Supplemental Material

Appendix B: Observer Effects

An observer effect is a situation in which the introduction of belief elicitation helps participants perform optimally in the underlying task. There is mixed evidence that belief elicitation affects game play in other settings. Rutström and Wilcox (2009) compare behavior when participants do not have beliefs elicited, when participants participate in unpaid Introspection, and when participants' beliefs are elicited using the Quadratic Scoring Rule (QSR). They test the idea that the cognitively "intrusive" QSR might drive a sharper wedge between the "affective process" of belief formation and the "deliberative judgement" reporting process. In the case of the QSR they ultimately reject the hypothesis that belief elicitation does not affect game play, although they note than these effects are concentrated in earlier periods. Nyarko and Schotter (2002) also find that belief elicitation has unintended consequences, and that participants exposed to belief elicitation are more likely to use mixed strategies than pure strategies. However, Guerra and Zizzo (2004) and others have found no evidence that elicitation affects decision-making.

Since the instructions for the SBDM require an explicit discussion of probability and incentives, we predicted that this mechanism would generate an observer effect and that this effect could reduce belief errors in the belief elicitation task. To test for observer effects, we use data from the initial experiment to compare the proportion of correct left/right choices in our no-elicitation treatment to the proportion of correct left/right choices in the SBDM mechanism and Introspection mechanism.

Result 8 *There is no statistically significant observer effect in the data.*

Support for Result 8 is provided in Table 7, which shows the proportion of correct left/right choices in blocks one and two of the experiment with the data split into subsets of 10 periods. We focus the analysis on Periods 11-30 since these are the ten periods directly before and after the introduction of beliefs.

An observer effect would create larger improvements in the proportion of correct left/right choices at the start of Block Two in the treatments with belief elicitation relative to the No-Elicitation treatment. There is no such pattern in the data: in the treatments with no belief elicitation, participants make mistakes in 35.1 percent of cases in Periods 11-20 and in 30.5 percent of cases in Periods 21-30. This difference of 4.63 percentage points is not significantly different to the difference of 4.7 percentage points observed in the SBDM mechanism in a difference-in-difference permutation test in which we restrict data to periods 11-30 (p-value = 0.898). It is also not different to the difference of 7.3 percentage points observed in the Introspection mechanism (p-value = 0.533).

In Block Two, the proportion of incorrect left/right decisions in the SBDM mechanism are not significantly different than the Introspection mechanism (p-value = 0.761).

	Block One		Block Two	
	Periods 1-10	**Periods 11-20**	**Periods 21-30**	**Periods 31-40**
No Belief Elicitation	0.405	0.351	0.305	0.346
SBDM	0.335	0.340	0.293	0.275
Introspection	0.400	0.347	0.274	0.274

Table 7: Proportion of incorrect left/right decisions in initial experiment

Appendix C: Additional Figures

Result 3 presented histograms of reported beliefs for consistent participants across all of the informative priors. Here we provide the histograms of reported beliefs for the other cases. Figure 6 shows the reported beliefs of consistent and inconsistent participants in the case of an uninformative signal for both the SBDM mechanism and the Introspection mechanism using data from both the high-information treatments with 14 black balls in the left side of Bucket A and the low-information treatments with 12 black balls. Figure 7 shows the reported beliefs for the inconsistent participants across the eight potential informative posteriors.

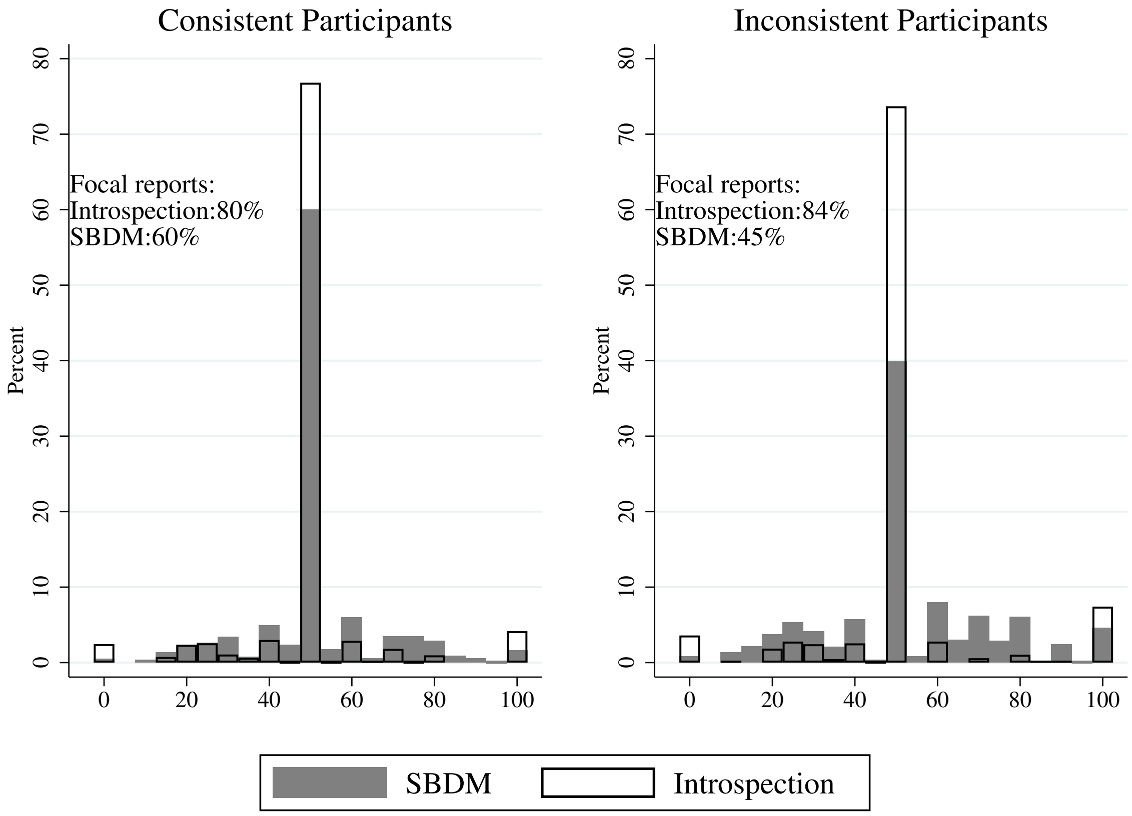
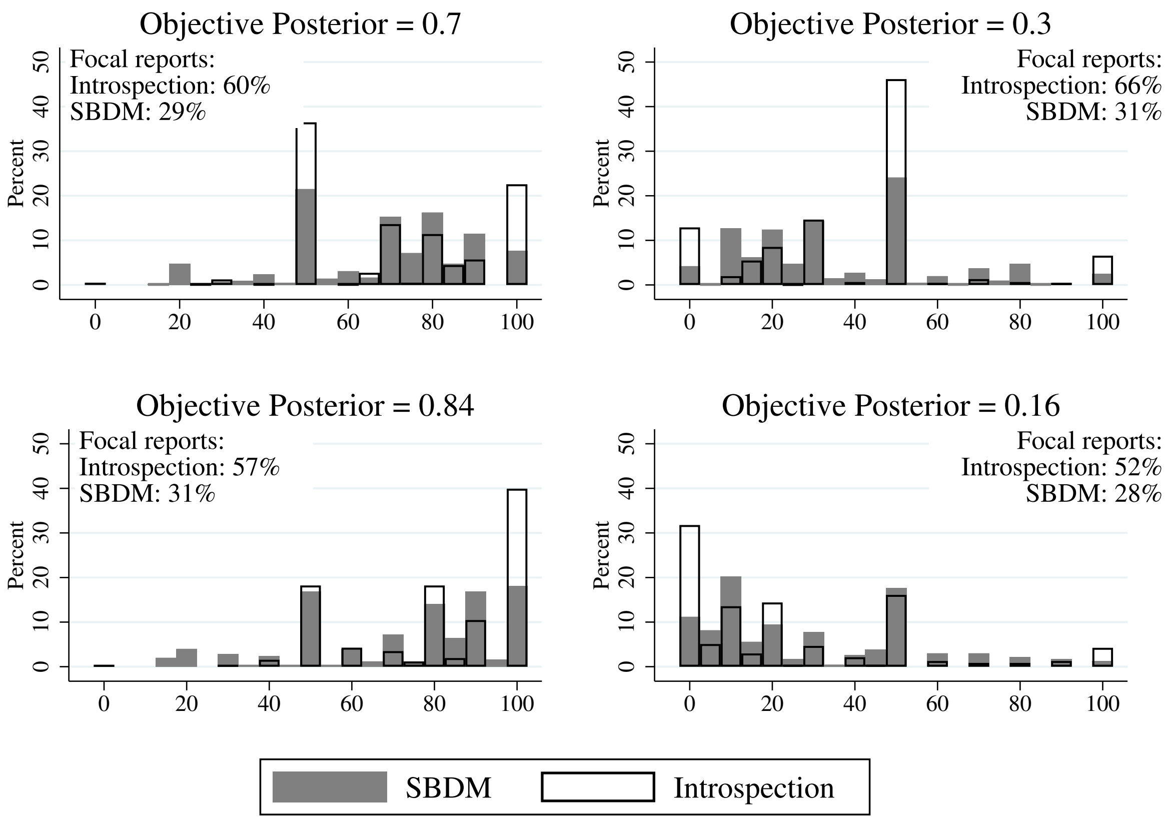
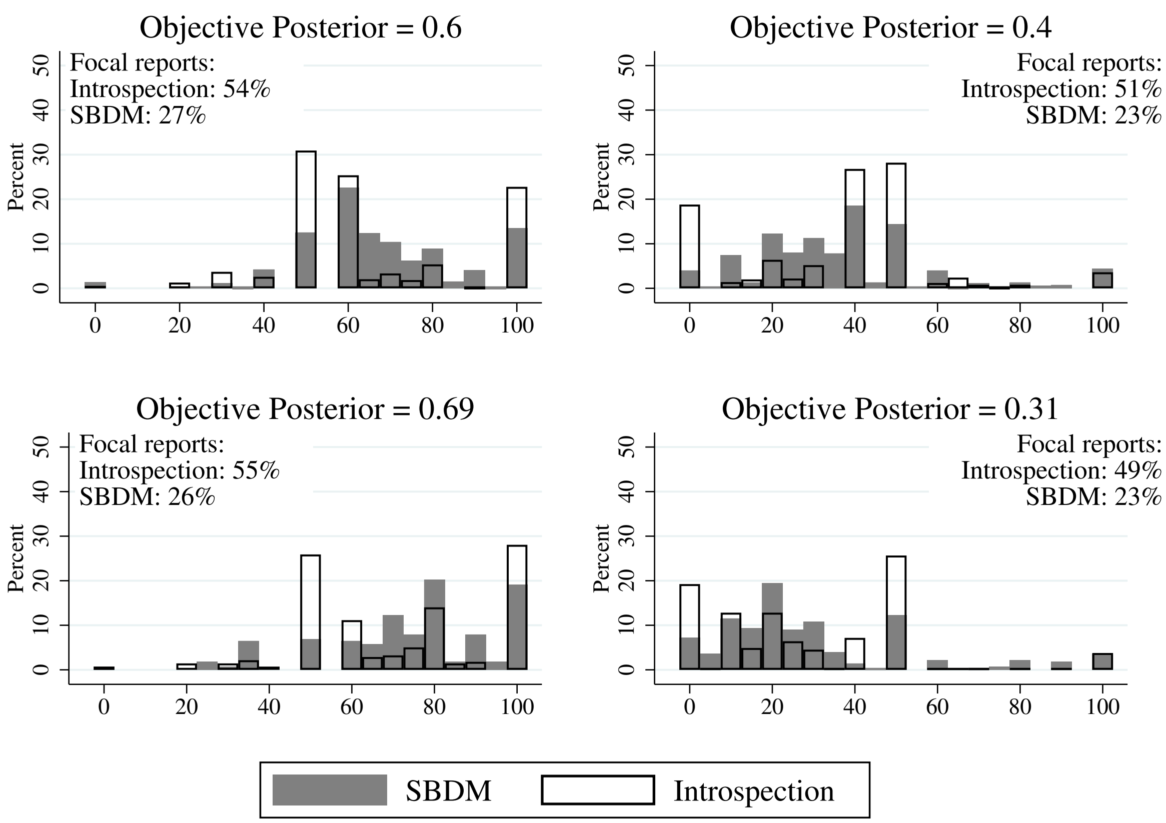


Figure 6: Distribution of reported beliefs when the posterior is 0.50



(a) High information treatments with 14 black balls in the left side of Bucket A



(b) Low information treatments with 12 black balls in the left side of Bucket A

Figure 7: Distribution of reported beliefs by inconsistent participants

Appendix D: Permutation Tests for Interactions

In this Appendix we briefly outline the Synchronized Permutation test of Perasin (2001) and Salmaso (2003) that we used to test for the interaction effect in Hypotheses 1 and 2. A more general introduction to permutation tests can be found in Good (2000) and Manly (2007). More details on Synchronized Permutation tests can be found in Basso et al. (2009) and Hahn and Salmaso (2017).

Hypotheses 1 and 2 both use a 2×2 factorial design. We are primarily interested in the interaction effect between factors. The standard approach would be to use a parametric ANOVA specification. However, as seen in the main text, the error distribution in the data is not normally distributed and thus the underlying assumption of parametric ANOVA is not satisfied. The permutation test represents an ideal alternative since it requires only a minimal assumptions about the errors, is exact in some cases, and has high power relative to other approaches.

The main assumption of permutation tests is that the data is exchangeable under the null hypothesis. Data is exchangeable if the probability of the observed data is invariant with respect to random permutations of the indexes (Basso et al., 2009). In the 2×2 factor design, the observations are typically not exchangeable since units assigned to different treatments have different expectations. This implies that approaches that freely permute data across cells may fail to separate main and interaction effects (Good, 2000). The synchronized permutation test of Perasin (2001) and Salmaso (2003) restricts permutations to the same level of a factor to generate test statistics for both main factors and, separately, interactions that depend only on the effect being tested and a combination of errors (Basso et al., 2009).

For clarity, we will concentrate the discussion on Hypothesis 1, in which each observation E_{ijk}, represents the mean error of an individual who has been assigned to mechanism $i = \{1, 2\}$ and who is of cognitive ability $j = \{1, 2\}$. We note that permutation tests will assign all observations of an individual to the same factor combination, which we refer to as a cell. Thus, there is no loss in power in using average effort as our dependent variable rather than treating each decision made by an individual as an observation.

Following the main text, we assume that each observation can be decomposed into a mean, two main effects, an interaction, and an error term:

$$E_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}, \tag{3}$$

in which $i = \{1, 2\}$ is the belief elicitation mechanism assigned to an individual, $j = \{1, 2\}$ is the cognitive ability of the individual, and $k = \{1, \dots, n_{ij}\}$ is the index of an observation within a treatment cell E_{ij}.

By including the additive constant μ, all main effects and interactions in the model can be defined to sum to zero. Thus, we assume that $\alpha_1+\alpha_2 = 0$, $\beta_1+\beta_2 = 0$, $(\alpha\beta)_{i1}+(\alpha\beta)_{i2} =$

0 for all i, and $(\alpha\beta)_{1j}+(\alpha\beta)_{2j}=0$ for all j. In this construction, $\alpha_1=-\alpha_2$ and thus, under the null of no effect of the mechanism on errors, each of the main effects $\alpha_1=\alpha_2=0$. Under the alternative, α_1 represents the difference from a zero average, and the interaction term $(\alpha\beta)_{ij}$ represents the deviation from the sum $\alpha_i+\beta_j$. The model assumes that errors ϵ_{ijk} are *exchangeable* and $\mathbb{E}(\epsilon_{ijk})=0$. Errors are exchangeable if the probability of the observed error is invariant with respect to random permutation of the data (Basso et. al, 2009).

We begin by considering a balanced design in which all cells have n observations and first construct a statistic for comparing the first factor (i.e., the mechanism) at each of the two levels of the second. Let $T_{A|1}=\sum_k E_{11k}-\sum_k E_{12k}$ and $T_{A|2}=\sum_k E_{21k}-\sum_k E_{22k}$. Further let $T_{AB}=T_{A|1}-T_{A|2}$ be a test for the interaction term. In a synchronized permutation, we select ν observations at random from the n observations in cell E_{11} and exchange them at random with observations from E_{12}. At the same time we select ν observations at random from E_{21} and exchange them at random with elements of E_{22}.

Noting that $\beta_1=-\beta_2$ and $(\alpha\beta)_{11}=-(\alpha\beta)_{12}$, a perturbation of $T_{A|1}$, will be equal to,

$$T^*_{A|1}=2(n-2\nu)\beta_1+2(n-2\nu)(\alpha\beta)_{11}+\sum_k \epsilon^*_{11k}-\sum_k \epsilon^*_{12k},$$

with the $*$ denoting a permutation of the data and ϵ^*_{ijk} denoting the permuted error. Likewise, a perturbation of $T_{A|2}$ is equal to

$$T^*_{A|2}=2(n-2\nu)\beta_1+2(n-2\nu)(\alpha\beta)_{21}+\sum_k \epsilon^*_{21k}-\sum_k \epsilon^*_{22k}.$$

Noting that $(\alpha\beta)_{11}=-(\alpha\beta)_{21}$, the expected value of the test statistic is

$$\mathbb{E}[T^*_{AB}]=4(n-2\nu)(\alpha\beta)_{11}.$$

This test statistic is independent of both main effects and relies only on the exchangeability of the errors.

We also calculate the test statistic T_{BA} where the second factor (i.e., cognitive type) is compared at each of the two levels of the first factor (i.e. the belief elicitation mechanism). Let $T_{B|1}=\sum_k E_{11k}-\sum_k E_{21k}$ and $T_{B|2}=\sum_k E_{12k}-\sum_k E_{22k}$. Then the test statistic $T^*_{BA}=T^*_{B|1}-T^*_{B|2}$ is also independent of both main effects. Since T_{AB} is obtained from synchronized permutations involving the row factor A and T_{BA} is obtained from permutations involving the column factor B, both are jointly and equally informative. It follows that their linear combination $T=T_{AB}+T_{BA}$ is a separate exact test for interaction. Following Basso et al. (2009), we use this linear combination as our main test statistic throughout the paper.

Note that in a balanced design, we can divide our test statistic by the number of observations in each cell without changing the relative value of the original test statistic and the value of the permutations. By doing so, both T_{AB} and T_{BA} are equal to the difference between (i) the difference in mean error in cells E_{11} and E_{12} and (ii) the difference in mean error in cells E_{21} and E_{22}. Thus, as described in the main text, the interaction term is based on the difference between (i) the difference in mean errors between consistent and inconsistent participants in the SBDM mechanism and (ii) the difference in mean errors between consistent and inconsistent participants in the Introspection mechanism.

We follow Basso et al. (2009) and use constrained synchronized permutations in which we exchange the observations in the same location within each cell on each iteration. This is done by permuting the observations in cells E_{11} and E_{12} and then using the same permutation of columns when shuffling observations in cells E_{21} and E_{22}, cells E_{11} and E_{21}, and cells E_{12} and E_{22}. The constrained synchronized permutation ensures that the same number of exchanges is made between each pair of cells. We perform an initial permutation of each cell to ensure that the original position of observations is irrelevant. This ensures that each permutation of the data is equally likely.

Finally, while we have aimed for a balance design, the median split of types was not always exactly 50:50 and our data is not balanced. As discussed in Good (2000), this has the potential of confounding the interaction and main effects. Basso et al. (2009) provides an approach of weighting observations that can be used to conduct synchronized permutations in an unbalanced 2×2 factor design. However, Hahn and Salmaso (2017) shows that these weights also influence the error terms and can lead to a test statistic that is too permissive. The alternative weights proposed by Hahn and Salmaso (2017), which can be used if there is balance in one direction, can be applied only in a subset of our analysis and restrict us to tests using only T_{AB} when it can be applied.

Rather than taking a weighting approach, we instead follow a suggestion in Montgomery (2017) of randomly dropping observations so that each cell has the same number of observations. Although we lose some power by reducing the size of the sample, the resulting data is a random sample of the original and the resulting test statistic is independent of the main effect. To ensure that our random subset of data is not driving our results, we use an outer loop in our testing procedure and perform our permutation test with 1000 sub samples. We report the average p-value over the 1000 samples in the main text. In Table 8 below we also report the percentage of iterations in which the individual p-value corresponds to the acceptance/rejection decision of the average p-value. For example, if the p-value of a test is 0.03 and we reject the null of no interaction, column 2 reports the percentage of sub samples in which the null was rejected.

The test for Hypothesis 2 is similar to the that of Hypothesis 1 with one major exception. In Hypothesis 2, we are comparing behavior of the same individual in informative and uninformative questions and thus the errors of observation E_{i1k} will be correlated with E_{i2k}. This correlation implies that we cannot randomly permute across informative and uninformative questions without changing the expected error distribution. In this case, we restrict attention only to the permutation test T_{BA} where we shuffle the same observations between cells E_{11} and E_{21} and then permute the same columns in cells E_{12} and E_{22}. This permutation keeps pairs of observations together and does not change the underlying error distribution.

As a robustness test, we also analyzed the data using the Wald-Type Permutation Statistic (WTPS) developed by Pauly et al. (2015). This procedure uses a free permutation of the dependent variable and is asymptotically valid in the case of heteroscedasticity in the errors across cells. In our experiment, this may be an issue if inconsistent participants have larger variation in errors. As the test is based on a Wald test, it is more sensitive to outliers. As such, we apply the test to the cleaned version of our dataset that drops outliers according to the criterion in Appendix F. As seen below, the acceptance/rejection decisions of the two tests coincide in all four of the main reported tests.

	Synchronized Permutation Test		**WTPS**
	***p*-value**	**Percentage of Iterations Matching Decision**	***p*-value**
Hypothesis 1: All Signals	0.027	89%	0.032
Hypothesis 1: Uninformative Signals	0.017	98%	0.023
Hypothesis 1: Informative Signals	0.075	74%	0.112
Hypothesis 2: All Signals	0.001	100%	0.001

Table 8: Comparison of p-values from Synchronized Permutation test and alternative Wald-Type Permutation Statistic

Appendix E: Robustness Check: Classification Criterion for Consistent and Inconsistent Participants in the Initial Experiment

In the main text we classified individuals into "consistent" and "inconsistent" types based on their decisions in the last ten periods of Block One of the experiment (Periods 11-20). This selection criterion was used to ensure that individuals were not being classified into type based on early experimentation. However, as this selection criterion could be interpreted as arbitrary, we also explored how variation of this criterion influences our results in the initial experiment where the criterion was not pre-specified.

We reported mean belief errors in the initial experiment in Appendix A in Table 5. Table 9 presents mean belief errors in the initial experiments using an alternative classification in which we do a median split using all 20 periods. As seen by comparing the two tables, mean belief errors are similar across the two classifications.

Concentrating on Table 9, which reports the mean errors for the alternative classification, the mean error for consistent participants in the SBDM mechanism is 10.76 while the mean error for inconsistent participants is 15.98. Thus, there is a -5.22 percentage point difference in means in the SBDM mechanism. The mean error for consistent participants in the Introspection mechanism is 14.57 while the mean error for inconsistent participants is 14.96. Thus there is a -0.39 percentage point difference in means in the Introspection mechanism. The difference-in-difference estimate of -4.83 is significant using the one-sided synchronized test used throughout the paper (p-value = .019). This is comparable to the difference-in-difference estimate of -5.40 that exists when we classify participants based on Periods 11-20, which is the measure we use throughout the paper (p-value = .009).

Belief Elicitation Method	Cognitive Type	Informative Signals $\rho' \in \{0.16, 0.84\}$	$\rho' \in \{0.30, 0.70\}$	$\rho' \in \{0.31, 0.69\}$	$\rho' \in \{0.40, 0.60\}$	All Informative Signals $\rho' \neq 0.5$	Uninformative Signals $\rho' = 0.5$	All Signals
SBDM	Consistent	8.41	11.25	13.74	9.72	10.57	11.39	10.76
Introspection	Consistent	16.13	18.03	15.43	13.97	16.05	9.31	14.57
- Permutation Test:		(p-value 0.001)	(p-value 0.061)	(p-value 0.672)	(p-value 0.331)	(p-value 0.015)	(p-value 0.516)	(p-value 0.078)
SBDM	Inconsistent	17.39	19.70	19.35	13.18	16.87	12.96	15.98
Introspection	Inconsistent	15.79	17.43	18.85	16.22	16.99	8.45	14.96
- Permutation Test:		(p-value 0.594)	(p-value 0.464)	(p-value 0.880)	(p-value 0.443)	(p-value 0.953)	(p-value 0.147)	(p-value 0.616)
SBDM	Full sample	12.16	14.84	16.80	11.58	13.60	12.15	13.27
Introspection	Full sample	15.97	17.73	17.34	15.22	16.54	8.84	14.78
- Permutation Test:		(p-value 0.053)	(p-value 0.236)	(p-value 0.833)	(p-value 0.203)	(p-value 0.066)	(p-value 0.134)	(p-value 0.306)

Table 9: Data from Initial Experiment Using Alternative Classification with all Observations From Block One: Mean belief errors in the SBDM mechanism and the Introspection mechanism for (i) consistent participants, (ii) inconsistent participants, and (iii) both consistent and inconsistent participants combined. The reported p-values are based on permutation tests using 10,000 iterations in which the subset of participants is held fixed and participants are randomly allocated to the SBDM or Introspection mechanism in each iteration of a regression on the treatment effect. The null hypothesis is that the treatment coefficient is equal to 0 (i.e. that there is no difference in accuracy between the SBDM and Introspection). The two-sided test statistic is reported.

Appendix F: Robustness Check: Results with Outliers Excluded

Due to Covid-19 restrictions our follow-up experiment was conducted online. The resulting data set was noisier than the data generated by the lab-based initial experiment. As part of our robustness check we removed outliers to ensure that these were not affecting results. We found that statistical tests on the reduced dataset led to statistically stronger conclusions that the results reported in the body of this paper.

When classifying participants as outliers, we began by counting the number of times that an individual reported a belief that was (i) less than or equal to 50 and (ii) greater than or equal to 50 in the final 10 periods of Blocks 2 and 3. Next, we classified an individual as an outlier if (i) either of the two counts were 19 or 20 and (ii) less than 50% of belief reports were exactly 50. This leads to the exclusion (for example) of participants who report a single number like 20 or 100 throughout the experiment, or whose probabilities are reported out of 40—the number of balls in the bucket—rather than 100.

This rule leads to the exclusion of 9 participants from the initial experiment (4 from the Introspection treatment; 5 from the SBDM treatment), and 16 from the follow-up experiment (5 from the Introspection treatment; 11 from the SBDM treatment).

Table 10 reports the mean errors from the pooled data when outliers are excluded. The difference-in-difference estimate for Hypothesis 1 is -3.23, which is statistically significant in a one-sided test using the estimator described in Appendix D (p-value = .009). Tables 11 and 12 report the mean errors for the initial experiments and follow-up experiments separately when the outliers are excluded.

Belief Elicitation Method	Cognitive Type	Informative Signals $\rho' \in \{0.16, 0.84\}$	$\rho' \in \{0.30, 0.70\}$	$\rho' \in \{0.31, 0.69\}$	$\rho' \in \{0.40, 0.60\}$	All Informative Signals $\rho' \neq 0.5$	Uninformative Signals $\rho' = 0.5$	All Signals
SBDM	Consistent	10.14	9.26	14.11	10.31	10.48	8.03	9.91
Introspection	Consistent	15.37	15.75	13.73	11.47	14.06	6.20	12.31
- Permutation Test:		(p-value: 0.001)	(p-value: 0.003)	(p-value: 0.843)	(p-value: 0.598)	(p-value: 0.004)	(p-value: 0.229)	(p-value: 0.033)
SBDM	Inconsistent	17.38	17.02	17.05	15.32	16.44	13.75	15.85
Introspection	Inconsistent	16.87	18.46	19.78	15.58	17.42	6.85	15.02
- Permutation Test:		(p-value: 0.837)	(p-value: 0.514)	(p-value: 0.156)	(p-value: 0.914)	(p-value: 0.496)	(p-value: 0.000)	(p-value: 0.538)
SBDM	Full sample	13.01	12.24	15.66	12.87	13.18	10.52	12.58
Introspection	Full sample	15.99	16.89	16.69	13.48	15.59	6.50	13.54
- Permutation Test:		(p-value: 0.035)	(p-value: 0.003)	(p-value: 0.459)	(p-value: 0.718)	(p-value: 0.013)	(p-value: 0.001)	(p-value: 0.280)

Table 10: Data with both experiments with outliers removed: mean belief errors under the SBDM mechanism and the Introspection mechanism for (i) consistent participants, (ii) inconsistent participants, and (iii) both consistent and inconsistent participants combined. The reported p-values are based on permutation tests using 10,000 iterations in which the subset of participants is held fixed and participants are randomly allocated to the SBDM or Introspection mechanism in each iteration of a regression on the treatment effect. The null hypothesis is that the treatment coefficient is equal to 0 (i.e. that there is no difference in accuracy between the SBDM and Introspection). The two-sided test statistic is reported.

Belief Elicitation Method	**Cognitive Type**	**Informative Signals** $\rho' \in \{0.16, 0.84\}$	$\rho' \in \{0.30, 0.70\}$	$\rho' \in \{0.31, 0.69\}$	$\rho' \in \{0.40, 0.60\}$	**All Informative Signals** $\rho' \neq 0.5$	**Uninformative Signals** $\rho' = 0.5$	**All Signals**
SBDM	High	8.23	10.67	14.49	9.10	10.34	9.09	10.06
Introspection	High	15.34	17.02	13.49	13.78	15.24	8.82	13.82
- Permutation Test:		(p-value: 0.004)	(p-value: 0.094)	(p-value: 0.775)	(p-value: 0.205)	(p-value: 0.020)	(p-value: 0.926)	(p-value 0.061)
SBDM	Low	16.69	18.46	18.74	14.38	16.93	14.57	16.41
Introspection	Low	16.97	17.84	20.84	14.56	17.09	7.17	14.72
- Permutation Test:		(p-value: 0.932)	(p-value: 0.846)	(p-value: 0.564)	(p-value 0.971)	(p-value: 0.949)	(p-value: 0.028)	(p-value: 0.449)
SBDM	Full sample	12.42	14.67	16.45	11.40	13.47	11.66	13.07
Introspection	Full sample	15.97	17.36	17.05	14.16	16.06	8.05	14.22
- Permutation Test:		(p-value: 0.079)	(p-value: 0.287)	(p-value: 0.817)	(p-value 0.358)	(p-value: 0.113)	(p-value: 0.096)	(p-value: 0.456)

Table 11: Data from initial experiment with outliers removed: Mean belief errors under the SBDM mechanism and the Introspection mechanism for (i) consistent participants, (ii) inconsistent participants, and (iii) all participants combined. The reported p-values are based on permutation tests using 10,000 iterations in which the subset of participants is held fixed and participants are randomly allocated to the SBDM or Introspection mechanism in each iteration of a regression on the treatment effect. The null hypothesis is that the treatment coefficient is equal to 0 (i.e. that there is no difference in belief error between the SBDM and Introspection). The two-sided test statistic is reported.

Belief Elicitation Method	**Cognitive Type**	**Informative Signals** $\rho' \in \{0.16, 0.84\}$	$\rho' \in \{0.30, 0.70\}$	$\rho' \in \{0.31, 0.69\}$	$\rho' \in \{0.40, 0.60\}$	**All Informative Signals** $\rho' \neq 0.5$	**Uninformative Signals** $\rho' = 0.5$	**All Signals**
SBDM	High	11.21	8.59	13.81	11.38	10.57	7.40	9.81
Introspection	High	15.38	14.84	13.86	10.13	13.29	4.51	11.32
- Permutation Test:		(p-value: 0.045)	(p-value: 0.016)	(p-value: 0.983)	(p-value: 0.645)	(p-value: 0.081)	(p-value: 0.091)	(p-value: 0.255)
SBDM	Low	18.18	15.31	16.15	15.81	16.08	13.14	15.43
Introspection	Low	16.81	18.87	19.21	16.16	17.62	6.63	15.20
- Permutation Test:		(p-value: 0.721)	(p-value: 0.257)	(p-value: 0.176)	(p-value: 0.903)	(p-value: 0.405)	(p-value: 0.006)	(p-value: 0.893)
SBDM	Full sample	13.46	10.58	15.14	13.87	12.99	9.77	12.24
Introspection	Full sample	16.01	16.57	16.50	13.10	15.29	5.48	13.11
- Permutation Test:		(p-value: 0.198)	(p-value: 0.004)	(p-value: 0.419)	(p-value: 0.704)	(p-value: 0.064)	(p-value: 0.002)	(p-value: 0.432)

Table 12: Data from follow-up experiment with outliers removed: Mean belief errors under the SBDM mechanism and the Introspection mechanism for (i) consistent participants, (ii) inconsistent participants, and (iii) all participants combined. The reported p-values are based on permutation tests using 10,000 iterations in which the subset of participants is held fixed and participants are randomly allocated to the SBDM or Introspection mechanism in each iteration of a regression on the treatment effect. The null hypothesis is that the treatment coefficient is equal to 0 (i.e. that there is no difference in belief error between the SBDM and Introspection). The two-sided test statistic is reported.

Appendix G: Instructions

Instructions and Quizzes

The experiment included 3 blocks of 20 periods, which were referred to in the instructions as Experiments 1, 2 and 3. Statements in parentheses and italics provide additional details or discuss differences between the treatments and do not form part of the experiment instructions.

Experiment One

Thank you for choosing to participate today. We appreciate your time. This experiment is an opportunity to earn money. You will be paid in cash at the end of the experiment. You will be paid a $10 attendance fee. You will also receive payments based on the outcome of three experiments. You will not learn your total payoff until the end of the experiment.

There is a very short, anonymous questionnaire at the end of the experiment. You will be paid when the questionnaire is completed.

If you have any questions during the experiment, please sit quietly and raise your hand. An experiment assistant will be with you as soon as possible.

Payment for the first experiment: You will play the first experiment 20 times. Each repetition is called a "period." In each period you will get a payoff of $0, $4, or $8. At the end of the experiment, 1 of the 20 periods will be chosen randomly by the computer. Each period is equally likely to be chosen. Your cash payment for the first experiment will be your payoff in the randomly chosen period.

*(In bold text:)*Although you will play 20 periods in the first experiment, you are only paid in cash for the payoff you earn in a single period.

You are going to participate in a decision-making task, which is referred to as the "Choose-A-Side Game." There are two buckets: Bucket A and Bucket B. Each bucket contains 40 balls. Each bucket is divided in half, with 20 balls in each side.

There is a 50-in-100 chance (50% chance) that you have been given Bucket A. The left side of Bucket A contains 12 black balls and 8 white balls. The right side of Bucket A contains 20 black balls and 0 white balls.

(Stylized illustration of Bucket A: a rectangle divided vertically in two, with black or white dots to illustrate the ratio of black and white balls in each half of the bucket.)

There is a 50-in-100 chance (50% chance) that you have been given Bucket B. The left side of Bucket A contains 8 black balls and 12 white balls. The right side of Bucket A contains 0 black balls and 20 white balls. (The buckets and balls are all computerized.)

(*Stylized illustration of Bucket B: a rectangle divided vertically in two, with black or white dots to illustrate the ratio of black and white balls in each half of the bucket.*)

One of the buckets will be randomly chosen by the computer. Both buckets have an equal chance of being chosen. This means that both buckets have a 50-in-100 chance of being chosen (50%). (You might imagine that the computer tosses a coin to decide which bucket will be used.) You will not be told which bucket has been chosen by the computer. The computer will randomly select a ball from the left hand side of your bucket. Each ball has an equal chance of being chosen. You will be told the colour of the ball. After you see the ball, it is put back in the left hand side of your bucket. If the ball is black, you receive $4. If it is white, you receive $0 (nothing). This is your Stage-1 payoff.

You then have a second chance to draw a ball from your bucket. As before, black balls are worth $4. White balls are worth $0 (nothing). You must decide whether you would like the computer to draw the ball from the left hand side of your bucket, or the right hand side. The computer randomly selects a ball from the side you choose. If it is black, you receive $4. If it is white, you receive $0 (nothing). This is your Stage-2 payoff.

Your payoff for the period is your Stage-1 payoff plus your Stage-2 payoff. In total you might have a payoff of $0, $4, or $8 across both stages of the Choose-A-Side Game.

In each period there is a 50-in-100 (50%) chance of being given Bucket A or Bucket B. Your bucket is randomly determined by the computer and is not affected by the bucket you have been given in previous periods.

Summary: Choose-A-Side Game

There are 2 buckets, Bucket A and Bucket B. Each bucket has a 50-in-100 chance (50%) of being chosen. Each bucket is divided in half, with 20 balls in each half. The computer randomly selects a bucket for you. You do not know which bucket you have been given. You will see a randomly chosen ball from the left-hand side of your bucket. If it is black, your Stage-1 payoff is $4. If it is white, your payoff is $0 (nothing). You then choose whether you want a second ball drawn from the left or right side of your bucket. The computer draws a ball from your chosen side. If it is black, your Stage-2 payoff is $4. If it is white, your payoff is $0 (nothing). Your period payoff is your Stage 1 plus your Stage 2 payoff. In each period you might get a payoff of $0 (nothing), $4, or $8 in the Choose-A-Side Game. 1 of the 20 periods will be randomly chosen. Each period has an equal (1-in-20) chance of being chosen. Your will be paid your earnings from that period in cash at the end of the experiment.

When you have finished Experiment 1 you will be given instructions for a second experiment.

Quiz

At the start of a period the computer randomly selects a bucket for you.

1. What is the chance-in-100 that you get Bucket A? (50)

2. What is the chance-in-100 that you get Bucket B? (50)

The bucket has 20 balls in each side; 40 in total. The computer shows you a ball from the left-hand side of your bucket, tells you its colour, and tells you whether it is worth $4 or $0. This is your payoff for Stage 1. The computer puts the ball back in your bucket. The computer asks whether you want the next ball drawn from the left-hand side or the right-hand side of the bucket.

3. How many balls are there in the left-hand side? (20)
4. How many balls are there in the right-hand side? (20)

The computer draws a ball from the side you choose, tells you its colour, and tells you if you have won $4 or $0. This is your payoff for Stage 2.

5. What is the minimum payoff possible in a period (Stage 1 + Stage 2)? (0)
6. What is the maximum payoff possible in a period (Stage 1 + Stage 2)? (8)

You then finish the period.

7. How many periods are there in this experiment? (20)
8. Do you receive a cash payment for your payoff in every period? (No)
9. Every period has a 1-in-? chance of being paid? (20)
10. When the next period starts, what is the chance-in-100 that you get Bucket A? (50)
11. What is the chance-in-100 that you get Bucket B? (50)

(Experiment begins when all questions are answered correctly. At the end of Experiment One:)

Thank you! You have now played 20 periods and finished the first experiment. At the end of the third experiment you will find out which period was randomly chosen. You will be paid your payoff from the randomly chosen period. You will be pain in cash. You will now read instructions for the second experiment.

Experiment Two

You will play the second experiment 20 times. Each repetition is called a 'period.' In each period you get a payoff of $0, $4, or $8. At the end of the second experiment, 1 of the 20 periods will be chosen randomly by the computer. Each period is equally likely to be chosen. Your cash payment for the second experiment will be your payoff in the randomly chosen period. Your total payment today will include:

- Your show-up fee of $10
- A cash payment for a randomly chosen period from the first experiment

- A cash payment for a randomly chosen period from the second experiment

- A cash payment for a randomly chosen period from the third experiment

Although you will play 20 periods in this second experiment, you only receive cash for your payoff from a single period.

The set-up for Experiment 2 is the same as Experiment 1. There are two buckets: Bucket A and Bucket B. Each bucket contains 40 balls. Each bucket is divided in half, with 20 balls in each side.

There is a 50-in-100 chance (50% chance) that you have been given Bucket A. The left side of Bucket A contains 12 black balls and 8 white balls. The right side of Bucket A contains 20 black balls and 0 white balls.

(*Stylized illustration of Bucket A: a rectangle divided vertically in two, with black or white dots to illustrate the ratio of black and white balls in each half of the bucket.*)

There is a 50-in-100 chance (50% chance) that you have been given Bucket B. The left side of Bucket A contains 8 black balls and 12 white balls. The right side of Bucket A contains 0 black balls and 20 white balls. (The buckets and balls are all computerized.)

(*Stylized illustration of Bucket B: a rectangle divided vertically in two, with black or white dots to illustrate the ratio of black and white balls in each half of the bucket.*)

One of the buckets will be randomly chosen by the computer. Both buckets have an equal (50-in-100) chance of being chosen. You will not be told which bucket has been chosen by the computer. The computer will randomly select a ball from the left hand side of your bucket. Each ball has an equal chance of being chosen. You will be told the colour of the ball. After you see the ball, it is put back in the left hand side of your bucket. If the ball is black, you receive $4. If it is white, you receive $0 (nothing). This is your Stage-1 payoff.

(The three treatments involve different instructions from this point.)

SBDM mechanism treatment

After seeing the colour of the ball, you need to think about the chance that the ball was drawn from Bucket A. This is your "belief" that the ball was drawn from Bucket A. Your "belief" is a number between 0 and 100, to indicate the chance-in-100 that the ball has been drawn from Bucket A.

For example: If you are sure that Bucket A is being used, your belief is that there is a 100-in-100 chance that Bucket A is being used. If you are sure that Bucket A is not being used, your belief is that there is a 0-in-100 chance that Bucket A is being used. If you believe that it is equally likely that Bucket A is being used as Bucket B, then your belief is that there is a 50-in-100 chance that Bucket A is being used. (These are just examples. You can enter any chance-in-100 belief between 0 and 100.)

You then answer 2 questions. Question 1: What is your belief that the ball was drawn from Bucket A? Question 2: Do you want the computer to draw a second ball from the left or right-hand side of your bucket?

The computer then tosses a coin to determine which question is used to determine your Stage-2 payoff. Tails: Question 1. Heads: Question 2. If the computer throws a Heads, your Stage-2 payoff will be determined the same way as Experiment 1 (the Choose-A-Side Game). The computer will draw a ball from the side of the bucket you choose. As before, black balls are worth $4. White balls are worth $0 (nothing).

We will now explain how Stage-2 payoffs are determined if the computer throws “Tails.”

In Question 1 you tell the computer your belief (the chance-in-100) that the first ball was drawn from Bucket A. If the computer throws “Tails”, this is how we determine your Stage-2 payoff:

The computer creates a Lottery Bag. The computer randomly chooses a number between 0 and 100. Each number is equally likely to be chosen. Although the computer knows this number, you do not. We call this randomly chosen number “?”. The computer fills a bag with 100 chips. ”?” chips are black, and the rest are white. ?-in-100 chips are black. There are now two ways to get a payoff of $4: The ‘Belief about Bucket A Game,” and the Lottery Bag Game.

(Table/illustration here.)

THE BELIEF ABOUT BUCKET A GAME:
Prize of $4 if the ball was from Bucket A.
Prize of $0 if the ball was from Bucket B.

Chance-in-100 of winning $4:

Belief (chance-in-100) that ball is from Bucket A

THE LOTTERY BAG GAME:
Prize of $4 if you draw a black chip.
Prize of $0 if you draw a white chip.

Chance-in-100 of winning $4:

“?”-in-100

The computer knows the chance of winning $4 in the Lottery Bag Game. Based on your reported belief that the ball was drawn from Bucket A, the computer will select the game that gives you the highest chance of winning $4. (If the games give you an equal chance of winning, you will play the Lottery Bag Game.)

You should think carefully about your belief that the ball has been drawn from Bucket A, as the computer will use your reported belief to decide whether you are paid according to your “Belief about Bucket A” or the ”Lottery Bag” Game. This experiment might feel very detailed and complicated, but it is set up this way so that it is in your best interest to report your beliefs honestly and carefully. If you make a report that is not your true belief, your payoff might be determined by the Lottery Bag Game when you would prefer to be paid based on your belief that the ball was drawn from Bucket A (or vice-versa).

The best thing you can do is report your belief honestly, so that you are given the game with the highest chance of a payoff of $4.

Summary: Experiment 2

You have a 50-in-100 chance of being given Bucket A or Bucket B in each period. You will be shown a ball from the left-hand side of your bucket. You will answer 2 questions. Question 1: What is your belief that the ball was drawn from Bucket A? Question 2: Do you want the computer to draw a second ball from the left or right side of your bucket? The computer then tosses a coin to determine which question is used to determine your Stage-2 payoff. Tails: Question 1. Heads: Question 2.

If the computer throws “Heads” your payoff is determined in the same way as Experiment 1. A second ball will be drawn from your bucket, from the side you choose. If the computer throws “Tails” your payoff will be determined by the ”Belief about Bucket A” Game or a Lottery Bag Game. The best thing you can do is report your belief honestly, so that you are given the game with the highest chance of a payoff of $4.

Your period payoff is your Stage-1 payoff plus your Stage-2 payoff. In each period you might get a payoff of $0, $4, or $8. 1 of the 20 periods will be randomly chosen. Each period has an equal (1-in-20) chance of being chosen. Your will be paid your earnings from that period in cash at the end of the experiment.

Quiz

Imagine that you are shown a ball. Based on its colour, you report your belief that there is a 20-in-100 chance that the ball is from Bucket A. The computer flips a coin and it lands on ”Tails.” The computer creates a Lottery Bag game, and randomly includes 25 black chips. It has a 25-in-100 chance of winning $4. Based on your report, the computer chooses the game that gives you a higher chance of winning $4.

1. Which game will be used to determine your payoff for the period? (Lottery Bag Game)

2. What is your chance-in-100 of winning $4? (25)

3. What is your chance-in-100 of winning $0? (75)

Imagine you start a new period. You are shown a new ball. This time, you believe there is an 81-in-100 chance that the ball was taken from Bucket A... but you make an error! You type 18 by mistake. This is your reported belief.

The computer doesn’t know your belief, only your reported belief. The computer thinks you believe there is an 18-in-100 chance of winning $4 in the Belief-About-Bucket-a Game.

The computer creates a Lottery Bag Game and randomly includes 36 black chips. It has a 36-in-100 chance of winning $4.

4. What do you believe is your chance-in-100 of winning $4 if you play the Belief-About-Bucket-A Game? (81)
5. What does the computer think you believe is the chance-in-100 of winning $4 if you play the Belief-About-Bucket-A Game? (18)
6. What is your chance-in-100 of winning $4 if you play the Lottery Bag Game? (36)

Based on your report, the computer chooses the game that it thinks will give you a higher chance of winning $4.

7. Which game will be used to determine your prize for the period? (Lottery Bag Game)

(Experiment begins when all questions are answered correctly.)

Unpaid Introspection Treatment

After seeing the colour of the ball, you need to think about the chance that the ball was drawn from Bucket A. This is your "belief" that the ball was drawn from Bucket A. Your "belief" is a number between 0 and 100 to indicate the chance in 100 that the ball has been drawn from Bucket A.You should think carefully about your belief that the ball has been drawn from Bucket A.

For example: If you are sure that Bucket A is being used, your belief is that there is a 100-in-100 chance that Bucket A is being used. If you are sure that Bucket A is not being used, your belief is that there is a 0-in-100 chance that Bucket A is being used. If you believe that it is equally likely that Bucket A is being used as Bucket B, then your belief is that there is a 50-in-100 chance that Bucket A is being used. (These are just examples. You can enter any chance-in-100 belief between 0 and 100.)

You then answer 2 questions.

Question 1: What is your belief that the ball was drawn from Bucket A

Question 2: Do you want the computer to draw a second ball from the left or right side of your bucket?

The computer randomly selects a ball from the side you choose. If it is black, you receive $4. If it is white, you receive $0 (nothing). This is your Stage-2 payoff. Your payoff for the period is your Stage-1 payoff plus your Stage-2 payoff. In total you might have a payoff of $0 (nothing), $4 or $8 across both stages of the Choose-A-Side Game. In each period there is a 50-in-100 (50%) chance of being given Bucket A or Bucket B.Your bucket is randomly determined by the computer and is not affected by the bucket you have been given in previous periods.

Summary: Experiment 2 You have a 50-in-100 (50%) chance of being given Bucket A or Bucket B. Each bucket is divided in half, with 20 balls in each half. You will be shown a ball from the left hand side of your bucket. If it is black, your Stage-1 payoff is \$4. If it is white, your payoff is \$0 (nothing). You will answer 2 questions:

Question 1: What is your belief that the ball was drawn from Bucket A?

Question 2: Do you want the computer to draw a second ball from the left or right side of your bucket?

You should think carefully about your belief that the ball has been drawn from Bucket A. If it is black, your Stage-2 payoff is \$4. If it is white, your payoff is \$0 (nothing). Your period payoff is your Stage 1 payoff plus your Stage 2 payoff. In each period you might get a payoff of \$0 (nothing), \$4 or \$8. 1 of the 20 periods will be randomly chosen. Each period has an equal (1-in-20) chance of being chosen. You will be paid your payoff from that period in cash at the end of the experiment. When you have finished Experiment 2 you will be given instructions for a third experiment.

Quiz

At the start of a period the computer randomly selects a bucket for you. The bucket has 20 balls in each side: 40 in total. The computer shows you a ball from the left-hand side of the bucket, tells you its colour, and tells you whether it is worth \$0 or \$4. This is your payoff for Stage 1. The computer puts the ball back in your bucket. The computer asks whether you want the next ball drawn from the left hand side or the right hand side of the bucket. The computer also asks your belief about the chance-in-100 that the ball was drawn from Bucket A.

Imagine that you're sure the ball is drawn from Bucket A. How would you report this as a chance-in-100 that the ball is drawn from Bucket A?

1. The chance-in-100 of the Ball being from Bucket A is: (100)

Imagine that you're sure the ball is not drawn from Bucket A. How would you report this as a chance-in-100 that the ball is drawn from Bucket A?

2. The chance-in-100 of the Ball being from Bucket A is: (0)

Imagine that you think there's an equal chance the ball is drawn from Bucket A. How would you report this as a chance-in-100 that the ball is drawn from Bucket A?

3. The chance-in-100 of the Ball being from Bucket A is: (50)

The computer draws a ball from the side you choose, tells you its colour, and tells you if you have won \$0 or \$4. This is your payoff for Stage 2.

4. What is the minimum payoff possible in a period (Stage 1 + 2)? (0)

5. What is the maximum payoff possible in a period (Stage 1 + 2)? (8)

You then finish the period.

6. How many periods are there in this experiment? (20)
7. Do you receive a cash payments for your payoff in every period? (No)
8. Every period has a 1-in-? chance of being paid? 1-in: (20)

(Experiment begins when all questions are answered correctly.)

No-Elicitation Treatment

You must then decide whether you would like the computer to draw a second ball from the left-hand side of your bucket, or the right-hand side.

The computer randomly selects a ball from the side you choose. If it is black, you receive \$4. If it is white, you receive \$0 (nothing). This is your Stage-2 payoff. Your payoff for the period is your Stage-1 payoff plus your Stage-2 payoff. In total you might have a payoff of \$0 (nothing), \$4, or \$8 across both stages of the Choose-A-Side Game. In each period there is a 50-in-100 (50%) chance of being given Bucket A or Bucket B. Your bucket is randomly chosen by the computer and is not affected by the bucket you have been given in previous rounds.

Summary: Experiment 2 You have a 50-in-100 (50%) chance of being given Bucket A or Bucket B. Each bucket is divided in half, with 20 balls in each half. You will be shown a ball from the left-hand side of your bucket. If it is black, your Stage-1 payoff is \$4. It if is white, your payoff is \$0 (nothing). You must then decide whether you would like the computer to draw a second ball from the left-hand side of your bucket, or the right-hand side. A second ball will be drawn from your bucket, from the side you choose. If it is black, your Stage-2 payoff is \$4. If it is white, your payoff is \$0 (nothing). Your period payoff is your Stage 1 playoff plus your Stage 2 payoff. In each period you might get a payoff of \$0 (nothing), \$4, or \$8 in the Choose-A-Side Game. 1 of the 20 periods will be randomly chosen. Each period has an equal (1-in-20) chance of being chosen. You will be paid your payoff from that period in cash at the end of the experiment. When you have finished Experiment 2 you will be given instructions for a third experiment.

Quiz

(The quiz for the No-Elicitation Treatment is the same as the quiz for Experiment 1—that is, the "Choose-A-Side" Experiment.)

(Experiment begins when all questions are answered correctly.)

Experiment Three

You are about to start Experiment 3. This is the final experiment. You will repeat the experiment 20 times. Each repetition is called a "period." In each period you get a payoff of \$0, \$4, \$8 or \$12. At the end of the third experiment, 1 of the 20 periods will be chosen randomly by the computer. Each period is equally likely to be chosen. Your cash payment for the third experiment will be your payoff in the randomly chosen period. Your total payment today will include:

- Your show-up-fee of \$10.
- A cash payment for a randomly chosen period from the first computerized experiment
- A cash payment for a randomly chosen period from the second computerized experiment
- A cash payment for a randomly chosen period from the third computerized experiment

Although you will play 20 periods in this third experiment, you are paid cash for your payoff in a single period. Experiment 3 is the same as Experiment 2, except you will see **two** balls drawn from your bucket

The computer will randomly select a ball from the left-hand side of your bucket. The computer will tell you the colour of the ball, and whether your payoff is \$0 or \$4. The computer will put the ball back in the left-hand side of your bucket. The computer will randomly select a **second** ball from the left-hand side of your bucket. Because the ball is randomly chosen, this might be the same ball (chosen a second time) or it might be a new ball. The computer will tell you the colour of the second ball, and whether your payoff is \$0 or \$4.

You have two chances to get a payoff of \$4 in Stage 1. This means you can secure a payoff of \$0, \$4 or \$8 in Stage 1.

(The three treatments involve different instructions from this point.)

SBDM Treatment

You then answer 2 questions:

- Question 1: What is your belief that the two balls were drawn from Bucket A?
- Question 2: Do you want the computer to draw a third ball from the left or right side of your bucket?

The computer then tosses a coin to determine which Question is used to determine your Stage-2 payment.

- Tails: Question 1

- Heads: Question 2.

Just like Experiment 2: If you throw a Heads, your Stage-2 payoff will be determined by the "Choose-A-Side Game." The computer will draw a ball from the side of the bucket you choose. You will get a payoff of $4 if the ball is black, and $0 (nothing) if it is white. If you throw a Tails, your Stage-2 payoff will be determined by the Lottery Bag Game, or the Belief-About-Bucket-A Game (whether the two balls were drawn from Bucket A). These games are played in exactly the same way as in Experiment 2. Based on your reported belief that the ball was drawn from Bucket A, the computer will select the game that gives you the highest chance of winning $4.

You should think carefully about your belief that the balls have been drawn from Bucket A, as the computer will use your reported belief to decide whether you are paid according to your "Belief-About-Bucket A" or the Lottery game. The experiment is set up so that it is in your best interest to report your belief honestly and carefully. If you make a report that is not your true belief your payoff might be determined by the Lottery Game when you would prefer to be paid based on your belief that the ball was drawn from Bucket A.

Summary: Experiment 3

You have a 50-in-100 (50%) chance of being given Bucket A or Bucket B. You will be shown **2 balls** from the left hand side of your bucket. You will answer 2 questions:

- Question 1: What is your belief that the 2 balls were drawn from Bucket A?
- Question 2: Do you want the computer to draw a third ball from the left or right side of your bucket?

The computer then tosses a coin to determine which question is used to determine your Stage-2 payment:

- Tails: Question 1
- Heads: Question 2

If the computer throws "Heads" your payoff is determined in the same way as Experiment 1. A second ball will be drawn from your bucket, based on the side you choose. If the computer throws "Tails" your payoff will be determined by the "Belief about Bucket A" game or a Lottery Game. The best thing you can do is report your belief honestly, so that you are given the game with the highest chance of a payoff of $4. Your period payoff is your Stage 1 payoff plus your Stage 2 payoff. In each period you might get a payoff of $0 (nothing), $4, $8 or $12 in the third experiment. 1 of the 20 periods will be randomly chosen. Each period has an equal (1-in-20) chance of being chosen. You will be paid your payoff from that period in cash at the end of the experiment.

Quiz

At the start of a period the computer randomly selects a bucket for you. Both buckets are equally likely to be chosen. The bucket has 20 balls in each side: 40 in total. The computer shows you a ball from the left side of your bucket, tells you its colour, and tells you whether your payoff is \$4 or \$0. The computer puts the ball back in the left-hand side of your bucket.

1. How many balls are there in the left hand side? (20)

The computer draws a second ball from the left-hand side of your bucket and tells you whether your second payoff is \$4 or \$0.

2. Is it possible that the computer drew the same ball twice? (Yes)

The computer puts the second ball back in your bucket.

3. How many balls are there in the left hand side?(20)

The computer asks whether you want the third ball drawn from the left hand side or the right hand side of the bucket. The computer also asks you to report your belief that the ball is from Bucket A.

4. What is the minimum payoff possible in a period (both balls from Stage 1 + ball from Stage 2)? (0)
5. What is the maximum payoff possible in a period (both balls from Stage 1 + ball from Stage 2)? (12)

(Experiment begins when all questions are answered correctly.)

Unpaid Introspection Treatment

After seeing the colour of the **two balls**, you need to think about the chance that the balls were drawn from Bucket A. This is your "belief" that the balls were drawn from Bucket A. Your "belief" is a number between 0 and 100 to indicate the chance-in-100 that the ball has been drawn from Bucket A. You should think carefully about your belief that the ball has been drawn from Bucket A. You then answer 2 questions:

- Question 1: What is your belief that the ball was drawn from Bucket A?
- Question 2: Do you want the computer to draw a second ball from the left or right side of your bucket?

The computer randomly selects a ball from the side you choose. If it is black, you receive \$4. If it is white, you receive \$0 (nothing). This is your Stage-2 payoff. Your payoff for the period is your Stage-1 payoff plus your Stage-2 payoff. In total you might have a payoff of $0, $4, $8 or $12 across both stages of the third experiment. In each period there is a 50-in-100 (50%) chance of being given Bucket A or Bucket B. Your bucket is randomly determined by the computer and is not affected by the bucket you have been given in previous periods

Summary: Experiment 3

You have a 50-in-100 (50%) chance of being given Bucket A or Bucket B. You will be shown **2 balls** from the left hand side of your bucket. You will answer 2 questions:

- Question 1: What is your belief that the 2 balls were drawn from Bucket A?
- Question 2: Do you want the computer to draw a third ball from the left or right side of your bucket?

You should think carefully about your belief that the ball has been drawn from Bucket A. A second ball will be drawn from your bucket, from the side you choose. If it is black, your Stage-2 payoff is $4. If it is white, your payoff is $0 (nothing). Your period payoff is your Stage 1 payoff plus your Stage 2 payoff. In each period you might get a payoff of $0 (nothing), $4, $8 or $8 in the third experiment. 1 of the 20 periods will be randomly chosen. Each period has an equal (1-in-20) chance of being chosen. You will be paid your payoff from that period in cash at the end of the experiment.

Quiz

At the start of a period the computer randomly selects a bucket for you. Both buckets are equally likely to be chosen. The bucket has 20 balls in each side: 40 in total. The computer shows you a ball from the left side of your bucket, tells you its colour, and tells you whether your payoff is $4 or $0. The computer puts the ball back in the left-hand side of your bucket.

1. How many balls are there in the left hand side? (2)

The computer draws a second ball from the left-hand side of your bucket and tells you whether your second payoff is $4 or $0.

2. Is it possible that the computer drew the same ball twice? (Yes)

The computer puts the second ball back in your bucket

3. How many balls are there in the left hand side? (20)

The computer asks whether you want the third ball drawn from the left hand side or the right hand side of the bucket. The computer also asks you to report your belief that the ball is from Bucket A.

4. What is the minimum payoff possible in a period (both balls from Stage 1 + ball from Stage 2)? (0)

5. What is the maximum payoff possible in a period (both balls from Stage 1 + ball from Stage 2)? (12)

(Experiment begins when all questions are answered correctly.)

No-elicitation Treatment

You then have a third chance to draw a ball from your bucket. As before, black balls are worth \$4. White balls are worth \$0 (nothing). You must decide whether you would like the computer to draw the ball from the left hand side of your bucket, or the right hand side. The computer randomly selects a ball from the side you choose.

If it is black, you receive \$4. If it is white, you receive \$0 (nothing). This is your Stage-2 payoff.

Your payoff for the period is your Stage-1 payoff plus your Stage-2 payoff. In total you might have a payoff of \$0, \$4, \$8, or \$12 across both stages of the third experiment.

In each period there is a 50-in-100 (50%) chance of being given Bucket A or Bucket B. Your bucket is randomly determined by the computer and is not affected by the bucket you have been given in previous periods.

Summary: Experiment 3

You have a 50-in-100 (50%) chance of being given Bucket A or Bucket B. You will be shown **2 balls** from the left hand side of your bucket You will answer a question: Question: Do you want the computer to draw a third ball from the left or right side of your bucket?

A third ball will be drawn from your bucket, from the side you choose. If it is black, your Stage-2 payoff is \$4. If it is white, your payoff is \$0 (nothing).Your period payoff is your Stage 1 payoff plus your Stage 2 payoff. In each period you might get a payoff of \$0 (nothing), \$4, \$8, or \$12 in the third experiment. 1 of the 20 periods will be randomly chosen. Each period has an equal (1-in-20) chance of being chosen. You will be paid your payoff from that period in cash at the end of the experiment

Quiz

At the start of a period the computer randomly selects a bucket for you. Both buckets are equally likely to be chosen. The bucket has 20 balls in each side: 40 in total. The computer shows you a ball from the left side of your bucket, tells you its colour, and tells you whether your payoff is \$4 or \$0. The computer puts the ball back in the left-hand side of your bucket.

1. How many balls are there in the left hand side? (20)

The computer draws a second ball from the left-hand side of your bucket and tells you whether your payoff is \$4 or \$0.

2. Is it possible that the computer drew the same ball twice? (Yes)

The computer puts the second ball back in your bucket.

3. How many balls are there in the left hand side? (20)

The computer asks whether you want the third ball drawn from the left hand side or the right hand side of the bucket.

4. What is the minimum payoff possible in a period (both balls from Stage 1 + ball from Stage 2)? (0)

5. What is the maximum payoff possible in a period (both balls from Stage 1 + ball from Stage 2)? (12)

(Experiment begins when all questions are answered correctly.)

Appendix H: Cognitive Response Questionnaire

Below is the Cognitive Reflection Test we used in the order in which questions appeared. Questions 2, 7, and 10 are from Frederick (2005). Questions 5, 6, and 9 are from Primi et al. (2016), which also uses Questions 2, 7, and 10. Placebo questions are from Thomson and Oppenheimer (2016). A participants' score was based on the number of questions correctly answered from Questions 2, 5, 6, 7, 9, and 10.

1. Sara, Emma, and Sophia embark on a river trip. Each of them brings one supply item for the trip: a kayak, a box of sandwiches, and a bag of apples. Sarah brought the apples and Emma didn't bring anything edible. What did Sophia bring? [Placebo A; correct answer = a box of sandwiches]

2. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? [correct answer = 47 days; heuristic answer = 24 days]

3. A mechanic shop had five silver cars, and one blue car in the garage. During the day, three silver cars and one blue car were picked up, and one black car was dropped off. How many silver cars are left at the end of the day? [Placebo B; correct answer = 2]

4. An expedition on a mountain climbing trip will be traveling with eleven horse packs. A horse can carry one, two, or three packs. What is the minimum number of horses that the expedition needs? [Placebo C; correct answer = 4]

5. If three elves can wrap three toys in hour, how many elves are needed to wrap six toys in 2 hours? [correct answer = 3 elves; heuristic answer = 6 elves]

6. Tall members of an athletics team are three times more likely to win a medal than short members. This year the team has won 60 medals so far. How many of these medals have been won by short athletes? [correct answer = 15 medals; heuristic answer = 20 medals]
7. If it takes 5 minutes for five machines to make five widgets, how long would it take for 100 machines to make 100 widgets? [correct answer = 5 minutes; heuristic answer = 100 minutes]
8. A ship has 500 crates of oranges. At the ship's first stop, 100 crates of oranges were unloaded. At the ship's second stop, 200 more crates were unloaded. How many crates of oranges were left on the ship after the second stop? [Placebo D: correct answer = 200]
9. Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are there in the class? [correct answer= 29 students; heuristic answer= 30 students]
10. A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost? [correct answer = 5 cents; heuristic answer = 10 cents]