Online Appendices for *What can sexual orientation reveal about the gender gap in confidence?*

by Bénédicte Droz

May 1, 2025

A.1 Questionnaire

Figure A.1: Welcome and Consent

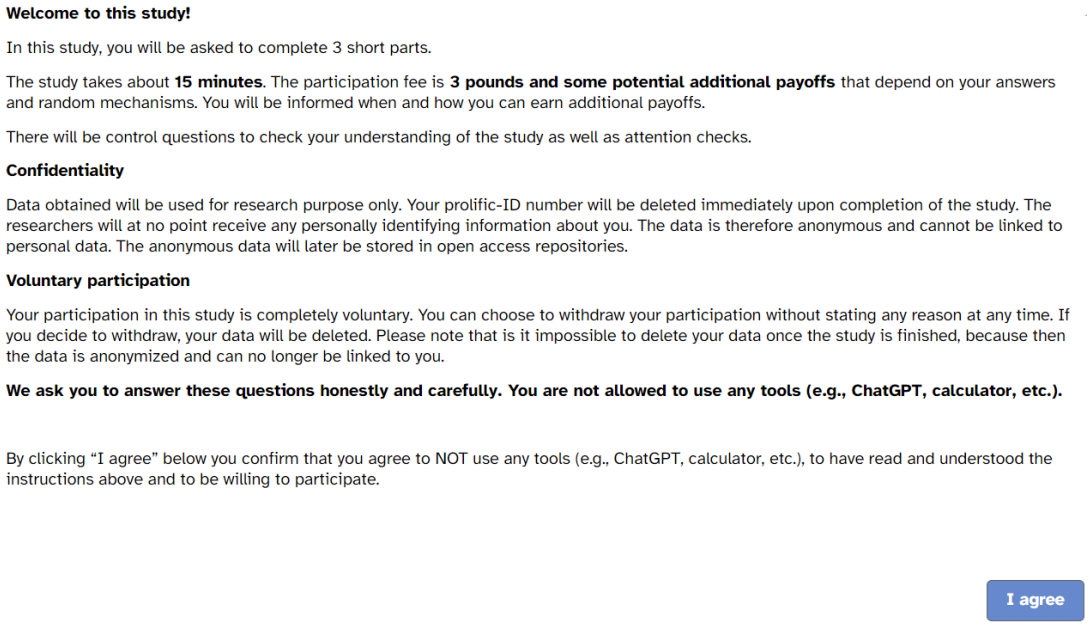


Figure A.2: Prolific ID

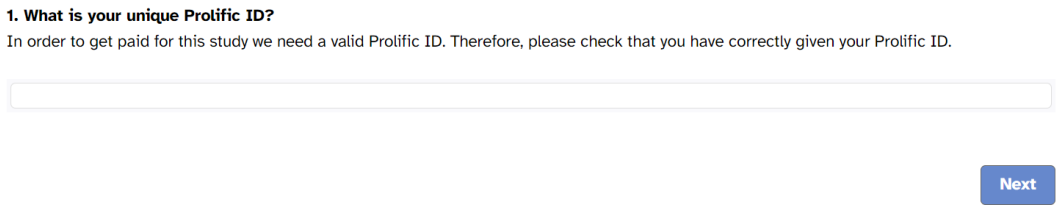


Figure A.3: Part 1 – Instructions

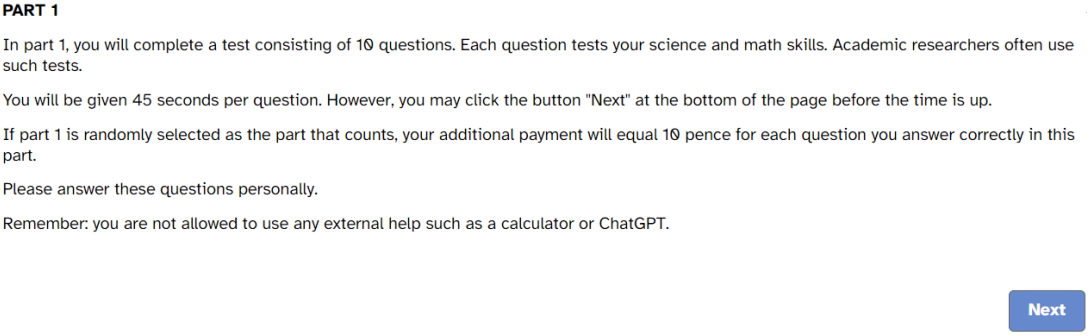


Figure A.4: Example Question

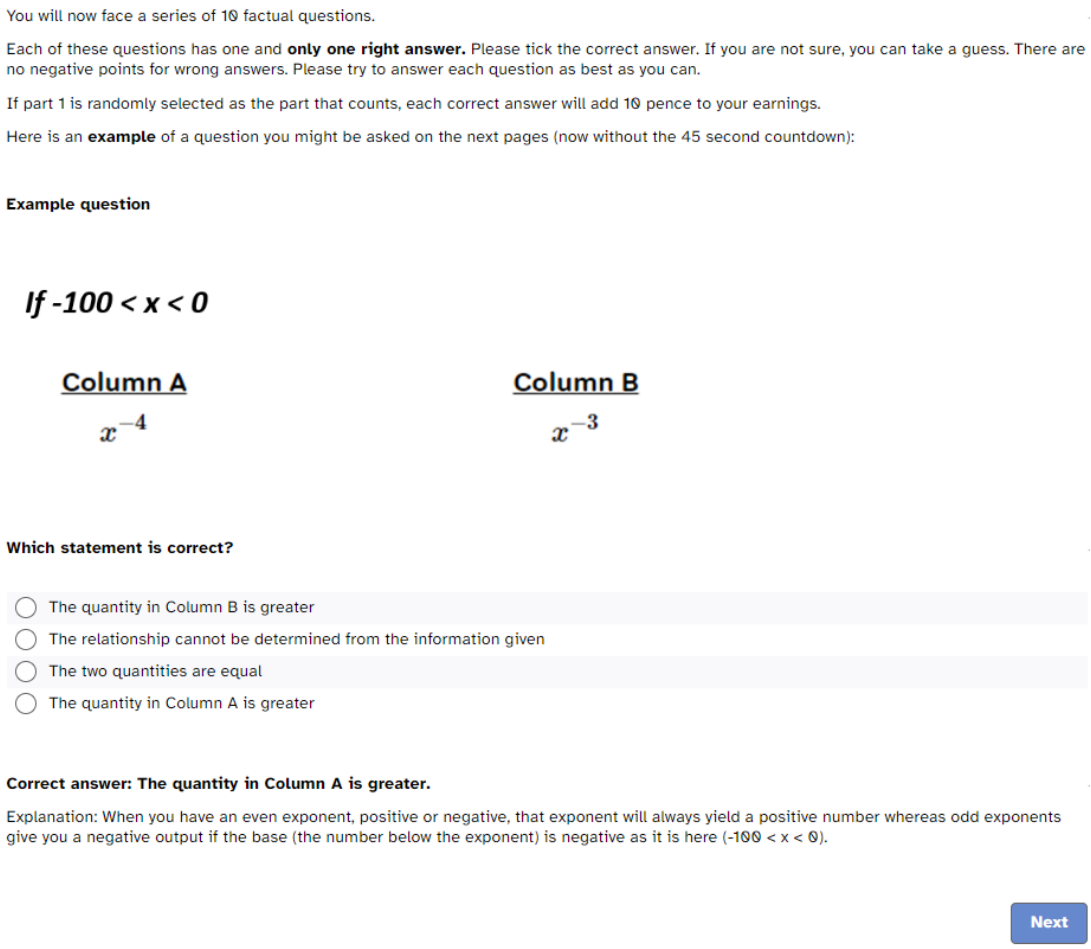
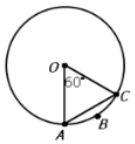


Figure A.5: Math and Science Quiz – 10 Questions



O is the center of the circle with radius 6.

Column A	Column B
Length of arc ABC	6

Question 1

Which statement is correct?

- ◯ The two quantities are equal
- ◯ The quantity in Column A is greater
- ◯ The relationship cannot be determined from the information given
- ◯ The quantity in Column B is greater

Remaining time: 0:39

Next

The revenue generated by Firm X is divided between Mary and Peter in a 6 to 5 ratio respectively.

Column A	Column B
Mary's share when the revenue generated by Firm X is 15'700 $	7'900 $

Question 2

Which statement is correct?

- ◯ The two quantities are equal
- ◯ The relationship cannot be determined from the information given
- ◯ The quantity in Column A is greater
- ◯ The quantity in Column B is greater

Remaining time: 0:42

Next

Question 3

In how many different ways can 3 identical green shirts and 3 identical red shirts be distributed among 6 children such that each child receives a shirt?

- 40
- 20
- 216
- 720

Remaining time: 0:16

Next

Question 4

The price of a pair of sneakers was $80 for the last six months of last year. On January first, the price increased 20%. After the price increase, an employee bought these sneakers with a 10% employee discount. What price did the employee pay?

- 86.40
- 88
- 83.33
- 82.00

Remaining time: 0:45

Next

The greatest prime factor of 144 is x

The greatest prime factor of 96 is y

Column A	Column B
x	y

Question 5

Which statement is correct?

- The quantity in Column B is greater
- The relationship cannot be determined from the information given
- The two quantities are equal
- The quantity in Column A is greater

Remaining time: 0:42

Next

ANIMAL DISTRIBUTION IN THE ZOO

Animal	Percent
Lions	32%
Leopards	16%
Ocelots	20%
Tigers	8%
Bobcats	24%

Question 6

If there are 44 leopards at the zoo, what is the zoo's total animal population?

- ◯ 225
- ◯ 350
- ◯ 325
- ◯ 275

Remaining time: 0:41

Next

STUDENT ENROLLMENT AT A SMALL COLLEGE

Distribution of Enrollment by Class and Gender
Total Enrollment: 1,400

Class	Males	Females
Freshmen	303	259
Sophomores	215	109
Juniors	182	88
Seniors	160	84
Total	860	540

Percent of Total Enrollment Majoring in Selected Academic Areas

Area	Percent
Humanities	33%
Social sciences	30%
Physical sciences	24%

Note: No student is majoring in more than one area.

Question 7

The ratio of the number of male freshmen to the number of female sophomores is approximately...

- ◯ 2 to 1
- ◯ 3 to 1
- ◯ 4 to 1
- ◯ 3 to 2

Remaining time: 0:34

Next

STUDENT ENROLLMENT AT A SMALL COLLEGE

Distribution of Enrollment by Class and Gender
Total Enrollment: 1,400

Class	Males	Females
Freshmen	303	259
Sophomores	215	109
Juniors	182	88
Seniors	160	84
Total	860	540

Percent of Total Enrollment Majoring in Selected Academic Areas

Area	Percent
Humanities	33%
Social sciences	30%
Physical sciences	24%

Note: No student is majoring in more than one area.

Question 8

How many students are either juniors or males or both?

- 1130
- 948
- 766
- 678

Remaining time: 0:38

Next

Question 9

Tony is four years younger than his brother Josh, and two years older than his sister Cindy. Tony also has a twin brother, Evan. All the ages of the children totaled together is 66. How old is Tony?

- 22
- 16
- 20
- 14

Remaining time: 0:41

Next

Question 10

The relationship between temperature C, in degrees Celsius, and temperature F, in degrees Fahrenheit, is given by the formula F= 9/5 x C + 32
If a recipe calls for an oven temperature of 210 degrees Celsius, what is the oven temperature in degrees Fahrenheit?
(x means "times" or "multiplied by")

- 320
- 350
- 410
- 420

Remaining time: 0:39

Next

Notes: Quiz questions taken from the GRE practice questions (magoosh.com, prepscholar.com, manhattanreview.com and the Book "Official GRE Quantitative Reasoning Practice Questions", Volume 1, Second Edition).

Figure A.6: Belief

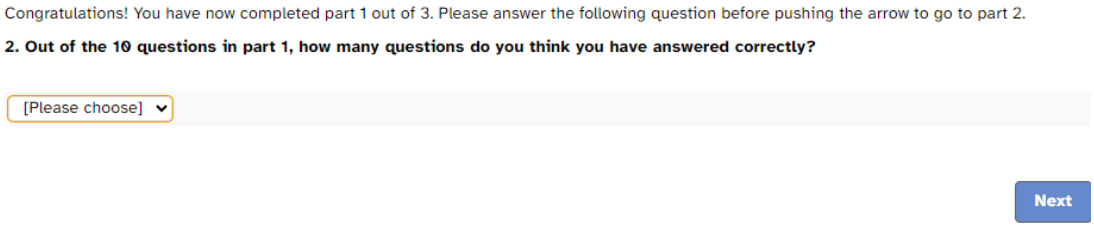


Figure A.7: Part 2 – Instructions

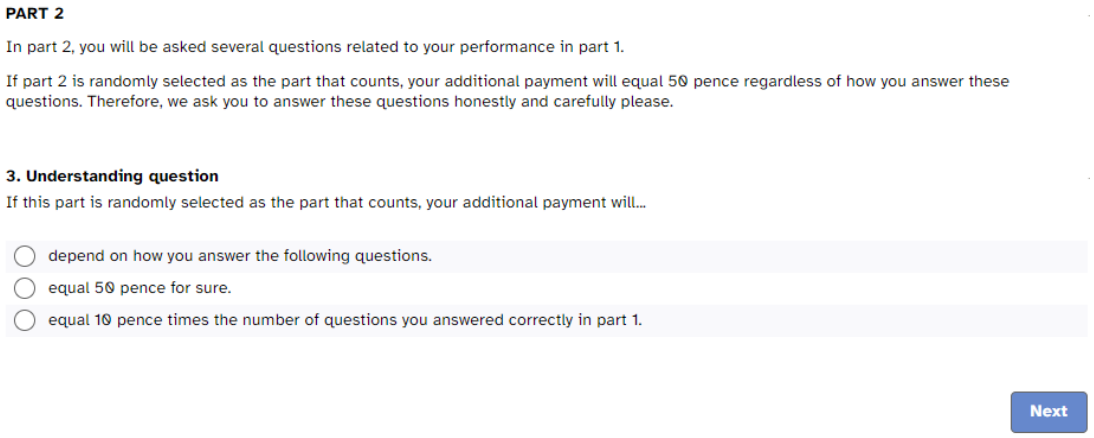


Figure A.8: Uninformed Self-Evaluations

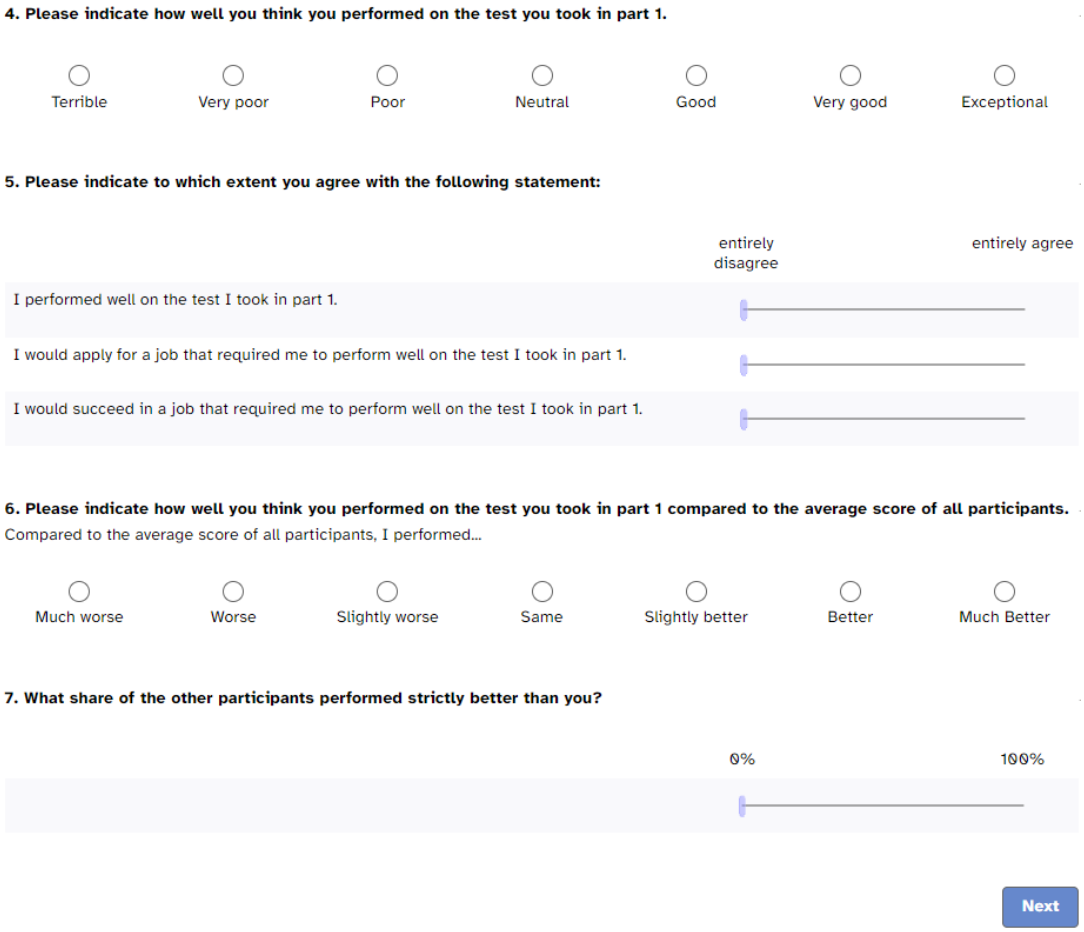


Figure A.9: Own Performance Revealed

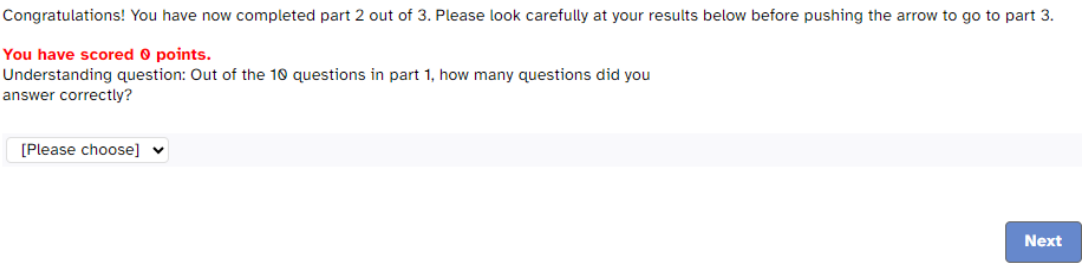


Figure A.10: Part 3 – Instructions

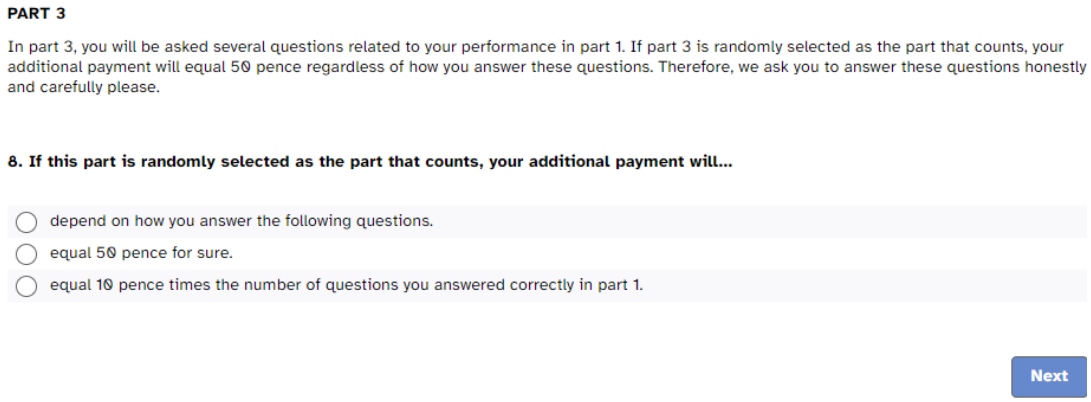


Figure A.11: Informed Self-Evaluations

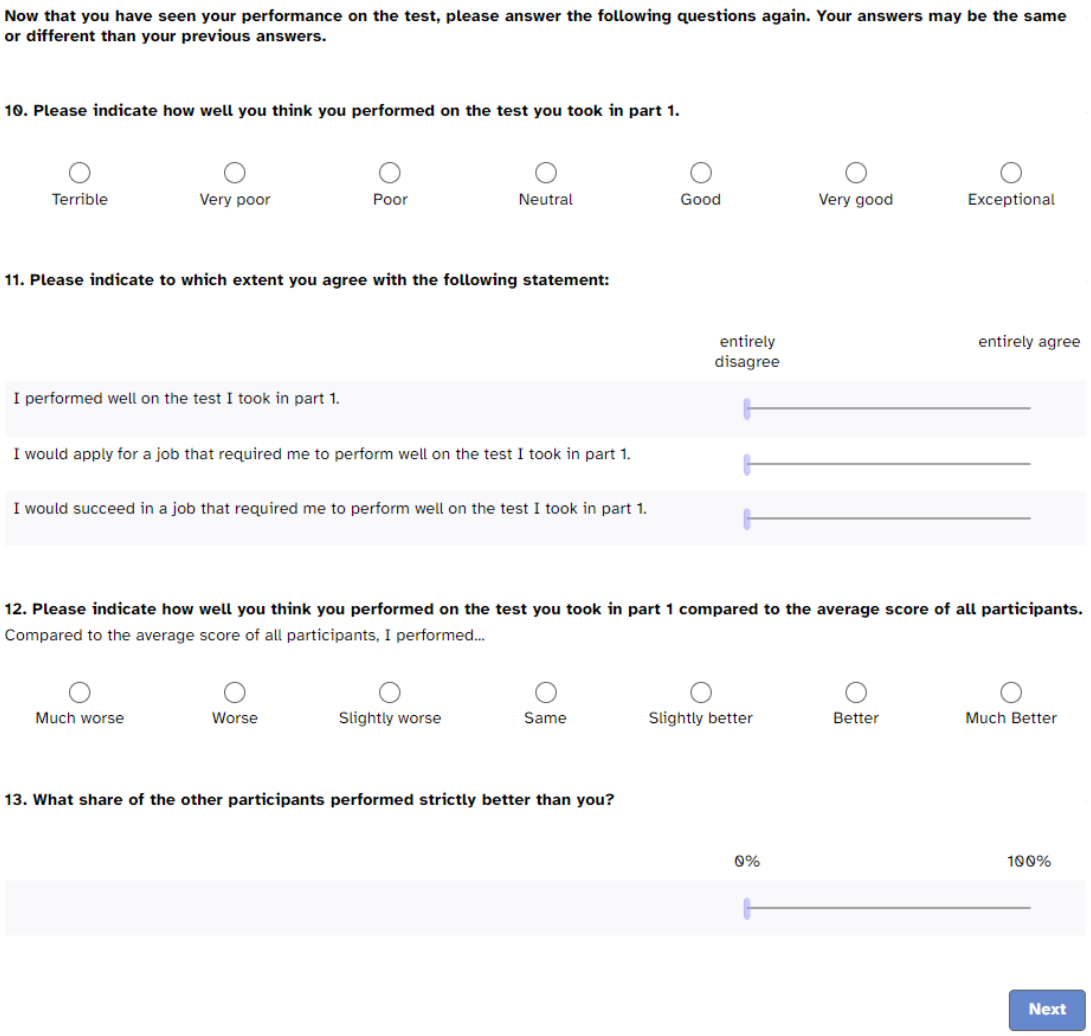


Figure A.12: Sociodemographics

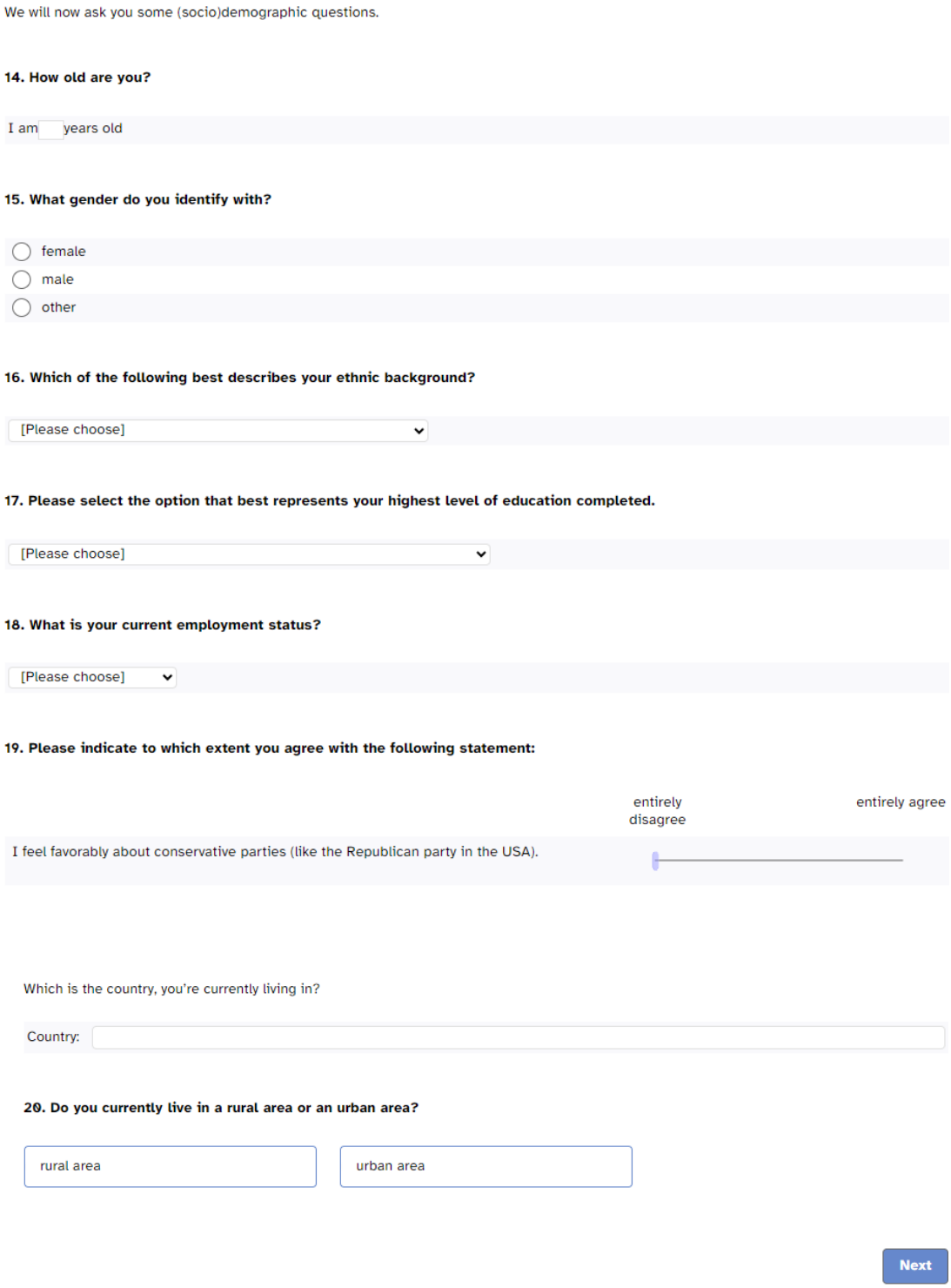


Figure A.13: Payoffs – Option 1

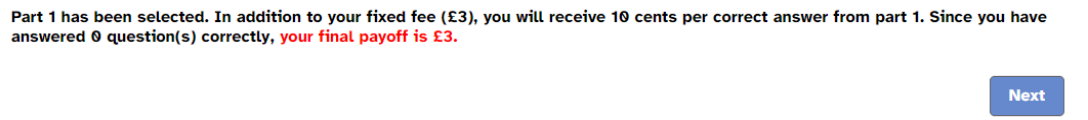


Figure A.14: Payoffs – Option 2

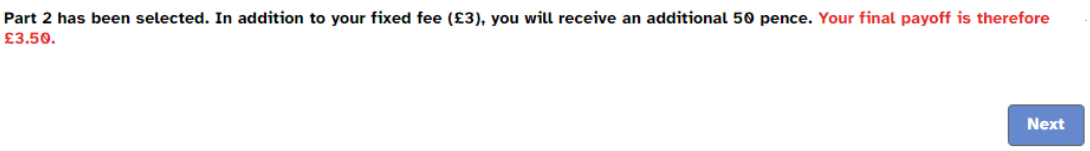


Figure A.15: Payoffs – Option 3

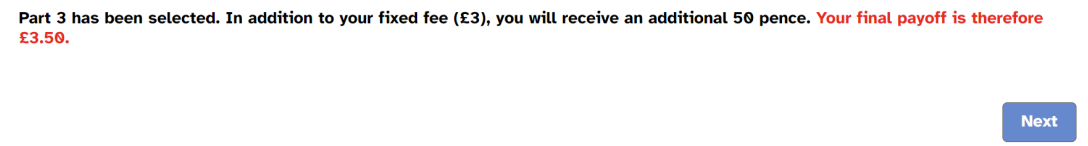


Figure A.16: Attention Check 2

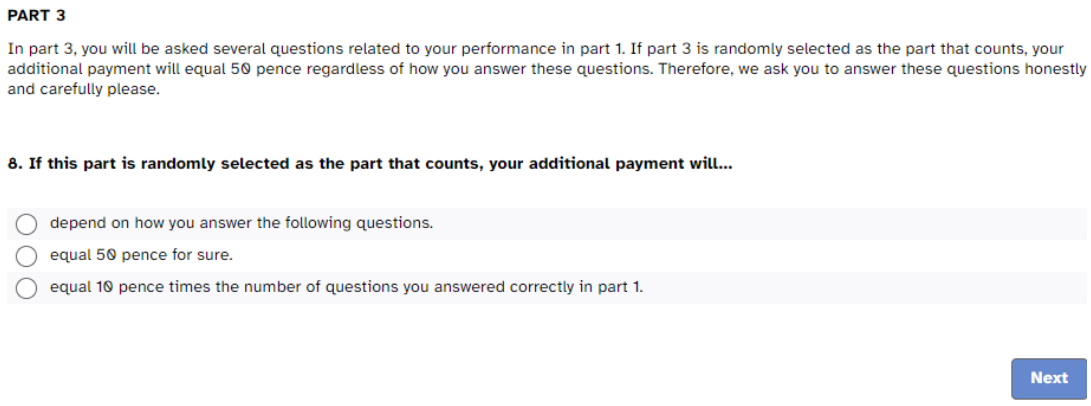


Figure A.17: Follow-up Attention Check 2

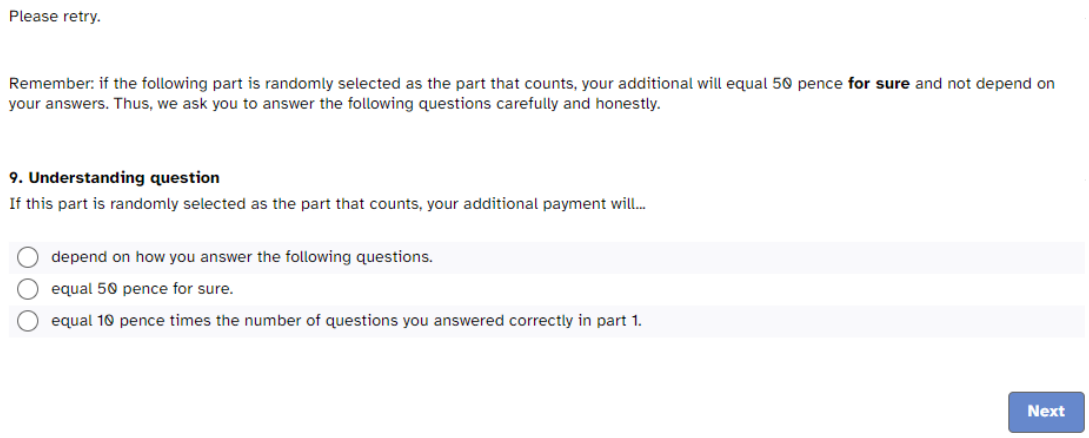
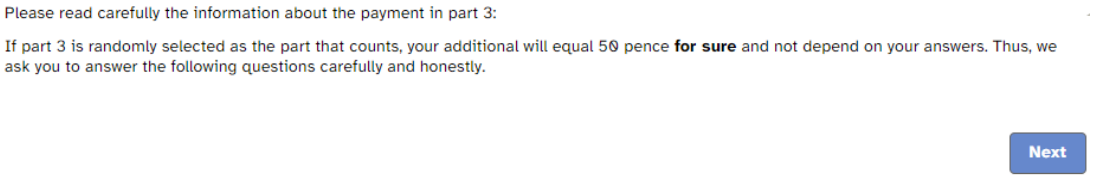
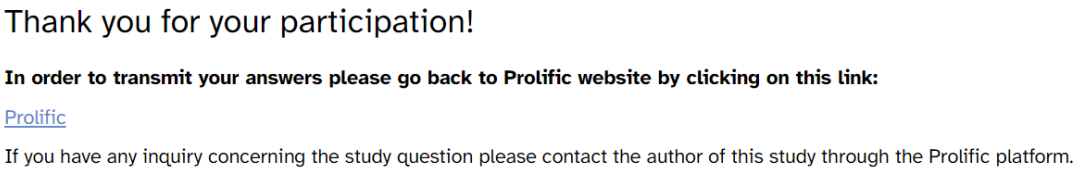


Figure A.18: Conclusion Attention Check 2



1

Figure A.19: Thanks



[1] *Attention check 1, Follow-up attention check 1* and *Conclusion attention check 1* are exactly the same but for Part 2 (instead of Part 3).

A.2 Homosexual vs. Heterosexual

To conduct separate analyses for women (comparing heterosexual women to lesbian women) and men (comparing heterosexual men to gay men), I created the dummy *lesbian woman* (*gay man)* which takes the value of 1 if the person was pre-screened as a lesbian woman (gay man) from the online platform and 0 if the person was pre-screened as heterosexual woman (heterosexual man). Model 1 (*Knowledge Score*) in Table A.1 shows that there is no statistically significant difference in knowledge scores between lesbian women and heterosexual women (Panel A); both groups perform similarly on the 10-question quiz. Same holds for the comparison between gay men and heterosexual men in Panel B. Model 2 (*Belief*) shows that equally-performing lesbian women and heterosexual women assess their exact number of correct answers similarly ($p = 0.186$). Same holds for the comparison between gay men and heterosexual men in Panel B.[2]

Table A.1: Knowledge Score and Belief: Lesbian Women (Gay Men) vs. Hetero Women (Men)

	(1) Knowledge Score	(2) Belief
Panel A: Lesbian Women vs. Hetero Women		
lesbian woman	-0.143	0.260
	(0.199)	(0.196)
Hetero Female Average	3.37	2.50
Observations	359	359
Panel B: Gay Men vs. Hetero Men		
gay man	0.0334	0.0205
	(0.208)	(0.183)
Hetero Male Average	3.73	3.32
Observations	360	360
Controls	Yes	Yes

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Rightwing, Ethnicity, Education, Employment Status, Age, and Rural as well as Knowledge Score for Model 2. *Hetero Female Average* and *Hetero Male Average* are the unadjusted means for heterosexual women (men), before controlling for covariates.

[2]Appendix A.3 of the manuscript shows that both lesbian women and heterosexual women are on average underconfident in their assessment, lesbian women being on average *less* underconfident but this difference is not statistically significant. Furthermore, Appendix A.4 of the manuscript shows that both gay men and heterosexual men are on average underconfident in their assessment, gay men being on average *more* underconfident but this difference is not statistically significant.

However, as Panel A in Table A.2 shows, equally-performing lesbian women and heterosexual women perceive their overall *performance* differently. Indeed, lesbian women provide answers which are 4.54 points higher than equally-performing heterosexual women when asked to agree on a scale from 0 to 100 to the following statement "I performed well on the test I took in part 1". Since the average female participant answers 18.91 out of 100 for this perceived performance question (see Table A.7 for the women's descriptive statistics), these 4.54 points represent 24% of the mean. Lesbian women further provide answers which are 5.908 points higher than equally-performing heterosexual women when asked to agree on a scale from 0 to 100 to the following statement "I would apply for a job that required me to perform well on the test I took in part 1". Since the average female participant answers 15.59 out of 100 for this question, these 5.908 points represent 37.9% of the mean. Similarly, for the statement, "I would succeed in a job that required me to perform well on the test I took in part 1", lesbian women score 6.112 points higher, which represents 32.65% of the mean score of 18.72. Although of a smaller size, I find similar results concerning *Performance bucket* (the coefficients for *Comparison* and BTY are not statistically significant). Interestingly, lesbian women do not tend to perceive their performance as superior to others more strongly compared to equally-performing heterosexual women (see BTY and *Comparison*); however, they do evaluate their performance more optimistically than equally-performing heterosexual women.

For the comparison among men, I do not find any statistically significant difference between gay men and heterosexual men in terms of uninformed self-evaluations as Panel B in Table A.2 shows.

Table A.2: Uninformed Self-Evaluations: Lesbian Women (Gay Men) vs. Hetero Women (Men)

	(1) Perceived Perf.	(2) Perf. Bucket	(3) Willing	(4) Success	(5) Comparison	(6) BTY
Panel A: Lesbian Women vs. Hetero Women						
lesbian woman	4.540**	0.218*	5.908**	6.112**	0.166	1.607
	(2.254)	(0.129)	(2.899)	(2.959)	(0.133)	(2.850)
Hetero Female Average	17.26	2.20	13.38	17.09	2.43	58.79
Observations	359	359	359	359	359	359
Panel B: Gay Men vs. Hetero Men						
gay man	-0.671	-0.0842	-0.398	-1.104	-0.0751	0.623
	(2.292)	(0.128)	(2.880)	(2.737)	(0.132)	(2.451)
Hetero Male Average	26.49	2.89	24.09	26.46	3.17	52.94
Observations	360	360	360	360	360	360
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls include Knowledge Score, Right-wing ideology, Ethnicity, Education, Employment Status, Age, and Rural residence. *Hetero Female Average* and *Hetero Male Average* represent unadjusted means for heterosexual women and men before controlling for covariates.

Result 1 (Self-evaluations (uninformed))**.** *There is a sexual orientation gap in self-evaluations.*

a. Lesbian women provide more optimistic self-evaluations of their (absolute) performance on a math and science quiz than equally-performing heterosexual women (when being uninformed of their true performance).

b. Gay men show no significant difference in self-evaluations of their performance on a math and science quiz compared to equally-performing heterosexual men (when being uninformed of their true performance).

One might argue that this result only holds when lesbian women are uninformed of their true performance. Upon learning their actual score, the difference in self-evaluations between equally-performing lesbian women and heterosexual women should disappear, as their self-evaluations should be based more on objective feedback rather than subjective beliefs.[3] Indeed, by informing participants of their true score, I mechanically eliminate

[3]Recall that participants receive feedback only on their own scores, not in relation to others. On average, lesbian participants might more strongly believe that they performed better than others or better than the average compared to equally-performing heterosexual women. Interestingly, the two variables capturing this comparison dimension (*Comparison* and *BTY*) remain not statistically significant after feedback. This suggests that lesbian women still do not compare themselves more positively to others and to the average than equally-performing heterosexual women after receiving feedback on their personal score. Yet, they remain more optimistic about their own performance.

any discrepancies in their beliefs about their absolute performance. However, Panel A in Table A.3 shows significant and substantial differences in self-evaluations even after their performance is revealed, suggesting that lesbian women maintain more optimistic self-evaluations than equally-performing heterosexual women. The gap in self-evaluations persists and is not primarily driven by any (potential) gap in beliefs about one's absolute score.

After receiving their absolute test scores, lesbian women rate themselves 6.044 points higher than equally-performing heterosexual women on the statement, "I performed well on the test I took in part 1". Since the average female participant answers 21.04 out of 100 for this perceived performance question, these 6.044 points represent 28.73% of the mean. Similarly, for the statement, "I would succeed in a job that required me to perform well on the test I took in part 1", lesbian women score 6.299 points higher, which represents 33.74% of the mean score of 18.67. I find similar results for *Performance bucket* (the coefficients for *Willing*, *Comparison* and *BTY* are not statistically significant). Again, lesbian women do not tend to perceive their performance as superior to others more strongly compared to equally-performing heterosexual women after feedback (see *BTY* and *Comparison*) but they evaluate themselves more optimistically than equally-performing heterosexual women even when their own absolute performance is revealed.

For the comparison among men, I do not find any statistically significant difference between gay men and heterosexual men in terms of informed self-evaluations as Panel B in Table A.3 shows.

Table A.3: Informed Self-Evaluations: Lesbian Women (Gay Men) vs. Hetero Women (Men)

	(1) Perceived Perf.	(2) Perf. Bucket	(3) Willing	(4) Success	(5) Comparison	(6) BTY
Panel A: Lesbian Women vs. Hetero Women						
lesbian woman	6.044***	0.371***	4.272	6.299**	0.181	-1.509
	(1.935)	(0.111)	(2.653)	(2.562)	(0.124)	(2.867)
Hetero Female Average	18.31	2.52	16.92	15.68	2.58	59.07
Observations	359	359	359	359	359	359
Panel B: Gay Men vs. Hetero Men						
gay man	-0.232	-0.0363	0.210	0.245	-0.0168	0.0182
	(1.824)	(0.0983)	(2.580)	(2.571)	(0.121)	(2.586)
Hetero Male Average	28.42	3.16	25.32	27.36	3.18	51.78
Observations	360	360	360	360	360	360
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls include Knowledge Score, Right-wing ideology, Ethnicity, Education, Employment Status, Age, and Rural residence. *Hetero Female Average* and *Hetero Male Average* represent unadjusted means for heterosexual women and men before controlling for covariates.

Result 2 (Self-evaluations (informed))**.** *The sexual orientation gap in self-evaluations persists even when informed of their own performance.*

a. Lesbian women still provide more optimistic self-evaluations of their (absolute) performance on a math and science quiz than equally-performing heterosexual women.

b. Gay men further show no significant difference in self-evaluations of their performance on a math and science quiz compared to equally-performing heterosexual men.

A.3 Women vs. Men

In addition to comparing lesbian women (gay men) with heterosexual women (heterosexual men), I also compare men to women. The results of this comparison – which replicates Exley and Kessler (2022)[4] – are displayed in this section. The results confirm the gender gap in confidence (and self-evaluations). Women report more pessimistic beliefs about their absolute performance than equally-performing men. Furthermore, women provide worse self-evaluations of their performance on a math and science quiz than equally-performing men (when being uninformed of their true performance). This gender gap in self-evaluations persists even when informed of their own performance. Women still provide worse self-evaluations of their performance on a math and science quiz than equally-performing men. However, it is important to note that my sample is evenly split between homosexual and heterosexual individuals, which is not representative of the real-world distribution. As a result, the gender gap in the general population is likely to be larger, given that the proportion of lesbian women among women is smaller in the broader population. Closer to reality, the first row (*Female*) in Table 1 of the manuscript shows the gender gap between heterosexual women and heterosexual men.

When comparing men and women Table A.4 shows that the gender gap in performance is clear and aligned with what the literature has shown in the past (Exley and Kessler, 2022); women on average perform worse than men. Women's scores are on average 0.43 points lower than those of men. Since the average score is 3.63 out of 10, these 0.43 points represent 11.85% of the mean. Women also hold more pessimistic beliefs about their absolute performance than equally-performing men. Indeed, to the question "Out of the 10 questions in part 1, how many questions do you think you have answered correctly?" women provide an answer which is 0.426 points lower than equally-performing men. Since the average belief answer is 3, these 0.426 points represent 14.2% of mean. These results are statistically significant at the 1% level, with appropriate controls applied.

[4]My study makes two adaptations: the inclusion of two additional questions related to relative performance, *Comparison* and *BTY*, and providing participants with feedback on their absolute score only, unlike Exley and Kessler (2022), who also include information about relative performance.

Table A.4: Knowledge Score and Belief: Men vs. Women

	(1) Knowledge Score	(2) Belief
Female	-0.430***	-0.426***
	(0.133)	(0.124)
Male Average	3.83	3.32
Controls	Yes	Yes
Observations	719	719

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Rightwing, Ethnicity, Education, Employment Status, Age and Rural as well as Knowledge Score for Model 2. *Male Average* is the unadjusted mean for all men, before controlling for covariates.

Result 3 (Beliefs about Absolute Performance)**.** *There is a gender gap in beliefs about one's absolute performance. Women report more pessimistic beliefs about their absolute performance than equally-performing men.*

Table A.5 shows that women provide answers which are 3.468 points lower than equally-performing men when asked to agree on a scale from 0 to 100 to the following statement "I performed well on the test I took in part 1". Since the average participant answers 22.23 out of 100 for this perceived performance question, these 3.468 points represent 15.6% of the mean. Women further provide answers which are 4.817 points lower than equally-performing men when asked to agree on a scale from 0 to 100 to the following statement "I would apply for a job that required me to perform well on the test I took in part 1". Since the average participant answers 19.45 out of 100 for this perceived performance question, these 4.817 points represent 24.77% of the mean. Women further provide answers which are 3.437 points lower than equally-performing men when asked to agree on a scale from 0 to 100 to the following statement "I would succeed in a job that required me to perform well on the test I took in part 1". Since the average participant answers 22.05 out of 100 for this perceived performance question, these 3.437 points represent over 15.59% of the mean. I find similar results concerning the remaining self-evaluations; *Performance bucket, Comparison* and *BTY*. Women evaluate their performance more pessimistically than equally-performing men.

Table A.5: Uninformed Self-Evaluations: Men vs. Women

	(1) Perceived Perf.	(2) Perf. bucket	(3) Willing	(4) Success	(5) Comparison	(6) BTY
Female	-3.468**	-0.325***	-4.817***	-3.437*	-0.398***	4.542***
	(1.537)	(0.0854)	(1.823)	(1.932)	(0.0902)	(1.733)
Male Average	25.54	2.82	23.30	25.37	3.13	53.14
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	719	719	719	719	719	719

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge score, Rightwing, Ethnicity, Education, Employment status, Age and Rural. *Male Average* is the unadjusted mean for all men, before controlling for covariates.

Result 4 (Self-evaluations (uninformed))**.** *There is a gender gap in self-evaluations. Women provide worse self-evaluations of their performance on a math and science quiz than equally-performing men (when being uninformed of their true performance).*

One could first argue that the result above only holds when women do not know their true performance and that when informed of their true score, the difference in self-evaluations between equally-performing women and men disappears because their self-evaluations should be based more on objective feedback rather than subjective beliefs. Indeed, by informing participants of their true score, I mechanically eliminate any discrepancies in their beliefs about their absolute performance. However, Table A.6 shows substantial and statistically significant differences in self-evaluations even after their performance is revealed which suggests that women still hold more pessimistic self-evaluations on their performance than equally-performing men and that the differences in self-evaluations are not primarily driven by any (potential) gap in beliefs about one's absolute score. The size of the differences is yet smaller for all informed self-evaluations except *Success*.

Women provide answers which are almost 2.794 points lower than equally-performing men when asked to agree on a scale from 0 to 100 to the following statement "I performed well on the test I took in part 1". Since the average participant answers 24.76 out of 100 for this perceived performance question, these 2.794 points represent 11.28% of the mean.

Women further provide answers which are 4.728 points lower than equally-performing men when asked to agree on a scale from 0 to 100 to the following statement "I would succeed in a job that required me to perform well on the test I took in part 1". Since the average participant answers 23.04 out of 100 for this perceived performance question, these 4.728 points represent over 20.52% of the mean. I find similar results concerning the remaining self-evaluations; *Performance bucket, Comparison* and *BTY* (the coefficient for *Willing* is not statistically significant anymore in the informed case). Women evaluate their performance more pessimistically than equally-performing men even after having received their own absolute score.

Table A.6: Informed Self-Evaluations: Men vs. Women

	(1) Perceived Perf.	(2) Perf. bucket	(3) Willing	(4) Success	(5) Comparison	(6) BTY
Female	-2.794**	-0.162**	-2.710	-4.728***	-0.230***	4.271**
	(1.206)	(0.0717)	(1.701)	(1.695)	(0.0820)	(1.748)
Male Average	28.48	3.16	25.08	27.40	3.19	51.38
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	719	719	719	719	719	719

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge score, Rightwing, Ethnicity, Education, Employment status, Age and Rural. *Male Average* is the unadjusted mean for all men, before controlling for covariates.

Result 5 (Self-evaluations (informed))**.** *The gender gap in self-evaluations persists even when informed of their own performance. Women still provide worse self-evaluations of their performance on a math and science quiz than equally-performing men.*

Table A.7: Descriptive Statistics – Women Only

Variable	Mean	SD	Min	Max
Knowledge score	3.42	1.66	0	8
Belief	2.67	1.78	0	10
Perceived performance (uninformed)	18.91	20.97	0	100
Performance bucket (uninformed)	2.34	1.17	1	6
Willing (uninformed)	15.59	23.63	0	100
Success (uninformed)	18.72	26.35	0	100
Comparison (uninformed)	2.57	1.22	1	6
BTY (uninformed)	58.93	23.80	1	100
Perceived performance (informed)	21.04	21.67	0	89
Performance bucket (informed)	2.72	1.34	1	6
Willing (informed)	18.02	24.84	0	100
Success (informed)	18.67	24.04	0	100
Comparison (informed)	2.71	1.35	1	7
BTY (informed)	58.26	25.08	1	100
Observations		359		

A.4 Exploratory Analysis: Difference between Informed and Uninformed Self-Evaluations

This section examines how individuals adjust their self-evaluations after receiving feedback. The outcome variables are calculated as the difference between the informed and uninformed measures (e.g., *informed Success* - *uninformed Success*). A positive coefficient indicates that self-evaluations, on average, increased after receiving feedback (what I would expect given the average underconfidence of the participants). Overall, I observe minimal differences between groups (men vs. women or gay men/lesbian women vs. heterosexual men/women), with a few notable exceptions in the men vs. women comparison. For instance, women increase their comparison scores more following feedback than equally-performing men. Heterosexual women also seem to increase their *Willingness* (to apply to a job that requires such skills) more after feedback than equally-performing heterosexual men.

Table A.8: Self-Evaluations Differences (Informed-Uninformed)

	(1) Perceived Perf. Difference	(2) Perf. Bucket Difference	(3) Willing Difference	(4) Success Difference	(5) Comparison Difference	(6) BTY Difference
Female	0.621	0.154	3.663*	-1.124	0.221**	-0.0471
	(2.171)	(0.122)	(1.969)	(1.954)	(0.112)	(1.750)
Homosexual	0.589	0.0703	0.692	1.332	0.0705	-0.999
	(1.936)	(0.117)	(1.773)	(1.574)	(0.112)	(1.714)
Female_Hom.	0.162	0.0256	-3.203	-0.239	-0.105	-0.553
	(2.825)	(0.163)	(2.661)	(2.668)	(0.151)	(2.452)
Hetero Male Average	1.92	0.27	1.23	0.91	0.02	-1.16
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	719	719	719	719	719	719

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge score, Rightwing, Ethnicity, Education, Employment status, Age and Rural. *Hetero Male Average* is the unadjusted mean for heterosexual men, before controlling for covariates. *Perceived Perf. Difference* is the difference between the informed *Perceived Perf.* and the uninformed *Perceived Perf.* Same holds for the other variables.

Table A.9: Self-Evaluations Differences (Informed-Uninformed) – Men vs. Women

	Perceived Perf. Difference	Perf. Bucket Difference	Willing Difference	Success Difference	Comparison Difference	BTY Difference
Female	0.674	0.163*	2.106	-1.291	0.168**	-0.271
	(1.436)	(0.0850)	(1.388)	(1.359)	(0.0786)	(1.256)
Male Average	2.94	0.34	1.78	2.03	0.07	-1.76
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	719	719	719	719	719	719

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge score, Rightwing, Ethnicity, Education, Employment status, Age, and Rural. *Male Average* is the unadjusted mean for all men, before controlling for covariates. *Perceived Perf. Difference* is the difference between the informed *Perceived Perf.* and the uninformed *Perceived Perf.* The same holds for the other variables.

Table A.10: Self-Evaluations Differences (Informed-Uninformed) – Lesbian Women vs. Hetero Women

	Perceived Perf. Difference	Perf. Bucket Difference	Willing Difference	Success Difference	Comparison Difference	BTY Difference
lesbian woman	1.505	0.153	-1.636	0.188	0.0145	-3.116
	(2.202)	(0.128)	(2.110)	(2.099)	(0.111)	(1.941)
Hetero Fem. Avg.	1.04	0.32	3.53	-1.41	0.14	0.28
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	359	359	359	359	359	359

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge score, Rightwing, Ethnicity, Education, Employment status, Age, and Rural. *Hetero Fem. Avg.* is the unadjusted mean for heterosexual women, before controlling for covariates. *Perceived Perf. Difference* is the difference between the informed *Perceived Perf.* and the uninformed *Perceived Perf.*. The same holds for the other variables.

Table A.11: Self-Evaluations Differences (Informed-Uninformed) – Gay Men vs. Hetero Men

	Perceived Perf. Difference	Perf. Bucket Difference	Willing Difference	Success Difference	Comparison Difference	BTY Difference
gay man	0.440	0.0479	0.608	1.349	0.0583	-0.605
	(1.973)	(0.120)	(1.786)	(1.519)	(0.115)	(1.728)
Hetero Male Avg.	1.92	0.27	1.23	0.91	0.02	-1.16
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	360	360	360	360	360	360

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge score, Rightwing, Ethnicity, Education, Employment status, Age, and Rural. *Hetero Male Avg.* is the unadjusted mean for heterosexual men, before controlling for covariates. *Perceived Perf. Difference* is the difference between the informed *Perceived Perf.* and the uninformed *Perceived Perf.*. The same holds for the other variables.

A.5 Robustness Checks

A.5.1 First Robustness Check (drop inattentive participants)

89 participants (12.4% of the whole sample) did not answer correctly the question "Out of the 10 questions in part 1, how many questions do you think you have answered correctly?" even though the answer was displayed on the same screen in red (see Figure A.9). In the next regression tables, these "inattentive" participants are withdrawn from the sample.

Compared to the main analysis involving all 719 participants in Table 1 of the manuscript (uninformed self-evaluations), Table A.12 shows that *Female* and *Female_Hom.* remain statistically significant across the 6 models (at least weakly) – except for model BTY where *Female_Hom.* was already not statistically significant in the main analysis. The difference between the two gaps (*Female* and the Combined Gap) is statistically significant in all models except for model BTY where the gap was already not statistically significant in the main analysis.

Compared to the main analysis involving all 719 participants in Table 2 of the manuscript (informed self-evaluations), Table A.13 on the informed self-evaluations shows that *Female* remains statistically significant across all models except for *Willing*. Concerning the interaction term *Female_Hom.*, it remains statistically significant for *Perceived Perf.*, *Perf. Bucket* and *Comparison* as in the main analysis but loses its significancy for model *Success*. Finally, the difference between the two gaps is statistically significant in all models except for model BTY where the gap was already not statistically significant in the main analysis and *Willing*, where it loses significance.

All in all, women exhibit more pessimistic self-evaluations than equally-performing men and the gender gap in self-evaluations is narrower for lesbian women. In contrast, no significant difference in self-evaluations is observed between gay and heterosexual men. This pattern holds true for both uninformed and informed scenarios and aligns with the results of the main analysis.

Table A.12: Uninformed Self-Evaluations

	Perceived Perf.	Perf. Bucket	Willing	Success	Comparison	BTY
Female	-6.257***	-0.555***	-7.691***	-6.260**	-0.654***	5.230**
	(2.396)	(0.129)	(2.696)	(2.980)	(0.132)	(2.638)
Homosexual	-1.892	-0.169	-1.279	-1.289	-0.169	0.994
	(2.289)	(0.126)	(2.947)	(2.842)	(0.137)	(2.478)
Female_Hom.	7.238**	0.535***	7.827**	7.888*	0.537***	-1.926
	(3.143)	(0.173)	(3.901)	(4.049)	(0.184)	(3.650)
Hetero Male Average	26.66	2.90	23.65	25.94	3.21	52.45
Female - Combined Gap (=Gap Hetero W vs Hetero M - Gap Lesbian W vs Hetero M)						
Difference	-5.346	-0.367	-6.548	-6.599	-0.367	0.932
P-value	0.022	0.005	0.024	0.033	0.007	0.746
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	630	630	630	630	630	630

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge score, Rightwing, Ethnicity, Education, Employment status, Age, and Rural. The baseline group is heterosexual men. *Hetero Male Average* is the unadjusted mean for heterosexual men, before controlling for covariates. *Difference* represents the difference between the gender gap, as indicated by the coefficient for *Female* (comparing heterosexual men and heterosexual women), and the Combined Gap, which is the sum of the coefficients for *Female*, *Homosexual*, and *Female_Hom.* (capturing the gap between lesbian women and heterosexual men). The P-value corresponds to the two-sided t-test for the difference between these two gap estimates.

Table A.13: Informed Self-Evaluations

	Perceived Perf.	Perf. Bucket	Willing	Success	Comparison	BTY
Female	-3.895**	-0.288***	-3.954	-6.477***	-0.335***	5.471**
	(1.511)	(0.0879)	(2.504)	(2.382)	(0.113)	(2.535)
Homosexual	0.212	-0.0152	0.126	0.942	-0.0459	0.397
	(1.654)	(0.0871)	(2.563)	(2.592)	(0.115)	(2.476)
Female_Hom.	5.075**	0.436***	3.749	4.944	0.340**	-2.515
	(2.284)	(0.127)	(3.528)	(3.418)	(0.158)	(3.592)
Hetero Male Average	28.81	3.20	25.28	27.12	3.25	50.21
Female - Combined Gap (=Gap Hetero W vs Hetero M - Gap Lesbian W vs Hetero M)						
Difference	-5.287	-0.421	-3.876	-5.886	-0.294	2.118
P-value	0.004	0.000	0.152	0.017	0.014	0.451
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	630	630	630	630	630	630

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge score, Rightwing, Ethnicity, Education, Employment status, Age, and Rural. The baseline group is heterosexual men. *Hetero Male Average* is the unadjusted mean for heterosexual men, before controlling for covariates. *Difference* represents the difference between the gender gap, as indicated by the coefficient for *Female* (comparing heterosexual men and heterosexual women), and the Combined Gap, which is the sum of the coefficients for *Female*, *Homosexual*, and *Female_Hom.* (capturing the gap between lesbian women and heterosexual men). The P-value corresponds to the two-sided t-test for the difference between these two gap estimates.

Compared to the main analysis among women only in Online Appendix A.2, Panel A in Table A.14 (uninformed self-evaluations) shows that *lesbian woman* in model *Comparison* is here significant at a 10% level (was not significant in the main analysis). In all other models, *lesbian woman* remains positive and statistically significant (in model *BTY*, *lesbian woman* was already not significant in the main analysis).

Looking at Panel A in Table A.15 (informed self-evaluations), we see that in model *Comparison*, *lesbian woman* is here significant at 10% level (was not significant in the main analysis). In all other models, significancy level and coefficient sign for *lesbian woman* do not change from the main analysis. All in all, consistent with the analysis in Online Appendix A.2, lesbian women provide more optimistic views of their performance compared to equally-performing heterosexual women.

Consistent with the main analysis among men only in Online Appendix A.2, Panel B in both Tables A.14 and A.15 show no difference between gay men and heterosexual men in terms of (un)informed self-evaluations.

Table A.14: Uninformed Self-Evaluations: Lesbian Women (Gay Men) vs. Hetero Women (Men)

	(1) Perceived Perf.	(2) Perf. Bucket	(3) Willing	(4) Success	(5) Comparison	(6) BTY
Panel A: Lesbian Women vs. Hetero Women						
lesbian woman	5.319**	0.295**	6.474**	6.876**	0.266*	0.745
	(2.421)	(0.139)	(3.254)	(3.331)	(0.143)	(3.110)
Hetero Female Average	17.08	2.17	13.27	17.06	2.40	58.76
Observations	312	312	312	312	312	312
Panel B: Gay Men vs. Hetero Men						
gay man	-2.136	-0.140	-1.355	-1.163	-0.123	0.0681
	(2.325)	(0.129)	(3.079)	(2.880)	(0.141)	(2.533)
Hetero Male Average	26.66	2.90	23.65	25.94	3.21	52.45
Observations	318	318	318	318	318	318
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge Score, Rightwing, Ethnicity, Education, Employment Status, Age, and Rural. *Hetero Female Average* and *Hetero Male Average* are the unadjusted means for heterosexual women (men), before controlling for covariates.

Table A.15: Informed Self-Evaluations: Lesbian Women (Gay Men) vs. Hetero Women (Men)

	(1) Perceived Perf.	(2) Perf. Bucket	(3) Willing	(4) Success	(5) Comparison	(6) BTY
Panel A: Lesbian Women vs. Hetero Women						
lesbian woman	6.073***	0.400***	4.037	5.710**	0.246*	-1.990
	(2.080)	(0.119)	(2.964)	(2.640)	(0.131)	(3.103)
Hetero Female Average	18.75	2.53	16.46	15.61	2.57	59.01
Observations	312	312	312	312	312	312
Panel B: Gay Men vs. Hetero Men						
gay man	-0.330	-0.0100	0.114	0.926	-0.0137	0.0404
	(1.644)	(0.0873)	(2.630)	(2.615)	(0.117)	(2.563)
Hetero Male Average	28.81	3.20	25.28	27.12	3.25	50.21
Observations	318	318	318	318	318	318
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge Score, Rightwing, Ethnicity, Education, Employment Status, Age, and Rural. *Hetero Female Average* and *Hetero Male Average* are the unadjusted means for heterosexual women (men), before controlling for covariates.

Compared to the analysis involving all 719 participants, which examines the differences between men and women in Online Appendix A.3, *Female* in model *Perceived Perf.* and *Success* is not statistically significant anymore in Table A.16 (uninformed self-evaluations). In all other models, *Female* remains statistically significant (at least weakly) and retains the same sign as in the main analysis. Finally, Table A.17 (informed self-evaluations) shows that *Female* in model *Perceived Perf.* and *Perf. bucket* is not statistically significant anymore. In all other models, *Female* remains statistically significant and retains the same sign as in the main analysis (in model *Willing* it was already not statistically significant in the main analysis).

Table A.16: Uninformed Self-Evaluations: Men vs. Women

	(1) Perceived Perf.	(2) Perf. bucket	(3) Willing	(4) Success	(5) Comparison	(6) BTY
Female	-2.621	-0.285***	-3.789*	-2.328	-0.383***	4.243**
	(1.620)	(0.0897)	(1.959)	(2.090)	(0.0964)	(1.850)
Male Average	25.18	2.80	22.72	25.00	3.15	52.72
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	630	630	630	630	630	630

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge score, Rightwing, Ethnicity, Education, Employment status, Age, and Rural. *Male Average* is the unadjusted mean for all men, before controlling for covariates.

Table A.17: Informed Self-Evaluations: Men vs. Women

	(1) Perceived Perf.	(2) Perf. bucket	(3) Willing	(4) Success	(5) Comparison	(6) BTY
Female	-1.407	-0.0722	-2.115	-4.082**	-0.166**	4.218**
	(1.158)	(0.0659)	(1.794)	(1.752)	(0.0798)	(1.802)
Male Average	29.03	3.20	25.16	27.56	3.25	50.16
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	630	630	630	630	630	630

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge score, Rightwing, Ethnicity, Education, Employment status, Age, and Rural. *Male Average* is the unadjusted mean for all men, before controlling for covariates.

A.5.2 Second Robustness Check (additionally drop participants who failed twice the attention checks in a row)

In the following section, I further exclude participants who failed to answer correctly the attention check twice in a row (see Figure A.16 for the attention check).

Compared to the analysis involving all 719 participants in Table 1 of the manuscript, Table A.18 on the uninformed self-evaluations shows that *Female* and *Female_Hom.* remain statistically significant (and retain the same sign) across the 6 models (at least weakly) – except for model BTY where *Female_Hom.* was already not statistically significant in the main analysis. The difference between the two gaps (*Female* and the Combined Gap) is statistically significant in all models, except for model BTY where the gap was already not statistically significant in the main analysis.

Compared to the analysis involving all 719 participants in Table 2 of the manuscript, Table A.19 on the informed self-evaluations shows that *Female* remains statistically significant across all models (at least weakly) and retains the same sign. Concerning the interaction term *Female_Hom.*, it remains statistically significant (and retains the same sign) for *Perceived Perf.*, *Perf. Bucket* and *Comparison* as in the main analysis but loses its significancy for model *Success*. The difference between the two gaps is statistically significant in all models, except for model BTY where the gap was already not statistically significant in the main analysis, and *Willing*, where it loses significance.

All in all, women exhibit more pessimistic self-evaluations than equally-performing men and the gender gap in self-evaluations is narrower for lesbian women. In contrast, no significant difference in self-evaluations is observed between gay and heterosexual men. This pattern holds true for both uninformed and informed scenarios and aligns with the results of the main analysis.

Table A.18: Uninformed Self-Evaluations

	Perceived Perf.	Perf. Bucket	Willing	Success	Comparison	BTY
Female	-6.275***	-0.559***	-7.358***	-6.643**	-0.649***	4.942*
	(2.424)	(0.131)	(2.722)	(2.973)	(0.134)	(2.665)
Homosexual	-1.873	-0.178	-1.235	-1.522	-0.187	0.741
	(2.305)	(0.127)	(2.961)	(2.865)	(0.137)	(2.496)
Female_Hom.	7.116**	0.522***	7.194*	7.840*	0.520***	-2.094
	(3.169)	(0.174)	(3.909)	(4.036)	(0.185)	(3.680)
Hetero Male Average	26.54	2.89	23.30	25.94	3.23	52.55
Female - Combined Gap (=Gap Hetero W vs Hetero M - Gap Lesbian W vs Hetero M)						
Difference	-5.243	-0.344	-5.960	-6.318	-0.333	1.354
P-value	0.027	0.008	0.040	0.036	0.014	0.641
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	620	620	620	620	620	620

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge score, Rightwing, Ethnicity, Education, Employment status, Age, and Rural. The baseline group is heterosexual men. *Hetero Male Average* is the unadjusted mean for heterosexual men, before controlling for covariates. *Difference* represents the difference between the gender gap, as indicated by the coefficient for *Female* (comparing heterosexual men and heterosexual women), and the Combined Gap, which is the sum of the coefficients for *Female*, *Homosexual*, and *Female_Hom.* (capturing the gap between lesbian women and heterosexual men). The P-value corresponds to the two-sided t-test for the difference between these two gap estimates.

Table A.19: Informed Self-Evaluations

	Perceived Perf.	Perf. Bucket	Willing	Success	Comparison	BTY
Female	-3.974***	-0.289***	-4.253*	-6.646***	-0.342***	5.029**
	(1.526)	(0.0885)	(2.524)	(2.418)	(0.113)	(2.553)
Homosexual	0.0719	-0.0213	0.00869	0.728	-0.0554	0.0276
	(1.621)	(0.0863)	(2.569)	(2.596)	(0.115)	(2.496)
Female_Hom.	4.559**	0.409***	3.477	4.484	0.331**	-2.665
	(2.216)	(0.125)	(3.510)	(3.405)	(0.158)	(3.615)
Hetero Male Average	28.85	3.21	25.25	27.16	3.26	50.38
Female - Combined Gap (=Gap Hetero W vs Hetero M - Gap Lesbian W vs Hetero M)						
Difference	-4.631	-0.388	-3.486	-5.212	-0.275	2.637
P-value	0.005	0.000	0.182	0.026	0.017	0.354
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	620	620	620	620	620	620

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge score, Rightwing, Ethnicity, Education, Employment status, Age, and Rural. The baseline group is heterosexual men. *Hetero Male Average* is the unadjusted mean for heterosexual men, before controlling for covariates. *Difference* represents the difference between the gender gap, as indicated by the coefficient for *Female* (comparing heterosexual men and heterosexual women), and the Combined Gap, which is the sum of the coefficients for *Female*, *Homosexual*, and *Female_Hom* (capturing the gap between lesbian women and heterosexual men). The P-value corresponds to the two-sided t-test for the difference between these two gap estimates.

Consistent with the analysis between lesbian women and heterosexual women (359 women) in Online Appendix A.2, Panel A in Table A.20 on the uninformed self-evaluations shows that *lesbian woman* remains statistically significant (at least weakly) and retains its sign across all models – except for model *Comparison* and model *BTY* where it was already not statistically significant in the main analysis. Panel A in Table A.21 on the informed self-evaluations shows again that for all models where *lesbian woman* was significant in the main analysis, it remains significant (and retains its sign) and even becomes weakly significant for model *Comparison* (same as first robustness check). All in all, consistent with the main analysis and the first robustness check, lesbian women provide more optimistic views of their (absolute) performance compared to equally-performing heterosexual women.

Consistent with the analysis between gay men and heterosexual men (360 men) in Online Appendix A.2, Panel B in Table A.20 and Panel B in Table A.21 show no difference to the main analysis (same as the first robustness check). In my sample, there seems to be no difference between gay men and heterosexual men in terms of (un)informed self-evaluations.

Table A.20: Uninformed Self-Evaluations: Lesbian Women (Gay Men) vs. Hetero Women (Men)

	(1) Perceived Perf.	(2) Perf. Bucket	(3) Willing	(4) Success	(5) Comparison	(6) BTY
Panel A: Lesbian Women vs. Hetero Women						
lesbian woman	5.204**	0.258*	5.669*	6.229*	0.219	0.257
	(2.477)	(0.136)	(3.230)	(3.198)	(0.141)	(3.170)
Hetero Female Average	16.81	2.16	13.15	16.66	2.40	58.75
Observations	305	305	305	305	305	305
Panel B: Gay Men vs. Hetero Men						
gay man	-2.052	-0.142	-1.185	-1.137	-0.135	-0.0368
	(2.340)	(0.130)	(3.101)	(2.912)	(0.141)	(2.555)
Hetero Male Average	26.54	2.89	23.30	25.94	3.23	52.55
Observations	315	315	315	315	315	315
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge Score, Rightwing, Ethnicity, Education, Employment Status, Age, and Rural. *Hetero Female Average* and *Hetero Male Average* are the unadjusted means for heterosexual women (men), before controlling for covariates.

Table A.21: Informed Self-Evaluations: Lesbian Women (Gay Men) vs. Hetero Women (Men)

	(1) Perceived Perf.	(2) Perf. Bucket	(3) Willing	(4) Success	(5) Comparison	(6) BTY
Panel A: Lesbian Women vs. Hetero Women						
lesbian woman	5.017***	0.341***	3.239	4.624*	0.205*	-2.713
	(1.748)	(0.103)	(2.806)	(2.435)	(0.123)	(3.163)
Hetero Female Average	18.49	2.52	15.97	15.28	2.56	58.93
Observations	305	305	305	305	305	305
Panel B: Gay Men vs. Hetero Men						
gay man	-0.173	-0.003	0.274	1.003	-0.012	-0.202
	(1.649)	(0.088)	(2.657)	(2.637)	(0.118)	(2.580)
Hetero Male Average	28.85	3.21	25.25	27.16	3.26	50.38
Observations	315	315	315	315	315	315
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge Score, Rightwing, Ethnicity, Education, Employment Status, Age, and Rural. *Hetero Female Average* and *Hetero Male Average* are the unadjusted means for heterosexual women (men), before controlling for covariates.

Compared to the analysis involving all 719 participants, which examines the differences between men and women in Online Appendix A.3, Table A.22 on the uninformed self-evaluations shows that *Female* retains its sign and remains statistically significant (at least weakly) across all models except for *Success* and *Perceived Perf.* as in the first robustness check. Table A.23 on the informed self-evaluations shows that in models *Perceived Perf.* and *Perf. bucket*, *Female* is not statistically significant anymore while in all other models it remains significant (and with the same sign), exactly as in the first robustness check (in model *Willing* the coefficient for *Female* was already not statistically significant in the main analysis).

Table A.22: Uninformed Self-Evaluations: Women vs. Men

	(1) Perceived Perf.	(2) Perf. bucket	(3) Willing	(4) Success	(5) Comparison	(6) BTY
Female	-2.678	-0.294***	-3.751*	-2.704	-0.385***	3.875**
	(1.628)	(0.0899)	(1.958)	(2.076)	(0.0966)	(1.858)
Male Average	25.10	2.80	22.55	24.99	3.15	52.77
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	620	620	620	620	620	620

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge score, Rightwing, Ethnicity, Education, Employment status, Age and Rural. *Male Average* is the unadjusted mean for all men, before controlling for covariates.

Table A.23: Informed Self-Evaluations: Women vs. Men

	(1) Perceived Perf.	(2) Perf. bucket	(3) Willing	(4) Success	(5) Comparison	(6) BTY
Female	-1.726	-0.0858	-2.536	-4.464**	-0.177**	3.712**
	(1.108)	(0.0639)	(1.779)	(1.737)	(0.0787)	(1.808)
Male Average	29.05	3.21	25.14	27.58	3.25	50.24
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	620	620	620	620	620	620

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge score, Rightwing, Ethnicity, Education, Employment status, Age and Rural. *Male Average* is the unadjusted mean for all men, before controlling for covariates.

A.6 Comparing Lesbian Women to Men

I created a dummy *Lesbian* which takes the value of 1 if the participant was pre-screened as a lesbian woman and 0 if the participant was pre-screened as a gay man or heterosexual man. Consequently, the sample size decreases from 719 to 539, as heterosexual women are excluded from the following regressions. Table A.24 shows that lesbian women tend to perform worse on average compared to men. Overall, this section shows that there is no statistically significant difference between equally-performing men and lesbian women regarding both uninformed and informed self-evaluations (with the exception of *BTY (uninformed),* $p = 0.089$).

Table A.24: Knowledge Score and Belief: Lesbian Women vs. Men

	(1) Knowledge score	(2) Belief
Lesbian	-0.356**	-0.165
	(0.161)	(0.149)
Male Average	3.83	3.32
Controls	Yes	Yes
Observations	539	539

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Rightwing, Ethnicity, Education, Employment Status, Age and Rural as well as Knowledge Score for Model 2. *Male Average* is the unadjusted mean for all men, before controlling for covariates.

Table A.25: Uninformed Self-Evaluations: Lesbian Women vs. Men

	(1) Perceived Perf.	(2) Perf. bucket	(3) Willing	(4) Success	(5) Comparison	(6) BTY
Lesbian	0.325	-0.0789	-0.424	0.259	-0.165	3.773*
	(1.846)	(0.102)	(2.457)	(2.491)	(0.112)	(2.214)
Male Average	25.54	2.82	23.20	25.37	3.13	53.14
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	539	539	539	539	539	539

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge score, Rightwing, Ethnicity, Education, Employment status, Age and Rural. *Male Average* is the unadjusted mean for all men, before controlling for covariates.

Table A.26: Informed Self-Evaluations: Lesbian Women vs. Men

	(1) Perceived Perf.	(2) Perf. bucket	(3) Willing	(4) Success	(5) Comparison	(6) BTY
Lesbian	1.266	0.109	-0.388	0.0596	-0.0248	3.113
	(1.609)	(0.0912)	(2.249)	(2.295)	(0.102)	(2.236)
Male Average	28.48	3.16	25.08	27.40	3.19	51.38
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	539	539	539	539	539	539

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results are based on OLS regressions (which include a constant term), with the dependent variable for each model indicated in the corresponding column. Controls are Knowledge score, Rightwing, Ethnicity, Education, Employment status, Age and Rural. *Male Average* is the unadjusted mean for all men, before controlling for covariates.

A.7 Pre-registration Document

ASPREDICTED

Wharton CREDIBILITY LAB
UNIVERSITY of PENNSYLVANIA

CONFIDENTIAL - FOR PEER-REVIEW ONLY

The sexual orientation minority gaps in confidence and self-evaluations (#180628)

Created: 06/25/2024 12:13 AM (PT)

This is an anonymized copy (without author names) of the pre-registration. It was created by the author(s) to use during peer-review.
A non-anonymized version (containing author names) should be made available by the authors when the work it supports is made public.

---

1) Have any data been collected for this study already?
It's complicated. We have already collected some data but explain in Question 8 why readers may consider this a valid pre-registration nevertheless.

2) What's the main question being asked or hypothesis being tested in this study?
We investigate whether there exist sexual orientation minority gaps in confidence and self-evaluation, i.e. if homosexual individuals differ from heterosexual in terms of confidence and self-evaluation when being informed of their personal score and when being uninformed of their personal score. We also investigate the gender gap in confidence and self-evaluation, i.e. if women differ from men in terms of confidence and self-evaluation when being informed of their personal score and when being uninformed of their personal score.

3) Describe the key dependent variable(s) specifying how they will be measured.
Key variables: in part 1 participants answer a 10 questions quiz from which we can create a score. Then we collect the following key dependent variables: first, we elicit the participants' beliefs about their absolute performance (how many questions they believe to have answered correctly) which is our confidence measure. Then we ask them 6 self-evaluation questions: performance bucket (how well they think they performed on the test in part 1 from terrible, very poor, poor, neutral, good, very good, to exceptional), perceived performance ("I performed well on the test I took in part 1" from entirely disagree 0 to entirely agree 100), willing to apply ("I would apply for a job that required me to perform well on the test I took in part 1" from entirely disagree 0 to entirely agree 100), success ("I would succeed in a job that required me to perform well on the test I took in part 1 " from entirely disagree 0 to entirely agree 100), comparison bucket ("Please indicate how well you think you performed on the test you took in part 1 compared to the average score of all participants" from Much worse, Worse, Slightly worse, Same, Slightly better, Better to Much Better) and better-than-you ("What share of the other participants performed strictly better than you?" from 0% to 100%). We ask the participants these 6 items twice: first when being uninformed of their own score (part 2) and second being informed of their own score (part 3).

4) How many and which conditions will participants be assigned to?
We will investigate potential behavioral differences between sexual orientations (homosexual vs heterosexual, sexual orientation minority gap) and also behavioral differences between men and women (gender gap). We will also be able to compare those two gaps. We will aim to recruit in equal numbers from the four groups: heterosexual men, heterosexual women, homosexual men, homosexual women.
There are four groups (heterosexual men, heterosexual women, homosexual men and homosexual women) which allow us to look at the gender gap between men and women as well as the sexual orientation minority gap (our main focus) between heterosexual people and homosexual people in the confidence question, and in these 6 items (self-evaluation questions) before being informed and also after being informed of their own score.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
We will compare our participants' confidence and self-evaluations answers between the 2 sexual orientation groups: homosexual and heterosexual as well as between the two groups: men and women.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
We will exclude participants who answered two times incorrectly our two attention checks or who have given two answers that strictly contradict each other.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
Our plan is to collect the questionnaires from 135 participants per group. Since we have 4 groups this means a total of 540 participants.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
We conducted a pilot in March 2024 with only a few participants per group (45) to test the Prolific platform and also to test our questionnaires and questions as it is usually done in pilots. We will not use this data for our main experiment.

References

Exley, C. L. and Kessler, J. B. (2022). The gender gap in self-promotion. *The Quarterly Journal of Economics*, 137(3):1345–1381.