Online Appendix

[Not intended for publication]

"We Should *Totally* Open a Restaurant"

In Appendix Appendix A we show that our main results hold subject to a number of robustness checks (Appendix A.1); remark on how netting out Baseline Beliefs helps to dispel concerns about individual-level unobserved heterogeneity (Appendix A.2); provide additional description of data generated by the experiment (Appendix A.3); and offer a formal conceptual framework for our definitions of optimism and overconfidence (Appendix A.4). In Appendix B, we we provide the full set of instructions read to subjects and screen shots from the experimental interface.

Appendix A Additional Results

Appendix A.1 Robustness of Main Results

In this section, we assess the robustness of our main results. First, we show that our main results are robust to order effects. Second, we show that our main results hold when we measure overconfidence as miscalibration. Third, we show that we can re-cast treatment effects as shifts in the expected number of white balls. Fourth, we include baseline beliefs as a control (rather than subtracting them from beliefs under other treatments). Fifth, we examine the problem of selection by averaging treatment effects within individuals and across distributions.

Appendix A.1.1 Robustness to Order

As the experiment was a within-subject design, we altered the order in which individual subjects faced the experimental treatments and created 5 different orders in which they could be administered. The first and most natural order for the treatments was: (1) Baseline Treatment, Payment Treatment, Performance Treatment, Combined Treatment. This allowed for each treatment to build off of the previous treatment. Of our 125 subjects, 68 subjects followed this order.

Our main concern was comprehension and thus we made a decision to keep the Baseline Treatment first in 4 of the 5 orders. We allowed each other treatment to be performed second, third and fourth in at least one order combination, while keeping the Baseline Treatment first. Thus, the next three orders were: (2) Performance, Payment, Combined (8 subjects); (3) Performance, Combined, Payment (10 subjects); and (4) Combined, Payment, Performance

(9 subjects). As a final robustness check, we conducted the experiment in the opposite order: (5) Combined, Performance, Payment, Baseline (30 subjects).

In Table S1 we show the correlation between overconfidence and optimism by order. In Col [1] we report the correlation for the first Order, in column [2] we pool orders (2), (3) and (4) together due to smaller sample sizes and show that the correlation is not statistically different from column [1] (p-value=.34).[1] Col [3] shows that the correlation between optimism and overconfidence is not statistically significant when the Baseline Treatment is performed last. We did find that subjects were significantly under-confident in Order (5), despite answering more IQ questions correctly. We found no significant differences in average optimism.

In Table S2 we show how optimism and overconfidence affects over-estimation in the Combined Treatment. Again Col [1] is restricted to subjects from Order (1), while Col [2] is restricted to subjects in Order (2), (3), and (4) and Col [4] is restricted to subjects in Order [5]. The results in each column are qualitatively equivalent to the main results in Table 5, although the estimate in Col [2] is less precisely measured, resulting in a p-value just greater than conventional levels of significance (p-value=.16).

Appendix A.1.2 Measuring Overconfidence as Miscalibration

Next, we show that our main results are robust if we measure overconfidence as miscalibration. This is in part motivated by the fact that our main definition of overconfidence is identified using individuals who answered an IQ question incorrectly. Recasting overconfidence as miscalibration relaxes this requirement. Within the context of our IQ task, Lichtenstein, Fischhoff, and Phillips (1977) define an individual as well-calibrated if their beliefs about the number of correctly IQ questions are on average correct, without requiring the individual to know exactly which of the IQ questions he answered correctly.

Our within-subject design means that we never directly ask subjects about their own assessment of their performance on the IQ questions. However, we can impute the subject's belief that he gave a correct answer, $\hat{p}$, in the single-draw distributions from equation S1, where $z_{i,d1,baseline}$ and $z_{i,d2,baseline}$ are the Baseline Beliefs in distribution 1 and distribution 2 respectively, and let $z_{i,d,performance}$ be his elicited belief in the Performance Treatment.[2]

[1] Because we cluster at the subject-level, restricting the regression only to orders 2, 3 or 4 results in sample sizes that are too small to cluster without additional adjustments. The results are qualitatively equivalent if we estimate the correlation on each of the orders separately.

[2] Suppose the subject starts by facing Distribution 1. If he answers the IQ question correctly, he moves to Distribution 2. If not, he remains in Distribution 1. From the Baseline Treatment, we know his belief about the probability of 1 white ball drawn from Distribution 1 and Distribution 2, which we denote $z^{d1}_{i,d1,baseline}$ and $z^{d2}_{baseline}$, respectively. Let $\hat{p}$ be the probability the subject places on having answered the IQ question

$$z_{i,d,performance} = \hat{p} \times z_{i,d2,baseline} + (1 - \hat{p}) \times z_{i,d1,baseline}. \tag{S1}$$

Figures S1a and S1b show the distribution of $\hat{p}$ in the single-draw distribution. Figure S1a shows that subjects who answered incorrectly have a mean belief of .53, but that the distribution is highly bimodal- the majority of subjects believe with certainty that their answer was correct ($\hat{p} = 1$) or incorrect ($\hat{p} = 0$). On the other hand, Figure S1b shows that among subjects who answered correctly, the mean belief is .81 while the majority of subjects believe they answered correctly ($\hat{p} = 1$).

Under our current definition, if a subject's probabilistic reports imply $\hat{p} = 0.55$ and, in reality, his answer was incorrect, then our measure classifies him as being overconfident. However, if the subject answers, on average, 55% of the IQ questions correctly, then $\hat{p} = 0.55$ is well-calibrated (Lichtenstein, Fischhoff, and Phillips, 1977). To measure calibration, we subtract the subject's overall proportion of correct IQ answers, q, from the imputed probabilistic report of a correct IQ answer, $\hat{p}$. If $\hat{p} > \; (<) \; q$ then the subject is over-calibrated (under-calibrated). Figure S2 shows the distribution of the calibration measure, which has a mean 0.10 and standard deviation 0.42.

Next, we replicate equation 2, but use the calibration measure as our measure of overconfidence. We find a significant positive correlation between miscalibration and optimism: subjects who are over-calibrated (under-calibrated) are also more likely to be optimistic (pessimistic).

Finally, in columns [2] and [3] we replicate regressions using equations (2) and (3) using our original measures of overconfidence and optimism, but we restrict the sample to those subjects who are significantly mis-calibrated. In particular, we restrict our sample to those subjects for whom $\mid \hat{p} - q \mid > 0.1$, although our results are robust to various thresholds. Results are reported in Table S3 and show qualitatively similar relationships to results in Tables 4 and 5.

Appendix A.1.3 Overconfidence and Optimism as the Expected Number of White Balls

Another alternative measure of optimism and overconfidence would amount to calculating shifts in the expected value, i.e., the expected number of white balls for a given distribution. To compute the average difference between the treatment belief and the baseline belief for each treatment and distribution, we would first use elicited beliefs in the Baseline Treatment

correctly and let $z_{performance}$ be his elicited belief in the performance-only treatment about the probability that 1 white ball is drawn from the jar. Then the only unknown is $\hat{p}$, which can be found using equation S1.

to compute the expected number of white balls. We would compare this to the expected number of white balls in the Payment, Performance and Combined Treatments. The problem with this approach is that it artificially over-weights treatment effects in three-draw distributions relative to single-draw distributions. One way to around this problem would be to divide the expected value of white balls in the three-draw distribution by three.

In what follows, we show that the variable $\overline{shift}$, the main variable used in our analysis, is equivalent to computing changes in the expected number of white balls and then dividing by three for complex distributions. In other words, the main variable used in our analysis can be interpreted as measuring individual-level shifts in the expected value attributable to experimental treatments, but suitably weighted to account for larger possible numbers of white balls in some distributions, which would artificially overweight their importance in our results.[3]

For a given complex distribution, suppose we elicit the following marginal probabilities (where, without loss of generality, we focus on shifts induced by the Payment Treatment).

$$\begin{aligned} \alpha_j^B &= \text{P[draw j white balls|treatment=Baseline]}, j = \{0,1,2,3\} \\ \alpha_j^M &= \text{P[draw j white balls|treatment=Payment]}, j = \{0,1,2,3\} \end{aligned} \tag{S2}$$

First, derive the expected shift. Define expected value for each treatment as:

$$\begin{aligned} EV^B &= [\alpha_0^B \cdot 0 + \alpha_1^B \cdot 1 + \alpha_2^B \cdot 2 + \alpha_3^B \cdot 3] \\ EV^M &= [\alpha_0^M \cdot 0 + \alpha_1^M \cdot 1 + \alpha_2^M \cdot 2 + \alpha_3^M \cdot 3] \end{aligned} \tag{S3}$$

Define the shift as:

$$\begin{aligned} EV^{shift} &= [(\alpha_0^M - \alpha_0^B) \cdot 0 + (\alpha_1^M - \alpha_1^B) \cdot 1 + (\alpha_2^M - \alpha_2^B) \cdot 2 + (\alpha_3^M - \alpha_3^B) \cdot 3] \\ &= [(\alpha_1^M - \alpha_1^B) + (\alpha_2^M - \alpha_2^B) \cdot 2 + (\alpha_3^M - \alpha_3^B) \cdot 3] \end{aligned} \tag{S4}$$

Next, we use the same notation to express $\overline{shift}$. Recall, $\overline{shift}$ takes the average of the difference of cumulative probabilities, where Payment, Performance or Combined cumulative probabilities are subtracted from the Baseline Treatment probabilities. It can be written:

$$\begin{aligned} \overline{shift} &= \tfrac{1}{3}[\alpha_0^B - \alpha_0^M + (\alpha_0^B + \alpha_1^B) - (\alpha_0^M + \alpha_1^M) + (\alpha_0^B + \alpha_1^B + \alpha_2^B) - (\alpha_0^M + \alpha_1^M + \alpha_2^M)] \\ &= \tfrac{1}{3}[3(\alpha_0^B - \alpha_0^M) + 2(\alpha_1^B - \alpha_1^M) + (\alpha_2^B - \alpha_2^M)] \end{aligned} \tag{S5}$$

[3] In results available from the authors, we analyze shifts in the expected value, thus allowing complex distribution shifts to be overweighted in the analysis. All results go through with this alternative measure.

where, by definition, $\alpha_0^M + \alpha_1^M + \alpha_2^M + \alpha_3^M = \alpha_0^B + \alpha_1^B + \alpha_2^B + \alpha_3^B = 1$. Next, replace

$$\begin{aligned}\alpha_0^B &= 1 - \alpha_0^B - \alpha_1^B - \alpha_2^B - \alpha_3^B \\ \alpha_0^M &= 1 - \alpha_0^M - \alpha_1^M - \alpha_2^M - \alpha_3^M\end{aligned} \tag{S6}$$

We obtain:

$$\begin{aligned}\overline{shift} &= \tfrac{1}{3}[3\alpha_1^M + 3\alpha_2^M + 3\alpha_3^M - 3\alpha_1^B - 3\alpha_2^B - 3\alpha_3^B + 2(\alpha_1^B - \alpha_1^M) + (\alpha_2^B - \alpha_2^M)] \\ &= \tfrac{1}{3}[\alpha_1^M + 2\alpha_2^M + 3\alpha_3^M - \alpha_1^B + 2\alpha_2^B + 3\alpha_3^B] \\ &= \tfrac{1}{3}[(\alpha_1^M - \alpha_1^B) + (\alpha_2^M - \alpha_2^B)\cdot 2 + (\alpha_3^M - \alpha_3^B)\cdot 3] \\ &= \tfrac{1}{3}EV^{shift}.\end{aligned} \tag{S7}$$

Appendix A.1.4 Alternative Specification of Optimism and Overconfidence

Using equation 1 to relate optimism and overconfidence is possibly over-restrictive by requiring a proportional relationship between the Baseline beliefs and the other righthand-side treatment beliefs. An alternative, more flexible, approach is to simply include Baseline beliefs ($z_{i,d,m,Baseline}$) as additional regressor to explain beliefs in the Payment, Performance and Combined Treatments. Using this approach, we compare treatment responses in the Payment and Performance Treatments and then assess how beliefs in the Combined Treatment relate to responses in the Payment and Performance Treatments. We estimate

$$z_{i,d,performance} = z_{i,d,payment}\phi_1^{A1} + z_{i,d,Baseline}\phi_2^{A1} + X_{i,d}\beta_1^{A1} + \epsilon_{i,d,\tau}^{A1} \tag{S8}$$

and

$$\begin{aligned}z_{i,d,combined} &= z_{i,d,payment}\phi_3^{A1} + z_{i,d,performance}\phi_4^{A1} \\ &+ z_{i,d,Baseline}\phi_5^{A1} + X_{i,d}\beta_2^{A1} + \eta_{i,d,\tau}^{A1}\end{aligned} \tag{S9}$$

Estimating equations (S8) and (S9) is comparable to estimating equations (2) and (3), respectively, where the goal of the former is to assess within-individual correlation between optimism and overconfidence and the goal of the latter is to assess how optimism and overconfidence jointly drive beliefs when individuals are potentially prone to both biases. To see the difference, notice that Equations (2) and (3) are equivalent to equations (S8) and (S9), but with the restriction that $\phi_1^{A1} = 1$ and $\phi_3^{A1} = 1$. Results using these less restrictive specifications are presented in Tables S4 and S5 and are qualitatively similar to our main results.[4]

[4] In results available from the corresponding author, we include the average baseline belief for each distribution as a control instead of each individual's baseline belief for that distribution and then re-estimate equation (S9). Estimates of coefficients on responses in the Performance and Payment treatments remain qualitatively similar.

In Tables S4 and S5, the unit of observation changes from the individual-distribution level to the individual-distribution-draw level since we are no longer taking an average difference between the Baseline belief and the other treatment beliefs, resulting in 1,849 observations at the individual-distribution-draw level. When comparing optimism and overconfidence in Table S4, we rely on observations where the same individual and distribution are observed under both the Payment and the Performance Treatments, which occurs for 870 observations. In column [5] of Table S4, we reduce the sample to the 527 observations also used in Table S5 to study how optimism and overconfidence predict beliefs in the Combined Treatment. These are the 527 individual-distribution-draw observations where the same individual and distribution are observed under the Payment, Performance and Combined Treatments.

Appendix A.1.5 Leveraging Multiple Observations for Each Individual

Next, we describe a robustness test that exploits how each individual is observed multiple times in each treatment since we elicit beliefs about multiple distributions. That is, we have multiple measures of the variable $\overline{shift}$ from equation (1) for each individual in our sample. Recall that due to our within-individual design, our main results are estimated using a sample of observations where the same individual faces the same distribution under multiple experimental treatments. This does not occur in all cases since the distribution subjects face in the Performance and Combined Treatments is a function of the number of correctly answered IQ questions.

One concern is that this is a selected subsample and that estimates are therefore biased. To assess robustness of our results using a larger sample of subjects, we average the $\overline{shift}$ variable for each individual in the single-draw distributions and also in the three-draw distributions. This results in two observations per individual in each treatment. In the three-draw distributions, two observations are dropped due to non-monotonicity in the Baseline Beliefs. This results in 125 Optimistic, Overconfident and Combined Treatment shifts for the single-draw distributions and 123 Optimistic, Overconfident and Combined Treatment Shifts for the three-draw distributions. Using these alternative shift variables, we replicate our main specifications from equations (2) and (3) and obtain qualitatively similar results, which are presented in Table S6. This helps to dispel concerns that our main results are biased since they are estimated on a relatively small sub-sample of individuals for whom we observe shifts for the same distribution across all treatments.[5]

[5] In a related robustness check, we address the possibility that our results are driven by measurement error. In a recent paper, Gillen, Snowberg, and Yariv (2017) show that measurement error can bias main experimental effects downward (attenuation bias) so that "new" secondary or interaction effects are inflated. To assess robustness or our results to measurement error, we again make use of the fact that we have multiple

Appendix A.2 Individual Heterogeneity

In this section we demonstate the importance of using the baseline belief as an individual-level control. First, we show that the subtraction of the baseline belief from the treatment belief helps reduce the impact of utility curvature on the elicited belief. Second, we show that there are systematic deviations in the Baseline belief from the Objective distribution based on IQ and gender. Thus, we would mis-measure the impact of our treatments on beliefs if we relied on the Objective distribution rather than the Baseline beliefs.

Let z_i be the elicited belief and x_i be the truth for $i = B$ and T, baseline and treatment respectively. Under the QSR, the curvature of the utility function breaks the incentive compatibility of the mechanism.

$$z_T = E[x_T] + \underbrace{\frac{cov(u'(z_T - x_T), z_T - x_T)}{E[u'(z_T - x_T)]}}_{\text{bias induced by curvature}}$$

However, our variables of interest throughout our analysis are $z_T - z_B$, which is given by

$$z_T - z_B = E[x_T] - E[x_B] + \frac{cov(u'(z_T - x_T), z_T - x_T)}{E[u'(z_T - x_T)]} - \frac{cov(u'(z_B - x_B), z_B - x_B)}{E[u'(z_B - x_B)]}$$

resulting in a bias reduction.

Appendix A.3 Additional Description of Experimental Data

In Section 3 of the main text, we describe the data generated by the experimented we conducted. Here, we provide some additional description and preliminary analyses of the data. In Table S7, we show the number of observations used to calculate shifts and within-subject relationships between shifts, both for the full sample and then separately for men and women and again separately for individuals who correctly answer above and below the median number of correctly answered IQ questions. In Table S8, we show average deviations from the objective. In Table S9 we show the average treatment effects for the full sample. Next, in Table S10 we show average treatment effects for the sample of subject-distribution pairs used for our main results (in Table 5), that is, for the subjects who report beliefs about the same distribution in each of the four treatments. We find that the average treatment effects are qualitatively equivalent to those in the full sample, shown in Table S9.

measures of $\overline{shift}$ for each individual in each treatment. For each treatment, we use multiple observed shifts per individual to identify underlying factors that are purged of measurement error. In results available from the authors, we show that our main results are robust to this type of factor analysis.

In Figure S3, we plot Baseline Treatment beliefs against Optimism (Beliefs in the Payment Treatment net of Baseline Treatment beliefs) separately for each of the six distributions. In Figure S4, we plot Baseline Treatment beliefs against Overconfidence (Beliefs in the Performance Treatment net of Baseline Treatment beliefs) for each distribution. These figures show that both Optimism and Overconfidence are independent of the Baseline Belief. If "low" Baseline Beliefs (i.e., under-estimation of high probability events) were inducing Optimism or Overconfidence, then we would expect a negative relationship between Optimism/Overconfidence and the Baseline Beliefs.

In Figure S5, we plot beliefs in the Payment Treatment against beliefs in the Performance Treatment for each distribution. The figure suggests a positive relationship. We think it is important to point out that our main result, the positive correlation between Optimism and Overconfidence, is a correlation between the errors in beliefs, whereas Figure S5 suggests there is also a correlation between the beliefs themselves. The exception is Distribution 1, where the positive correlation is less apparent, largely because optimism occurs rarely, as illustrated by the fact that most observations lie on the vertical line that crosses the x-axis at 0.5.

Appendix A.4 Conceptual Framework

In this section, we formally define what we mean by optimism and overconfidence. The agent in our setting may face two dimensions of uncertainty: uncertainty about his performance represented by a random type $\theta \in \Theta$ and uncertainty about the outcome of a lottery X with discrete integer realizations $x = 0, 1, ..., M$. The outcome of the lottery, x, may depend on the agent's performance type and is described by a distribution, which we denote:

$$G_X(x|\theta) = P(X \leq x|\theta). \tag{S10}$$

We assume that the measure of performance and the outcome of the lottery are ordered and discrete. The link between performance type θ and the distribution $G_X(x|\theta)$ is given by the following: if $\theta' > \theta''$ then $G_X(x|\theta = \theta')$ first order stochastically dominates $G_X(x|\theta = \theta'')$. In other words, better performance is associated with a "better" distribution, in the sense of stochastic dominance. The agent has a true performance type, $\theta_o \in \theta$, and the distribution Θ is degenerate at θ_o, but θ_o is unknown to the agent. The agent has beliefs about his true performance type, θ_o, given by the distribution $\tilde{\Theta}$ over θ.

There is also a non-decreasing map from lottery outcomes to monetary payoffs, $m(\cdot)$, meaning that higher values of x have weakly larger monetary payoffs. Moreover, we assume that the agent's preferences can be described by a utility function, such that if $x' > x''$ then $u(m(x')) > u(m(x''))$. The agent has subjective beliefs about G that are given by

$$F\left(x|m(\cdot);\tilde{\Theta}\right), \tag{S11}$$

where we explicitly condition on $m(\cdot)$ to capture that an agent's beliefs may be affected by the payoff function.

Next, we will characterize beliefs in four settings. First, consider the case where there is no performance uncertainty and the monetary payoff is constant (does not depend on x), thus we write $m(x) = m$. Then the agent has subjective beliefs given by

$$F\left(x|m(x) = m;\tilde{\Theta}\right) = F(x), \tag{S12}$$

where we do not condition on $m(\cdot)$ since the payoff is independent of x and we do not condition on $\tilde{\Theta}$ since there is no performance uncertainty and thus no uncertainty regarding the distribution the agent faces. Call this set of beliefs the agent's Baseline Beliefs.

Second, if there is no performance uncertainty and the monetary payoff is weakly increasing in x, then the agent has subjective beliefs given by

$$F\left(x|m(\cdot);\tilde{\Theta}\right) = F(x|m(\cdot)), \tag{S13}$$

where we explicitly condition on $m(\cdot)$ since the payoff changes with x and might therefore influence beliefs about x. Call this set of beliefs the agent's Payment Treatment Beliefs.

For the third and fourth beliefs, we drop the assumption that there is a single performance type and allow for subjective uncertainty regarding performance type. Recall that θ_o denotes the agent's true type, which is unknown to the agent. $\tilde{\Theta}$ denotes the agent's beliefs about the degenerate distribution over θ, Θ. The agent is tasked with forming beliefs over:

$$G(x|\theta_o) \tag{S14}$$

In this case we allow G to change with performance as described above: if $\theta' > \theta''$ then $G(x|\theta = \theta')$ stochastically dominates $G(x|\theta = \theta'')$. Again, when an agent performs better, he faces a better distribution in the sense of stochastic dominance. The distribution is therefore endogenously determined by the agent's level of performance. The third setting for which we characterize beliefs is when the agent faces performance uncertainty and a monetary payoff function that is constant. His beliefs are given by

$$F\left(x|m\left(x\right)=m;\tilde{\Theta}\right)=F\left(x|\tilde{\Theta}\right) \tag{S15}$$

where we do not condition on $m\left(\cdot\right)$ as payoffs are constant and do not depend on x, but we do allow F to be a function of the agent's performance beliefs $\tilde{\Theta}$. Call this set of beliefs the agent's Performance Treatment Beliefs.

Fourth, when the agent simultaneously faces performance uncertainty and a monetary payoff that is weakly increasing in x, then his beliefs are given by

$$F\left(x|m\left(\cdot\right),\tilde{\Theta}\right). \tag{S16}$$

These beliefs are called Combined Treatment Beliefs since they reflect the agent's beliefs when both performance uncertainty and a preference for larger realizations of x are present. It is important to note that in this setting, the agent can affect the lottery he faces with his performance and that larger values of x are payoff-favorable. This means that the agent increases the likelihood of a larger monetary payoff through better performance.

Definition 1. *An agent displays an optimistic shift (pessimistic shift) when*

$$\frac{1}{M}\Sigma_{x=0}^{M}[F\left(x|m\left(x\right)=m\right)-F\left(x|m\left(\cdot\right)\right)]>(<)\,0. \tag{S17}$$

Thus, the agent reports a larger expected value of X when larger outcomes are payoff-favorable relative to the expected value of X when larger outcomes are not payoff-favorable. Notice that, for each individual, optimism and pessimism are measured relative to the individual's own Baseline Beliefs. We will sometimes refer to these types of shifts simply as optimism or pessimism. Similarly,

Definition 2. *An agent displays an overconfident shift (under-confident shift) when*

$$\frac{1}{M}\Sigma_{x=0}^{M}[F\left(x|m\left(x\right)=m\right)-F\left(x|\tilde{\Theta}\right)]>(<)\,0. \tag{S18}$$

When our meaning is clear, we will refer to these shifts simply as overconfidence or under-confidence.

Definition 3. *An agent displays an overestimation shift (underestimation shift) in the Combined Treatment when*

$$\frac{1}{M}\Sigma_{x=0}^{M}[F\left(x|m\left(x\right)=m\right)-F\left(x|m\left(\cdot\right)\right);\tilde{\Theta}]>(<)\,0. \tag{S19}$$

References

Gillen, Ben, Erik Snowberg, and Leeat Yariv. 2017. "Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study." NBER Working Paper.

Lichtenstein, Sarah, Baruch Fischhoff, and Lawrence D Phillips. 1977. *Calibration of Probabilities: The State of the Art*. Springer.

Appendix A Tables and Figures

Appendix Table S1: Correlation between Optimism and Overconfidence, Order Effects

	[1]	[2]	[3]
Optimistic Shift	0.7***	0.57***	0.09
	(0.08)	(0.13)	(0.27)
Male	-0.005	0.03	0.01
	(0.01)	(0.03)	(0.03)
Correct IQ Answers	-0.002	-0.004	-0.007
	(0.004)	(0.006)	(0.008)
Constant	0.14***	0.12***	0.03
	(0.05)	(0.04)	(0.08)
Observations	204	91	88
R^2	0.55	0.46	0.08
Distribution Dummies	[Y]	[Y]	[Y]

This table shows estimates from OLS regressions, run separately by order, where the outcome variable is Baseline Beliefs minus Performance Treatment Beliefs ($\overline{shift}_{i,d,performance}$) and the explanatory variable of interest is Baseline Beliefs minus Payment Treatment Beliefs ($\overline{shift}_{i,d,payment}$), which we denote "Optimistic Shift". *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. Standard errors are clustered by individual.

Appendix Table S2: The Role of Optimism and Overconfidence, Order Effects

	[1]	[2]	[3]
Overconfident Shift	0.65***	0.26	0.34***
	(0.11)	(0.18)	(0.11)
Optimistic Shift	0.24*	0.33***	0.53
	(0.13)	(0.09)	(0.41)
Male	-0.0000208	-0.02	0.09**
	(0.02)	(0.03)	(0.04)
Correct IQ Answers	0.001	-0.00009	0.006
	(0.006)	(0.01)	(0.007)
Constant	0.05	0.08	-0.09
	(0.05)	(0.09)	(0.08)
Observations	132	56	59
R^2	0.65	0.63	0.38
Distribution Dummies	[Y]	[Y]	[Y]

This table shows estimates from OLS regressions, run separately for each order, where the outcome variable is Baseline Beliefs subtracted from the Combined Treatment Beliefs ($\overline{shift}_{i,d,combined}$) and the explanatory variables of interest are Baseline Beliefs subtracted from the Payment Treatment Beliefs ($\overline{shift}_{i,d,payment}$) which we denote "Optimistic Shift," and Baseline Beliefs subtracted from the Performance Treatment Beliefs ($\overline{shift}_{i,d,performance}$) which we denote "Overconfident Shift." *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. Standard errors are clustered by individual.

Appendix Table S3: Robustness to Miscalibration

	[1]	[2]	[3]
Optimistic Shift	0.35*	0.57***	0.26**
	(0.19)	(0.1)	(0.11)
Overconfident Shift	.	.	0.5***
			(0.11)
Male	0.05	0.04*	0.02
	(0.07)	(0.02)	(0.02)
Correct IQ Answers	-0.06***	-0.02***	-0.003
	(0.01)	(0.005)	(0.005)
Constant	0.74***	0.25***	0.06
	(0.13)	(0.05)	(0.05)
Observations	159	170	137
R^2	0.23	0.39	0.54
Independent Obs	112	122	111
Order Dummies	[Y]	[Y]	[Y]

This table shows estimates from an OLS regression where the outcome variable is Baseline Beliefs subtracted from the Combined Treatment Beliefs ($\overline{shift}_{i,d,combined}$) and the explanatory variables of interest are Baseline Beliefs subtracted from the Payment Treatment Beliefs ($\overline{shift}_{i,d,payment}$) which we denote "Optimistic Shift". The other explanatory replaces our measures of overconfidence ($\overline{shift}_{i,d,performance}$) with a variable that measures miscalibration as described. *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. Standard errors are clustered by individual.

Appendix Table S4: Robustness: Correlation between Optimism and Overconfidence

	[1]	[2]	[3]	[4]	[5]
Payment Beliefs	0.8***	0.52***	0.52***	0.41***	0.4***
	(0.03)	(0.06)	(0.06)	(0.07)	(0.07)
Baseline Beliefs	.	0.34***	0.35***	0.25***	0.21***
		(0.05)	(0.05)	(0.06)	(0.06)
Male	.	.	-0.006	-0.006	-0.02*
			(0.01)	(0.01)	(0.01)
Correct IQ Answers	.	.	0.003	-0.002	-0.001
			(0.003)	(0.003)	(0.003)
Observations	870	870	870	870	527
R^2	0.66	0.71	0.71	0.74	0.76
Independent Obs	125	125	125	125	121
Order Dummies	[Y]	[Y]	[Y]	[Y]	[Y]
Distribution Dummies	[N]	[N]	[N]	[Y]	[Y]
Realization Dummies	[N]	[N]	[N]	[Y]	[Y]
Smaller Sample	[N]	[N]	[N]	[N]	[Y]

This table shows estimates from an OLS regression where the outcome variable is Performance Treatment Beliefs ($z_{i,d,performance}$) and the explanatory variables of interest are Payment Treatment Beliefs ($z_{i,d,payment}$) and Baseline Beliefs ($z_{i,d,baseline}$) . *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. Standard errors are clustered by individual.

Appendix Table S5: Robustness: The Role of Optimism and Overconfidence

	[1]	[2]	[3]	[4]	[5]
Performance Beliefs	0.6***	.	0.46***	0.45***	0.35***
	(0.06)		(0.07)	(0.07)	(0.07)
Payment Beliefs	.	0.53***	0.28***	0.28***	0.24***
		(0.08)	(0.08)	(0.08)	(0.08)
Baseline Beliefs	0.26***	0.29***	0.14**	0.14**	0.09
	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)
Male	.	-0.04**	.	-0.03	-0.03
		(0.02)		(0.02)	(0.02)
Correct IQ Answers	.	.	.	0.0006	-0.003
				(0.004)	(0.004)
Constant	0.06***	0.09***	0.05***	0.06	0.15***
	(0.02)	(0.02)	(0.02)	(0.04)	(0.06)
Observations	527	527	527	527	527
R^2	0.68	0.65	0.7	0.7	0.73
Independent Obs	121	121	121	121	121
Order Dummies	[Y]	[Y]	[Y]	[Y]	[Y]
Distribution Dummies	[N]	[N]	[N]	[N]	[Y]
Realization Dummies	[N]	[N]	[N]	[N]	[Y]

This table shows estimates from an OLS regression where the outcome variable is Combined Treatment Beliefs ($z_{i,d,combined}$). Robust standard errors clustered at the subject-level in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Appendix Table S6: Treatment Effects and Main Results, Averaged over Distribution Class

	[1]	[2]	[3]	[4]	[5]
Average Optimistic Shift	0.008	0.04***	0.12***	0.58***	0.55***
	(0.006)	(0.01)	(0.03)	(0.15)	(0.12)
Average Overconfident Shift	0.03***	0.06***	0.14***	.	0.38***
	(0.009)	(0.01)	(0.03)		(0.08)
MN	0.02**	0.05***	0.13***	.	.
	(0.009)	(0.01)	(0.03)		
BNshift-ave	.	.	.	.	0.38***
					(0.08)
MXshift-ave	.	.	.	0.58***	0.55***
				(0.15)	(0.12)
Three-draw Distribution	.	-0.04***	-0.04***	-0.03**	-0.003
		(0.01)	(0.01)	(0.02)	(0.02)
Male	.	.	0.005	0.005	0.003
			(0.01)	(0.01)	(0.01)
Correct IQ Answers	.	.	-0.009***	-0.01***	-0.003
			(0.002)	(0.004)	(0.003)
Observations	747	747	747	248	248
R^2	0.02	0.07	0.1	0.28	0.39
Independent Obs	125	125	125	125	125
Order Dummies	[N]	[Y]	[Y]	[Y]	[Y]
Distribution Dummies	[N]	[N]	[Y]	[Y]	[Y]

Col (1) - (3) regress the $\overline{shift}$ variable on the three treatment dummies. In Col (4), the outcome variable is the average Overconfidence shift and in Col (5) the outcome variable is the average Combined Shift. Robust standard errors are clustered at the subject-level in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. Standard errors are clustered by individual.

Appendix Table S7: Sample Size and Characteristics for Within Subject Shifts

	Total	Male	Female	IQ Low	IQ High
Subjects	125	59	66	61	64
Observations (Treatment-Induced Shifts)					
Total	1,518	730	788	746	772
Payment	738	352	386	360	378
Performance	385	184	201	191	194
Combined	395	194	201	195	200
Sample for Main Analysis					
Payment and Performance	**383**	184	199	189	194
Combined, Payment and Performance	**247**	129	118	118	129

The sample sizes for the main analysis are determined by the number of subject-distribution pairs that are observed for the relevant treatments. For the analysis in Table 4 in the main text, we only require that subject-distribution pairs be present in the Payment and Performance Treatments, resulting in 383 observations. For the analysis in Table 5 in the main text, we apply an additional restriction and require subject-distribution pairs to also be observed in the Combined Treatment, resulting in 247 observations.

Appendix Table S8: Average Deviation from the Objective Distribution

	[1]	[2]	[3]	[4]
Correct IQ Answers	-0.007***	-0.007***	-0.006***	-0.006***
	(0.002)	(0.002)	(0.002)	(0.002)
Three-Draw Distribution	.	0.09***	0.09***	0.09***
		(0.009)	(0.009)	(0.009)
Male	.	.	-0.02*	-0.02**
			(0.009)	(0.009)
Observations	726	726	726	726
R^2	0.03	0.19	0.19	0.2
Independent Obs	125	125	125	125
Order Dummies	[N]	[N]	[N]	[Y]

This table shows estimates from OLS regressions where the outcome variable is the absolute difference between the Objective distribution and the Baseline Treatment distribution. *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. Standard errors are clustered by individual.

Appendix Table S9: Average Treatment Effects: Full Sample

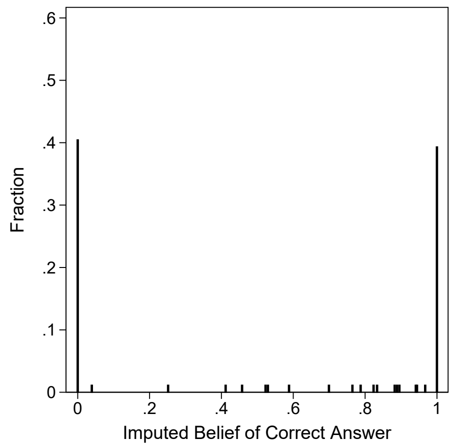
	[1]	[2]	[3]	[4]	[5]	[6]
Payment Treatment	0.003	0.03^{***}	0.04^{***}	0.04^{***}	0.11^{***}	0.07^{***}
	(0.006)	(0.009)	(0.01)	(0.01)	(0.02)	(0.03)
Performance Treatment	0.03^{***}	0.05^{***}	0.06^{***}	0.06^{***}	0.13^{***}	0.15^{***}
	(0.008)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)
Combined Treatment	0.02^{***}	0.04^{***}	0.06^{***}	0.06^{***}	0.12^{***}	0.15^{***}
	(0.009)	(0.01)	(0.01)	(0.01)	(0.03)	(0.04)
Three-draw Distribution	.	-0.04^{***}	-0.04^{***}	-0.04^{***}	-0.04^{***}	-0.04^{***}
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Male	.	.	.	-0.009	0.0005	0.0007
				(0.01)	(0.01)	(0.01)
Correct IQ Answers	.	.	.	.	-0.007^{***}	.
					(0.002)	
IQ × Pay.	.	.	.	.	.	-0.004^{*}
						(0.002)
IQ × Perf.	.	.	.	.	.	-0.01^{***}
						(0.003)
IQ × Combined	.	.	.	.	.	-0.01^{***}
						(0.003)
Observations	1518	1518	1518	1518	1518	1518
R^2	0.02	0.03	0.04	0.04	0.05	0.06
Independent Obs	125	125	125	125	125	125
Order Dummies	[N]	[Y]	[Y]	[Y]	[Y]	[Y]
Distribution Dummies	[N]	[N]	[Y]	[Y]	[Y]	[Y]

This table shows estimates from OLS regressions where the outcome variable is $\overline{shift}$, which we have pooled across the Payment, Performance and Combined Treatments. *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. Standard errors are clustered by individual.

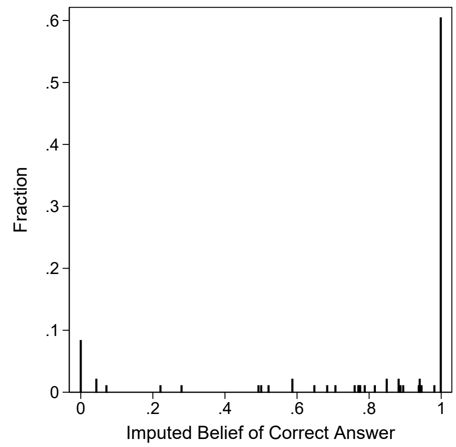
Appendix Table S10: Average Treatment Effects: Smaller Sample

	[1]	[2]	[3]	[4]	[5]	[6]
Payment Treat. Beliefs	0.009	0.03**	0.09***	0.08***	0.11***	0.08*
	(0.009)	(0.01)	(0.02)	(0.02)	(0.04)	(0.04)
Performance Treat. Beliefs	0.02**	0.05***	0.1***	0.09***	0.13***	0.16***
	(0.01)	(0.01)	(0.02)	(0.02)	(0.04)	(0.04)
Combined Treat. Beliefs	0.02**	0.04***	0.1***	0.09***	0.13***	0.13***
	(0.01)	(0.01)	(0.02)	(0.02)	(0.04)	(0.05)
Male	.	.	.	0.02	0.03	0.03
				(0.02)	(0.02)	(0.02)
Correct IQ Answers	.	.	.	.	-0.005	.
					(0.003)	
IQ × Pay. Treat. Beliefs	.	.	.	.	.	-0.001
						(0.004)
IQ × Perf. Treat. Beliefs	.	.	.	.	.	-0.008**
						(0.003)
IQ × Combined Treat. Beliefs	.	.	.	.	.	-0.005
						(0.004)
Observations	741	741	741	741	741	741
R^2	0.02	0.05	0.12	0.13	0.13	0.13
Independent Obs	121	121	121	121	121	121
Order Dummies	[N]	[Y]	[Y]	[Y]	[Y]	[Y]
Distribution Dummies	[N]	[N]	[Y]	[Y]	[Y]	[Y]

This table shows estimates from an OLS regression where the outcome variable is Baseline Beliefs minus beliefs reported in the remaining three treatments, i.e., the $\overline{shift}$ variable pooled across the Payment, Performance and Combined Treatments. Here, we limit ourselves to the 741 observations where the individual is observed for the same distribution in all experimental treatments. *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. Standard errors are clustered by individual.

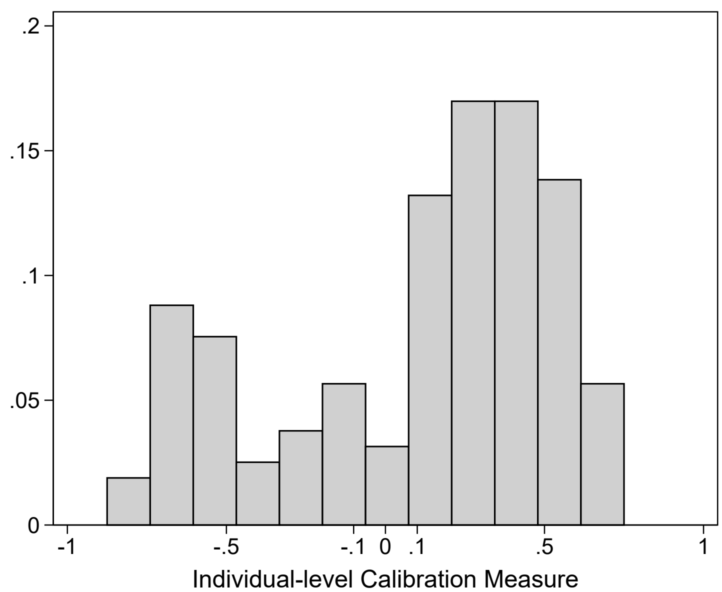


(A) Incorrect Answer

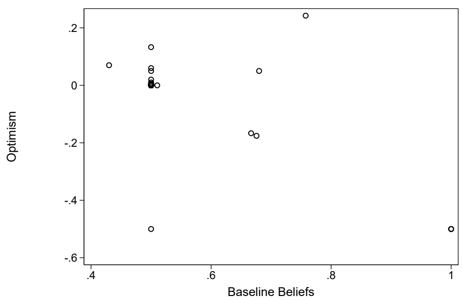


(B) Correct Answer

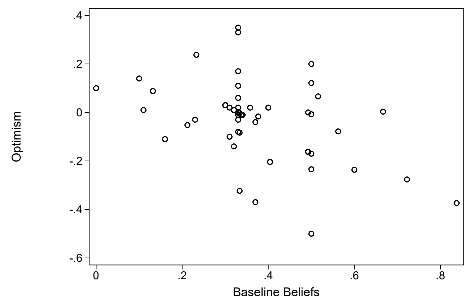
Appendix Figure S1: Imputed Belief.



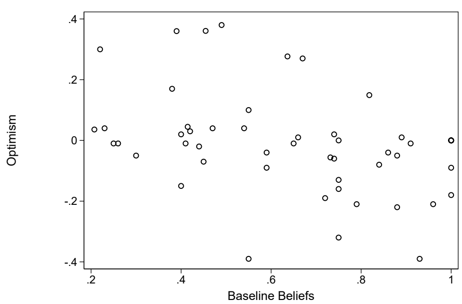
Appendix Figure S2: Calibration. This figure shows the average calibration at the subject-level, which is measured as the difference between the subject's imputed belief and the subject's proportion of correctly answered IQ questions.



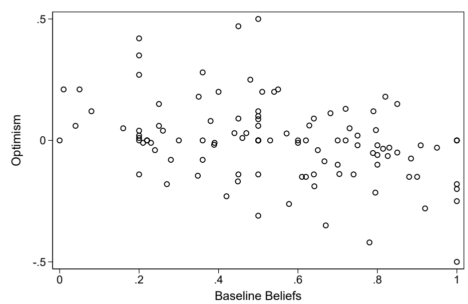
(A) Distribution 1



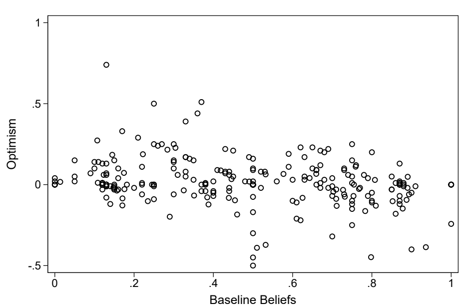
(B) Distribution 2



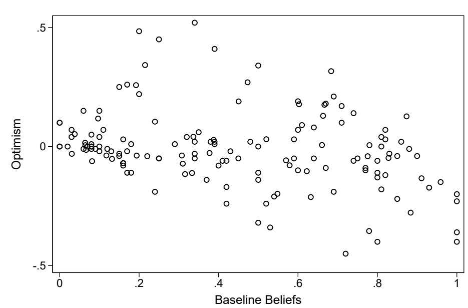
(C) Distribution 3



(D) Distribution 4

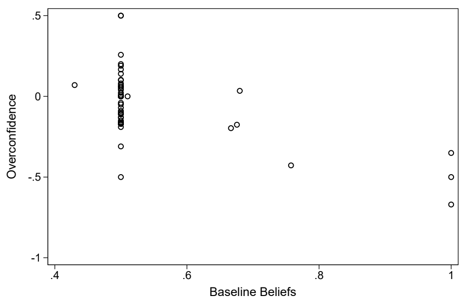


(E) Distribution 5

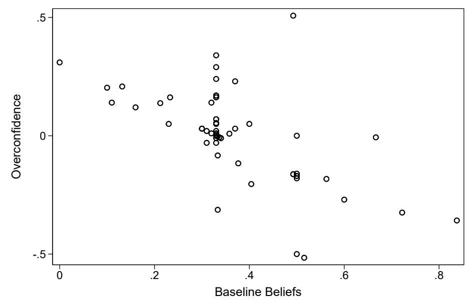


(F) Distribution 6

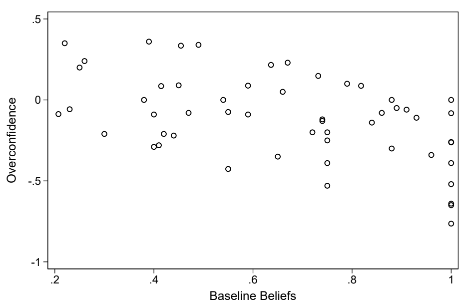
Appendix Figure S3: Baseline Beliefs and Optimism.



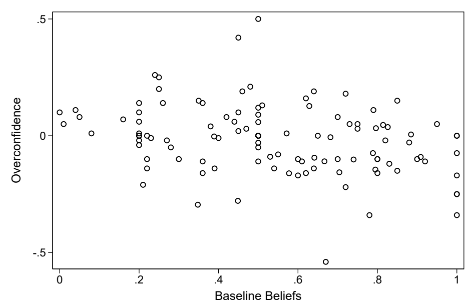
(A) Distribution 1



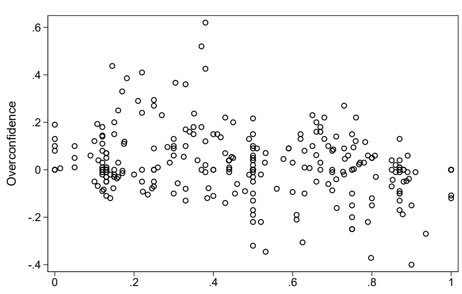
(B) Distribution 2



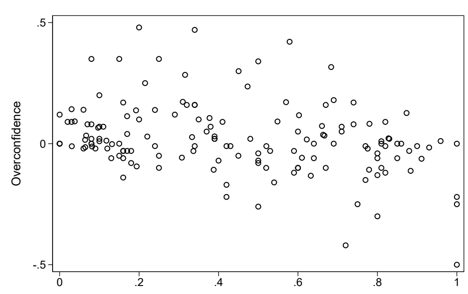
(C) Distribution 3



(D) Distribution 4

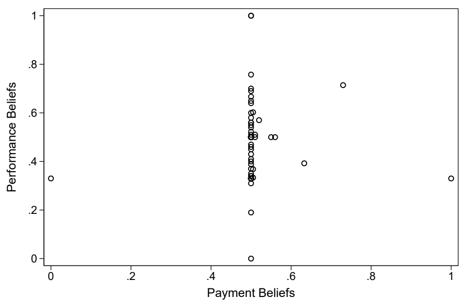


(E) Distribution 5

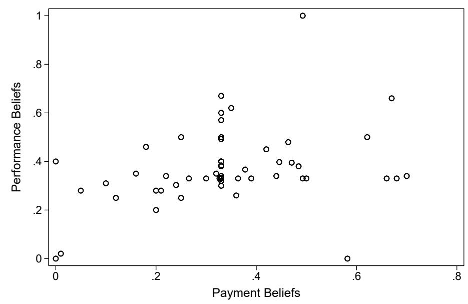


(F) Distribution 6

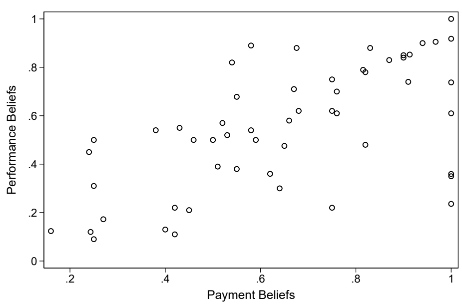
Appendix Figure S4: Baseline Beliefs and Overconfidence.



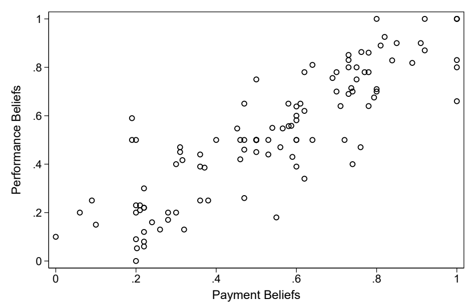
(A) Distribution 1



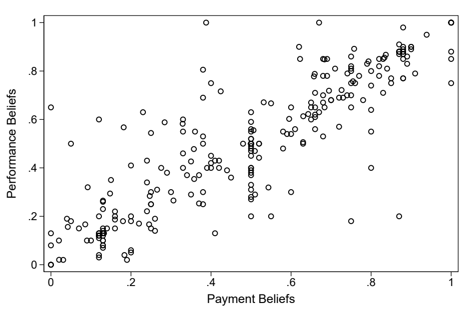
(B) Distribution 2



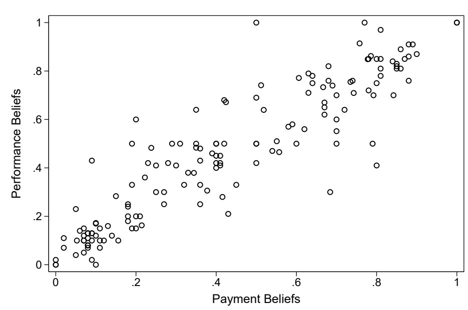
(C) Distribution 3



(D) Distribution 4



(E) Distribution 5



(F) Distribution 6

Appendix Figure S5: Optimism and Overconfidence.

Appendix B Experimental Instructions and Screenshots

Here, we provide the script that was read to participants in the experiment used to generate the data used in this paper. Before beginning a new treatment, subjects were required to correctly answer the set of review questions shown in Figure S6. During the experiment subjects were given a pay-off sheet that was meant to assist them in understanding the QSR. A copy of the pay-off sheet can be found in Figure S7. We have also included screenshots from the Z-Tree program in Figures S8-S10.

General Setting

Welcome! You will receive 5 dollars for showing up, regardless of the results in this session.

This session is scheduled for 2 hours. While we may finish before 2 hours, if you cannot participate for the full 2 hours then please raise your hand now. Please remain seated at your station throughout the session. If you have any questions please raise your hand and wait for one of us to come to your station. Otherwise, please remain silent and refrain from distracting behavior. Please put away all cell phones and store all personal items under your desk. This lab has a strict no-deception policy.

I will read the instructions aloud. Please follow along using the set of instructions at your station and the projector screen at the front of the room.

These instructions outline how the decisions you make will determine your total cash payment at the end of the session. The amount of money you earn in this session depends on your answers and on chance. You cannot lose money. At the end of the session, you will receive all earnings and the show-up fee in cash and confidentially.

The first part of the session consists of 4 Tasks. Your main objective in each Task is to accurately guess the percent chance that a WHITE ball is drawn from a jar that contains white and black balls. You will earn points based on the accuracy of your answers. You will be told the total number of each color of ball in the jar and the number of balls that the computer will randomly draw from the jar. When a ball is drawn from the jar, it is placed back in the jar before the next ball is drawn. The balls are randomly placed in the jar and the computer will randomly draw balls from the jar.

Once you enter the security code, you will see an example of the General Task. Please enter code ___.

Please use the sliding ball to indicate a number between 0 and 100 that reflects your belief of the percent chance of each event occurring.

Paper and pen are provided for your use. As you can see, the composition of the jar is given on the left side of the screen and the event in question is stated beneath the number line. The box to the right of each number line shows the percent chance indicated by the position of the white ball. The number in the box will move when you move the white ball. You will have 2 minutes to answer the questions on each screen. A timer at the top of the screen will inform you of the time remaining to submit your answer. If you do not click "OK" before two minutes have passed, then your answer will be given by the current position of the white ball.

When you are satisfied with your decisions you should click on the button "OK." Your answer is your private information and does not have any effect on later rounds.

Now I will explain how to earn points in this Task. Everyone has a "Pay-Off Sheet" that describes the number of points you will earn for each possible Reported Belief if an Event Occurs or if an Event Does Not Occur.

You start with 10 points for each question. We then subtract points depending on how close your Reported Belief is to the Outcome.

The previous question asked about your belief that ONE or FEWER WHITE balls are drawn from the jar. Suppose you report 70% (.7) and ONE OR FEWER WHITE balls are drawn. This means the Outcome equals 100% (1). And your Reported Belief is 30% (.3) away from the Outcome. If MORE than ONE WHITE balls are drawn, then the Outcome equals 0 and your reported belief is 70% (.7) away from the Outcome.

This difference from the Outcome is squared and multiplied by 10. This amount is then subtracted from your 10 points. So, in the case that you reported 70% and it turns out that ONE or FEWER WHITE BALLS are drawn then you earn $10 - 10(.3)^2 = 9.1$ points. This was determined by this equation:

$$10 - 10 * [\text{Reported Belief} - 1]^2$$

On the other hand, if MORE than ONE WHITE BALL is drawn, then you earn $10-10(.7)^2 = 5.1$ points. This was determined by this equation:

$$10 - 10 * [\text{Reported Belief} - 0]^2$$

To Summarize: With this equation you earn the most money by correctly guessing the percent chance that ONE or fewer WHITE balls are drawn. Since you do not know how many WHITE balls will be drawn from the jar, this equation ensures that you can expect to get the most points by making your best guess.

In the next few paragraphs, we provide some guidance on how to answer questions about probabilities. This guidance is not meant to influence your best guess, but will help you to avoid common inconsistencies. Suppose, for example, that you are asked the following question:

Question 1: "What do you expect is the chance of drawing ONE or FEWER WHITE balls from the jar?"

Suppose that you answer "50%". You are then asked a second question about the same jar:

Question 2: "What do you expect is the chance of drawing TWO or FEWER WHITE balls from the jar?"

Given your answer to question 1, it would not make sense to answer anything less than 50% to question 2. Why is this?

Let's examine your answer to question 1 more closely. By answering 50%, you said that you think there is a 50% chance that 0, or 1 balls are drawn from the jar. The next question asks you the chance that 0, 1, *or* 2 balls are drawn from the jar. If 0 or 1 ball is drawn from the jar, then it it is also true that 0, 1, or 2 balls are drawn from the jar. In other words, if the event of one or fewer WHITE balls occurs, then the event of two or fewer WHITE balls has also occurred. Since you answered that there is a 50% chance that 0, or 1 balls are drawn from the jar, there must be *at least* a 50% chance that 0, 1 *or* 2 balls are drawn from the jar. Are there any questions?

Now, suppose you are facing a jar that contains 1 WHITE ball and 1 BLACK ball and the computer is going to draw 4 balls from the jar. You are asked the following question:
Question 1: "What do you expect is the chance of drawing FOUR or FEWER WHITE balls from the jar?"

The answer to this question must be 100%. Why is this?

This question asks about the percent chance of FOUR or FEWER WHITE balls. We know that MORE than 4 WHITE balls cannot be drawn from the jar since only 4 balls will be drawn in total. We know that one of the following events MUST occur–0 white balls are drawn, 1 white ball drawn, 2 white balls drawn, 3 white balls drawn, or 4 white balls are drawn. This means that the percent chance of one of these events occurring is 100%. Are there any questions?

For each choice you make in the four Tasks you will earn points. Points will then be converted into dollars. 10 points are worth 1 dollar. At the end of the last Task, we will draw a number: a 1, a 2, a 3, or a 4 to decide which of the Tasks will determine your monetary payment. For example, if a 1 is drawn, you will be paid based on the number of points you earned in the Task 1. You will not know which of the Tasks will be used to determine your monetary payment until the end of the session.

Each of the 4 Tasks differs slightly from the other three Tasks. We will explain those differences before starting a new Task. Are there any questions?

Now we are going to go through a few questions to ensure that everyone has understood the General Setting. Your answers do not affect your payment. However, we will not begin the Task until all questions have been answered correctly. If you have any questions, please raise your hand. Please enter code ___.

ADMINISTRATOR WILL GIVE ANSWERS.

Task 1

We will now discuss Task 1. In Task 1, your objective is to accurately guess the percent chance that a certain color of ball is drawn from a jar. You will only be paid for the accuracy of your answers as described during the first set of instructions. In fact, Task 1 does not differ in any way from the first set of instructions given.

Once you enter the security code, you will see an example of the Task 1. Please enter code ___.

Please use the sliding ball to indicate a number between 0 and 100 that reflects your belief of the percent chance of each event occurring.

As you can see, the composition of the jar is given is on the left side of the screen and the event in question is stated beneath the number line. The box to the right of each number line indicates the percent chance indicated by the position of the white ball.

When you are satisfied with your decisions you should click on the button "OK." Your answer is your private information and does not have any effect on later rounds.

We will now begin Task 1. Whenever you have finished reading a screen, please click on the button "OK" to advance to the next screen. Please enter code ___.

Task 2

Now we will move on to Task 2. In Task 2, you will answer questions about the color of ball that will be drawn from a jar and will be paid for the accuracy of your answers as described in the first set of instructions. However, in Task 2, you will have the opportunity to win additional points when more WHITE balls are drawn.

The additional points you earn will be in the form of a lottery ticket. I will use an example to explain how the additional points are earned.

Suppose there are 2 balls in the jar: 1 WHITE and 1 BLACK. Suppose we draw 2 balls from the jar for every question and there are 3 questions. This means that the maximum number of WHITE balls that can be drawn from the jar is 6: 2 for every question. Now, suppose 2 WHITE balls are actually drawn. Then, you earn a lottery ticket that pays 100 additional points with $100 \times \frac{2}{6} = 33.33\%$ chance and pays 0 with a 66.67% chance.

The percent chance of winning the additional 100 points depends on the total number of WHITE balls drawn during Task 2. The more WHITE balls that are drawn, the higher chance you have of winning the additional 100 points.

The next screen provides an example. Please enter code ___.

When you are satisfied with your decisions you should click on the button "OK." Your answer is your private information and does not have any effect on later rounds.

Now we are going to go through a few questions to ensure that everyone has understood the instructions. Your answers do not affect your payment. However, we will not begin the Task until all questions have been answered correctly. If you have any questions, please raise your hand. Please enter code ___.

ADMINISTRATOR WILL GIVE ANSWERS.

We will now begin Task 2. Whenever you have finished reading a screen, please click on the button "OK" to advance to the next screen. Please enter code ___.

Task 3

Now we will discuss Task 3. In addition to guessing the percent chance that a certain number of white balls are drawn from a jar, your objective in Task 3 will also include answering various trivia questions. You will be paid for the accuracy of your answers about the color of balls drawn and for correctly answering the trivia questions. You will earn an additional 20 points for each correct trivia answer. You will have 1 minute to answer a single question. If you leave a question unanswered, you will receive 0 points.

In addition to earning an extra 20 points, a white ball will be added to your jar for each correct trivia answer. However, you will not be told whether you answered the question correctly.

Let's go through an example on the computer. Please enter code ___.

In Task 3, you will face two screens. On the first screen you will answer trivia questions. If there is only one trivia question, then you can add at most ONE white ball to your jar by answering correctly. If there are TWO trivia questions, you can add at most TWO white balls to your jar if you answer BOTH questions correctly. You will have 1 minute to answer a trivia question. If there are 2 trivia questions, then you have two minutes to complete those questions. If you do not enter an answer before your time is up, then you will not receive any points for the unanswered trivia questions. After answering click “OK” and enter code ___.

Now, you will see a screen that is similar to the screen faced in the General Setting. However, now you will also see a list of all of the possible compositions of balls in your jar. In this example, if you answer both questions correctly, then you will face a jar that has 4 WHITE balls and 2 BLACK balls. If you do not answer the question correctly, then you will face a jar with 2 WHITE balls and 2 BLACK balls. If you are 100% sure that you answered both questions correctly, then you are facing a jar that contains 4 WHITE balls. If you are 100% sure that you did NOT answer either the question correctly, then you are facing a jar that contains 2 WHITE balls.

Please use the sliding bar to indicate a number between 0 and 100 that reflects your belief of the percent chance of each event occurring.

When you are satisfied with your decisions you should click on the button “OK.” Your answer is your private information and does not have any effect on later rounds.

Now we are going to go through a few questions to ensure that everyone has understood the instructions. Your answers do not affect your payment. However, we will not begin the Task until all questions have been answered correctly. If you have any questions, please raise your hand. Please enter code ___.

ADMINISTRATOR WILL GIVE ANSWERS.

We will now begin Task 3. Whenever you have finished reading a screen, please click on the button “OK” to advance to the next screen. Please enter code ___.

Task 4

Now we will move on to Task 4. In addition to guessing the percent chance that a certain number of white balls are drawn from a jar, your objective in Task 4 will also include answering various trivia questions. You will be paid for the accuracy of your answers about the color of balls drawn and for correctly answering the trivia questions. You will earn an additional 20 points for each correct answer. You will have 1 minute to answer a single question. If you leave a question unanswered, you will receive 0 points.

In addition to earning an extra 20 points, a white ball will be added to your jar for each correct trivia answer. You will earn points towards an additional lottery ticket based on the number of WHITE balls drawn from all of the jars during Task 4. This means, that the more trivia questions you answer correctly, the more white balls will be to your jar and the higher chance you have of winning the additional lottery ticket that pays 100 points.

However, you will not be told whether you answered the question correctly.

Let's go through an example on the computer. Please enter code ___.

In Task 4, you will face two screens. On the first screen you will answer trivia questions. If there is only one trivia question, then you can add at most ONE white ball to your jar by answering correctly. If there are TWO trivia questions, you can add at most TWO white balls to your jar if you answer BOTH questions correctly. You will have 1 minute to answer a trivia question. If there are 2 trivia questions, then you have two minutes to complete those questions. If you do not enter an answer before your time is up, then you will not receive any points for the unanswered trivia questions. After answering click "OK" and enter code ___.

Now, you will see a screen that is similar to the screen faced in the General Setting. In this example, if you answer both questions correctly, then you will face a jar that has 4 WHITE balls and 2 BLACK balls. If you do not answer the question correctly, then you will face a jar with 2 WHITE balls and 2 BLACK balls. If you are 100% sure that you answered both questions correctly, then you are facing jar that contains 4 WHITE balls. If you are 100% sure that you did NOT answer either the question correctly, then you are facing a jar that contains 2 WHITE balls.

Please use the sliding bar to indicate a number between 0 and 100 that reflects your belief of the percent chance of each event occurring.

When you are satisfied with your decisions you should click on the button "OK." Your answer is your private information and does not have any effect on later rounds.

Now we are going to go through a few questions to ensure that everyone has understood the instructions. Your answers do not affect your payment. However, we will not begin the Task until all questions have been answered correctly. If you have any questions, please raise your hand. Please enter code ___.

ADMINISTRATOR WILL GIVE ANSWERS.

We will now begin Task 4. Whenever you have finished reading a screen, please click on the button "OK" to advance to the next screen. Please enter code ___.

General Review Questions After Introduction:

1. How many tasks will be completed during the session?
2. Will you earn points in every task (Y/N)?
3. Will you be paid for the point you earn in every task (Y/N)?
4. Suppose there are 2 white balls, 1 black ball and 2 balls will be drawn from the jar. Please answer the following questions.
    a. What is the minimum number of white balls that can be drawn from the jar?
    b. What is the maximum number of white balls that can be drawn from the jar?
    c. You believe there is a 54% change that 0 white balls will be draw. Does it make sense to answer that there is a 43% chance that ONE or FEWER white balls are drawn (Y/N)?
5. You are asked the following question: What do you expect is the change that ONE or FEWER white balls are drawn from the jar? Suppose you state 34%. (Subjects are prompted to use their pay-off sheet to answer.)
    a. How many points will you earn if 2 white balls are drawn?
    b. How many points will you earn if 1 white ball is drawn?
    c. How many points will you earn if 0 white balls are drawn?

Review Questions before the Baseline Task:

1. Do you earn points for the accuracy of your answers about the chance that white balls are drawn from the jar (Y/N)?
2. Can you affect the number of white balls in your jar (Y/N)?
3. Can you earn more points if more white balls are drawn (Y/N)?
4. If you win the extra lottery ticket, how many points does it pay? (Payment Treatment Only)

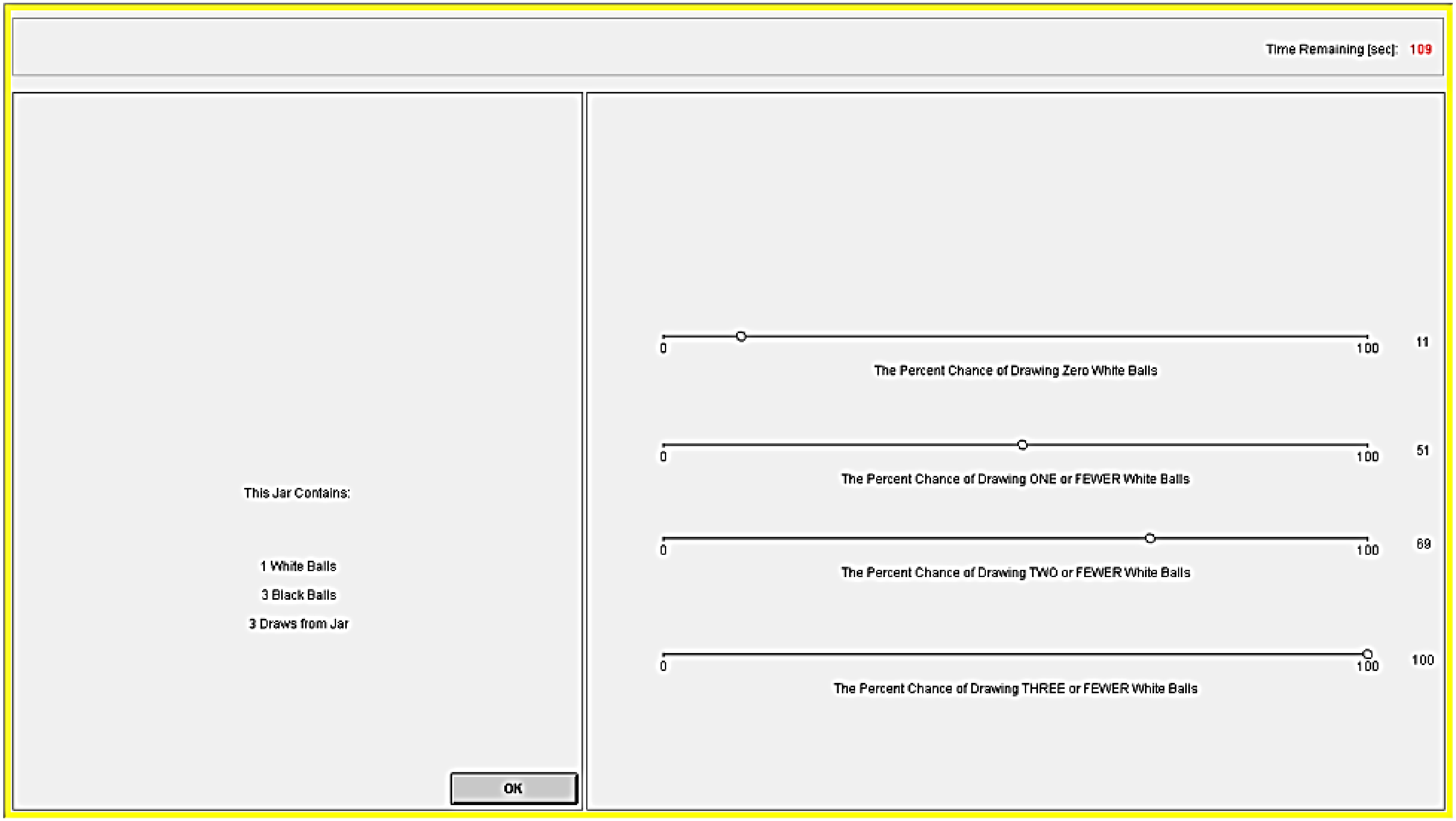
Review Questions before the Performance and Combined Treatments:

1. Do you earn points for the accuracy of your answers about the chance that white balls are drawn from the jar (Y/N)?
2. Can you affect the number of white balls in your jar (Y/N)?
3. Can you earn more points if more white balls are drawn (Y/N)?
4. Suppose there are 2 white balls and 2 black balls in the jar and 2 balls are going to be drawn. You answer 1 IQ question.
    a. If you answer the question CORRECTLY, then how many white balls are now in your jar?
    b. If you answer the question INCORRECTLY, then how many white balls are no in your jar?
    c. How many points do you earn for a correct answer to an IQ question?
5. If there are more white balls in your jar then the higher chance you have of winning an additional 100 points in the this Task (Y/N)?

Appendix Figure S6: Comprehension Test. Before a subject could begin a new treatment they were required to answer all review questions correctly. This allowed us to ensure that there was full comprehension about the differences between each treatment.

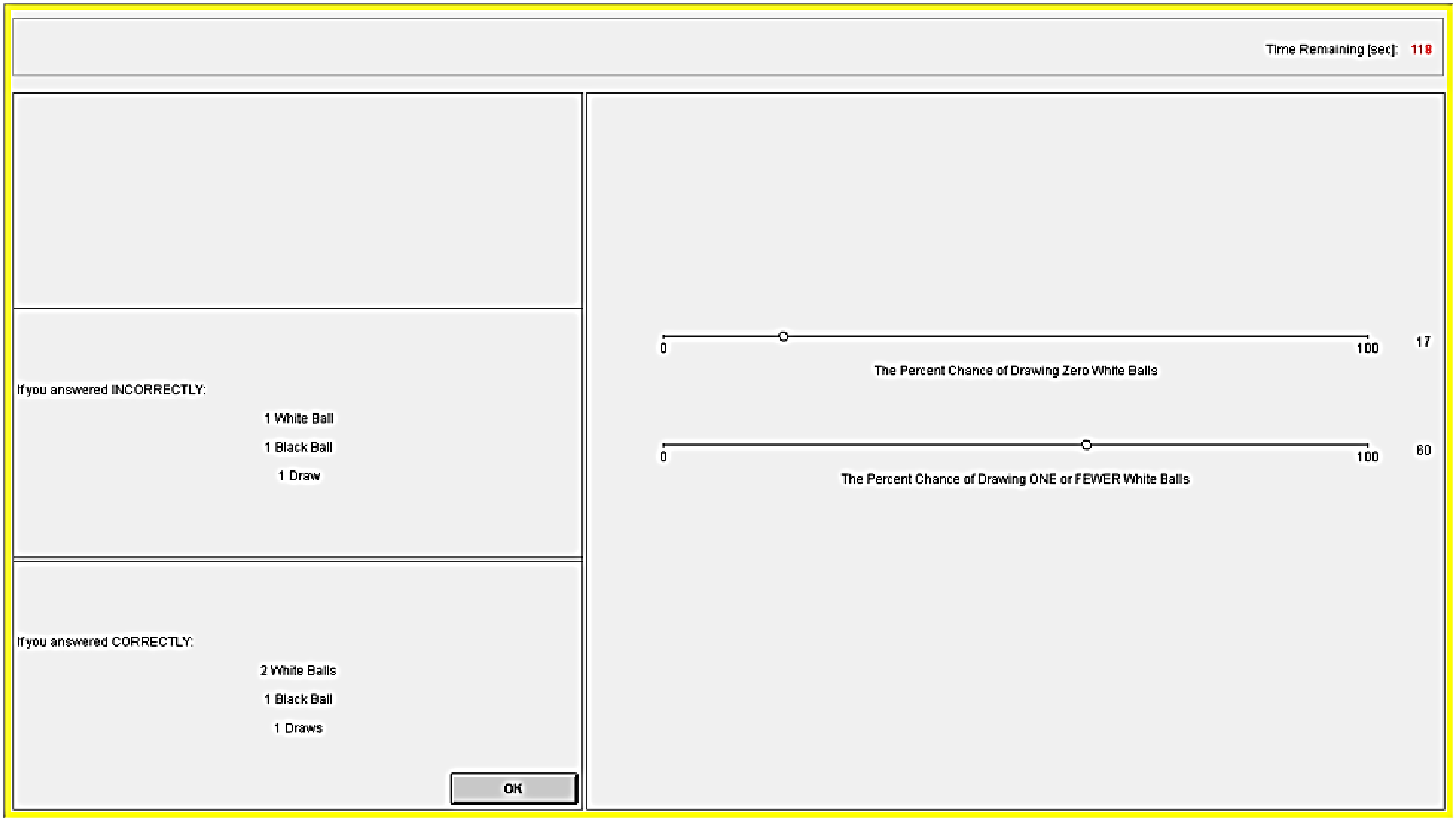
Pay-Off Sheet					
	Pay-Off			Pay-Off	
Reported Belief	Event Occurs	Event Does Not Occur	Reported Belief	Event Occurs	Event Does Not Occur
Percent Chance	For Example: ONE or FEWER WHITE BALLS is Drawn	For Example: MORE than ONE WHITE BALL is Drawn	Percent Chance	For Example: ONE or FEWER WHITE BALLS is Drawn	For Example: MORE than ONE WHITE BALL is Drawn
0.00	0.00	10.00	51.00	7.60	7.40
1.00	0.20	10.00	52.00	7.70	7.30
2.00	0.40	10.00	53.00	7.79	7.19
3.00	0.59	9.99	54.00	7.88	7.08
4.00	0.78	9.98	55.00	7.98	6.98
5.00	0.98	9.98	56.00	8.06	6.86
6.00	1.16	9.96	57.00	8.15	6.75
7.00	1.35	9.95	58.00	8.24	6.64
8.00	1.54	9.94	59.00	8.32	6.52
9.00	1.72	9.92	60.00	8.40	6.40
10.00	1.90	9.90	61.00	8.48	6.28
11.00	2.08	9.88	62.00	8.56	6.16
12.00	2.26	9.86	63.00	8.63	6.03
13.00	2.43	9.83	64.00	8.70	5.90
14.00	2.60	9.80	65.00	8.78	5.78
15.00	2.78	9.78	66.00	8.84	5.64
16.00	2.94	9.74	67.00	8.91	5.51
17.00	3.11	9.71	68.00	8.98	5.38
18.00	3.28	9.68	69.00	9.04	5.24
19.00	3.44	9.64	70.00	9.10	5.10
20.00	3.60	9.60	71.00	9.16	4.96
21.00	3.76	9.56	72.00	9.22	4.82
22.00	3.92	9.52	73.00	9.27	4.67
23.00	4.07	9.47	74.00	9.32	4.52
24.00	4.22	9.42	75.00	9.38	4.38
25.00	4.38	9.38	76.00	9.42	4.22
26.00	4.52	9.32	77.00	9.47	4.07
27.00	4.67	9.27	78.00	9.52	3.92
28.00	4.82	9.22	79.00	9.56	3.76
29.00	4.96	9.16	80.00	9.60	3.60
30.00	5.10	9.10	81.00	9.64	3.44
31.00	5.24	9.04	82.00	9.68	3.28
32.00	5.38	8.98	83.00	9.71	3.11
33.00	5.51	8.91	84.00	9.74	2.94
34.00	5.64	8.84	85.00	9.78	2.78
35.00	5.78	8.78	86.00	9.80	2.60
36.00	5.90	8.70	87.00	9.83	2.43
37.00	6.03	8.63	88.00	9.86	2.26
38.00	6.16	8.56	89.00	9.88	2.08
39.00	6.28	8.48	90.00	9.90	1.90
40.00	6.40	8.40	91.00	9.92	1.72
41.00	6.52	8.32	92.00	9.94	1.54
42.00	6.64	8.24	93.00	9.95	1.35
43.00	6.75	8.15	94.00	9.96	1.16
44.00	6.86	8.06	95.00	9.98	0.98
45.00	6.98	7.98	96.00	9.98	0.78
46.00	7.08	7.88	97.00	9.99	0.59
47.00	7.19	7.79	98.00	10.00	0.40
48.00	7.30	7.70	99.00	10.00	0.20
49.00	7.40	7.60	100.00	10.00	0.00
50.00	7.50	7.50			

Appendix Figure S7: Pay-Off Sheet. Subjects were given this pay-off sheet during the experiment in order to assist them with pay-off calculations.



Appendix Figure S8: Screenshot. Belief elicitation in the Baseline and Payment Treatment (distribution 3).

Appendix Figure S9: Screenshot. IQ questions in the Performance and Combined Treatments (starting Distribution 1).



Appendix Figure S10: Screenshot. Belief elicitation following IQ stage in the Performance and Combined Treatments.