>: **If your worker in this prediction had an evaluator who described their performance as indicative of poor math and science skills, what do you think is the percent chance that your worker is overconfident because they predicted that they had an evaluator who did NOT describe their performance as indicative of poor math and science skills?**

```
┌─────────────────┐
│                 │
└─────────────────┘
```

<u>Underconfidence Prediction</u>: **If your worker in this prediction had an evaluator who did NOT describe their performance as indicative of poor math and science skills, what do you think is the percent chance that your worker is underconfident because they predicted that they had an evaluator who described their performance as indicative of poor math and science skills?**

```
┌─────────────────┐
│                 │
└─────────────────┘
```