Online Appendices:

Gender, Confidence, and the Mismeasure of Intelligence, Competitiveness and Literacy

by

Glenn W. Harrison, Don Ross and J. Todd Swarthout †

January 2025

† Maurice R. Greenberg School of Risk Sciences and Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, USA (Harrison). School of Society, Politics, and Ethics, University College Cork, Cork, Ireland; School of Economics, University of Cape Town, Cape Town, South Africa; and Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, USA (Ross). Department of Economics and Experimental Economics Center, Andrew Young School of Policy Studies, Georgia State University, USA; and Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, USA (Swarthout). Harrison is also affiliated with the School of Economics, University of Cape Town. E-mail contacts: gharrison@gsu.edu, don.ross931@gmail.com and swarthout@gsu.edu. We are grateful to three referees, the editor, and participants at numerous presentations for constructive comments.

Table of Contents

Appendix A: Instructions (Online Working Paper) . . . . . -A1-
1. Non-Salient "Pen and Paper" Treatment. . . . . -A1-
2. Eighty Tokens Treatment . . . . . -A4-
3. Single Token Treatment . . . . . -A9-
4. Scrambled Treatment . . . . . -A13-

Appendix B: Subjective Beliefs and Risk Preferences (Online Working Paper) . . . . . -A14-

Appendix C: Gender and Competitiveness (Online Working Paper) . . . . . -A19-

Appendix D: Pilot Results with Variants on the Raven Task (Online Working Paper) . . . . . -A50-
1. Time Constraints . . . . . -A50-
2. Progressively Increasing Incentives . . . . . -A51-
3. Instructions: Time Constraint Treatment. . . . . -A57-
4. Instructions: Increased Incentives Treatment . . . . . -A63-

Appendix A: Instructions (Online Working Paper)

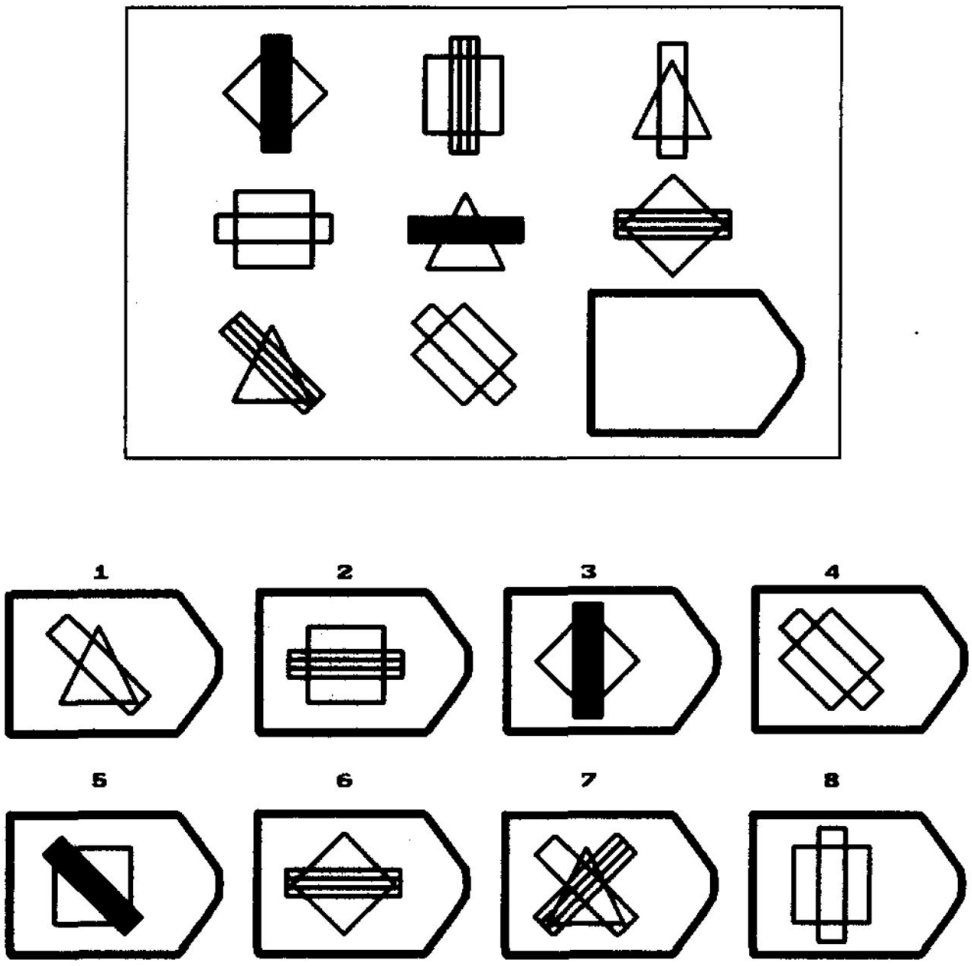
1. Non-Salient "Pen and Paper" Treatment

C

Your Instructions

This task is a test of perception and clear-thinking. You have already completed the first part of the task, in a previous session, when you were given 12 similar problems. We will now consider a fresh set of 36 problems.

Consider a similar problem, shown here.



The top part of this problem is a pattern with a bit cut out of it. Look at the pattern, and think what the piece needed to complete the pattern correctly both across and down must be like. Then find the right piece out of the eight bits shown, and numbered, in the bottom part of this problem.

Only one of these pieces is perfectly correct. Pieces #2, #5 and #8 complete the problem correctly going down, in the sense that they have a square as the larger piece. Pieces #1, #4 and #5 are correct going across, in the sense that they have the right slope for the rectangle.

Piece #5 is the correct bit, isn't it? It is correct going down as well as going across. So the answer is #5. In this case, if you were certain that piece #5 is the single best answer, you should write the number 5 on the answer sheet you have. You should write your answers in this part of the answer sheet:

Set II								
1			13			25		
2			14			26		
3			15			27		
4			16			28		
5			17			29		
6			18			30		
7			19			31		
8			20			32		
9			21			33		
10			22			34		
11			23			35		
12			24			36		

So you write the answer for the problem marked 1 in the box to the right of the number 1, and so on for all 36 problems.

This illustrative problem is relatively easy, but some of the later problems will not be so easy, and it might be harder to identify the one correct answer. You would still benefit by ruling out certain pieces so that you are more likely to select the correct answer from the pieces you have not ruled out.

You will find that the problems soon get difficult. Whether the problems are easy or difficult, you will notice that to solve them you have to use the same method all the time. The important thing is to notice how the problems develop and to learn the method of solving them. In each problem you look across each row, or down each column, and decide what the missing figure is like. Then look for the answer that is right both ways, across **and** down, from the options you have to select from.

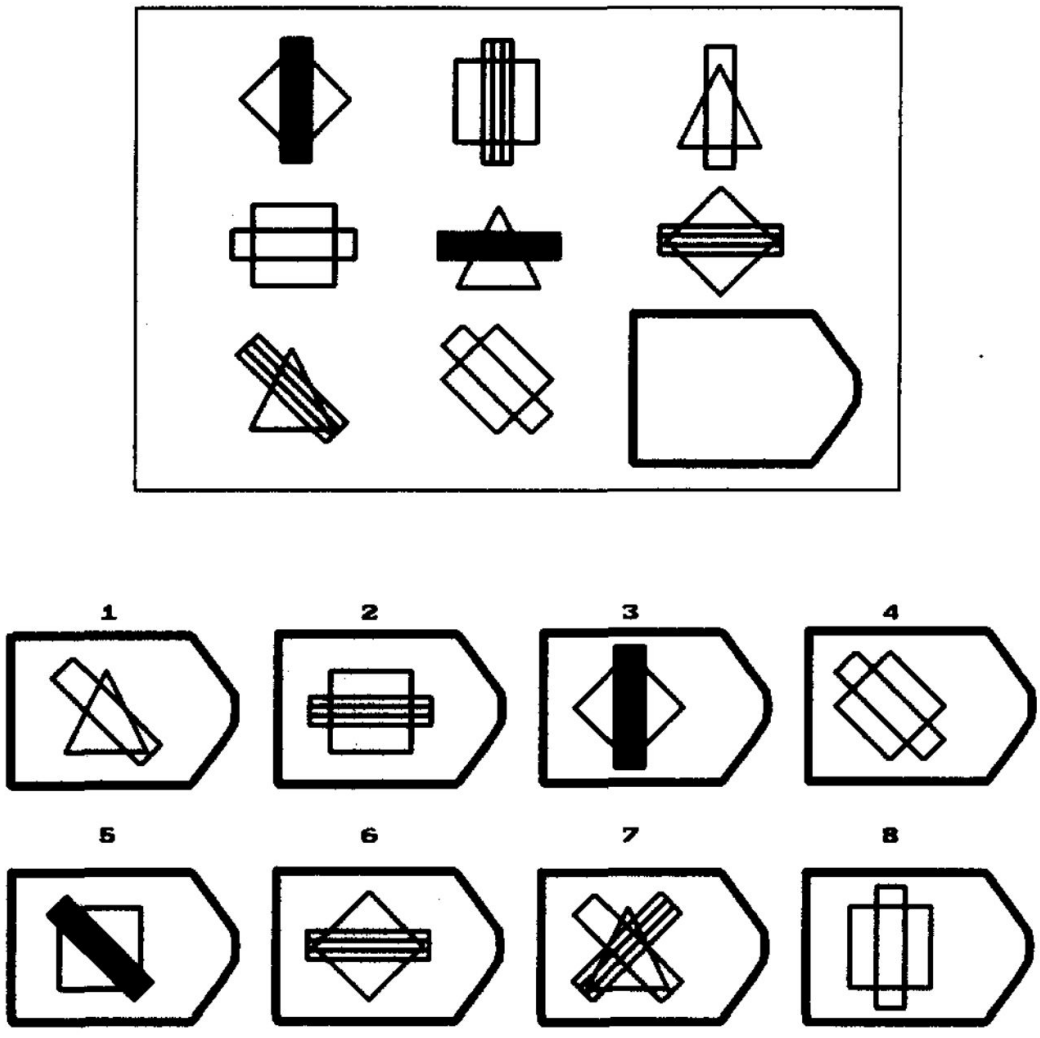
You can have as much time as you want, although we have to be out of the room in 90 minutes. Remember it is accurate work that counts. Attempt each problem in turn. Do your best to eliminate the incorrect pieces and find the correct piece to complete it before going on to the next problem.

You will receive $5 for completing this task.

Your Instructions

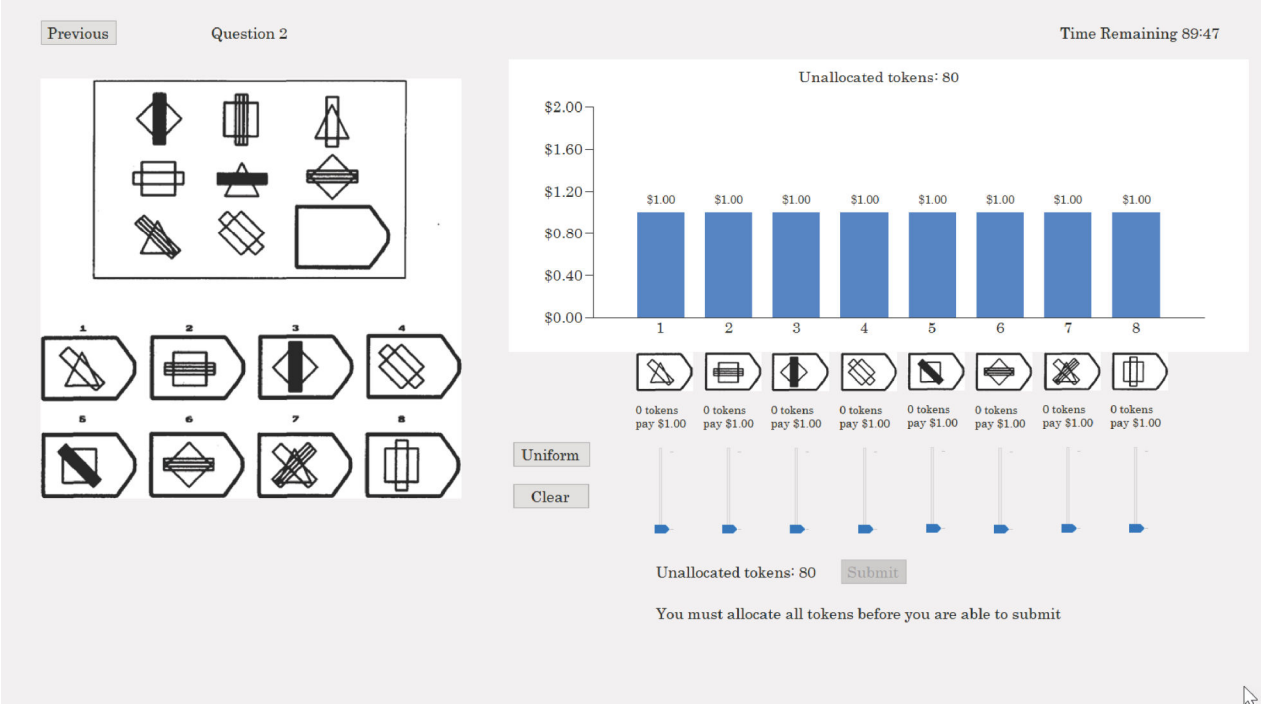
This task is a test of perception and clear-thinking. You have already completed the first part of the task, in a previous session, when you were given 12 similar problems. We will now consider a fresh set of 36 problems.

Consider a similar problem, shown here.

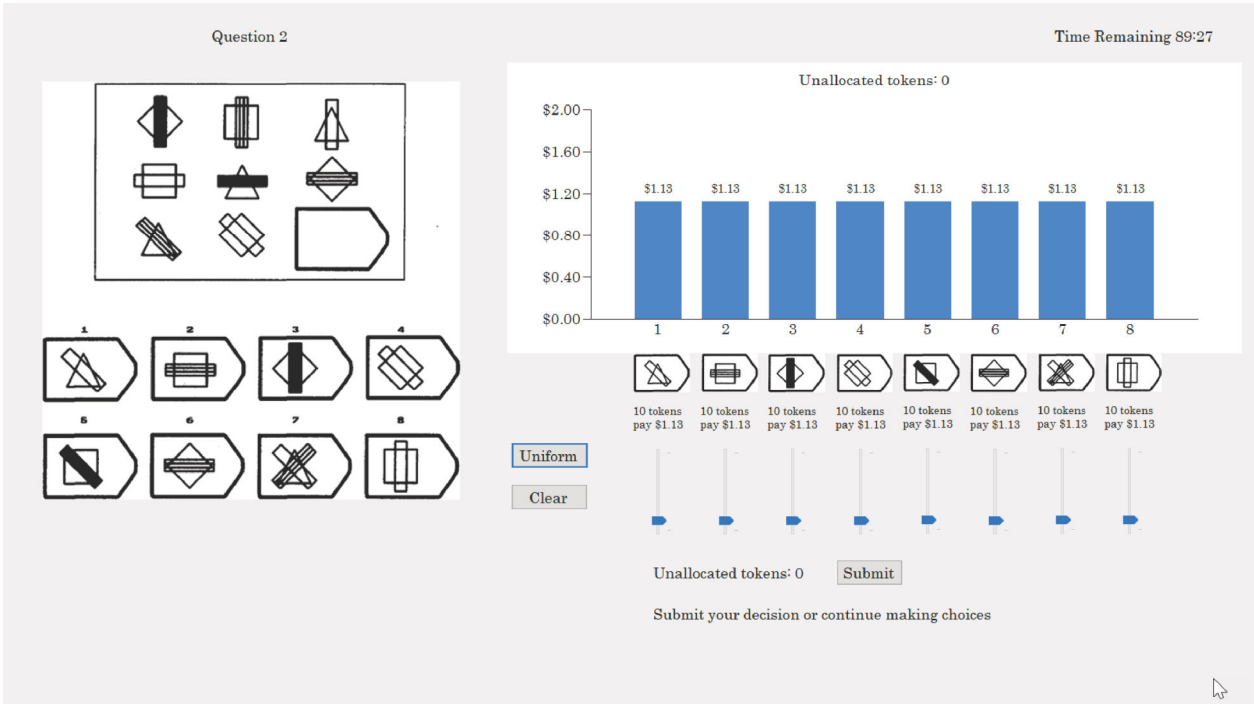


The top part of this problem is a pattern with a bit cut out of it. Look at the pattern, and think what the piece needed to complete the pattern correctly both across and down must be like. Then find the right piece out of the eight bits shown, and numbered, in the bottom part of this problem.

You will be asked to report your beliefs about the correct answer using an interface like this one, which is also generally familiar to you from a previous session.



The version you will see on the computer will be larger and easier to read. The problem and possible solutions are shown on the left of the screen, in the usual manner. On the right of the screen you have 80 tokens to allocate across the 8 possible answers. We start off with 0 tokens allocated to each of the possible answers. If you wanted to change this initial allocation so that there were 10 tokens allocated to each possible answer, just click on the **Uniform** button, and you will see this display:

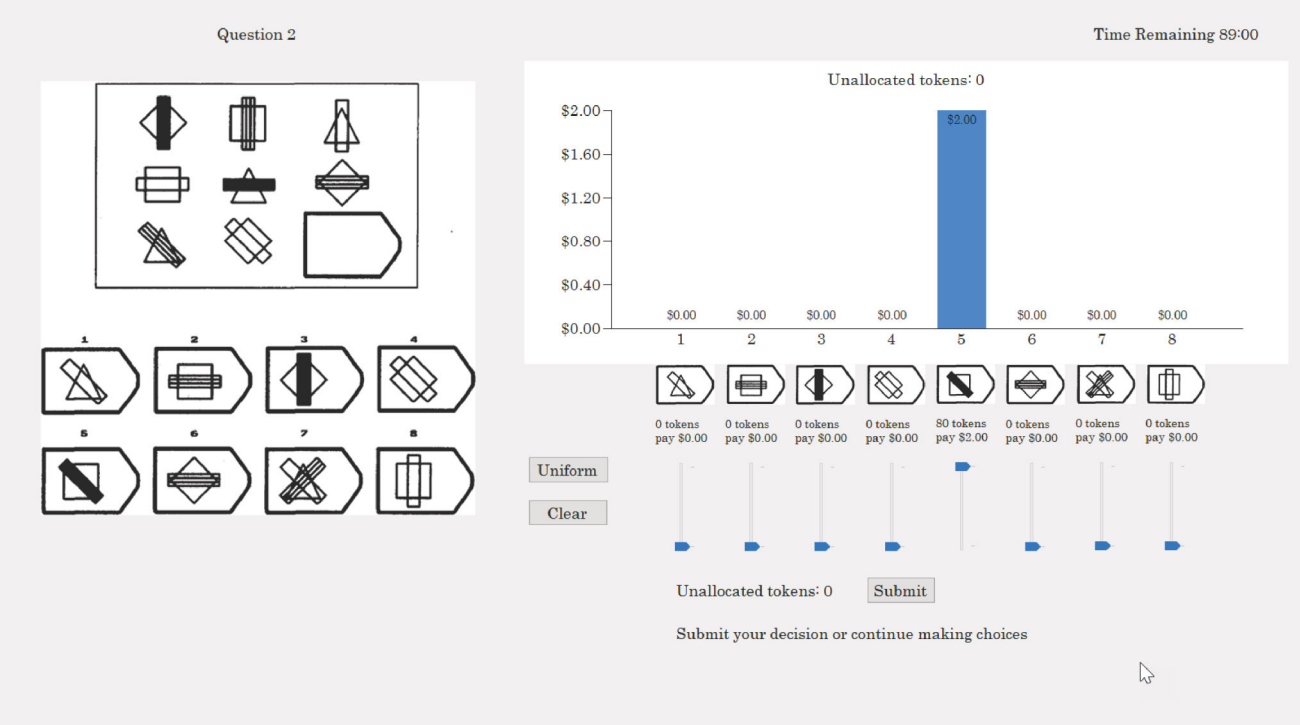


Here you have allocated 10 tokens to each possible answer, and you would earn $1.13 if you reported this allocation of tokens, since only one of the 8 possible answers is correct. You can return to the initial allocation of 0 tokens for each possible answer by clicking on the **Clear** button.

As you allocate tokens, by moving the sliders up or down, the earnings will change above each bar. These are the earnings that you will receive for this problem if that bar refers to the correct answer to the problem. **You will be paid for all 36 problems, and each problem will pay between $0 and $2 depending on your answer.**

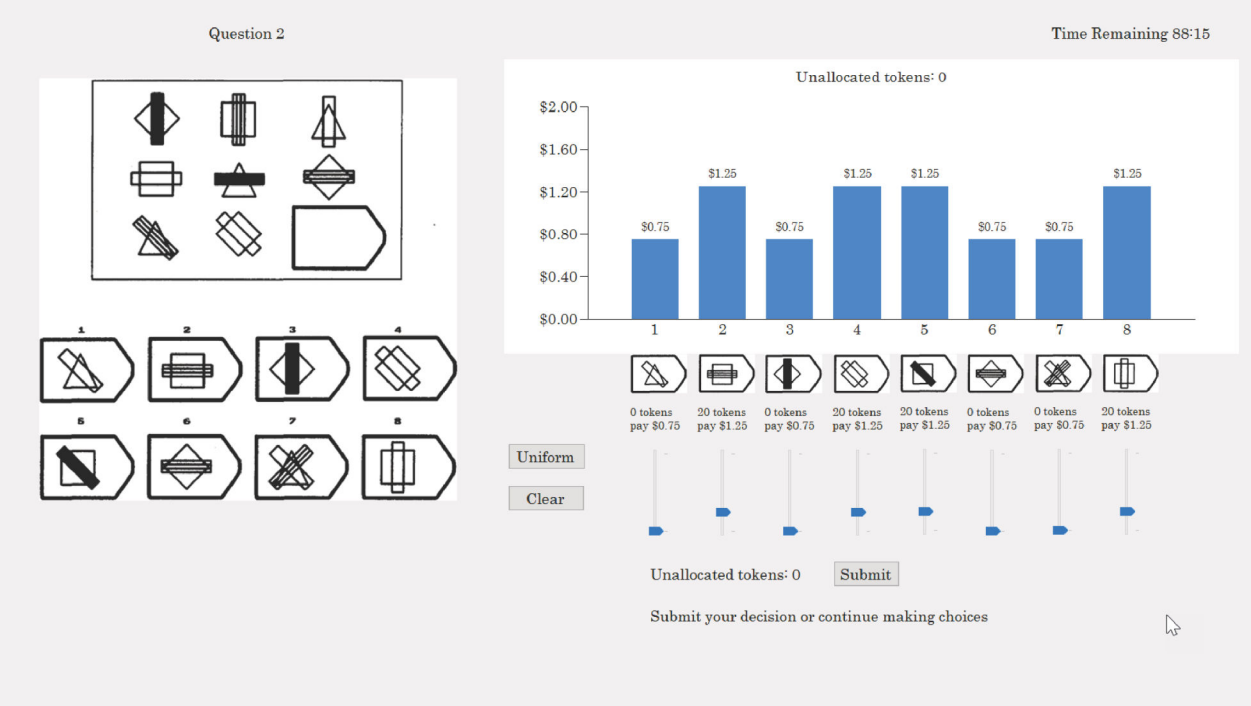
Return now to the problem itself. Only one of these pieces is perfectly correct. Pieces #2, #5 and #8 complete the problem correctly going down, in the sense that they have a square as the larger piece. Pieces #1, #4 and #5 are correct going across, in the sense that they have the right slope for the rectangle.

Piece #5 is the correct bit, isn't it? It is correct going down as well as going across. So the answer is #5. In this case, if you were certain that piece #5 is the single best answer, you should allocate all 80 tokens to the bin representing piece #5 as in this display:



So in this case you would earn $2.00 if indeed the correct answer was #5. Of course, if any of the other pieces turned out to be the correct answer you would, in this case, earn $0.

If you had decided that the correct answer was one of #2, #4, #5 or #8, but had not decided that #5 was actually the correct answer of these four possibilities, you might decide to allocate your tokens equally across the bars representing pieces #2, #4, #5 and #8 like this:



You can see that even if you eliminate some pieces, such as #1, #3, #6 and #7, that are clearly wrong, you give yourself a 1-in-4 chance of earning more money than if you guessed across all 8 pieces. In this case you would expect to earn $1.25 if indeed one of pieces #2, #4, #5 or #8 had been correct. Recall that if you had allocated the tokens roughly equally across all 8 bars, thinking that any of the 8 pieces might be correct, you would only earn $1.13 for this problem.

This illustrative problem is relatively easy, but some of the later problems will not be so easy, and it might be harder to identify the one correct answer. You would still benefit by ruling out certain pieces so that you can allocate more tokens to the ones that you have not ruled out.

You will find that the problems soon get difficult. Whether the problems are easy or difficult, you will notice that to solve them you have to use the same method all the time. The important thing is to notice how the problems develop and to learn the method of solving them. In each problem you look across each row, or down each column, and decide what the missing figure is like. Then look for the answer that is right both ways, across **and** down, from the options you have to select from.

As you recall from allocating tokens in a similar task in a previous session, you can change your token allocation as many times as you like before you hit the **Submit** button and **Confirm** your token allocation. You can also review and modify your prior decisions. You will have navigation buttons in the top left corner of the screen. These navigation buttons are only available before you start moving the sliders for a problem. Once you move any sliders for a problem, the navigation buttons disappear and you must submit your answer and move to the next problem in order to see the navigation buttons again.

You can work at your own speed, although we have to be out of the room in 90 minutes. Your screen displays how many minutes (and seconds) are remaining in the top, right corner. Remember it is accurate work that counts. Attempt each problem in turn. Do your best to eliminate the incorrect pieces and find the correct piece to complete it before going on to the next problem. You will have the option to go back to previous problems if you want to. After you have completed the last question you must confirm if you are finished. When you are finished you will be told your total earnings, and will not have the option to go back to any of the problems. If you run out of time, you will automatically be treated as having finished, and told your total earnings.

Where you position each slider depends on your beliefs about the correct answer to the question. Again, each bar shows the amount of money you could earn if the true outcome corresponds to the possible solution shown under the bar.

You will be **paid for each of the 36 problems**, so you should think carefully about each problem. Since you can earn up to $2 for each problem, you could earn up to $72 over all 36 problems. You will not earn anything on any problems for which you have not confirmed an allocation of tokens.

It is up to you to balance the strength of your personal beliefs with the possibility of them being wrong. There are several important points for you to keep in mind when making your decisions:

- First, your belief about the correct answer to each problem is your personal judgment.
- Second, you have up to 90 minutes to complete this task.
- Third, you will not earn anything on any problems for which you have not confirmed an allocation of tokens.
- Fourth, depending on your choices and the correct answer you can earn up to $2 for each problem, or up to $72 over all 36 problems.
- Finally, your choices might also depend on your willingness to take risks or to gamble.

The decisions you make are a matter of personal choice. Please work silently, and make your choices by thinking carefully about the questions you are presented with.

When you are finished you will see a summary of your responses and your total earnings from this task. Your earnings are in addition to the show-up payment you receive for participating.

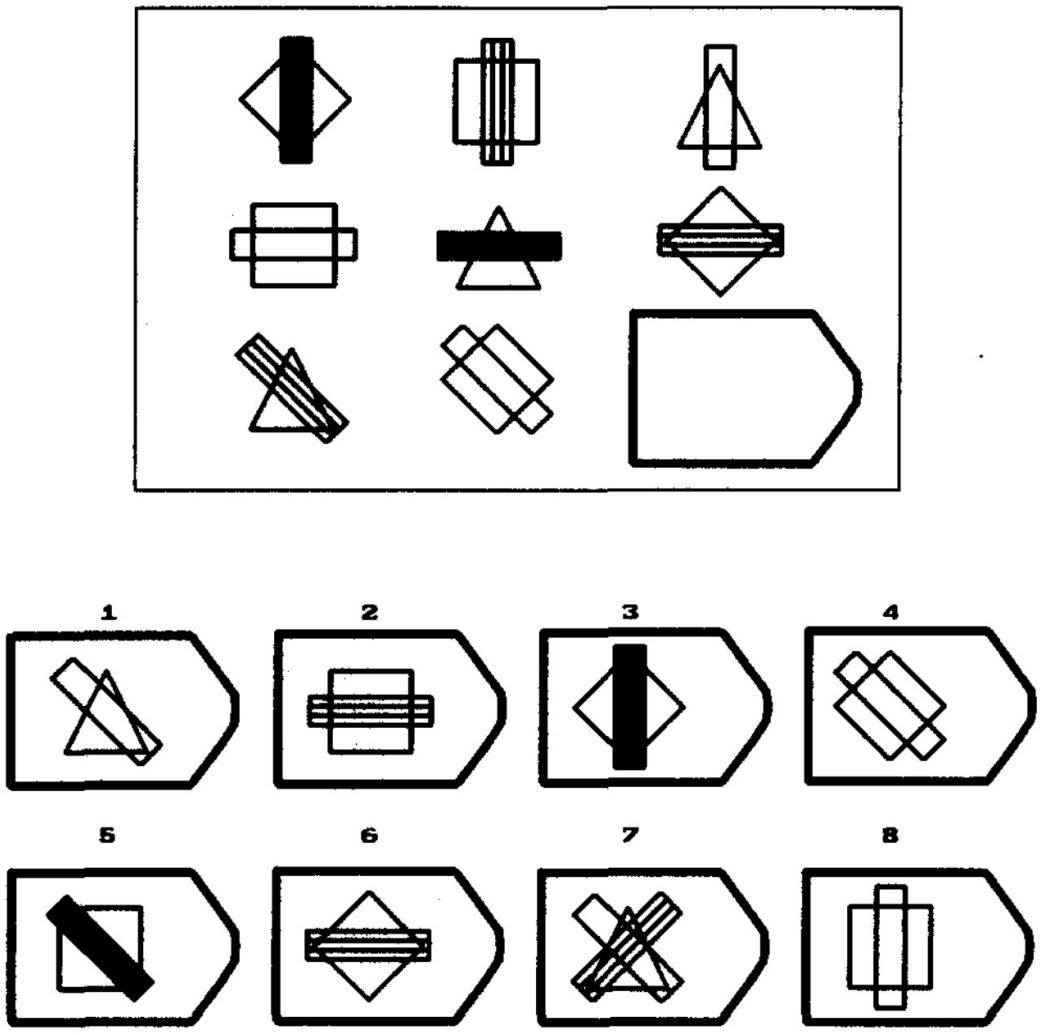
3. Single Token Treatment

Rone

Your Instructions

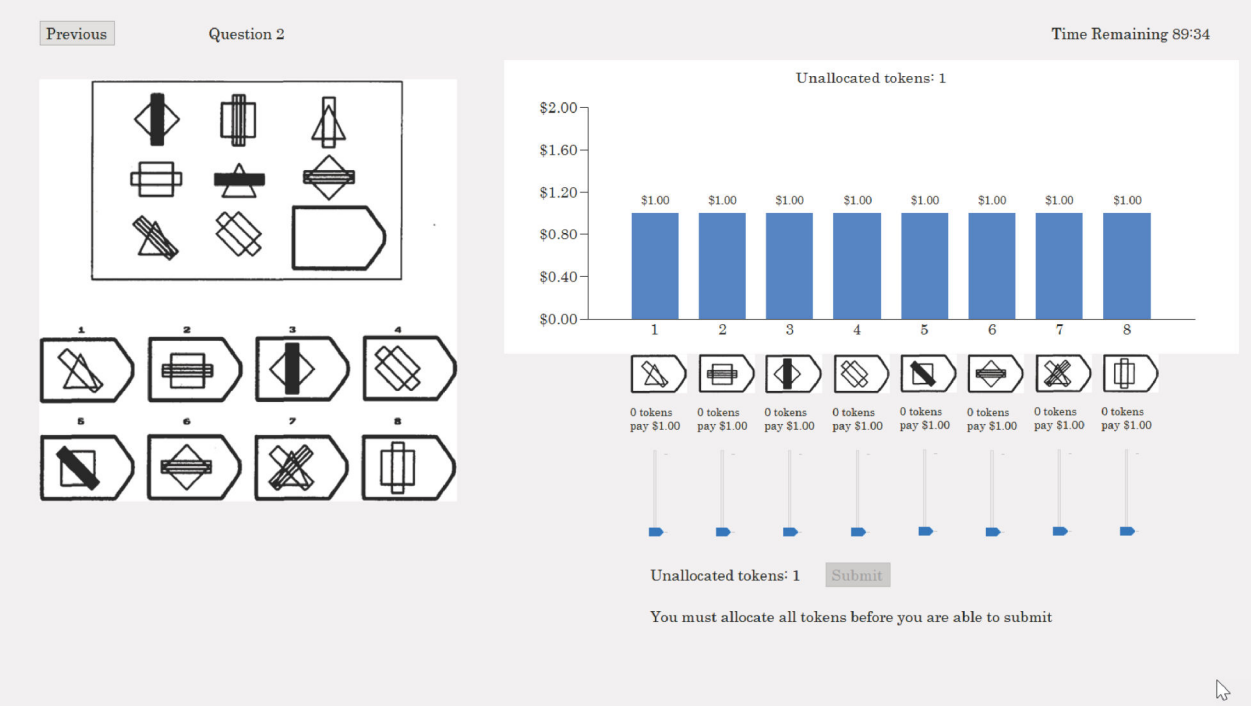
This task is a test of perception and clear-thinking. You have already completed the first part of the task, in a previous session, when you were given 12 similar problems. We will now consider a fresh set of 36 problems.

Consider a similar problem, shown here.



The top part of this problem is a pattern with a bit cut out of it. Look at the pattern, and think what the piece needed to complete the pattern correctly both across and down must be like. Then find the right piece out of the eight bits shown, and numbered, in the bottom part of this problem.

You will be asked to report your beliefs about the correct answer using an interface like this one, which is also generally familiar to you from a previous session. The version you will see on the computer will be larger and easier to read.

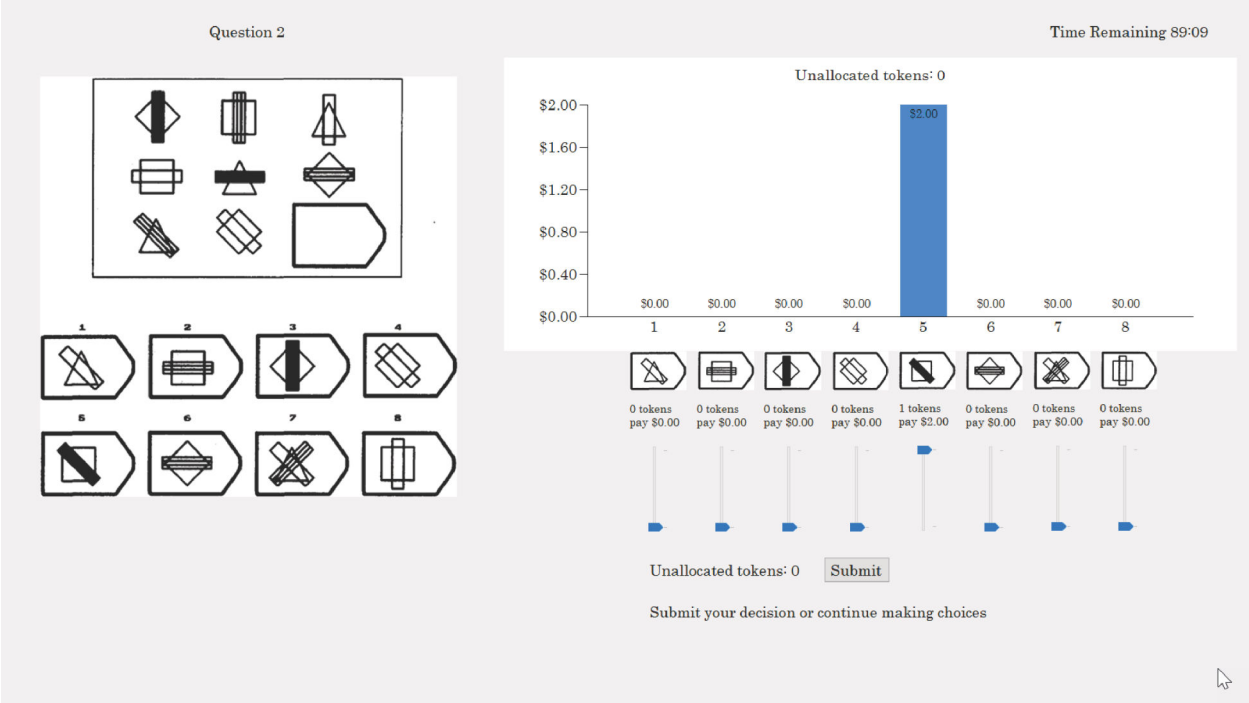


The problem and possible solutions are shown on the left of the screen, in the usual manner. On the right of the screen you have 1 token to allocate across the 8 possible answers. We start off with 0 tokens allocated to each of the possible answers.

As you allocate the token, by moving the sliders up or down, the earnings will change above each bar. These are the earnings that you will receive for this problem if that bar refers to the correct answer to the problem. **You will be paid for all 36 problems, and each problem will pay between $0 and $2 depending on your answer.**

Return now to the problem itself. Only one of these pieces is perfectly correct. Pieces #2, #5 and #8 complete the problem correctly going down, in the sense that they have a square as the larger piece. Pieces #1, #4 and #5 are correct going across, in the sense that they have the right slope for the rectangle.

Piece #5 is the correct bit, isn't it? It is correct going down as well as going across. So the answer is #5. In this case, if you were certain that piece #5 is the single best answer, you should allocate the token to the bin representing piece #5 as in this display:



So in this case you would earn $2.00 if indeed the correct answer was #5. Of course, if any of the other pieces turned out to be the correct answer you would, in this case, earn $0.

This illustrative problem is relatively easy, but some of the later problems will not be so easy, and it might be harder to identify the one correct answer. You would still benefit by ruling out certain pieces so that you can allocate your token to one of the possible answers that you have not ruled out.

You will find that the problems soon get difficult. Whether the problems are easy or difficult, you will notice that to solve them you have to use the same method all the time. The important thing is to notice how the problems develop and to learn the method of solving them. In each problem you look across each row, or down each column, and decide what the missing figure is like. Then look for the answer that is right both ways, across **and** down, from the options you have to select from.

As you recall from allocating tokens in a similar task in a previous session, you can change your token allocation as many times as you like before you hit the **Submit** button and **Confirm** your token allocation. You can also review and modify your prior decisions. You will have navigation buttons in the top left corner of the screen. These navigation buttons are only available before you start moving the sliders for a problem. Once you move any sliders for a problem, the navigation buttons disappear and you must submit your answer and move to the next problem in order to see the navigation buttons again.

You can work at your own speed, although we have to be out of the room in 90 minutes. Your screen displays how many minutes (and seconds) are remaining in the top, right corner. Remember it is accurate work that counts. Attempt each problem in turn. Do your best to eliminate the incorrect pieces and find the correct piece to complete it before going on to the next problem. You will have the option to go back to previous problems if you want to. After you have completed the last question you must confirm if you are finished. When you are finished you will be told your total earnings, and will not have the option to go back to any of the problems. If you run out of time, you will automatically be treated as having finished, and told your total earnings.

Where you allocate the token depends on your beliefs about the correct answer to the question. Again, each bar shows the amount of money you could earn if the true outcome corresponds to the possible solution shown under the bar.

You will be **paid for each of the 36 problems**, so you should think carefully about each problem. Since you can earn up to $2 for each problem, you could earn up to $72 over all 36 problems. You will not earn anything on any problems for which you have not confirmed an allocation of your token.

It is up to you to balance the strength of your personal beliefs with the possibility of them being wrong. There are several important points for you to keep in mind when making your decisions:

- First, your belief about the correct answer to each problem is your personal judgment.
- Second, you have up to 90 minutes to complete this task.
- Third, you will not earn anything on any problems for which you have not confirmed an allocation of your token.
- Fourth, depending on your choices and the correct answer you can earn up to $2 for each problem, or up to $72 over all 36 problems.
- Finally, your choices might also depend on your willingness to take risks or to gamble.

The decisions you make are a matter of personal choice. Please work silently, and make your choices by thinking carefully about the questions you are presented with.

When you are finished you will see a summary of your responses and your total earnings from this task. Your earnings are in addition to the show-up payment you receive for participating.

4. Scrambled Treatment

The instructions for the Scrambled treatments were the same as their one-token and eight-token counterparts, with this text changed:

This illustrative problem is relatively easy, but some of the later problems will not be so easy, and it might be harder to identify the one correct answer. You would still benefit by ruling out certain pieces so that you can allocate more tokens to the ones that you have not ruled out.

You will find that the problems soon get difficult. Whether the problems are easy or difficult, you will notice that to solve them you have to use the same method all the time. The important thing is to notice how the problems develop and to learn the method of solving them. In each problem you look across each row, or down each column, and decide what the missing figure is like. Then look for the answer that is right both ways, across **and** down, from the options you have to select from.

As you recall from allocating tokens in a similar task in a previous session, you can change your token allocation as many times as you like before you hit the **Submit** button and **Confirm** your token allocation. You can also review and modify your prior decisions. You will have navigation buttons in the top left corner of the screen. These navigation buttons are only available before you start moving the sliders for a problem. Once you move any sliders for a problem, the navigation buttons disappear and you must submit your answer and move to the next problem in order to see the navigation buttons again.

The new text completely replaces the second paragraph, deletes the word "later" from the first line of the first paragraph, and replaces "problems develop" with "problems can be solved" in the third paragraph:

This illustrative problem is relatively easy, but some of the problems will not be so easy, and it might be harder to identify the one correct answer. You would still benefit by ruling out certain pieces so that you can allocate more tokens to the ones that you have not ruled out.

The order in which you see the easier and harder problems, as well as the ones in between, will be selected at random for each person. So the problems do not generally become easier or harder as you work through them.

Whether the problems are easy or difficult, you will notice that to solve them you have to use the same method all the time. The important thing is to notice how the problems can be solved, and to learn the method of solving them. In each problem you look across each row, or down each column, and decide what the missing figure is like. Then look for the answer that is right both ways, across **and** down, from the options you have to select from.

Appendix B: Subjective Beliefs and Risk Preferences (Online Working Paper)

In the main text all results use reported beliefs. In this appendix we present the logic behind the calculation of recovered beliefs that take into account estimated risk preferences of individual subjects. We only reproduce the displays for which the distinction between reports and recovered beliefs could matter. To ensure that the exposition here can be read alone, we repeat some text that is included in the main text.

There are three general approaches to eliciting beliefs, which we review in in §5.D. The approach we employ is to directly pay subjects for their reports using financial incentives, and then recover beliefs if needed using estimates of the risk preferences of individuals. In order to infer subjective beliefs we therefore need two experimental tasks: one in which we elicit risk preferences defined over objective lotteries, and one in which we elicit subjective beliefs using the QSR defined over monetary payoffs. We ensure that the scale of payoffs in each task is comparable, to avoid extrapolation. We have each subject undertake both tasks to allow estimation of risk preferences, and recovery of subjective beliefs at the level of the individual.

Every subject completed a risk battery of 30 or 100 binary choices. In the 2019 experiments each subject received the same 100 lottery pairs, randomly ordered for each subject, from the battery used by Harrison et al. [2022c]. In the 2022 experiments each subject received 30 lottery pairs, selected at random from a larger battery of 67 gain-frame lottery choices from Harrison and Swarthout [2023]. We estimate risk preferences for each subject using a Bayesian Hierarchical Model that allows us to have informative priors for inferences about the risk preferences of each individual, following Gao, Harrison and Tchernis [2023]. We estimate risk preferences for each individual assuming EUT or assuming RDU. Although RDU nests EUT, for some normative applications it is useful to have EUT estimates. A major advantage of the Bayesian approach to estimating risk preferences is that each of the simulated values of the posterior distribution of parameters can be *directly* used to infer subjective beliefs as a matter of the theory developed in Harrison, Monroe and Ulm [2022] .

It is a straightforward matter to evaluate the expected welfare cost to a respondent of being forced to report their subjective beliefs in a way that requires that they only report their modal belief. In effect, this is the welfare cost to respondents of forcing them to use our **One Token** instrument instead of the **Eighty Tokens** instruments. This calculation uses the recovered subjective beliefs of each subject in each question, the QSR payoffs implied by their token allocation, and their risk preferences to evaluate the risky lottery posed to each subject by such an intervention. Using this information we can calculate the Certainty Equivalent (CE) to the subject, and then report that as a percent of the CE of using the unconstrained instrument.

In the main text we used the RDU risk preferences of the subject to calculate the welfare cost, and reported them in Figure 13. Some find EUT a more attractive normative metric for such calculations, so we report in Figure B13 the same calculations when one uses EUT risk preferences for the subject. The qualitative pattern is virtually the same as if one used RDU risk preferences (this is not, in general, the case).

In the main text we used recovered beliefs using the RDU risk preferences of the subjects in Figure 16, for the belief of winning the competitions. Figure B16 show the same results for these subjects, but using their reported beliefs. We draw the same conclusions, apart from women being more optimistic of winning the tournament in the 80-token case; they were already modestly optimistic in

terms of their recovered beliefs.

In the main text we used recovered beliefs using the RDU risk preferences of the subjects in Figures 17 and 18, for the core financial literacy questions. Figures B17 and B18 show the same results for these subjects, but using their reported beliefs. We draw the same conclusions.

Figure B13: Expected Welfare Cost to Respondent of Being Required to Only Report Modal Belief

Evaluated using EUT risk preferences of each subject
Using recovered beliefs from **Eighty Tokens** task

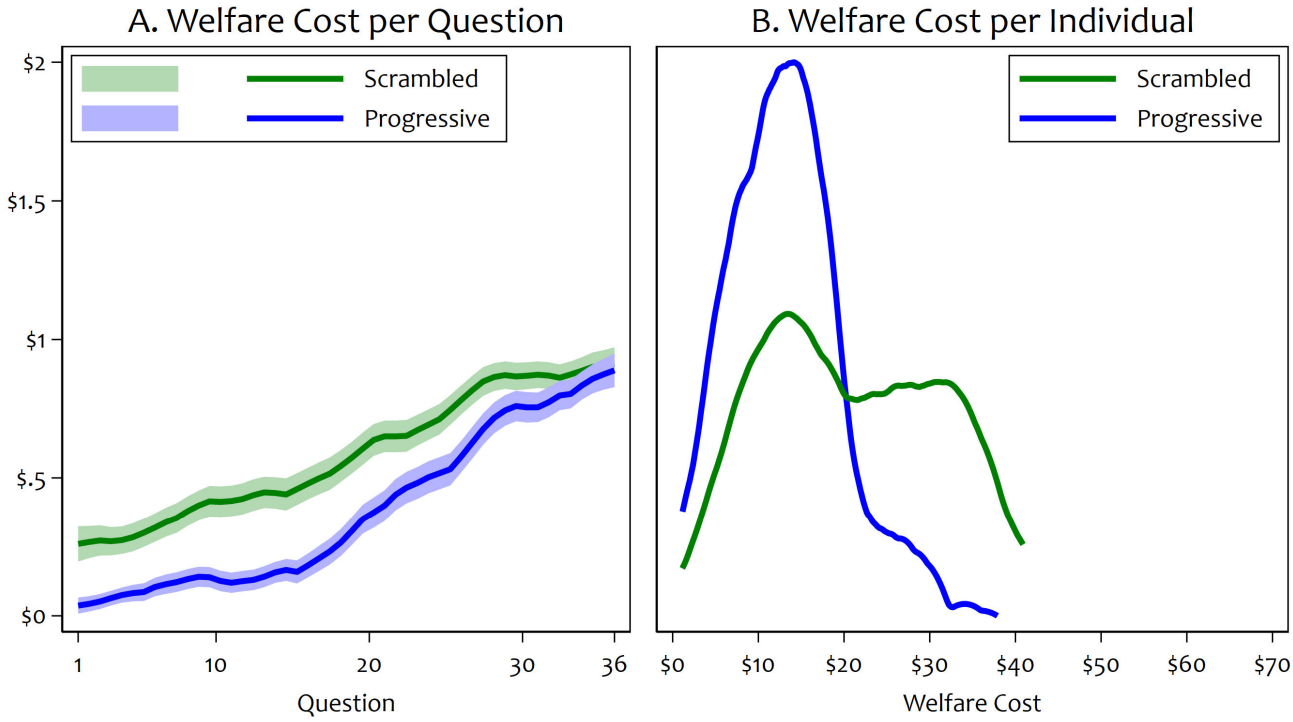


Figure B16: Reported Beliefs on the Probability of Personal Performance Ranks by Men and Women

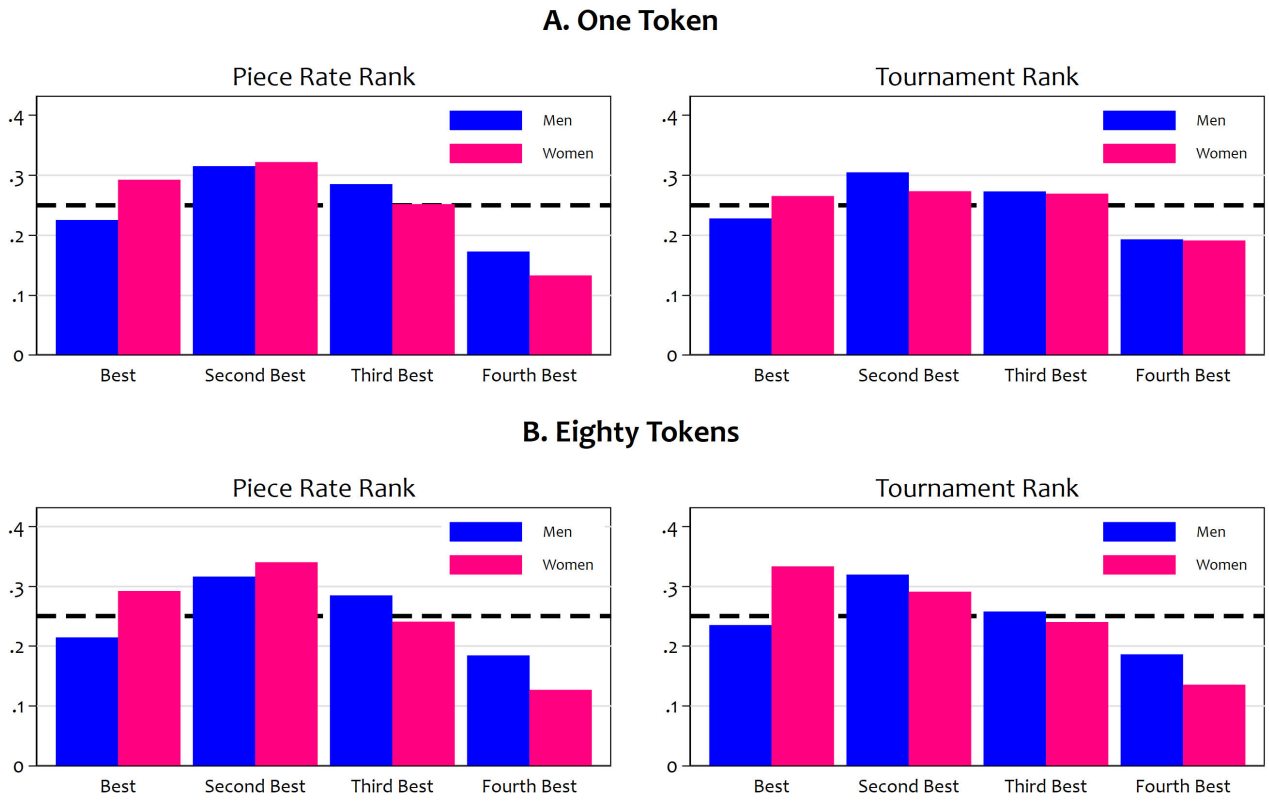


Figure B17: Bias and Confidence in the Literacy of Men and Women on the Inflation Question

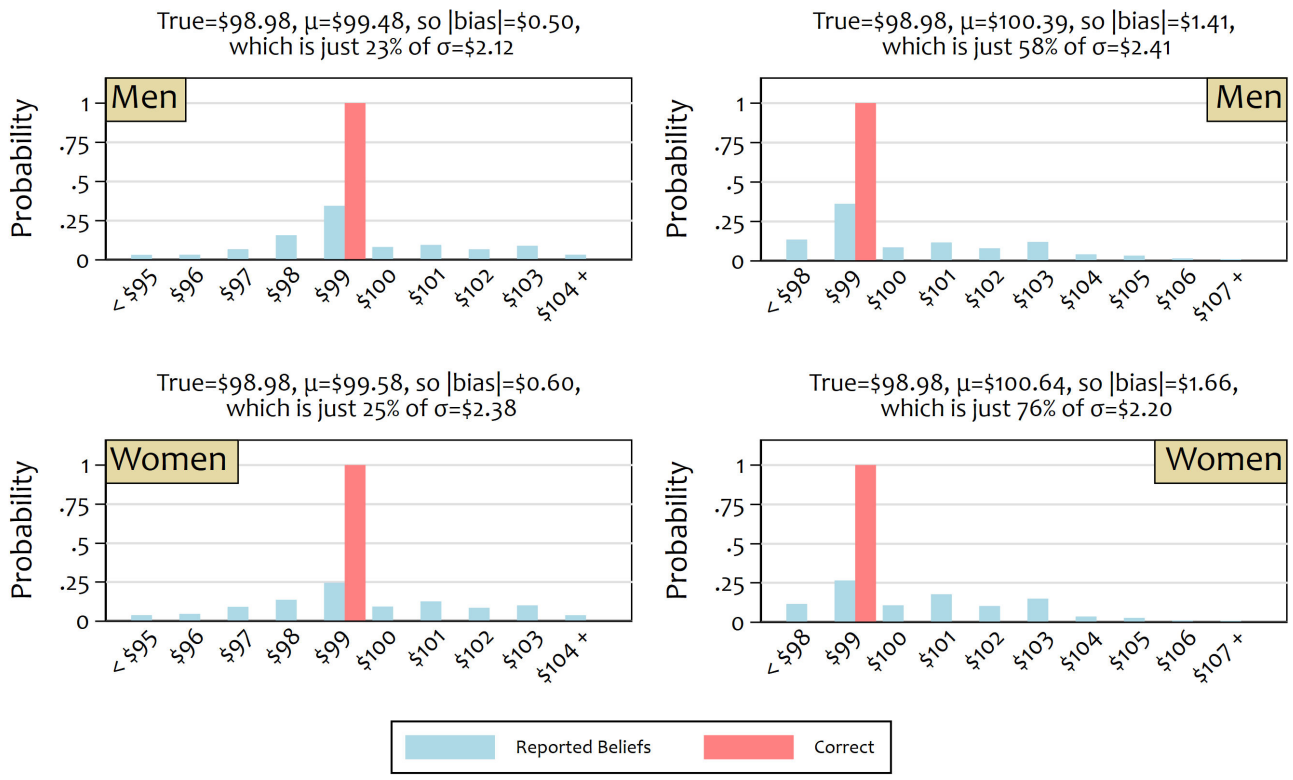
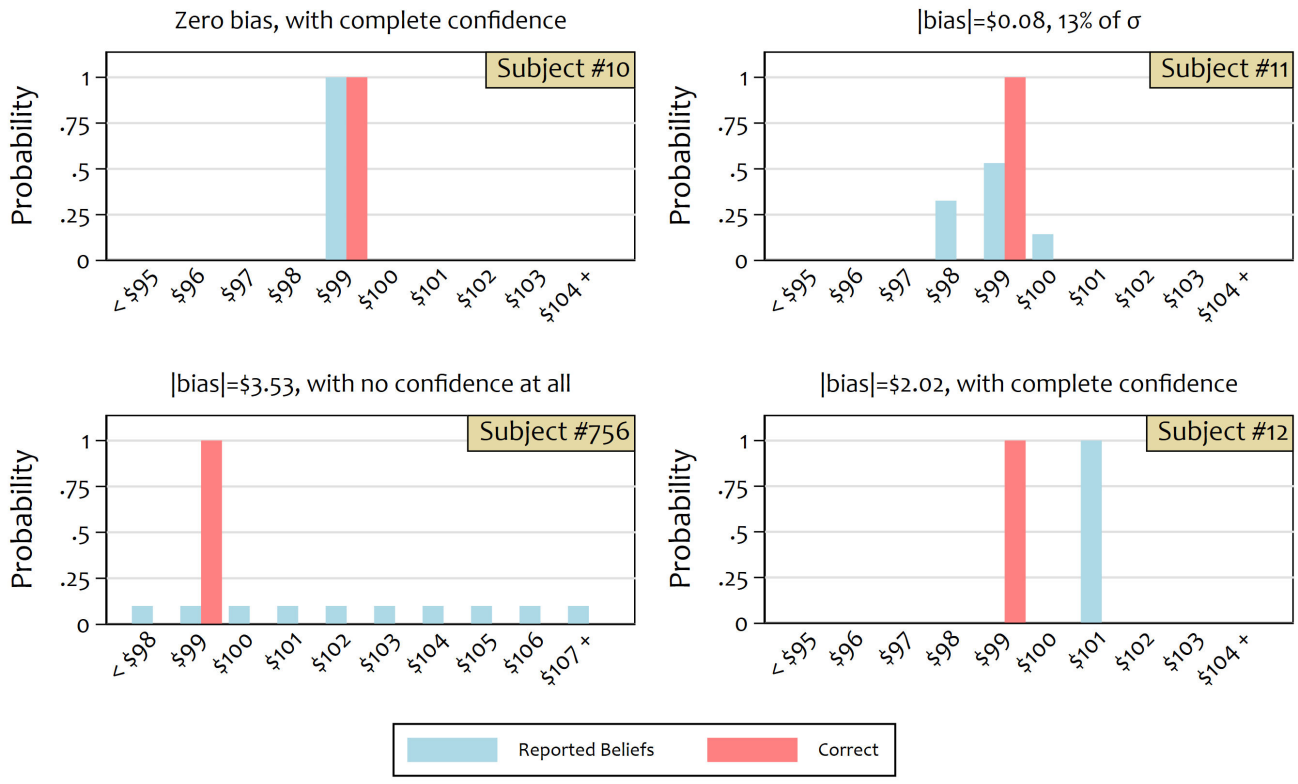


Figure B18: Literacy and Confidence: The Good, the Bad, and the Ugly



Appendix C: Gender and Competitiveness (Online Working Paper)

In Section 1 we provide a numerical example of the welfare calculations of behavior using data drawn from the experiment of Niederle and Vesterlund [2007]. We refer to the core results in the text, and use the same logic to evaluate behavior in our own experiment. In Section 2 we provide the instructions used in our extension of their design. These instructions refer to the 80 token condition, and were different in obvious ways for the 1 token condition. In Section 3 we show additional results from our experiments, referred to in the text, as well as used in Table 2.

1. Numerical Examples

The *Excel* display in Table C1 uses data on average solutions from Niederle and Vesterlund [2007; Table I, p. 1081]. These values are in columns C and D. The payoffs in cells E, F and G then reflect the payoffs described in the experiment. Cell E4 is the average female production under a piece rate scheme, multiplied by $0.50 per unit. Similarly, cell F5 assumes that the average woman produced 11.77 units in their Task 2 with a tournament would win, and hence receives $2 times 11.77 units. Sadly, if she loses, she receives the $0 in cell G4, no matter how many units she produces. Cells H4, I4 and J4 just fill in the probabilities explained in the text, under the null hypothesis that subjective probabilities of winning equal empirical probabilities of winning.

We assume EUT for the moment. Nothing hinges on that, other than ease of exposition of the basic point. Further assuming a CRRA utility function $u(x) = x^{1-r}/(1-r)$, in column K three possible values of the CRRA parameter r are considered. Risk neutrality is when $r = 0$, very modest risk aversion is when $r = 0.1$, and a more realistic risk aversion is considered when $r = 0.53$ for women and $r = 0.43$ for men. These estimates come from a sample if 614 GSU undergraduates, in experiments reported in Harrison, Morsink and Schneider [2022], with roughly 60% of the sample being female.

It is then a simple matter to calculate the utility of each possible payoff, the expected utility (EU) of the safe piece rate lottery and the EU of the risky tournament lottery. These calculations are in columns S, T and U and W and X, respectively, of the Excel (not shown in Table D1). Armed with the EU of each compensation scheme, and the CRRA parameter, we can calculate the Certainty Equivalent (CE) of selecting each compensation scheme, since $u(CE) = EU$ solves for $CE = [EU \times (1-r)]^{(1/(1-r))}$ for this CRRA utility function. The CE of each scheme is displayed in columns L amd M. The difference in the CE, assuming the tournament scheme is chosen, is then displayed in column N. This is the Expected Consumer Surplus (ECS) of selecting the tournament scheme.

We observe confirmation of the intuition stated in the text. If subjects are all risk neutral, and $r = 0$, then it does make sense to select the tournament scheme. Even for women that actually selected the piece rate scheme, if they had been risk neutral they would have gained in welfare terms by the equivalent of receiving $0.71 for certain. For risk-neutral women that actually selected the tournament scheme, the welfare gain would have been $1.07. For men, in each case, the welfare gains would have been $0.59 and $0.77, respectively.

The rub comes when we consider risk aversion. In all but one case, women that chose the tournament, being modestly risk averse with $r = 0.1$ means a welfare loss from selecting the tournament. These welfare losses are only around $0.13, and $0.20, but it is the negative sign that matters.

If we consider more realistic risk aversion estimates of the CRRA parameter r, the welfare losses from selecting the tournament become noticeable. In fact, we can state the welfare effects precisely:

- The women that chose the piece rate scheme *gained* in welfare terms by +$3.94.
- The women that chose the tournament scheme *lost* in welfare terms by $3.65.
- The men that chose the piece rate scheme *gained* in welfare terms by +$3.01.
- The men that chose the tournament scheme *lost* in welfare terms by $3.01.

Hence it is important to recognize that women tended to do the right thing, in their Task 3 and assuming that EUT describes their risk preferences, by selected the tournament less than 50% of the time: only 35% of the time. But men tended to do the wrong thing, by selecting the tournament more than 50% of the time: 73% of the time, in fact.

The *Excel* example in Table C2 corresponds to the RDU example summarized in Table 2 in the text, using data from our experiments.

Table C1: Some EUT Finger-Math Applied to the Niederle and Vesterlund [2007] Data

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		Selected	Average Performance		Payoffs			Probabilities			CRRA	Certainty Equivalents		ECS
2	Gender	Scheme	Piece Rate	Tournament	Piece Rate	Tournament if Win	Tournament if Lose	Piece Rate	Tournament Win	Tournament Lose	r	Piece Rate	Tournament	Tournament
3														
4	Female	Piece Rate	10.35	11.77	$5.18	$23.54	$0.00	1	0.25	0.75	0	$5.18	$5.89	$0.71
5			10.35	11.77	$5.18	$23.54	$0.00	1	0.25	0.75	0.1	$5.18	$5.05	-$0.13
6			10.35	11.77	$5.18	$23.54	$0.00	1	0.25	0.75	0.53	$5.18	$1.26	-$3.92
7														
8	Female	Tournament	9.79	11.93	$4.90	$23.86	$0.00	1	0.25	0.75	0	$4.90	$5.97	$1.07
9			9.79	11.93	$4.90	$23.86	$0.00	1	0.25	0.75	0.1	$4.90	$5.11	$0.22
10			9.79	11.93	$4.90	$23.86	$0.00	1	0.25	0.75	0.53	$4.90	$1.27	-$3.62
11														
12	Male	Piece Rate	9.91	11.09	$4.96	$22.18	$0.00	1	0.25	0.75	0	$4.96	$5.55	$0.59
13			9.91	11.09	$4.96	$22.18	$0.00	1	0.25	0.75	0.1	$4.96	$4.75	-$0.20
14			9.91	11.09	$4.96	$22.18	$0.00	1	0.25	0.75	0.43	$4.96	$1.96	-$3.00
15														
16	Male	Tournament	10.97	12.52	$5.49	$25.04	$0.00	1	0.25	0.75	0	$5.49	$6.26	$0.78
17			10.97	12.52	$5.49	$25.04	$0.00	1	0.25	0.75	0.1	$5.49	$5.37	-$0.12
18			10.97	12.52	$5.49	$25.04	$0.00	1	0.25	0.75	0.43	$5.49	$2.21	-$3.28

Table C2: Some RDU Finger-Math Applied to Our Experimental Data

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1			Payoffs				Probabilities		CRRA	PWF		Certainty Equivalents		ECS
2	Gender		Piece Rate	Tournament if Win	Tournament if Lose	Piece Rate	Tournament Win	Tournament Lose	r	η	φ	Piece Rate	Tournament	Tournament
3														
4	Female		$1.36	$5.72	$0.00	1	0.33	0.67	0	1	1	$1.36	$1.89	$0.53
5			$1.36	$5.72	$0.00	1	0.33	0.67	0.1	1.25	1	$1.36	$1.23	-$0.13
6			$1.36	$5.72	$0.00	1	0.33	0.67	0.70	0.90	0.92	$1.36	$0.26	-$1.10
7														
8	Male		$1.37	$5.66	$0.00	1	0.23	0.77	0	1	1	$1.37	$1.30	-$0.07
9			$1.37	$5.66	$0.00	1	0.23	0.77	0.1	1.25	1	$1.37	$0.74	-$0.63
10			$1.37	$5.66	$0.00	1	0.23	0.77	0.71	0.97	1.00	$1.37	$0.07	-$1.30

2. Experimental Instructions

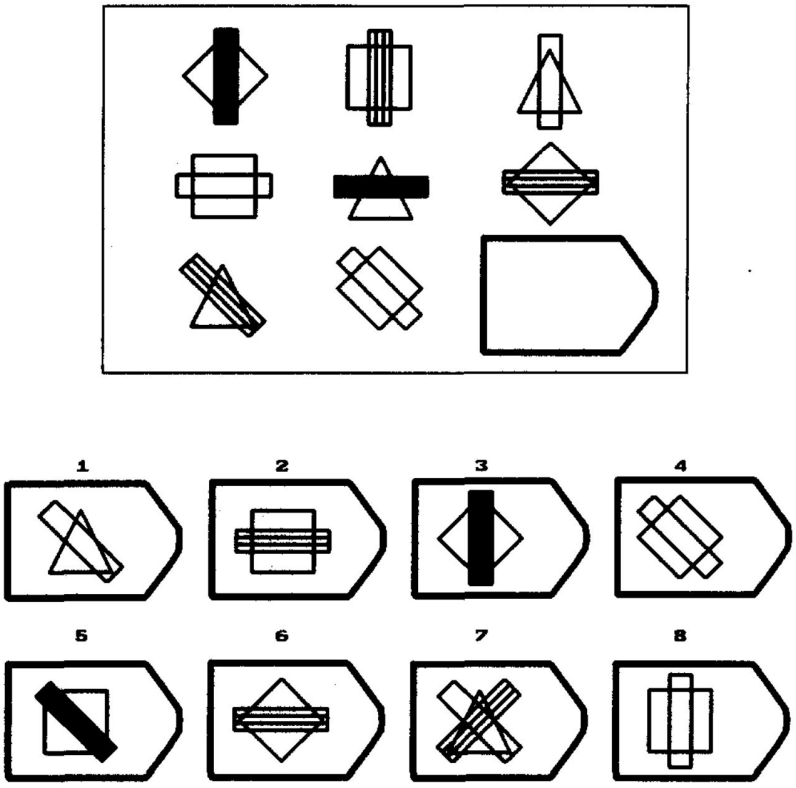
Comp80

WELCOME!

The final task today consists of six parts. The first part will take a bit longer than the rest, since it provides an introduction that will apply to all parts. We will randomly select one of the six parts of this task and pay you based on your performance in that part. Once you have completed the six parts we determine which part counts for payment by drawing a number between 1 and 6. The method we use to determine your earnings varies across parts. Before each part we will describe in detail how your payment is determined.

Part 1 – Piece Rate

This part of the task is a test of perception and clear thinking. In a previous session you were given 12 similar problems. We will now consider a fresh set of 12 problems, like this one:

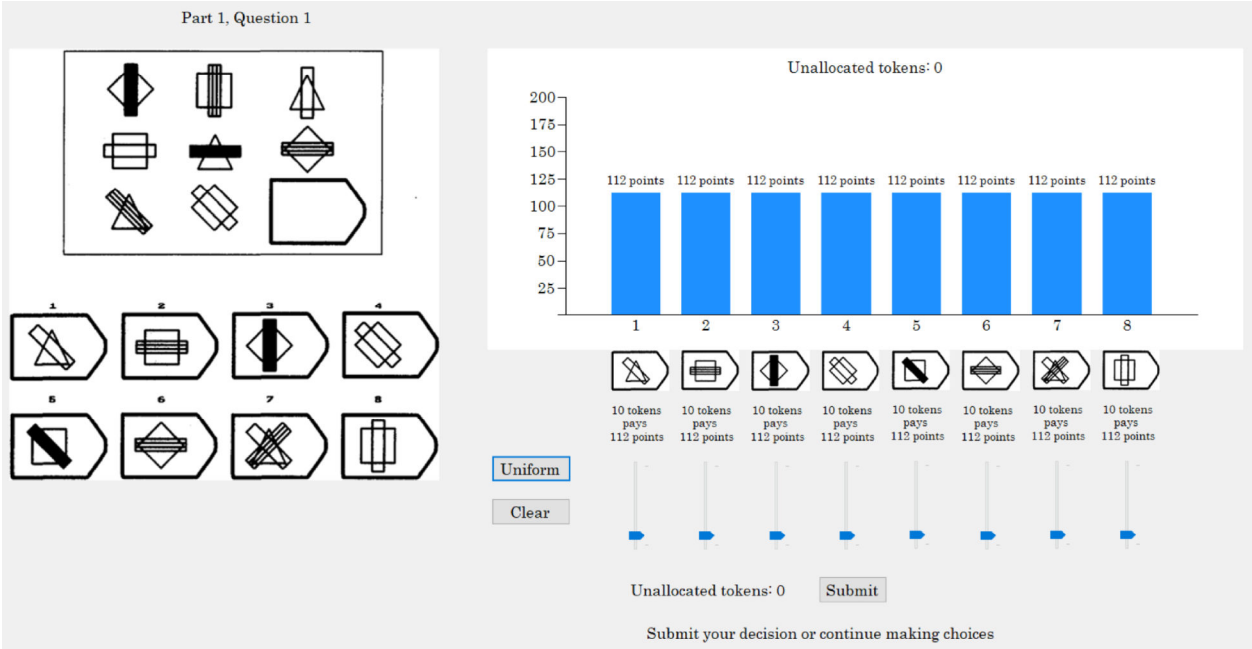


The top part of this problem is a pattern with a piece cut out of it. Look at the pattern, and think what the piece needed to complete the pattern correctly both across and down must be like. Then find the right piece out of the eight pieces shown, and numbered, in the bottom part of this problem.

For **Part 1** you will be asked to report your beliefs about 12 problems like this one, and you will have up to 10 minutes to do so. If **Part 1** is the part randomly selected for payment, then you get up to 200 points per problem you solve correctly in the 10 minutes. If **Part 1** is the part randomly selected for payment, then each point you earn gets converted to money, and 100 points is converted to $1. So if you earn 200 points per problem you solve correctly, in **Part 1** you will earn $2 per problem you solve correctly. Your payment does not decrease if you provide an incorrect answer to a problem. We refer to this payment as the **Piece Rate** payment.

You will be asked to report your beliefs about the correct answer using an interface like this one, which is also generally familiar to you from a previous task. The version you will see on the computer will be larger and easier to read.

The problem and possible solutions are shown on the left of the screen, in the usual manner. On the right of the screen you have 80 tokens to allocate across the 8 possible answers. We start off with 0 tokens allocated to each of the possible answers. If you wanted to change this initial allocation so that there were 10 tokens allocated to each possible answer, just click on the **Uniform** button, and you will see this display:

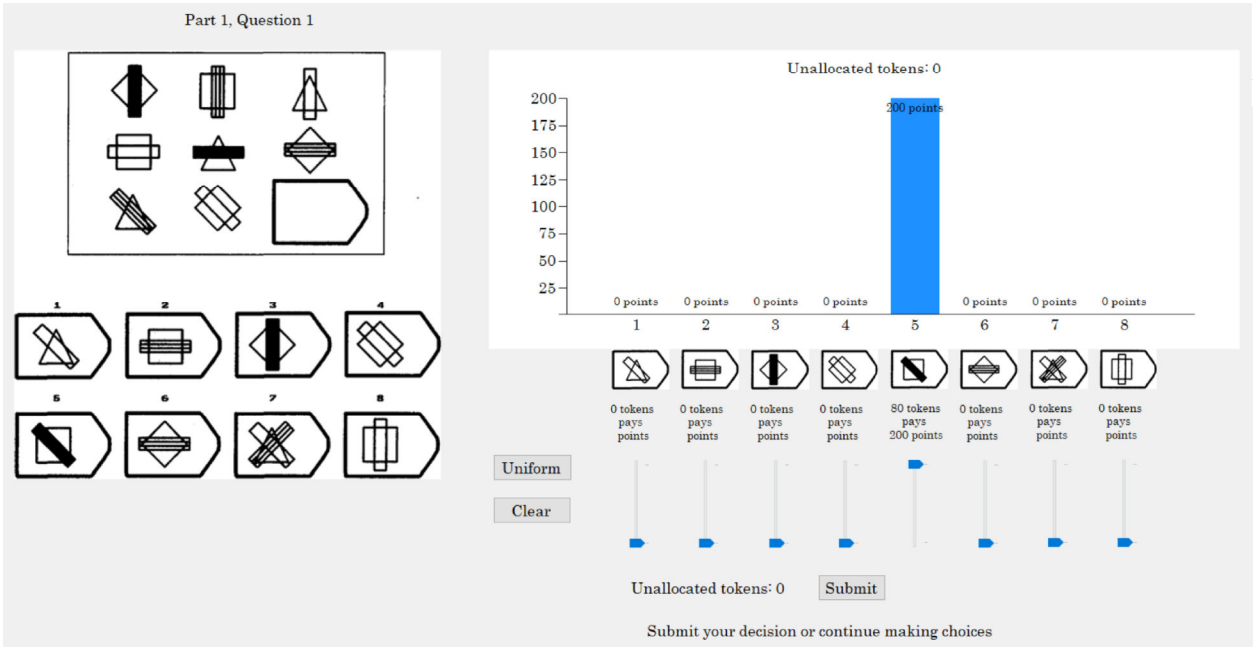


Here you have allocated 10 tokens to each possible answer, and you would earn 112 points if you reported this allocation of tokens, since only one of the 8 possible answers is correct. You can return to the initial allocation of 0 tokens for each possible answer by clicking on the **Clear** button.

As you allocate tokens, by moving the sliders up or down, the points will change above each bar. These are the points that you will receive for this problem if that bar corresponds to the correct answer to the problem. **You will be paid for all 12 problems, and each problem will pay between 0 points and 200 points depending on your answer.**

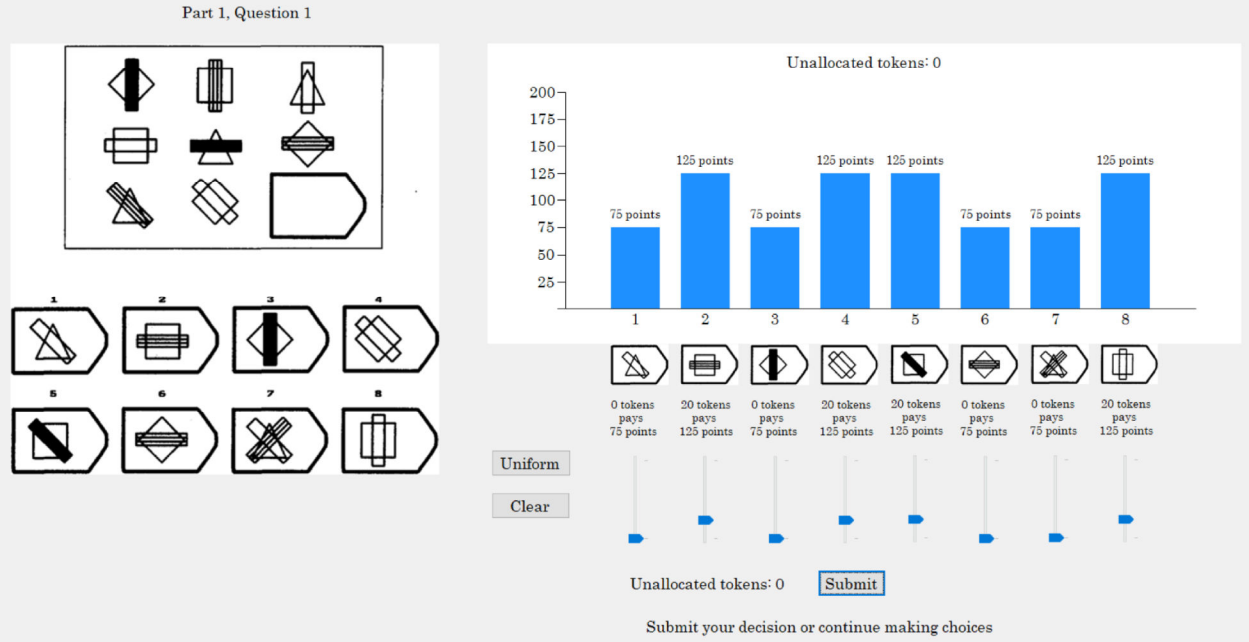
Return now to the problem itself. Only one of these pieces is perfectly correct. Pieces #2, #5 and #8 complete the problem correctly going down, in the sense that they have a square as the larger piece. Pieces #1, #4 and #5 are correct going across, in the sense that they have the correct slope for the rectangle.

Piece #5 is the correct piece both ways, isn't it? It is correct going down as well as going across. So the answer is #5. In this case, if you were certain that piece #5 is the single best answer, you should allocate all 80 tokens to the bin representing piece #5 as in this display:



So in this case you would earn 200 points if indeed the correct answer was #5. Of course, if any of the other pieces turned out to be the correct answer you would, in this case, earn 0 points.

If you had decided that the correct answer was one of #2, #4, #5 or #8, but had not decided that #5 was actually the correct answer of these four possibilities, you might decide to allocate your tokens equally across the bars representing pieces #2, #4, #5 and #8 like this:



You can see that even if you eliminate some pieces, such as #1, #3, #6 and #7, that are clearly wrong, you give yourself a 1-in-4 chance of earning more money than if you guessed across all 8 pieces. In this case you would expect to earn 125 points if indeed one of pieces #2, #4, #5 or #8 had been correct. Recall that if you had allocated the tokens roughly equally across all 8 bars, thinking that any of the 8 pieces might be correct, you would only earn 112 points for this problem.

This illustrative problem is relatively easy, but some of the problems will not be so easy, and it might be harder to identify the one correct answer. You would still benefit by ruling out certain pieces so that you can allocate more tokens to the ones that you have not ruled out.

The order in which you see the easier and harder problems, as well as the ones in between, will be selected at random. So the problems do not generally become easier or harder as you work through them.

Whether the problems are easy or difficult, you will notice that to solve them you have to use the same method all the time. The important thing is to notice how the problems can be solved, and to learn the method of solving them. In each problem you look across each row, or down each column, and decide what the missing piece should be like. Then look for the answer that is correct both ways, across **and** down, from the options you have to select from.

As you recall from allocating tokens in a similar part in a previous session, you can change your token allocation as many times as you like before you hit the **Submit** button and **Confirm** your token allocation. You can also review and modify your prior decisions. You will have navigation buttons to the left of the sliders. These navigation buttons are only available before you start moving the sliders for a problem.

You will be allowed 10 minutes to complete this part. Your screen displays how many minutes (and seconds) are remaining in the top, right corner. Remember it is accurate work that counts. Attempt each problem in turn. Do your best to eliminate the incorrect pieces and find the correct piece to complete it before going on to the next problem. You will have the option to go back to previous problems if you want to. After you have completed the last question you must confirm if you are finished. When you are finished you will be told your total points **if Part 1 is selected for payment**, and will not have the option to go back to any of the problems. If you run out of time, you will automatically be treated as having finished, and told your total points if Part 1 is selected for payment.

Where you position each slider depends on your beliefs about the correct answer to the question. Again, each bar shows the number of points you could earn if the true outcome corresponds to the possible solution shown under the bar.

You will be **paid for each of the 12 problems**, so you should think carefully about each problem. Since you can earn up to 200 points for each problem, you could earn up to 2,400 points over all 12 problems. And **if Part 1 is selected for payment**, these 2,400 points would be converted to $24. You will not earn anything on any problems for which you have not confirmed an allocation of tokens.

It is up to you to balance the strength of your personal beliefs with the possibility of them being wrong. There are several important points for you to keep in mind when making your decisions:

- First, your belief about the correct answer to each problem is your personal judgment.
- Second, you have up to 10 minutes to complete this part.
- Third, you will not earn anything on any problems for which you have not confirmed an allocation of tokens.
- Fourth, depending on your choices and the correct answer you can earn up to 200 points for each problem, or up to 2,400 points over all 12 problems.
- Finally, your choices might also depend on your willingness to take risks or to gamble.

The decisions you make are a matter of personal choice. Please work silently, and make your choices by thinking carefully about the questions you are presented with.

When you are finished you will see a summary of your responses and your total earnings from this part of the task if it is selected for payment.

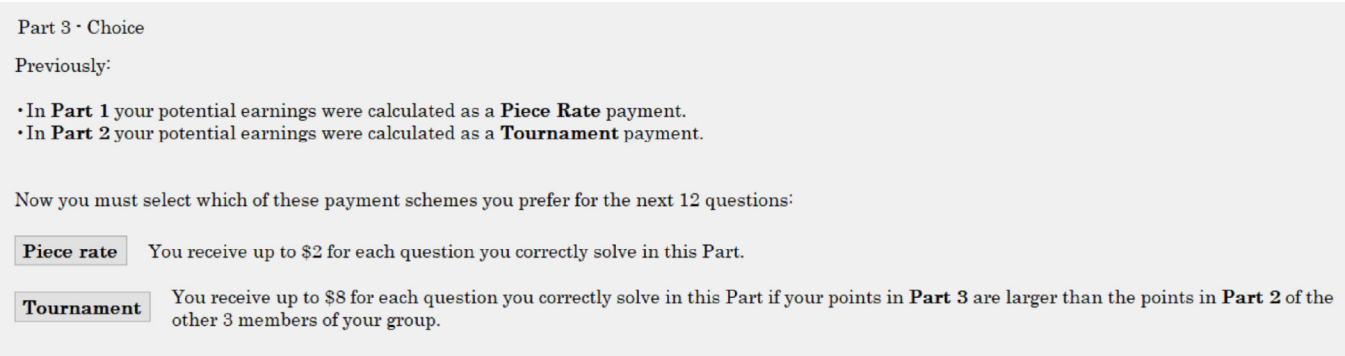
Part 2 – Tournament

As in **Part 1** you will be given 10 minutes to report your beliefs about the solution to up to 12 problems like the ones in **Part 1**. However, for this part your payment depends on the points earned from your reports compared to the points of three other participants in your group. Each group consists of four people, selected at random from other participants in **Part 2**.

If **Part 2** is the part of the task randomly selected for payment, then your cash earnings depend on the points you get from your reports compared to the points of the three other people in your group. The individual with the most points over all 12 problems will receive up to $8 per problem, while the other participants receive no cash payment for **Part 2**. In other words, every 100 points you get from your reports gets converted to $4 in cash, and every 200 points you get from your reports gets converted to $8 in cash. We refer to this as the **Tournament** payment. You will not be told how you did in the tournament until all six parts have been completed, although you will know your own cash earnings in **Part 2** if it is selected for payment and you were the winner. If there are ties the winner will be randomly selected.

Part 3 – Choose Payment Scheme

As in the previous two parts, you will be given 10 minutes to report your beliefs about the solution to up to 12 problems like the ones in **Part 1** and **Part 2**. However, in **Part 3** you will now get to choose which of the two previous payment schemes you prefer to apply to your points. Here is the screen you will face to make your choice.



If **Part 3** is the part of the task randomly selected for payment, then your cash earnings for this task are determined as follows.

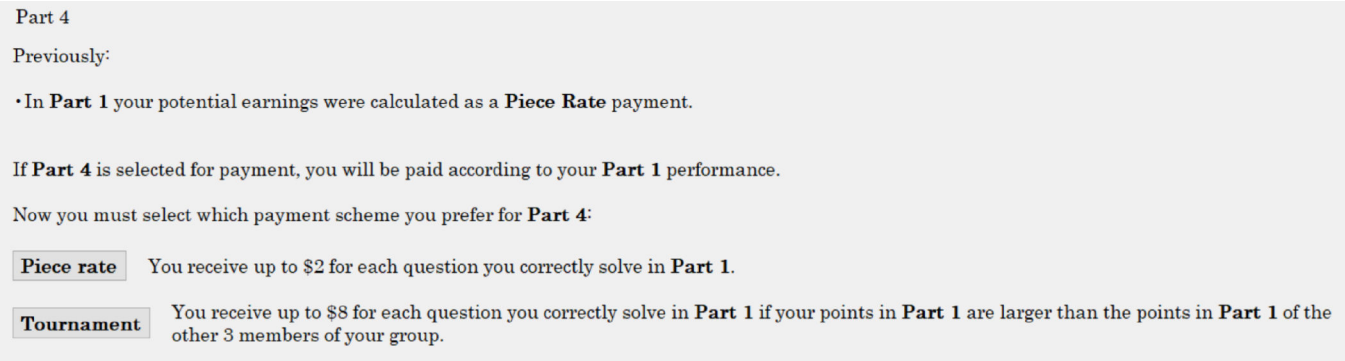
- If you choose the **Piece Rate** payment you receive up to $2 per problem you solve correctly. So, as in Part 1, every 100 points you get from your reports converts to $1 in cash, and every 200 points you get from your reports converts to $2 in cash.

- If you choose the **Tournament** payment your points will be compared to the points of the other three participants in your group in **Part 2 – Tournament**. The **Part 2 – Tournament** is the one you just completed. If your points in **Part 3** are larger than their reports in **Part 2**, then you receive four times the payment from the piece rate: you receive up to $8 per problem you solve correctly. So, as in **Part 2**, every 100 points you get from your reports converts to $4 in cash, and every 200 points you get from your reports converts to $8 in cash. You will receive no earnings for this part if you choose the **Tournament** and points from your reports in **Part 3** are smaller than the points from the reports of the three others in your group in **Part 2**. You will not be informed of how you did in the **Tournament** until all six parts have been completed. If there are ties the winner will be randomly determined.

You are asked to choose whether you want the **Piece Rate** or the **Tournament** applied to your performance. You will then be given 10 minutes to report your beliefs on the solution to up to 12 problems.

Part 4 – Submit Your Piece Rate Performance

You do not have to report your beliefs on any more problems for **Part 4**. Instead, you may be paid one more time for your earnings in **Part 1 – Piece Rate** of the task. However, you now must choose which payment scheme you want applied to your points from **Part 1**. You can either choose to be paid according to the **Piece Rate**, or according to the **Tournament**. Here is the screen you will face to make your choice.



If **Part 4** is the part of the task selected for payment, then your earnings for this task are determined as follows.

- If you choose the **Piece Rate** payment you receive up to $2 per problem you correctly solved in **Part 1**, so 100 points again converts to $1 and 200 points again converts to $2.

- If you choose the **Tournament** payment your performance will be compared to the performance of the other three participants in your group in **Part 1 – Piece Rate**. If the points from your reports in **Part 1** are larger than the points from their reports in **Part 1**, then you receive four times the payment compared to the piece rate: you receive up to $8 per problem you solve correctly, since in this case 100 points converts to $4 and 200 points converts to $8. You will receive no earnings for this part if you choose the **Tournament** and your points from **Part 1** are smaller than the points of the three others in your group in **Part 1**.

Part 5 – Beliefs About Your Part 1 Rank

In **Part 5** you are asked to report your beliefs about the rank that *your* points in the **Part 1 – Piece Rate** task had, when compared to the points in **Part 1** of the other three participants of your group. Here is the screen you will face to make your choice.

You have 100 tokens to allocate. And you must allocate them over the four possible outcomes for the rank of your points in the **Part 1 – Piece Rate** task compared to the rank of the points of the other three participants in your group. If **Part 5** is selected for payment you will be paid the cash earnings indicated on this screen for the actual rank of your points, which is based on the number of tokens you allocate to that points rank outcome.

Part 6 – Beliefs About Your Part 2 Rank

In **Part 6** you are asked to report your beliefs about the rank that *your* points in the **Part 2 – Tournament** task had, when compared to the points in **Part 2** of the other three participants of your group. Here is the screen you will face to make your choice.

You again have 100 tokens to allocate. And you must allocate them over the four possible outcomes for the rank of your points in the **Part 2 – Tournament** task compared to the rank of the points of the other three participants in your group. If **Part 6** is selected for payment you will be paid the cash earnings indicated on this screen for the actual rank of your points, which is based on the number of tokens you allocate to that points rank outcome.

3. Additional Results

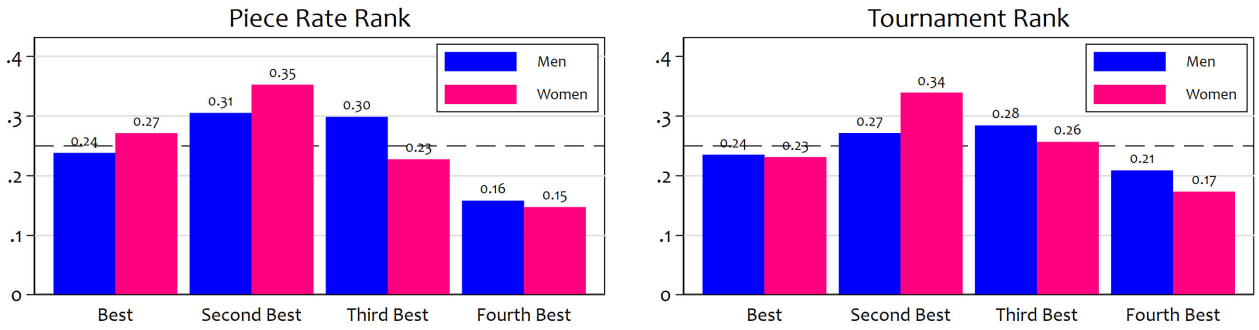
Figures C1 and C2 provide detailed results on the reported beliefs of men and women, or Blacks and non-Blacks, referred to in the text.

Figures C3 and C4 provide displays of the reported beliefs and recovered beliefs of subject #2. This subject had a concave utility function and probability weighting implying very slight, global probability pessimism, so relatively close to EUT risk preferences. Her reports for the QSR are therefore slightly smoother than her beliefs. She also exhibits strikingly different beliefs over her piece rate ranks and her tournament ranks, with considerable confidence of winning the tournament.

Figures C5 and C6 provide displays of the reported beliefs and recovered beliefs of subject #3. This subject also had a concave utility function, but his probability weighting implying significant, global probability pessimism, so clearly not close to EUT risk preferences. His reports for the QSR are again smoother than his beliefs, but the difference between beliefs and reports is much more substantial than for subject #2. He reported the same beliefs over his piece rate ranks and his tournament ranks, with considerable confidence of being ranked second or third in both.

Figure C1: Recovered Beliefs on the Probability of Personal Performance Ranks by Men and Women

A. One Token



B. Eighty Tokens

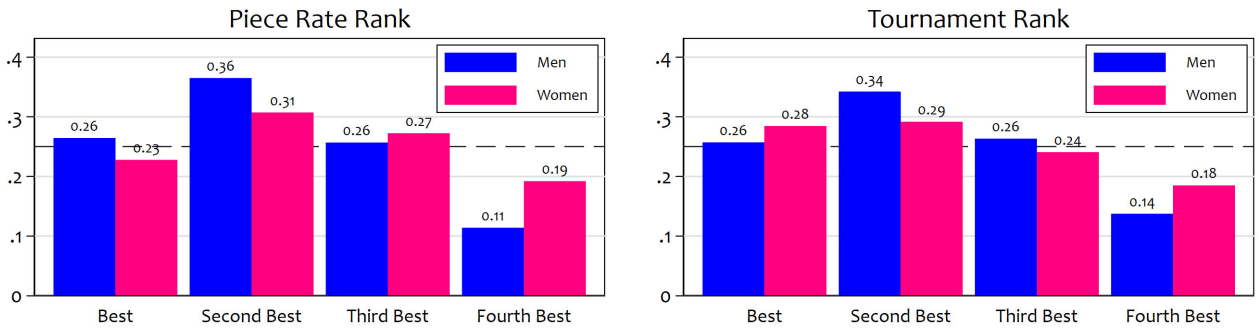
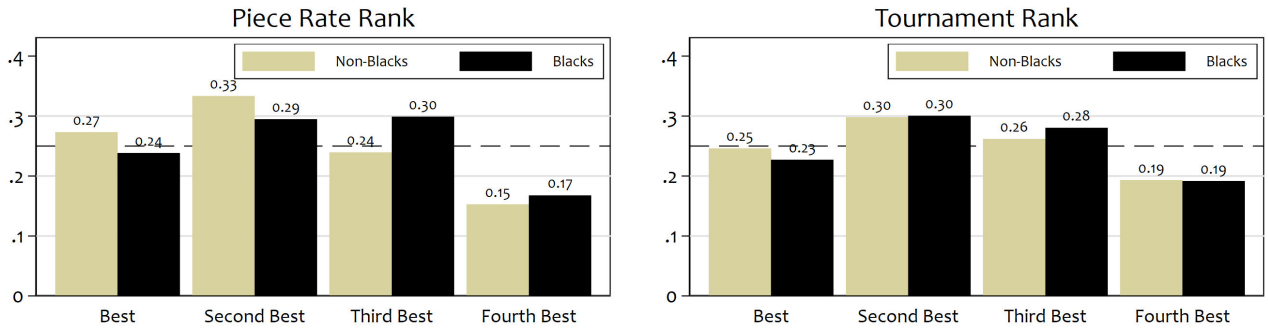


Figure C2: Recovered Beliefs on the Probability of Personal Performance Ranks by Blacks and Non-Blacks

A. One Token



B. Eighty Tokens

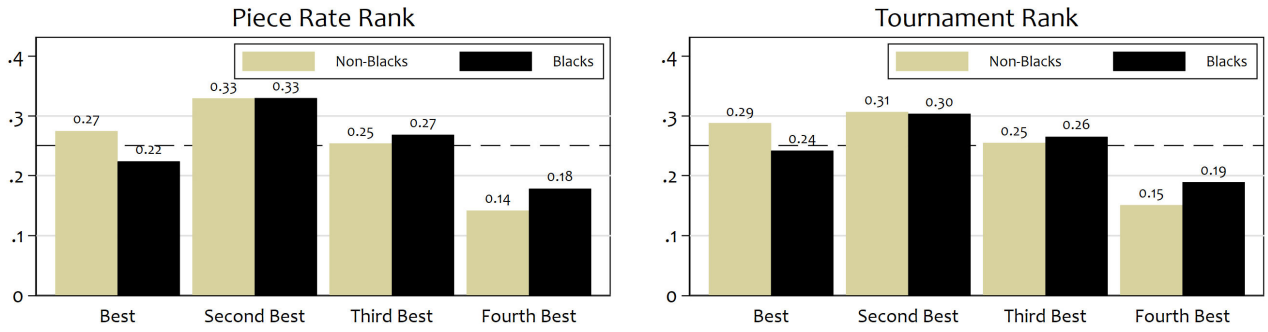


Figure C3: Beliefs of Subject #2 for the Piece Rate Rank Question

Posterior mean RDU risk preferences are r = 0.68, η = 1.22 and φ = 1.02

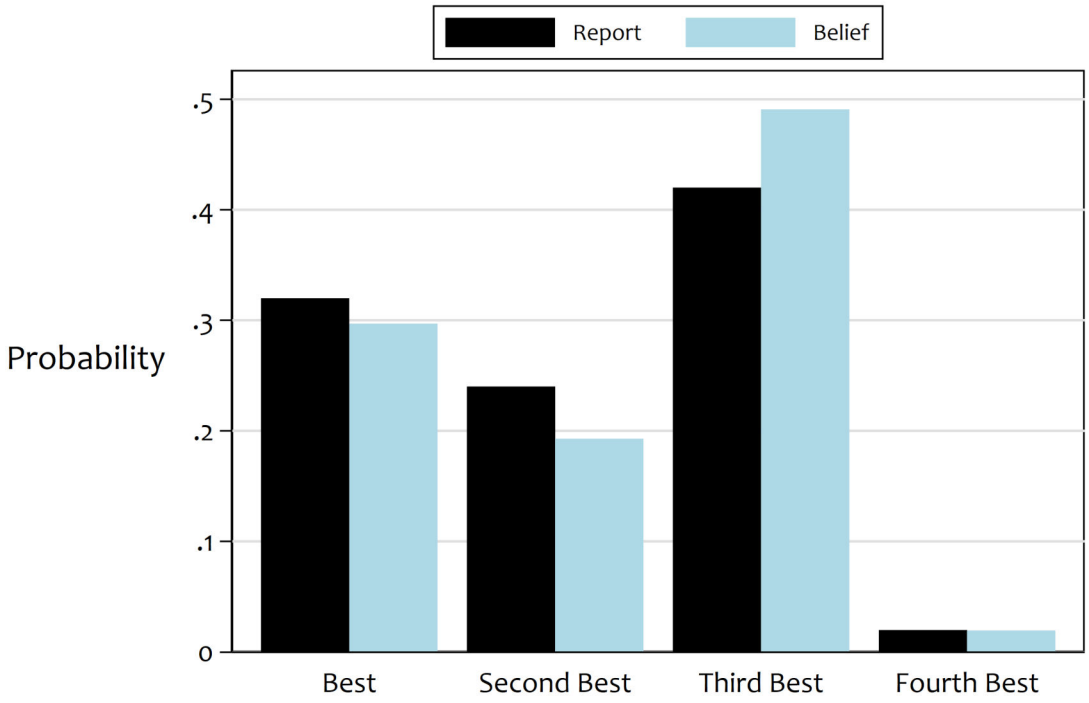


Figure C4: Beliefs of Subject #2 for the Tournament Rank Question

Posterior mean RDU risk preferences are r = 0.68, η = 1.22 and φ = 1.02

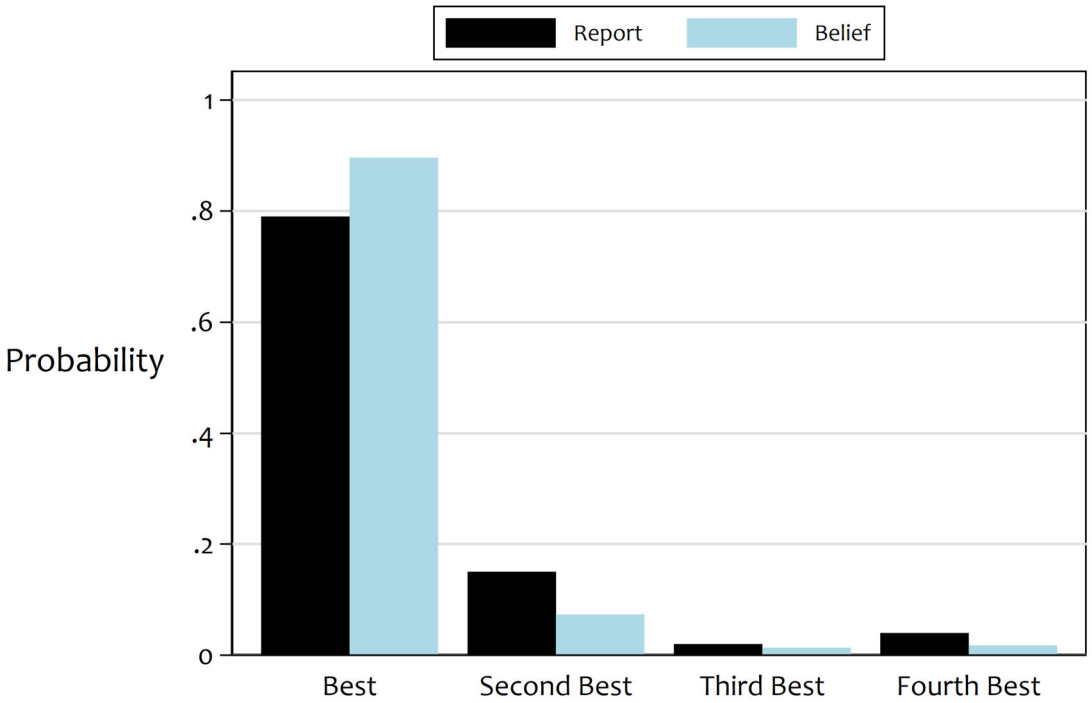


Figure C5: Beliefs of Subject #3 for the Piece Rate Rank Question

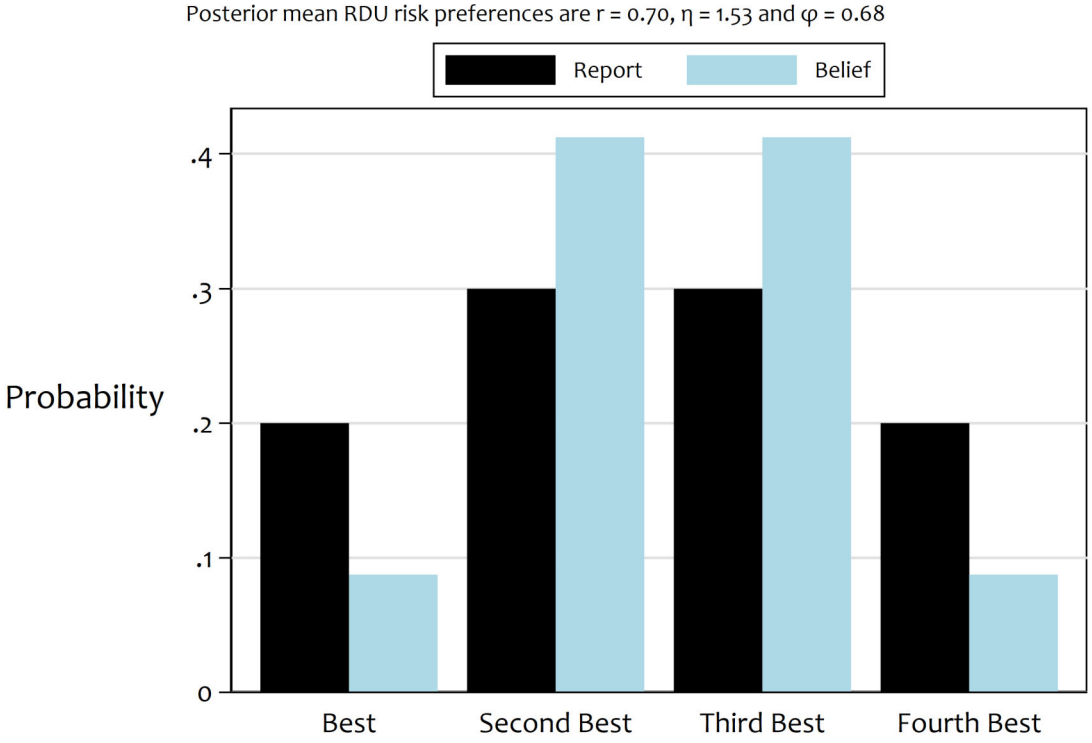
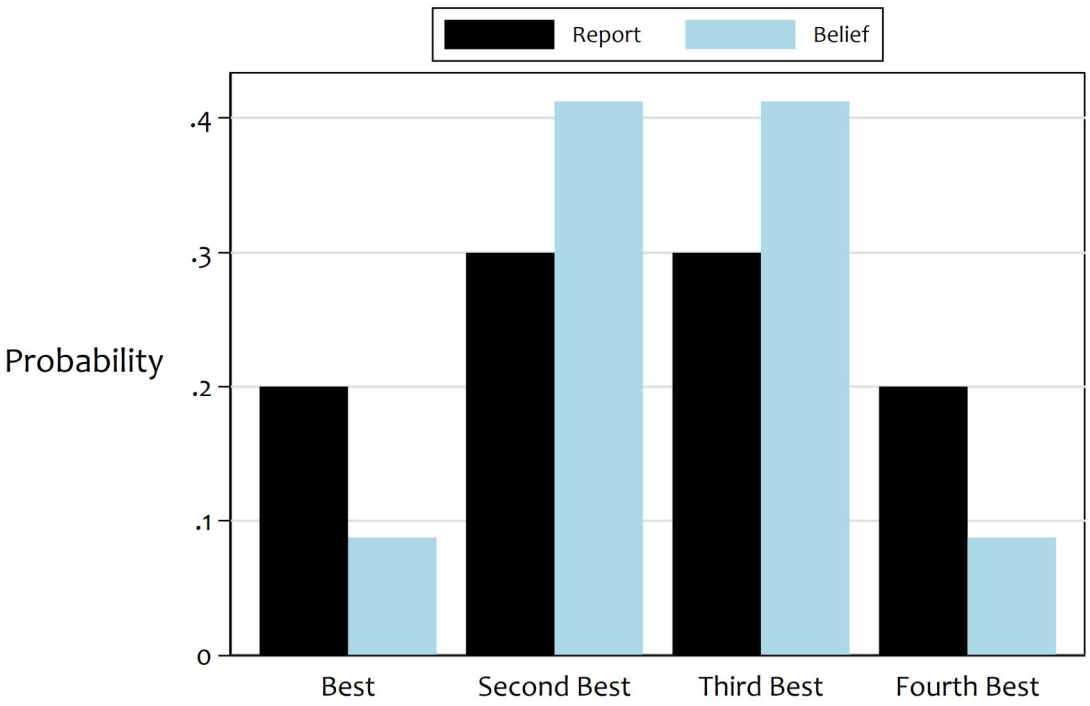


Figure C6: Beliefs of Subject #3 for the Tournament Rank Question



The listing below documents the conclusions drawn from the statistical analysis of qualitative welfare effects, and reported in the text. All beliefs reported are recovered from the observed reports, using the RDU risk preferences of the individual. Similarly, all welfare results reflect those recovered beliefs and the RDU risk preferences of the individual. Comparable results using the EUT risk preferences of the individual are also listed. The parameters used in Table 2 are also documented.

Variable name	Storage type	Display format	Value label	Variable label
welfare_gain	float	%9.0g		Actual welfare gain (=1) or loss (=0)
female	byte	%8.0g	flab	Female
age	byte	%10.0g		Age
black	byte	%8.0g		Black
business	byte	%8.0g		Business major
noreligion	byte	%8.0g		Atheist, Nonreligious, or Agnostic
bfi_extra	byte	%8.0g		BFI Extraversion
bfi_agree	byte	%8.0g		BFI Agreeableness
bfi_cons	byte	%8.0g		BFI Conscientiousness
bfi_neur	byte	%8.0g		BFI Neuroticism
bfi_open	byte	%8.0g		BFI Openness
optimistT	float	%9.0g		Subjective probability of winning greater than 0.25 with tournament performance
optimistP	float	%9.0g		Subjective probability of winning greater than 0.25 with piece-rate performance
earnT	float	%9.0g		Piece-Rate earnings in top 25%
earnP	float	%9.0g		

```
. * result 1: overall welfare gain RDU

. summ welfare_gain if rdu==1

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
welfare_gain |        178    .6910112      .46338          0          1

. * comparison to EUT
. summ welfare_gain if rdu==0

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
welfare_gain |        178    .7022472    .4585601          0          1

. * result 2: welfare gain RDU with stress or no stress

. summ welfare_gain if rdu==1 & submit==0

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
welfare_gain |         89    .6516854    .4791357          0          1

. summ welfare_gain if rdu==1 & submit==1

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
welfare_gain |         89    .7303371     .446299          0          1

. * comparison to EUT
. summ welfare_gain if rdu==0 & submit==0

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
welfare_gain |         89    .6853933      .46699          0          1

. summ welfare_gain if rdu==0 & submit==1

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
welfare_gain |         89    .7191011    .4519846          0          1

. * result 3: unconditional welfare gain RDU for women

. tab welfare_gain female if rdu==1, col nofreq exact

    Actual |
   welfare |
 gain (=1) |
   or loss |        Female
      (=0) |      Male     Female |     Total
-----------+----------------------+----------
         0 |     30.00      32.89 |     31.25
         1 |     70.00      67.11 |     68.75
-----------+----------------------+----------
     Total |    100.00     100.00 |    100.00

           Fisher's exact =                 0.744
   1-sided Fisher's exact =                 0.402

. * comparison to EUT
. tab welfare_gain female if rdu==0, col nofreq exact

    Actual |
   welfare |
 gain (=1) |
   or loss |        Female
      (=0) |      Male     Female |     Total
-----------+----------------------+----------
         0 |     30.00      30.26 |     30.11
         1 |     70.00      69.74 |     69.89
-----------+----------------------+----------
     Total |    100.00     100.00 |    100.00

           Fisher's exact =                 1.000
   1-sided Fisher's exact =                 0.550
```

```
. * result 4: conditioning on demographics, BFI and optimism and earnings

. probit welfare_gain `demog_i' `bfi' i.optimistT i.earnT i.optimistT#i.earnT
i.female#i.eighty_tokens if rdu==1, vce(cluster sid)


Probit regression                                       Number of obs =    176
                                                        Wald chi2(15) =  22.59
                                                        Prob > chi2   = 0.0932
Log pseudolikelihood = -98.784529                       Pseudo R2     = 0.0963

                                          (Std. err. adjusted for 88 clusters in sid)
------------------------------------------------------------------------------------
                   |               Robust
       welfare_gain | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------------+----------------------------------------------------------------
              female |
              Female  |   .3165435   .3648816     0.87   0.386    -.3986112    1.031698
                 age |  -.0154132   .0820441    -0.19   0.851    -.1762167    .1453904
               1.black |   .3335434   .2242917     1.49   0.137    -.1060602     .773147
            1.business |   .1071957   .2740318     0.39   0.696    -.4298968    .6442883
          1.noreligion |   .1896377   .2511163     0.76   0.450    -.3025412    .6818165
             bfi_extra |  -.1215833    .165013    -0.74   0.461    -.4450028    .2018361
             bfi_agree |   .2718446   .1945039     1.40   0.162    -.1093761    .6530652
              bfi_cons |  -.2898139   .1871696    -1.55   0.122    -.6566596    .0770318
              bfi_neur |  -.0455818   .1930262    -0.24   0.813    -.4239062    .3327426
              bfi_open |   .2556879   .2162158     1.18   0.237    -.1680872     .679463
           1.optimistT |   -.788307   .5783565    -1.36   0.173    -1.921865     .345251
               1.earnT |  -.9688624   .3484725    -2.78   0.005    -1.651856   -.2858689
                       |
     optimistT#earnT |
                 1 1  |   .7409653   .6366913     1.16   0.245    -.5069268    1.988857
                       |
female#eighty_tokens |
                Male#1  |   .4930705    .316505     1.56   0.119     -.127268    1.113409
              Female#1  |  -.1092502   .3331375    -0.33   0.743    -.7621877    .5436873
                       |
                 _cons |   .5339297   1.550506     0.34   0.731    -2.505005    3.572865
------------------------------------------------------------------------------------

Average marginal effects                                    Number of obs = 176
Model VCE: Robust

Expression: Pr(welfare_gain), predict()
dy/dx wrt:  1.female age 1.black 1.business 1.noreligion bfi_extra bfi_agree bfi_cons bfi_neur
bfi_open 1.optimistT 1.earnT 1.eighty_tokens

------------------------------------------------------------------------------
                   |            Delta-method
                   |      dy/dx   std. err.      z    P>|z|     [95% conf. interval]
-------------------+----------------------------------------------------------------
            female |
            Female  |  -.0005873   .0830183    -0.01   0.994    -.1633001    .1621255
               age |  -.0048969   .0260977    -0.19   0.851    -.0560474    .0462536
             1.black |   .1066111   .0719475     1.48   0.138    -.0344034    .2476257
          1.business |   .0337188   .0850712     0.40   0.692    -.1330178    .2004554
        1.noreligion |   .0590086   .0766894     0.77   0.442    -.0912998    .2093171
           bfi_extra |  -.0386283    .052335    -0.74   0.460     -.141203    .0639465
           bfi_agree |   .0863678    .060927     1.42   0.156    -.0330468    .2057824
            bfi_cons |  -.0920768   .0575288    -1.60   0.109    -.2048312    .0206776
            bfi_neur |  -.0144818   .0612178    -0.24   0.813    -.1344664    .1055028
            bfi_open |   .0812346   .0685575     1.18   0.236    -.0531355    .2156048
         1.optimistT |  -.0694125   .0895648    -0.77   0.438    -.2449563    .1061314
             1.earnT |  -.2058842   .0736567    -2.80   0.005    -.3502487   -.0615197
1.eighty_tokens |   .0713186    .073475     0.97   0.332    -.0726897    .2153269
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.
```

```
. * comparison to EUT

. probit welfare_gain `demog_i' `bfi' i.optimistT i.earnT i.optimistT#i.earnT
i.female#i.eighty_tokens if rdu==0, vce(cluster sid)

Probit regression                                     Number of obs =    176
                                                      Wald chi2(15) =  23.28
                                                      Prob > chi2   = 0.0783
Log pseudolikelihood = -97.239548                     Pseudo R2     = 0.0970

                                          (Std. err. adjusted for 88 clusters in sid)
------------------------------------------------------------------------------------
                   |               Robust
       welfare_gain | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------------+----------------------------------------------------------------
              female |
              Female  |   .2437876   .3744595     0.65   0.515    -.4901395    .9777148
                 age |  -.0337342   .0800274    -0.42   0.673    -.1905851    .1231166
            1.black |   .2994134   .2185508     1.37   0.171    -.1289383    .7277651
         1.business |   .3843431    .268255     1.43   0.152    -.1414271    .9101133
       1.noreligion |   .1969226   .2326391     0.85   0.397    -.2590418    .6528869
          bfi_extra |  -.0036687   .1547591    -0.02   0.981    -.3069909    .2996535
          bfi_agree |   .3142669   .1899836     1.65   0.098    -.0580941    .6866279
           bfi_cons |  -.3144162   .1761357    -1.79   0.074    -.6596359    .0308034
           bfi_neur |  -.0313468   .1989629    -0.16   0.875    -.4213069    .3586133
           bfi_open |   .3098938   .1987146     1.56   0.119    -.0795796    .6993672
         1.optimistT |  -.6743993   .5726012    -1.18   0.239    -1.796677    .4478783
            1.earnT |  -.9263598   .3615647    -2.56   0.010    -1.635014   -.2177061
                   |
     optimistT#earnT |
               1 1  |   .6714149   .6306384     1.06   0.287    -.5646138    1.907443
                   |
female#eighty_tokens |
             Male#1  |   .2731307   .3069343     0.89   0.374    -.3284495    .8747109
           Female#1  |  -.0178173   .3267929    -0.05   0.957    -.6583195    .6226849
                   |
               _cons |   -.004545   1.534636    -0.00   0.998    -3.012376    3.003286
------------------------------------------------------------------------------------


Average marginal effects                                  Number of obs = 176
Model VCE: Robust

Expression: Pr(welfare_gain), predict()
dy/dx wrt:  1.female age 1.black 1.business 1.noreligion bfi_extra bfi_agree bfi_cons bfi_neur
bfi_open 1.optimistT 1.earnT 1.eighty_tokens

------------------------------------------------------------------------------
                |            Delta-method
                |      dy/dx   std. err.      z    P>|z|     [95% conf. interval]
----------------+-------------------------------------------------------------
         female |
         Female  |   .0281981   .0825764     0.34   0.733    -.1336487     .190045
            age |   -.010552   .0250349    -0.42   0.673    -.0596195    .0385154
          1.black |   .0939512   .0692787     1.36   0.175    -.0418327     .229735
       1.business |   .1153841    .075268     1.53   0.125    -.0321386    .2629067
     1.noreligion |   .0602455   .0701918     0.86   0.391    -.0773278    .1978188
        bfi_extra |  -.0011476   .0484119    -0.02   0.981    -.0960332     .093738
        bfi_agree |   .0983022    .058512     1.68   0.093    -.0163792    .2129836
         bfi_cons |  -.0983489   .0522622    -1.88   0.060    -.2007809    .0040831
         bfi_neur |  -.0098052   .0621397    -0.16   0.875    -.1315968    .1119864
         bfi_open |   .0969343   .0617667     1.57   0.117    -.0241262    .2179948
       1.optimistT |   -.048208   .0841208    -0.57   0.567    -.2130817    .1166657
          1.earnT |  -.1973012   .0714649    -2.76   0.006    -.3373699   -.0572325
1.eighty_tokens |   .0460322   .0704389     0.65   0.513    -.0920255      .18409
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.
```

```
. * result(s) 5: no-stress in part 4, conditioning on demographics, BFI, optimism and earnings

. probit welfare_gain `demog_i' `bfi' i.optimistP i.earnP i.earnP i.optimistP#i.earnP
i.female#i.eighty_tokens if rdu==1 & submit==1, vce(cluster sid)

Probit regression                                       Number of obs =     88
                                                        Wald chi2(15) =  21.16
                                                        Prob > chi2   = 0.1317
Log pseudolikelihood = -42.372448                       Pseudo R2     = 0.1783

                                          (Std. err. adjusted for 88 clusters in sid)
------------------------------------------------------------------------------------
                   |               Robust
       welfare_gain | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------------+----------------------------------------------------------------
              female |
              Female  |   .6480445   .5032921     1.29   0.198    -.3383899    1.634479
                 age |  -.0164485   .1112026    -0.15   0.882    -.2344016    .2015046
               1.black |     .65977    .334209     1.97   0.048     .0047325    1.314808
            1.business |  -.0054182   .3547768    -0.02   0.988     -.700768    .6899317
          1.noreligion |   .5764353   .3700241     1.56   0.119    -.1487986    1.301669
             bfi_extra |  -.0403915   .2041438    -0.20   0.843     -.440506    .3597229
             bfi_agree |   .1974102   .2838089     0.70   0.487    -.3588449    .7536654
              bfi_cons |  -.1107276   .2631052    -0.42   0.674    -.6264044    .4049491
              bfi_neur |   .0655545    .262408     0.25   0.803    -.4487557    .5798648
              bfi_open |  -.1677185   .3043648    -0.55   0.582    -.7642626    .4288255
           1.optimistP |  -1.013151   .7683918    -1.32   0.187    -2.519171    .4928694
               1.earnP |  -1.435437   .5071451    -2.83   0.005    -2.429423   -.4414504
                       |
     optimistP#earnP |
                 1 1   |   .4728323   .8618582     0.55   0.583    -1.216379    2.162043
                       |
female#eighty_tokens |
               Male#1  |   .3209009   .4130765     0.78   0.437    -.4887141    1.130516
             Female#1  |  -.3235379   .5332501    -0.61   0.544    -1.368689     .721613
                       |
                 _cons |   1.469008   2.033738     0.72   0.470    -2.517045    5.455061
------------------------------------------------------------------------------------


Average marginal effects                                Number of obs = 88
Model VCE: Robust

Expression: Pr(welfare_gain), predict()
dy/dx wrt:  1.female age 1.black 1.business 1.noreligion bfi_extra bfi_agree bfi_cons bfi_neur
bfi_open 1.optimistP 1.earnP 1.eighty_tokens

------------------------------------------------------------------------------
                   |            Delta-method
                   |      dy/dx   std. err.      z    P>|z|     [95% conf. interval]
-------------------+----------------------------------------------------------------
            female |
            Female  |   .0735548   .0881184     0.83   0.404     -.099154    .2462636
               age |  -.0044415   .0300776    -0.15   0.883    -.0633926    .0545096
             1.black |    .178667   .0862419     2.07   0.038      .009636    .3476981
          1.business |  -.0014641   .0959314    -0.02   0.988    -.1894862     .186558
        1.noreligion |    .145903   .0867272     1.68   0.093    -.0240791    .3158852
           bfi_extra |  -.0109068   .0549687    -0.20   0.843    -.1186434    .0968298
           bfi_agree |    .053306   .0758586     0.70   0.482    -.0953742    .2019862
            bfi_cons |  -.0298994   .0708303    -0.42   0.673    -.1687243    .1089255
            bfi_neur |   .0177015   .0706158     0.25   0.802     -.120703    .1561059
            bfi_open |  -.0452884   .0814813    -0.56   0.578    -.2049888    .1144119
         1.optimistP |  -.1792824   .0990404    -1.81   0.070    -.3733981    .0148332
             1.earnP |  -.2868691   .0808513    -3.55   0.000    -.4453348   -.1284035
1.eighty_tokens |    .012591    .088551     0.14   0.887    -.1609657    .1861477
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.
```

```
. * comparison to EUT

. probit welfare_gain `demog_i' `bfi' i.optimistP i.earnP i.earnP i.optimistP#i.earnP
i.female#i.eighty_tokens if rdu==0 & submit==1, vce(cluster sid)

Probit regression                                       Number of obs =     88
                                                        Wald chi2(15) =  23.31
                                                        Prob > chi2   = 0.0778
Log pseudolikelihood = -42.367424                       Pseudo R2     = 0.1933

                                          (Std. err. adjusted for 88 clusters in sid)
------------------------------------------------------------------------------------
                   |               Robust
       welfare_gain | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------------+----------------------------------------------------------------
              female |
              Female  |   .6705979   .5071849     1.32   0.186    -.3234662    1.664662
                 age |  -.0326749   .1073441    -0.30   0.761    -.2430656    .1777157
               1.black |   .4054507   .3453843     1.17   0.240    -.2714902    1.082392
            1.business |   .3919002   .3702614     1.06   0.290    -.3337989    1.117599
          1.noreligion |   .6433324   .3776918     1.70   0.089      -.09693    1.383595
             bfi_extra |   .0172992   .2023596     0.09   0.932    -.3793182    .4139167
             bfi_agree |   .1985773   .2844121     0.70   0.485    -.3588601    .7560147
              bfi_cons |  -.2067891   .2751567    -0.75   0.452    -.7460864    .3325083
              bfi_neur |   .0944357   .2749979     0.34   0.731    -.4445502    .6334217
              bfi_open |   .1360013   .3325084     0.41   0.683    -.5157031    .7877058
           1.optimistP |   -.894054   .7503942    -1.19   0.233      -2.3648    .5766915
               1.earnP |  -1.505667    .513225    -2.93   0.003     -2.51157   -.4997648
                       |
     optimistP#earnP |
                 1 1  |   .2464838   .8354424     0.30   0.768    -1.390953    1.883921
                       |
female#eighty_tokens |
               Male#1  |   .0046846   .4043314     0.01   0.991    -.7877905    .7971597
            Female#1  |  -.1707547   .5486411    -0.31   0.756    -1.246071    .9045621
                       |
                 _cons |   .7091177   2.148426     0.33   0.741    -3.501719    4.919955
------------------------------------------------------------------------------------

Average marginal effects                                Number of obs = 88
Model VCE: Robust

Expression: Pr(welfare_gain), predict()
dy/dx wrt:  1.female age 1.black 1.business 1.noreligion bfi_extra bfi_agree bfi_cons bfi_neur
bfi_open 1.optimistP 1.earnP 1.eighty_tokens

------------------------------------------------------------------------------
                 |            Delta-method
                 |      dy/dx   std. err.      z    P>|z|     [95% conf. interval]
-----------------+------------------------------------------------------------
           female |
           Female  |   .1484242   .0877143     1.69   0.091    -.0234928    .3203412
              age |  -.0088015    .028891    -0.30   0.761    -.0654268    .0478238
            1.black |   .1087823   .0903986     1.20   0.229    -.0683957    .2859602
         1.business |   .1004807   .0891177     1.13   0.260    -.0741869    .2751482
       1.noreligion |   .1622946   .0872815     1.86   0.063     -.008774    .3333633
          bfi_extra |   .0046598   .0545352     0.09   0.932    -.1022272    .1115468
          bfi_agree |   .0534898   .0760134     0.70   0.482    -.0954938    .2024734
           bfi_cons |  -.0557018    .073797    -0.75   0.450    -.2003413    .0889377
           bfi_neur |   .0254377   .0737507     0.34   0.730    -.1191111    .1699865
           bfi_open |    .036634   .0892619     0.41   0.682     -.138316    .2115841
        1.optimistP |  -.1975527   .0980081    -2.02   0.044     -.389645   -.0054604
            1.earnP |  -.3203305   .0780357    -4.10   0.000    -.4732778   -.1673833
1.eighty_tokens |  -.0170335   .0921194    -0.18   0.853    -.1975842    .1635172
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.
```

```
. * Table 2 risk parameters

. bysort female: summ r_eut r_rdu eta phi if record==1 & female ~= .

-> female = 0

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
       r_eut |         50    .5960055    .1778426   .2659489   .8953205
       r_rdu |         50    .7075162    .0379635   .6257992   .8038139
         eta |         50    .9709536    .5589444   .3421361   2.794703
         phi |         50    .9975626    .4088162   .3699581   2.381267

-> female = 1

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
       r_eut |         38     .562654    .1637901   .1287987   .8315247
       r_rdu |         38    .6988082    .0329121    .626736   .7685972
         eta |         38    .9009178    .4462629   .2821337   1.803559
         phi |         38    .9193542    .3137058   .4147758   1.611156


. * Table 2 payouts and probabilities

For the 80 token case, the reported probability of winning the Tournament under Stress for men
       was 0.26, and was 0.28 for women.
For the 80 token case, the reported probability of winning the Tournament with no Stress for men
       was 0.26, and was 0.23 for women.
For the 80 token case, the Piece Rate payout for men was $1.37, and was $1.36 for women.
For the 80 token case, the Tournament payout for men, conditional on winning, was $5.66, and was
       $5.72 for women.


. * Used for the Table 2 welfare calculations

For the 1 token case, the reported probability of winning the Tournament under Stress for men
       was 0.24, and was 0.23 for women.
For the 1 token case, the reported probability of winning the Tournament with no Stress for men
       was 0.24, and was 0.27 for women.
For the 1 token case, the Piece Rate payout for men was $1.04, and was $1.17 for women.
For the 1 token case, the Tournament payout for men, conditional on winning, was $3.87, and was
       $4.23 for women.
```

4. Evaluating the Costs of Competition Interventions

As noted in the text, Niederle, Segal and Vesterlund [2013] and Balafoutas and Sutter [2012] consider a variety of "affirmative action" policies to reduce the gender gap in choosing to compete, but offer no evidence that it improves anyone's welfare. Alan and Ertac [2019] consider interventions to teach young children about the "value of grit" and determination, and that seems to reduce a gender gap in choosing to compete, but it is far from obvious that it improves welfare for risk averse agents. Here we review the methods used in each of these studies to evaluate the cost of the behavior they observe.

We start in reverse order, since **Alan and Ertac [2019]** provides the cleanest example. They evaluate Efficiency, as they refer to it in the title of §5.4, as follows:

> Based on performance, for some children it is payoff-maximizing to compete, whereas for others it is better to stay out. One can be concerned that interventions such as the one we evaluate in this paper may lead to unintended inferior outcomes for some children, by inducing decisions that turn out to be bad for payoffs ex post. Analyzing how children's decisions fare in terms of expected material payoffs can

> shed light on these issues. To do this, we first simulate the probabilities of winning and tying in competition for any given performance level, using the empirical distribution of performances. For the first stage, we use the actual performance of the whole sample when calculating the empirical win probabilities. That is, we use simulations to form random matches from the whole empirical distribution in the first choice stage, and compute the empirical win, loss and tie probabilities that correspond to each realized performance level. [p.1168]

The need to simulate probabilities is simply due to their design not including belief elicitation stages, understandable in the context of a vast field experiments with children at school. In their second stage subjects were told their own performance, so the efficiency analysis was modified by calcualtiung empirical win probabilities compared to others that had strictly different scores.

The next step was to infer the cost of selecting or not selecting the profit-maximizing choice:

> Using these probabilities along with realized performances, we calculate each child's expected payoff from competition, and analyze whether the child's actual choice was payoff maximizing ex post. We then estimate the treatment effects on this outcome to see whether the treatment leads to suboptimal choices from a payoff maximization perspective. It should be noted that this analysis does not make utility comparisons and therefore it is not an optimality analysis per se, as it disregards effort costs, which are unobservable. [p.1169]

The final sentence is problematic. It is correct in stating that the analysis does not make utility comparisons, but implies that the sole purpose of using utility functions is to evaluate the difference between expected benefits of competing and expected costs of competing, where costs are solely due to cognitive effort levels needed. But it ignores the role of the utility function (and probability weighting function, if using RDU) in evaluating the risk premium entailed in making the choice to compete in the tournament. As Table D1 and Table 2 illustrate, one can immediately see the importance of allowing for a risk premium even if there is some reasonable guess at levels of risk aversion. Ideally one would like estimates of actual risk aversion, as we illustrate, but the very design of the task, through the relative payoffs of the piece rate compensation scheme and the tournament compensation scheme, assumes risk neutrality as a benchmark, alerting one to the need to consider aversion to risk.

One attractive feature of the evaluation of Alan and Ertac [2019; Figure 4, p.1171] is attention to the full distributional effects on all girls and all boys. The vertical axes show treatment effects in monetary units, and the horizontal axes display quantiles (presumably from the baseline):

One cannot just evaluate welfare, even if assuming risk neutrality, but just seeing if the "best performing" girls did better. The left panel shows that all girls appear to have done statistically significantly better from the treatment in terms of expected payoffs.

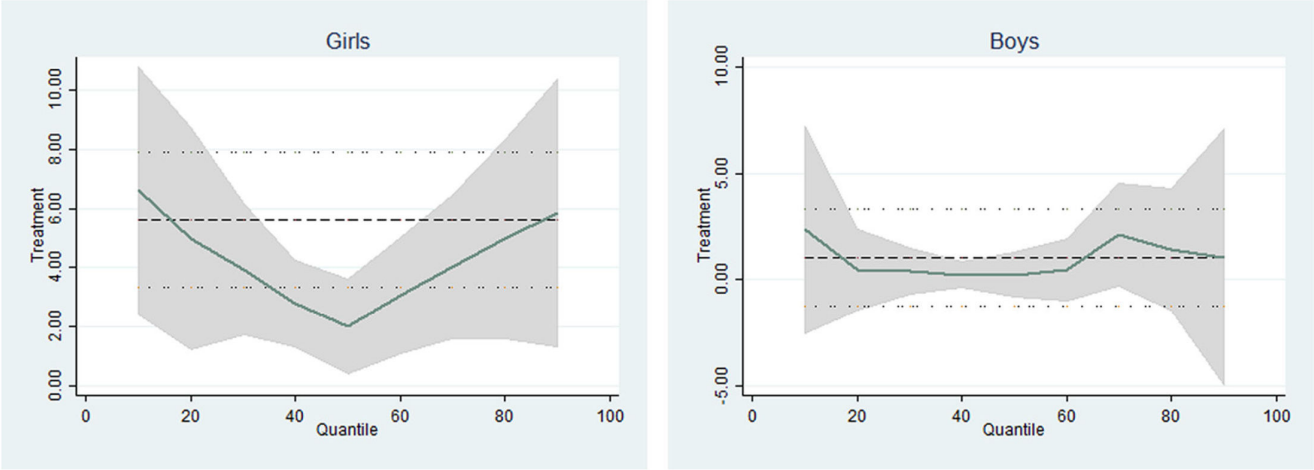


FIGURE 4. Distribution of treatment effects on expected payoff gains. Figures plot the quantile treatment effects on expected payoff gains and 95% confidence bands. The dashed lines indicate the estimated average treatment effects with the dotted lines showing the confidence bands for the average treatment effect estimates.

However, their results for boys need to be evaluated carefully. They make the following claim (p.1171/2) from these displays: "At first glance, the figures show that treatment effects are generally non-negative across quantiles, for both genders, reassuring us that no one was hurt by the treatment in terms of payoffs." This is not what one should infer from the right panel for boys. Note the 0-gain line for boys on the vertical axis, just below the dashed line showing the average treatment effect. The fact that the 95% confidence band falls *below* the 0-gain line *for all quantiles* says, literally, that some farction of the sample suffered losses in terms of the expected payoffs effect of the treatment. The confusion arises from an intended inference about the average effect rather than an inference about the sample as a whole. That is, one might infer that there is no evidence of a significant loss at any quantile *for the average boy at each quantile*, but that is not what is being claimed here. With very large samples of 1281 boys across all stages, these confidence intervals paint a clear picture of some losses in for boys in terms of expected payoffs.

We completely set aside a utilitarian social welfare function which would add up expected gains and losses and declare a welfare improvement if the sum (and hence average) is greater than zero. At least initially, the welfare analysis should, as suggested, look at the whole distribution.

The focus of **Balafoutas and Sutter [2012]** is immediately on efficiency, from the title itself, "Affirmative Action Policies Promote Women *and Do Not Harm Efficiency* in the Laboratory" (emphasis added). Their design is a replication of Niederle, Segal and Vesterlund [2013], with extensions to consider alternative versions of Affirmative Action policies. They motivate their concern with effficiency as follows:

> Policy interventions to support the promotion of women often face the criticism that they are inefficient in [not] assigning the best available candidate, irrespective of gender, to a particular job when several candidates compete for it [...]. While this is

> difficult to measure in the field because it is hard to exactly identify a candidate's qualifications, laboratory-based economic experiments allow for an unambiguous assessment of the efficiency of affirmative action programs in promoting the best candidates, although this is measured in the artificial context of a well-controlled, quantitative task. [p.579].

To make the meaning explicit, we add "[not]" to the opening sentance of this text. Furthermore,

> Interventions that promote the entry of women may have two opposing effects on the overall efficiency in selecting te best candidates as winners. On the one hand, any intervention that gives an advantage to women may yield efficiency losses by passing over better-performing men for the same of promoting women. On the other hand, interventions may induce more high-performing women to choose competition instead of going for the piece rate, leading to efficiency gains. [p.580]

This is defining efficiency solely in terms of production efficiency, measuring whether output is being produced at least cost by the appropriate mix of factors of production. This is one necessary condition for the economy to be efficient, but not sufficient. In particular, it ignores consumption efficiency, measuring whether a given bundle of goods ends up with the individuals that value it the most.

It is easy to see how the outcomes evaluated here could violate consumption efficiency. Imagine that the most productive women were the most risk averse, and place a relatively large premium on avoiding the risk associated with competing in the tournament. *Ceteris paribus*, it could easily be efficient for less productive women, or even less productive men, to take on the tournament competion. Or, more to the point from Table D1 and Table 2, it could easily be efficient *for none of the production to occur with tournament compensation*.

The details of the calculations of production efficiency gains for men and women are not relevant, since they fail to consider consumption efficiency.

Finally, **Niederle, Segal and Vesterlund [2013]** developed the variant of the Niederle and Vesterlund [2007] design to study one implementation of affirmative action policies to encourage women to choose to compete. Again, their title immediately focusses on issues of cost: "How Costly Is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness". They ask

> ... whether affirmative action can encourage more high-performing women to enter competitions. We focus not only on determining how the policy changes the decision to compete, but also on how it changes the gender composition of the pool of competitors. Accounting for changes in entry, we ask how costly it is to secure that women be equally represented among those who win competitions. How much lower will the performance be for winners under the policy? How many better performing men will have to be passed by to secure equal representation of those hired? [p.2]

So the focus is again solely on production efficiency. With some simplification, they find that better-performing women tend to displace poorer-performing women in choosing to compete in

the tournament compensation scheme. Hence there tends to be less displacement of better performing men than one might have expected, *a priori*.

An important subtlety is addressed when evaluating "How Costly is Affirmative Action", the title of §4:

> A concern when introducing affirmative action is how costly it will be to achieve a more diverse set of winners. Looking directly at the performance of those who enter and win the tournaments, we do not see deterioration in performance. [...] This difference is not significant [...] and suggests that, in contrast to expectations affirmative action need not decrease performance. Although it is tempting to focus solely on the performance of actual winners, it is important to note that whether an individual is or is not identified as a winner depends on the performance of the group she is randomly assigned to. Thus, examination of actual winners only provides limited insight on the effect of affirmative action. [p.7]

The general way in which they propose to evaluate this point is elegant:

> To assess the cost of the policy we view all the participants in the experiment as (potential) candidates and those who enter competitions as applicants for jobs. We ask what the minimum performance requirement would be if a firm wanted to hire a certain number of applicants and wanted to secure that only the best available applicants were hired. We then ask how much lower the requirement has to be if we want to hire the same number of applicants under the equal representation rule. To evaluate the degree of reverse discrimination, we determine how many strictly better performing men will be passed by to secure that women are at least equally represented among those hired. Passing by better performing applicants is inequitable and costly for the firm, as it no longer can hire the best available applicants. Crucial for assessing these two adverse effects is the performance and gender composition of those who decide to enter the competition. [p.7]

The same general point has been referred to in a related setting of gender discrimination, wage determination in a university, as "efficient equity," the idea that one can search over alternative ways to reach some given equity goal and select the one that does so at least cost: see Coller, Harrison and Rutherford [1998]. It is also used in trade policy evaluation, to facilitate liberalization while ensuring that certain equity goals are met at least cost: see Harrison, Rutherford and Tarr [2003].

Again, as with Balafoutas and Sutter [2012], the details of the calculations of production efficiency gains for men and women are not relevant, since they fail to address consumption efficiency.

Additional References

Coller, Maribeth; Harrison, Glenn W., and Rutherford, Thomas F., "Efficient Equity: Removing Salary Discrimination By Meeting Statistical Legal Constraints at Least Cost," *Economics Letters*, 52, 1996, 81-88.

Harrison, Glenn W.; Rutherford, Thomas F., and Tarr, David G., "Trade Liberalization, Poverty and Efficient Equity," *Journal of Development Economics*, 71, June 2003, 97-128.

Appendix D: Pilot Results with Variants on the Raven Task (Online Working Paper)

1. Time Constraints

One popular variant of the Raven task involves the *imposition of time constraints* on completing the full task. The most common time constraint is 40 minutes, and we modify the Progressive Eighty Tokens condition to examine the effect of that constraint.[1] Of course, our default experiment formally had a time constraint of 90 minutes, but that was not particularly binding, as illustrated by the solid black line in Figure D1. The average time required was 48 minutes, and most subjects were completed after about an hour. A constraint of 40 minutes happens to be at the 40th percentile of this distribution, so is expected *a priori* to be binding for some individuals.[2]

As mentioned earlier, intelligence measures with binding time constraints are generally viewed as measuring intellectual efficiency rather than some "pure intelligence."[3] Since we view it as natural to think of cognition and intelligence as contextual, and time is clearly something with an opportunity cost, the manner in which individuals apply their cognitive production functions under varying constraints is of great practical interest to economists.[4] This pilot experiment also allows us to see the pure effect of a time constraint, which plays a key role in the "gender and competitiveness" designs we examine in §3.

When we impose a 40 minute time constraint in the time-constrained condition we find no no evidence of a change in average Accuracy or average Efficiency. The dashed line in Figure D1 shows the re-allocation of time spent on the task, with a sharp peak just immediately prior to the time constraint. Figure D2 displays the effect of the time constraint in terms of Accuracy and Efficiency. The average marginal effect of the time constraint on Accuracy is -4.0 percentage points, but not statistically significant (p-value = 0.20), and the effect on Efficiency is -0.7 percentage points and also not statistically significant (p-value = 0.78). The effect of demographics on the time constraint and Efficiency are shown in Figure D3. The time constraint does significantly lower Accuracy and Efficiency for women, by 8.4 and 4.1 percentage points (p-values

---

[1] Raven, Raven and Court [1998] present norms for 30-minute and 40-minute timed versions of the RAPM. Hamel and Schmittmann [2006] evaluated a 20-minute timed version of the test with 397 pseudo-volunteers (viz., psychology students taking the test as a course requirement) against an untimed version with 59 subjects drawn from the same population. They report a correlation of 0.75 between the scores of the untimed and timed versions.

[2] A 20-minute constraint is at the 3rd percentile of this distribution, and a 30-minute constraint at the 20th percentile.

[3] For example, by Raven, Raven and Court [1993].

[4] Borghans, Meijers and Ter Weel [2008] find a striking interaction between the effect of financial incentives and time constraints in an experiment in which university students were presented with a battery of cognitive tests. Extra financial incentives led to substantially more *time* being spent on answering those questions. But whether or not that extra time effort generated better scores depended on time constraints, amongst other things. When subjects were time constrained, there was less scope for the extra effort to be rewarded with better performance, and hence less scope for financial incentives to show an effect on performance. In addition, the link between greater effort and greater performance was mediated by non-cognitive factors, such as personality traits, and the specific set of cognitive questions considered. Only two of the ten cognitive questions were "Raven-like," and it is inappropriate to view this set of 10 as a test of the effect of incentives and time constraints on fluid intelligence (let alone IQ, as they claim). But the general results are, nonetheless, important.

of 0.013 and 0.084) respectively. So even if the average effect of this time constraint are not statistically significant at conventional levels, there are significant distributional effects correlated with one important demographic for policy.

2. Progressively Increasing Incentives

Apart from imposing a binding time constraint to performance, another question for economists to ask is whether Accuracy and/or Efficiency would improve if *financial incentives were increased for the harder questions*.[5] We therefore increased the earnings for allocating all 80 tokens to the correct solution from the default $2 for all questions, to be $2 for questions 1 through 12, $3 for questions 13 through 24, and $5 for questions 25 through $36. This increased potential aggregate earnings from $72 up to $120.

We observe a marked *improvement* in Accuracy *and* Efficiency when financial incentives within the RAPM are progressively increased. Figure D4 displays the effect of the increasing financial incentives in terms of Accuracy and Efficiency. The average marginal effect of increasing incentives on average Accuracy is +11.8 percentage points (p-value = 0.002), and the effect on average Efficiency is +7.3 percentage points (p-value = 0.013). Figure D5 shows effects of the increasing financial incentives for certain demographics, There are significant improvements in Efficiency for Blacks (+6.2 pp, p-value = 0.03) and Business majors (+16.9 pp, p-value = 0.01), accompanied by comparably significant improvements in Accuracy. For women there are clear improvements in Accuracy of +7.5 pp (p-value = 0.04), and *some* evidence of improvements in Efficiency of +4.7 pp (p-value = 0.11).

Additional References

Hamel, Ronald, and Schmittmann, Verena, D., "The 20-Minute Version as a Predictor of the Raven Advanced Progressive Matrices Test," *Educational and Psychological Measurement*, 66(6), December 2006, 1039-1046.

[5] McDaniel and Rutström [2001] considered a similar *between-subjects* treatment in an incentivized task in which subjects attempted to solve the Tower of Hanoi puzzle in the least number of moves. The gave 25 subjects "low" rewards for a better score and 30 subjects "high" rewards for a better score, and found (p.153) no statistically significant difference in performance. Consistent with some of the literature on RAPM, such as Gignac [2018], they did report an increase in "effort." Our treatment is *within-subject*.

Figure D1: Distribution of Time Taken

Time taken to complete the Progressive 80 Tokens task
with a 90-minute constraint or a 40-minute constraint

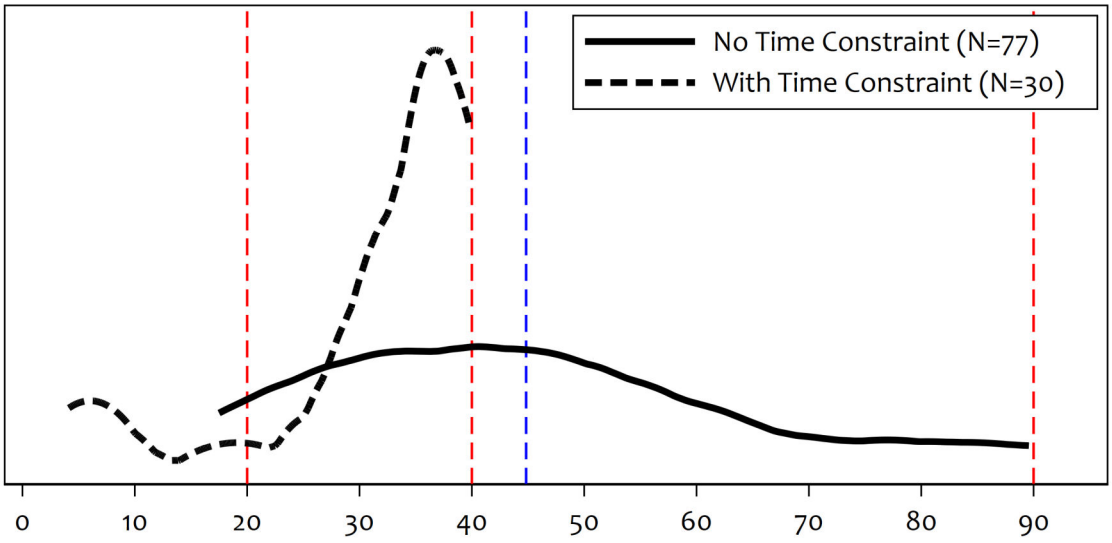


Figure D2: Effect of 40-Minute Time Constraint on Accuracy and Efficiency

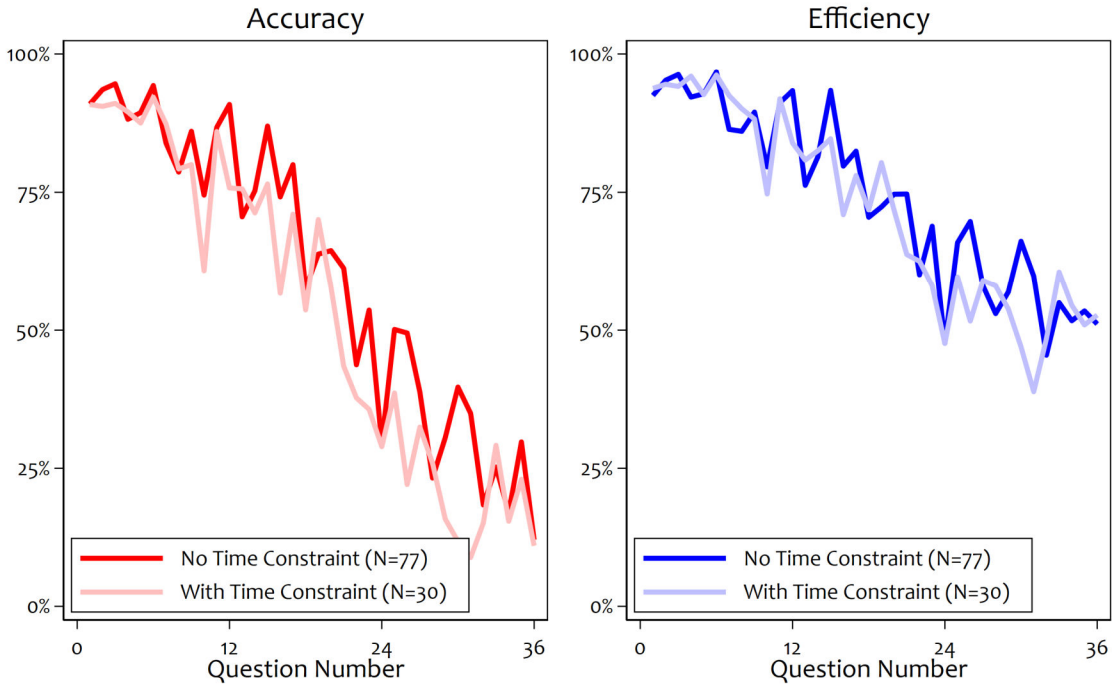


Figure D3: Effects of Time Constraint on Efficiency

Average marginal effects of 40-minute time constraint from Fractional Regression
Solely comparing results within the Progressive Treatment
subjects with a time-constraint, and 77 unconstrained

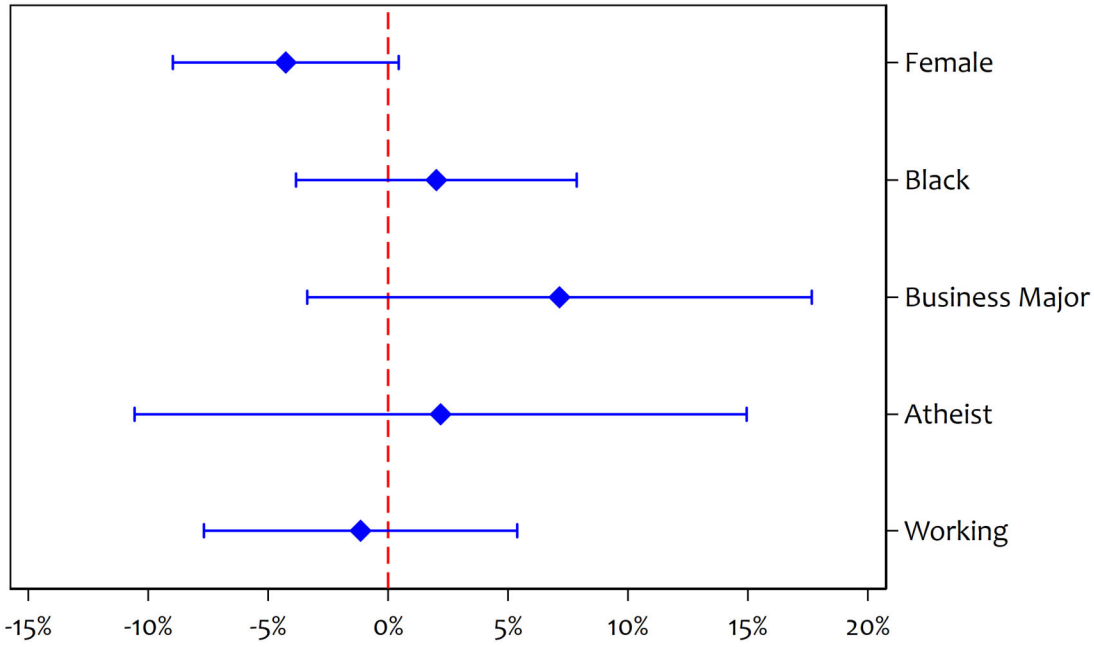


Figure D4: Effect of Increasing Financial Incentives on Accuracy and Efficiency

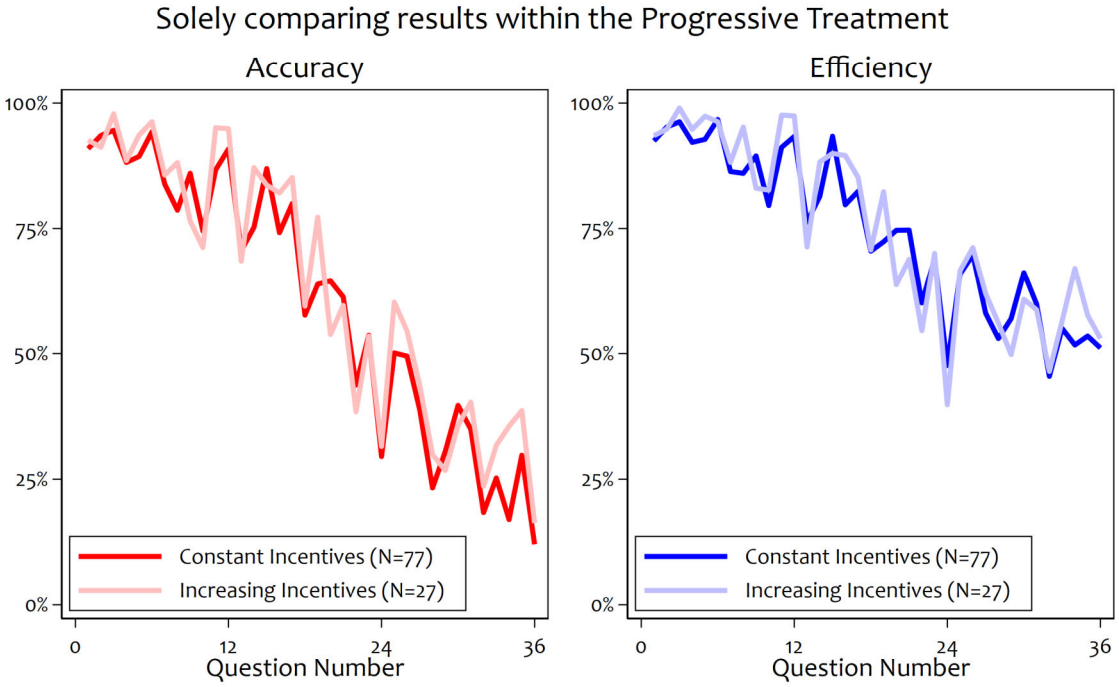
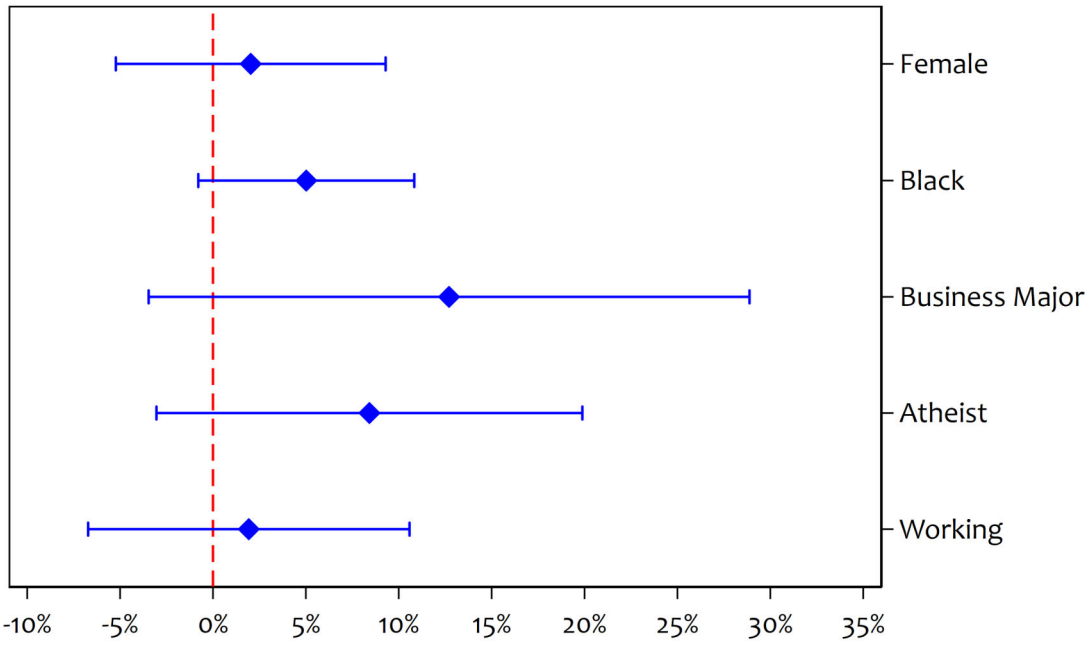


Figure D5: Effects of Increasing Incentives on Efficiency

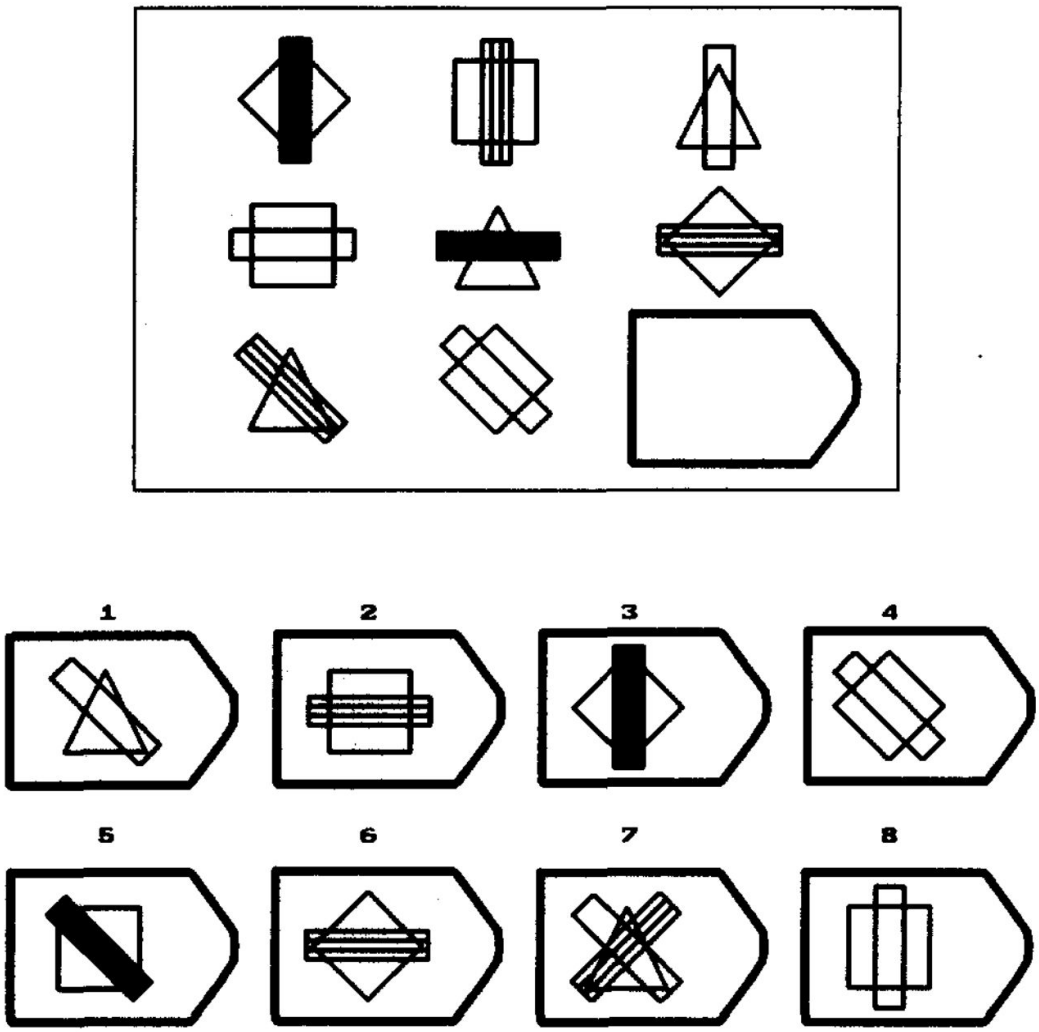
Average marginal effects of progressively increasing incentives from Fractional Regression
Solely comparing results within the Progressive Treatment
27 subjects with increasing incentives, and 77 with constant incentives



Your Instructions

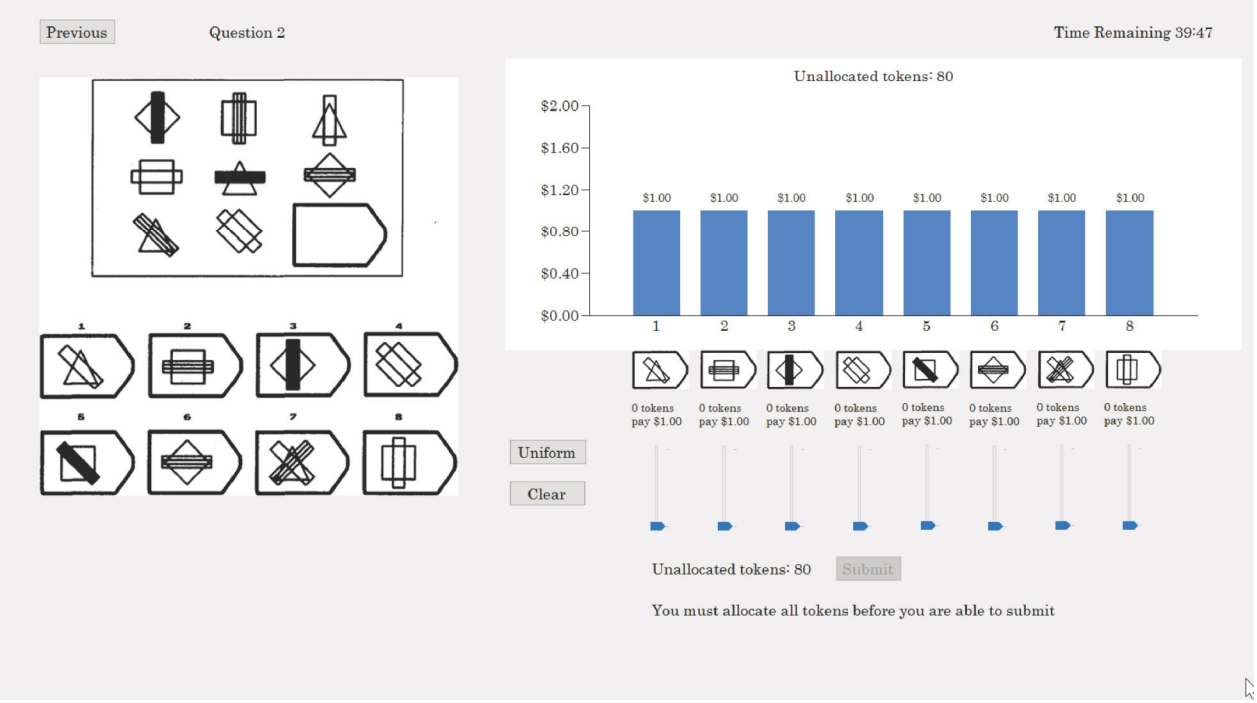
This task is a test of perception and clear-thinking. You have already completed the first part of the task, in a previous session, when you were given 12 similar problems. We will now consider a fresh set of 36 problems.

Consider a similar problem, shown here.

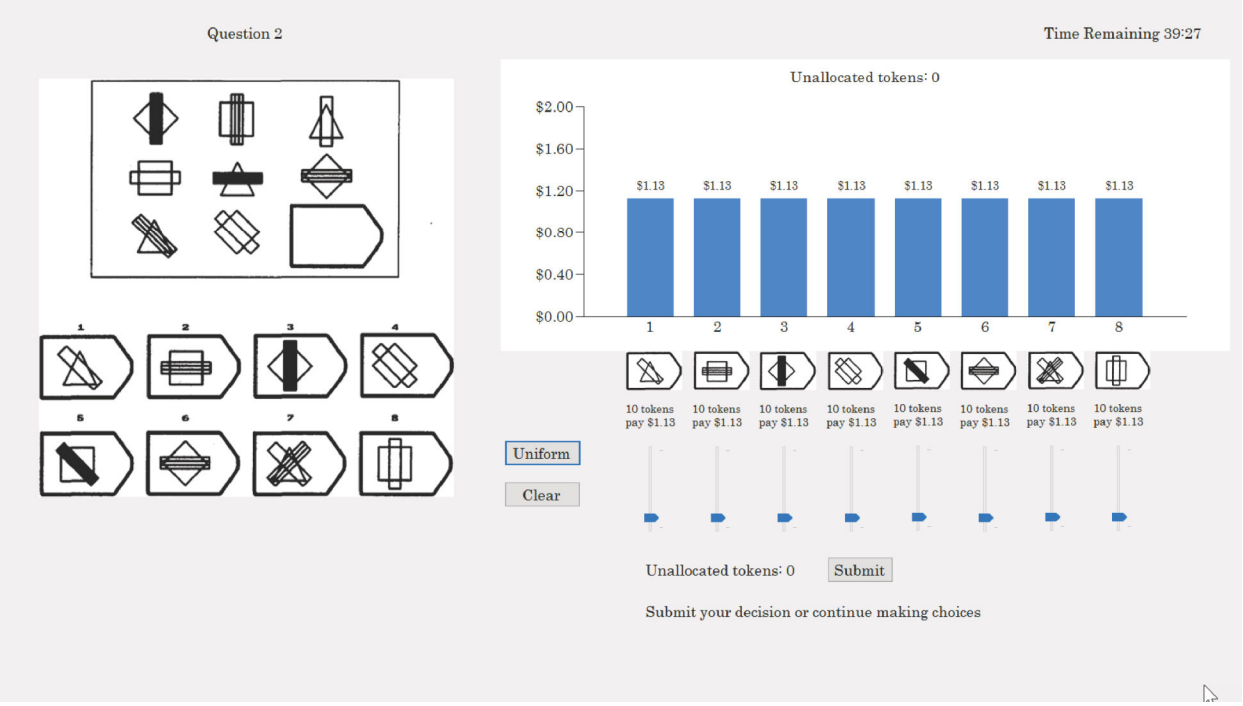


The top part of this problem is a pattern with a bit cut out of it. Look at the pattern, and think what the piece needed to complete the pattern correctly both across and down must be like. Then find the right piece out of the eight bits shown, and numbered, in the bottom part of this problem.

You will be asked to report your beliefs about the correct answer using an interface like this one, which is also generally familiar to you from a previous session. The version you will see on the computer will be larger and easier to read.



The problem and possible solutions are shown on the left of the screen, in the usual manner. On the right of the screen you have 80 tokens to allocate across the 8 possible answers. We start off with 0 tokens allocated to each of the possible answers. If you wanted to change this initial allocation so that there were 10 tokens allocated to each possible answer, just click on the **Uniform** button, and you will see this display:

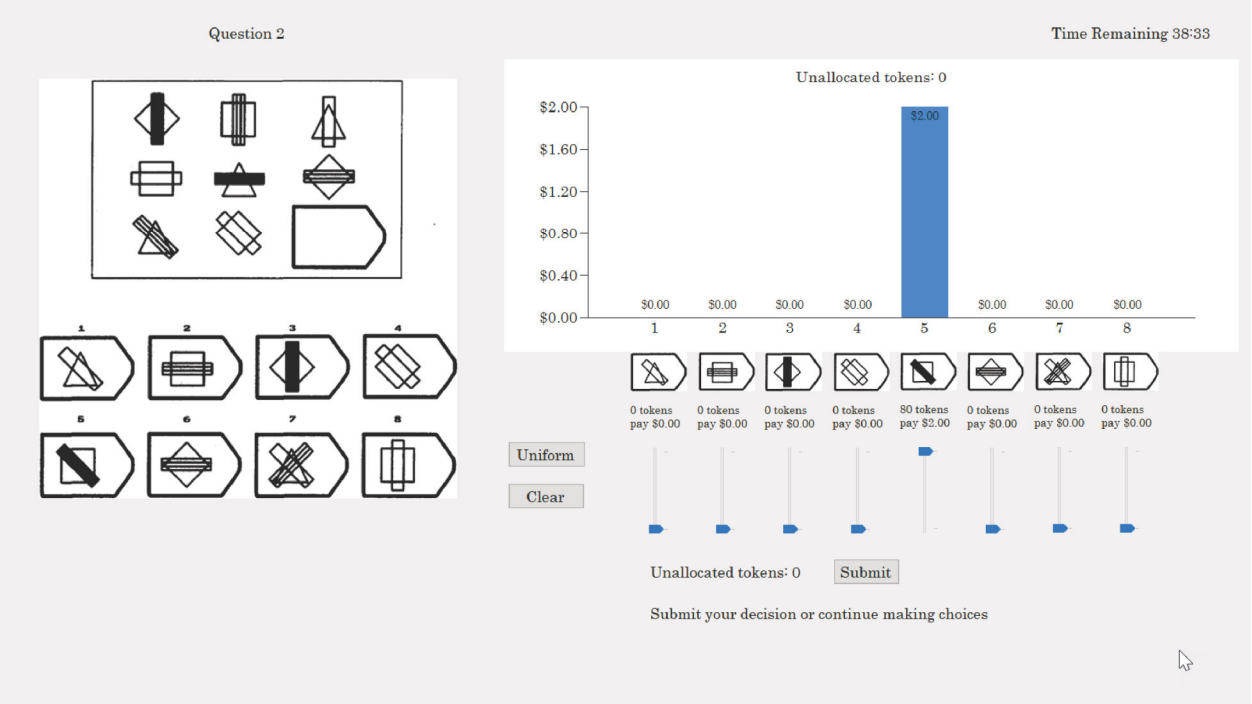


Here you have allocated 10 tokens to each possible answer, and you would earn $1.13 if you reported this allocation of tokens, since only one of the 8 possible answers is correct. You can return to the initial allocation of 0 tokens for each possible answer by clicking on the **Clear** button.

As you allocate tokens, by moving the sliders up or down, the earnings will change above each bar. These are the earnings that you will receive for this problem if that bar refers to the correct answer to the problem. **You will be paid for all 36 problems, and each problem will pay between $0 and $2 depending on your answer.**

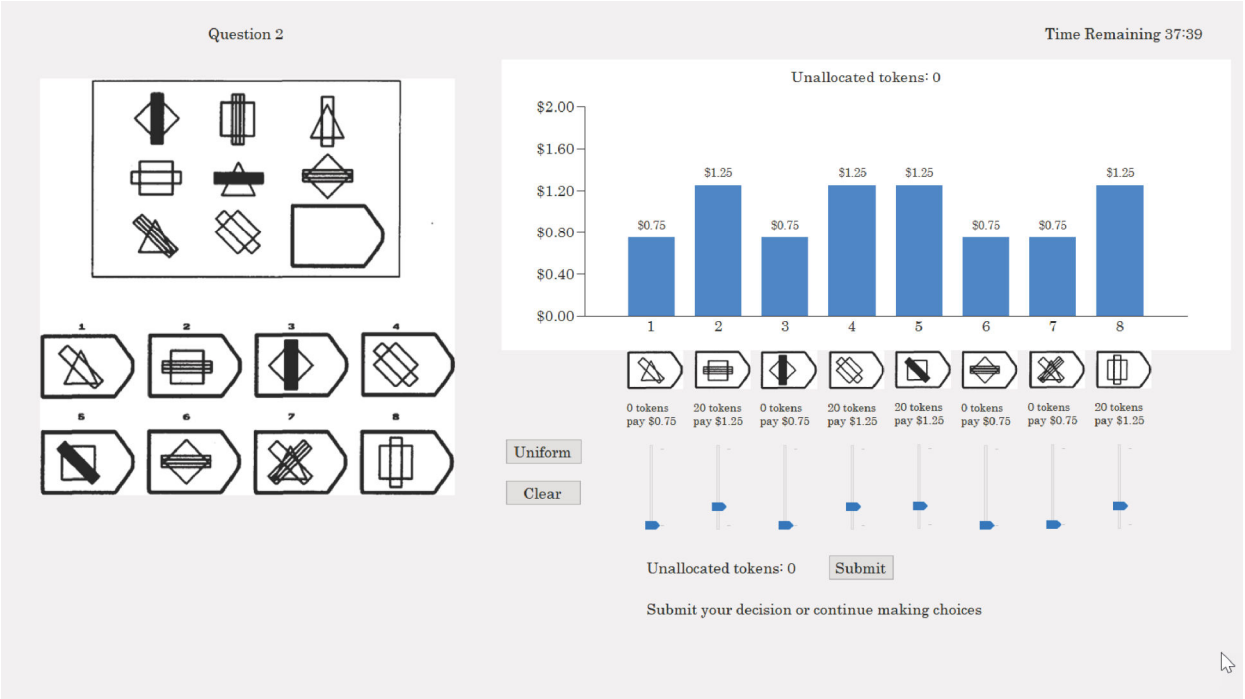
Return now to the problem itself. Only one of these pieces is perfectly correct. Pieces #2, #5 and #8 complete the problem correctly going down, in the sense that they have a square as the larger piece. Pieces #1, #4 and #5 are correct going across, in the sense that they have the right slope for the rectangle.

Piece #5 is the correct bit, isn't it? It is correct going down as well as going across. So the answer is #5. In this case, if you were certain that piece #5 is the single best answer, you should allocate all 80 tokens to the bin representing piece #5 as in this display:



So in this case you would earn $2.00 if indeed the correct answer was #5. Of course, if any of the other pieces turned out to be the correct answer you would, in this case, earn $0.

If you had decided that the correct answer was one of #2, #4, #5 or #8, but had not decided that #5 was actually the correct answer of these four possibilities, you might decide to allocate your tokens equally across the bars representing pieces #2, #4, #5 and #8 like this:



You can see that even if you eliminate some pieces, such as #1, #3, #6 and #7, that are clearly wrong, you give yourself a 1-in-4 chance of earning more money than if you guessed across all 8 pieces. In this case you would expect to earn $1.25 if indeed one of pieces #2, #4, #5 or #8 had been correct. Recall that if you had allocated the tokens roughly equally across all 8 bars, thinking that any of the 8 pieces might be correct, you would only earn $1.13 for this problem.

This illustrative problem is relatively easy, but some of the later problems will not be so easy, and it might be harder to identify the one correct answer. You would still benefit by ruling out certain pieces so that you can allocate more tokens to the ones that you have not ruled out.

You will find that the problems soon get difficult. Whether the problems are easy or difficult, you will notice that to solve them you have to use the same method all the time. The important thing is to notice how the problems develop and to learn the method of solving them. In each problem you look across each row, or down each column, and decide what the missing figure is like. Then look for the answer that is right both ways, across **and** down, from the options you have to select from.

As you recall from allocating tokens in a similar task in a previous session, you can change your token allocation as many times as you like before you hit the **Submit** button and **Confirm** your token allocation. You can also review and modify your prior decisions. You will have navigation buttons in the top left corner of the screen. These navigation buttons are only available before you start moving the sliders for a problem. Once you move any sliders for a problem, the navigation buttons disappear and you must submit your answer and move to the next problem in order to see the navigation buttons again.

You will be allowed 40 minutes to complete this task. Your screen displays how many

minutes (and seconds) are remaining in the top, right corner. Remember it is accurate work that counts. Attempt each problem in turn. Do your best to eliminate the incorrect pieces and find the correct piece to complete it before going on to the next problem. You will have the option to go back to previous problems if you want to. After you have completed the last question you must confirm if you are finished. When you are finished you will be told your total earnings, and will not have the option to go back to any of the problems. If you run out of time, you will automatically be treated as having finished, and told your total earnings.

Where you position each slider depends on your beliefs about the correct answer to the question. Again, each bar shows the amount of money you could earn if the true outcome corresponds to the possible solution shown under the bar.

You will be **paid for each of the 36 problems**, so you should think carefully about each problem. Since you can earn up to $2 for each problem, you could earn up to $72 over all 36 problems. You will not earn anything on any problems for which you have not confirmed an allocation of tokens.

It is up to you to balance the strength of your personal beliefs with the possibility of them being wrong. There are several important points for you to keep in mind when making your decisions:

- First, your belief about the correct answer to each problem is your personal judgment.
- Second, you have up to 40 minutes to complete this task.
- Third, you will not earn anything on any problems for which you have not confirmed an allocation of tokens.
- Fourth, depending on your choices and the correct answer you can earn up to $2 for each problem, or up to $72 over all 36 problems.
- Finally, your choices might also depend on your willingness to take risks or to gamble.

The decisions you make are a matter of personal choice. Please work silently, and make your choices by thinking carefully about the questions you are presented with.

When you are finished you will see a summary of your responses and your total earnings from this task. Your earnings are in addition to the show-up payment you receive for participating.

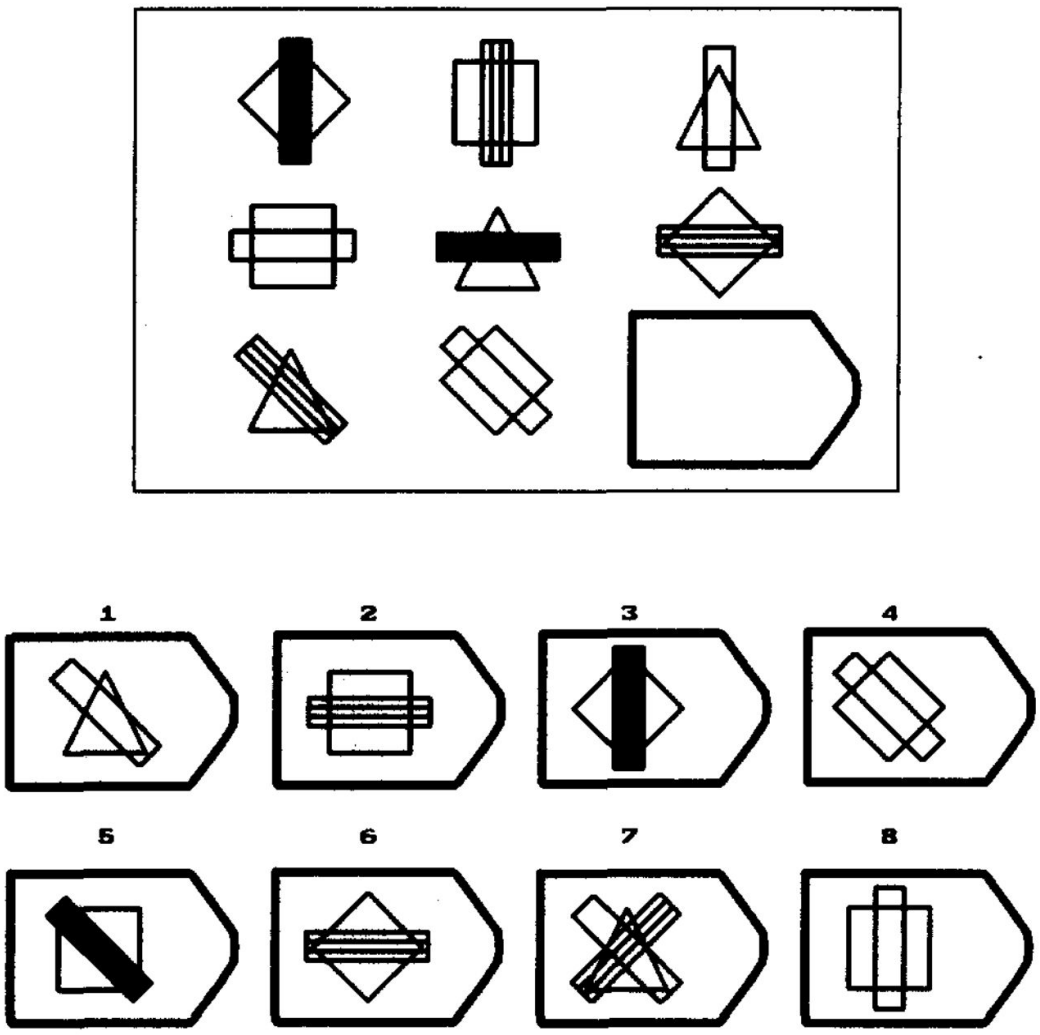
4. Instructions: Increased Incentives Treatment

R80inc

Your Instructions

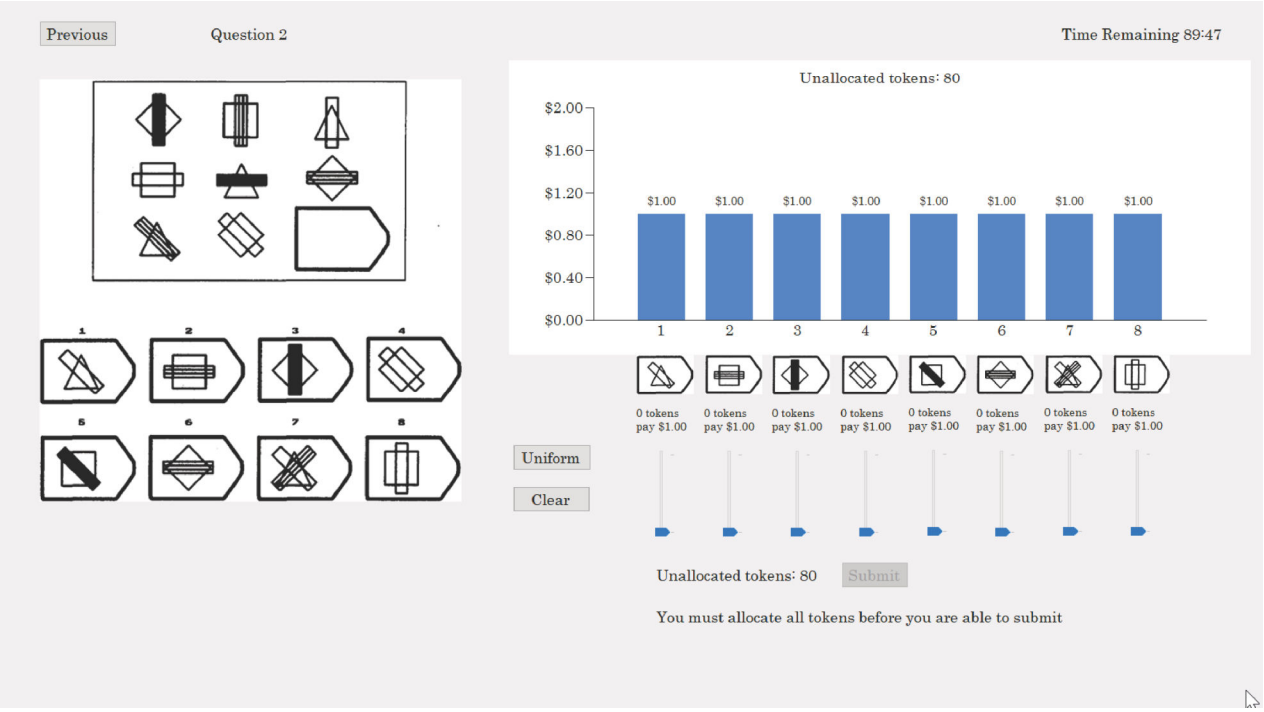
This task is a test of perception and clear-thinking. You have already completed the first part of the task, in a previous session, when you were given 12 similar problems. We will now consider a fresh set of 36 problems.

Consider a similar problem, shown here.

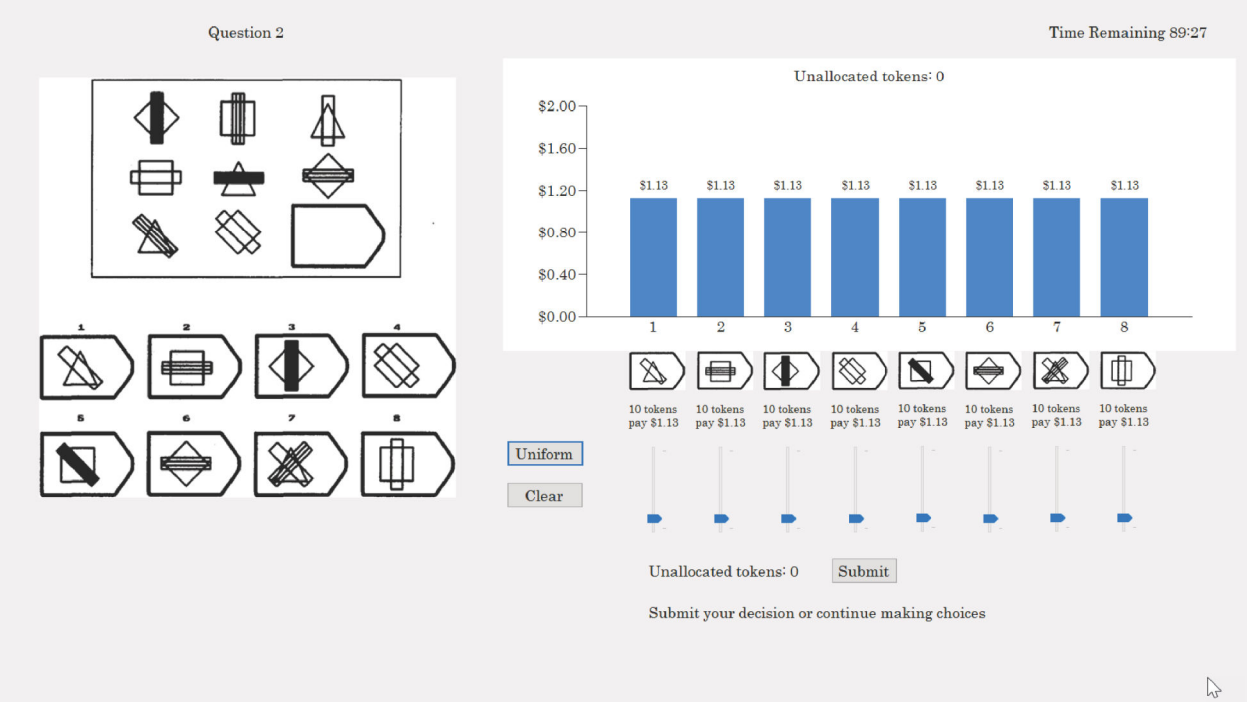


The top part of this problem is a pattern with a bit cut out of it. Look at the pattern, and think what the piece needed to complete the pattern correctly both across and down must be like. Then find the right piece out of the eight bits shown, and numbered, in the bottom part of this problem.

You will be asked to report your beliefs about the correct answer using an interface like this one, which is also generally familiar to you from a previous session. The version you will see on the computer will be larger and easier to read.



The problem and possible solutions are shown on the left of the screen, in the usual manner. On the right of the screen you have 80 tokens to allocate across the 8 possible answers. We start off with 0 tokens allocated to each of the possible answers. If you wanted to change this initial allocation so that there were 10 tokens allocated to each possible answer, just click on the **Uniform** button, and you will see this display:

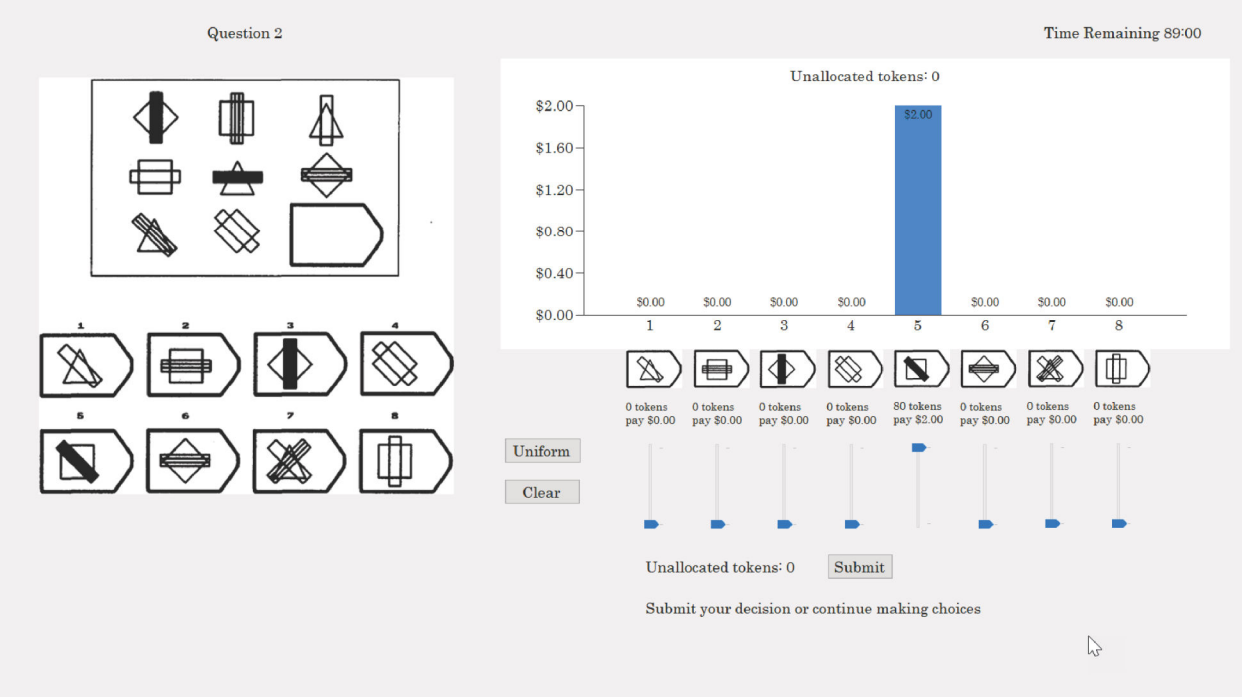


Here you have allocated 10 tokens to each possible answer, and you would earn $1.13 if you reported this allocation of tokens, since only one of the 8 possible answers is correct. You can return to the initial allocation of 0 tokens for each possible answer by clicking on the **Clear** button.

As you allocate tokens, by moving the sliders up or down, the earnings will change above each bar. These are the earnings that you will receive for this problem if that bar refers to the correct answer to the problem. **You will be paid for all 36 problems. The first 12 problems will pay between $0 and $2 depending on your answer, the next 12 problems will pay between $0 and $3 depending on our answer, and the last 12 problems will pay between $0 and $5 depending on your answer.**

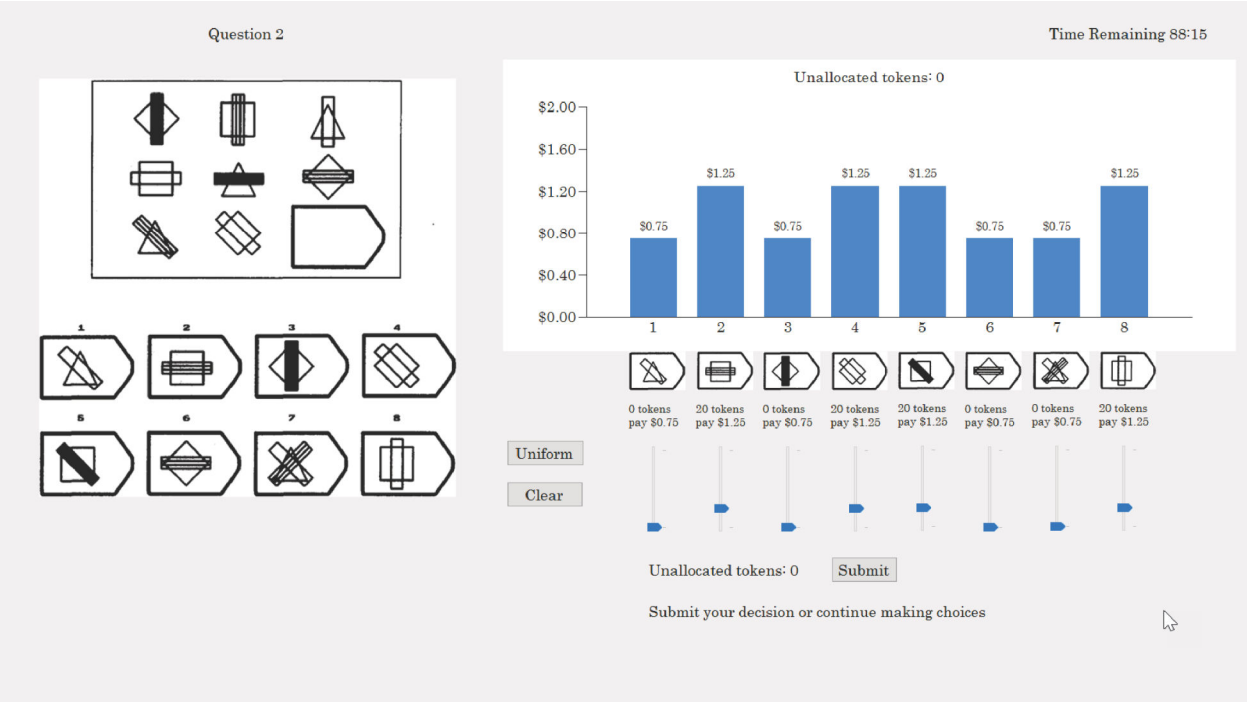
Return now to the problem itself. Only one of these pieces is perfectly correct. Pieces #2, #5 and #8 complete the problem correctly going down, in the sense that they have a square as the larger piece. Pieces #1, #4 and #5 are correct going across, in the sense that they have the right slope for the rectangle.

Piece #5 is the correct bit, isn't it? It is correct going down as well as going across. So the answer is #5. In this case, if you were certain that piece #5 is the single best answer, you should allocate all 80 tokens to the bin representing piece #5 as in this display:



So in this case you would earn $2.00 if indeed the correct answer was #5 and this was one of the first 12 problems. Of course, if any of the other pieces turned out to be the correct answer you would, in this case, earn $0. If this was one of problems 13 through 24 you would earn $3.00 if indeed the correct answer was #5. And if this was one of problems 25 through 36 you would earn $5.00 if indeed the correct answer was #5.

Assume that this was one of the first 12 problems. If you had decided that the correct answer was one of #2, #4, #5 or #8, but had not decided that #5 was actually the correct answer of these four possibilities, you might decide to allocate your tokens equally across the bars representing pieces #2, #4, #5 and #8 like this:



You can see that even if you eliminate some pieces, such as #1, #3, #6 and #7, that are clearly wrong, you give yourself a 1-in-4 chance of earning more money than if you guessed across all 8 pieces. In this case you would expect to earn $1.25 if indeed one of pieces #2, #4, #5 or #8 had been correct. Recall that if you had allocated the tokens roughly equally across all 8 bars, thinking that any of the 8 pieces might be correct, you would only earn $1.13 for this problem.

This illustrative problem is relatively easy, but some of the later problems will not be so easy, and it might be harder to identify the one correct answer. You would still benefit by ruling out certain pieces so that you can allocate more tokens to the ones that you have not ruled out.

You will find that the problems soon get difficult. Whether the problems are easy or difficult, you will notice that to solve them you have to use the same method all the time. The important thing is to notice how the problems develop and to learn the method of solving them. In each problem you look across each row, or down each column, and decide what the missing figure is like. Then look for the answer that is right both ways, across **and** down, from the options you have to select from.

As you recall from allocating tokens in a similar task in a previous session, you can change your token allocation as many times as you like before you hit the **Submit** button and **Confirm** your token allocation. You can also review and modify your prior decisions. You will have navigation buttons in the top left corner of the screen. These navigation buttons are only available before you start moving the sliders for a problem. Once you move any sliders for a problem, the navigation buttons disappear and you must submit your answer and move to the next problem in order to see the navigation buttons again.

You can work at your own speed, although we have to be out of the room in 90 minutes. Your screen displays how many minutes (and seconds) are remaining in the top, right corner. Remember it is accurate work that counts. Attempt each problem in turn. Do your best to eliminate the incorrect pieces and find the correct piece to complete it before going on to the next problem. You will have the option to go back to previous problems if you want to. After you have completed the last question you must confirm if you are finished. When you are finished you will be told your total earnings, and will not have the option to go back to any of the problems. If you run out of time, you will automatically be treated as having finished, and told your total earnings.

Where you position each slider depends on your beliefs about the correct answer to the question. Again, each bar shows the amount of money you could earn if the true outcome corresponds to the possible solution shown under the bar.

You will be **paid for each of the 36 problems**, so you should think carefully about each problem. Since you can earn up to $2 for each of problems 1 through 12, $3 for each of problems 13 through 24, and $5 for each of problems 25 through 36, you could earn up to $120 over all 36 problems. You will not earn anything on any problems for which you have not confirmed an allocation of tokens.

It is up to you to balance the strength of your personal beliefs with the possibility of them being wrong. There are several important points for you to keep in mind when making your decisions:

- First, your belief about the correct answer to each problem is your personal judgment.
- Second, you have up to 90 minutes to complete this task.
- Third, you will not earn anything on any problems for which you have not confirmed an allocation of tokens.
- Fourth, depending on your choices and the correct answer you can earn up to $2, $3 or $5 for each problem, or up to $120 over all 36 problems.
- Fifth, your potential earnings are higher as you progress through the problems.
- Finally, your choices might also depend on your willingness to take risks or to gamble.

The decisions you make are a matter of personal choice. Please work silently, and make your choices by thinking carefully about the questions you are presented with.

When you are finished you will see a summary of your responses and your total earnings from this task. Your earnings are in addition to the show-up payment you receive for participating.