ONLINE APPENDICES

Appendix A: Theory — some further details

A.1 The naive extrapolation equilibrium in each scenario

Scenario 1: Learning from others

In this case, a symmetric pure strategy equilibrium would require that an investor of type X invests if she gets a signal $s_X \geq s^{SN1}$. Thus $q_{inv}^{SN1}(s^*, s_B, s_C) = 1$ if $s_B, s_C \geq s^{SN1}$, $q_{inv}^{SN1}(s^*, s_B, s_C) = 0.5$ if $s_B \geq s^{SN1} > s_C$ or $s_C \geq s^{SN1} > s_B$ and $q_{inv}^{SN1}(s^*, s_B, s_C) = 0$ otherwise where s^{SN1} should be such that $\widehat{P}(\bar{x}|s^{SN1}; q_{inv}^{SN1}) \cdot \overline{x} > c$ and $\widehat{P}(\bar{x}|s^{SN1}-1; q_{inv}^{SN1}) \cdot \overline{x} < c$ if $s^{SN1} > 1$.

Given the symmetry of the problem, it is readily verified that $\widehat{P}(\bar{x}|s_A; q_{inv}^{SN1})$ simplifies into $P(\bar{x}|s_A; s_B \geq s^{SN1})$ which is clearly larger than $P(\bar{x}|s_A)$ (because the extra conditioning on $s_B \geq s^{SN1}$ shifts the probability of success upwards). This in turn implies that there is more investment in the equilibrium with naive investors in scenario 1 (in comparison to the Bayesian benchmark). We have $s^{SN1} = 6$ when $\bar{x} = 3.40$ and $c = 1$.

Scenario 2: Learning from others with correlated signals

When all investors are of type A, if agents follow the threshold strategy to invest if s_A is no smaller than s^{SN2}, we would have $q_{inv}^{SN2}(s^*, s_B, s_C) = 1$ if $s^* \geq s^{SN2}$ and 0 otherwise. This implies that for all $s_A \geq s^{SN2}$, $\widehat{P}(\bar{x}|s_A; q_{inv}^{SN2}) = P(\bar{x}|s_A)$ and thus it cannot be that $s^{SN2} < s^{Bayes}$ (given that $P(\bar{x}|s_A)\bar{x} < c$ for any $s_A < s^{Bayes}$). When signals are perfectly correlated among investors, there cannot be overinvestment in equilibrium, and imposing some exogenous trembling would force $s^{SN2} = s^{Bayes}$.

Scenario 3: Learning from better-informed individuals

In this case investors of a type A face in their feedback projects that were handled either by naive investors of type B or C, or, by omniscient investors that would invest only if the project is successful (the latter are implemented through machines). Letting λ denote the proportion of non-omniscient investors, and letting s^{SN3} denote their equilibrium threshold, one would have $q_{inv}^{SN3}(s^*, s_B, s_C) = \lambda q^C(s^*, s_B, s_C) + (1-\lambda) q^R(s^*, s_B, s_C)$ where $q^C(s^*, s_B, s_C) = 1$ if $s_B, s_C \geq s^{SN3}$, $q^C(s^*, s_B, s_C) = 0.5$ if $s_B \geq s^{SN3} > s_C$ or $s_C \geq s^{SN3} > s_B$ and $q^C(s^*, s_B, s_C) = 0$ otherwise; $q^R(s^*, s_B, s_C) = 1$ if $s^* + s_B + s_C \geq W$ and 0 otherwise, and s^{SN3} should be such that $\widehat{P}(\bar{x}|s^{SN3}; q_{inv}^{SN3}) \cdot \overline{x} > c$ and $\widehat{P}(\bar{x}|s^{SN3}-1; q_{inv}^{SN3}) \cdot \overline{x} < c$ if $s^{SN3} > 1$.

It can be shown that when there are more fully informed omniscient investors around (when λ is smaller), the overinvestment bias increases, i.e., s^{SN3} gets smaller. When $\lambda = 1/2$,

$\bar{x} = 3.40$ and $c = 1$ (as in our experiment), we have a symmetric equilibrium in mixed strategies, where players mix between playing $s^{SN3} = 5$ with probability $\mu = 0.8$ and playing $s^{SN3} = 6$ with probability $1 - \mu = 0.2$.[42]

Scenario 4: Learning from others when also observing the counterfactual

Here, type A investors observe feedback about all projects faced by type B or C investors, irrespective of whether they invested or not. This implies that while Naive Extrapolators still form their belief about the mapping from signals to success probabilities, $\widehat{P}(\cdot|\cdot)$, according to equation 2, they now observe the outcomes of past projects with probability one. Therefore, q_{inv} should be replaced by $q = 1$ in equation 2, and Naive Extrapolators form a belief $\widehat{P}(\bar{x}|s^*; q = 1)$, which in expectation is equal to the Bayesian posterior, $P(\bar{x}|s^*)$. Consequently, $s^{SN4} = s^{Bayes}$.

[42]Note, λ refers to the fraction of fully informed omniscient investors amongst those who are generating the feedback. This is why $\lambda = 1/2$ is relevant in relation to our experimental design, and not $\lambda = 1/3$.

A.2.1 Calculating the equilibrium in scenario 1

As noted above, without loss of generality, due to the symmetry of the game, we can consider the perspective of a type A agent. Making use of the two observations that: (i) $P(\bar{x}|s_A; s_B \geq s^{SN1}) = \sum_{s_B} P(\bar{x}|s_A; s_B)$, and (ii) $P(\bar{x}|s_A; s_B) = \frac{s_A+s_B-11}{10}$ if $s_A+s_B \geq 12$ and 0 otherwise, equation 2 can be expanded using the following analytical expression:

$$\begin{aligned} \widehat{P}(\bar{x}|s_A; q^{SN1}_{inv}) &= P(\bar{x}|s_A; s_B \geq s^{SN1}) \\ &= \frac{(\widetilde{s^{SN1}} - 11) \cdot (2s_A + \widetilde{s^{SN1}} - 12)}{20 \cdot (s^{SN1} - 11)} \end{aligned} \tag{4}$$

where

$$\widetilde{s^{SN1}} := \begin{cases} s^{SN1} & \text{if } s_A + s^{SN1} \geq 12 \\ 12 - s_A & \text{if } s_A + s^{SN1} < 12 \end{cases}$$

This allows us to construct a table consisting of the expected fraction of successful projects in a type A agent's feedback, conditional on signal s_A and the threshold strategy being used by the other agents, s^{SN1} (see Table 5 below). Therefore, each column of the table reports the success fractions that a type A agent would observe in her feedback if all other players were following a specific strategy s^{SN1}, as the number of projects in her feedback gets large.

To illustrate this, let us consider how the feedback of a type A agent changes, depending on the strategy followed by the agents generating her feedback. The right-most column shows the success fractions observed if the feedback is generated by players who only invest after receiving a signal of 10. For a type A agent with this type of feedback, on average 50% of the projects that she observes with $s_A = 6$ are successful projects. In contrast, the left-most column shows the success fractions observed if the feedback is generated by players who invest for a signal of 1 or higher (i.e. they always invest). For a type A agent with this type of feedback, on average 15% of the projects that she observes with $s_A = 6$ are successful projects. The success fractions in this left-most column coincide with the Bayesian success probabilities (since there is no selection in the feedback).

Table 5: Expected success fractions for observed projects

	$s^{SN1} =$									
	1	2	3	4	5	6	7	8	9	10
$s_A =$										
1	0	0	0	0	0	0	0	0	0	0
2	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.03	0.05	0.1
3	0.03	0.03	0.04	0.04	0.05	0.06	0.08	0.1	0.15	0.2
4	0.06	0.07	0.08	0.09	0.1	0.12	0.15	0.2	0.25	0.3
5	0.1	0.11	0.13	0.14	0.17	0.2	0.25	0.3	0.35	0.4
6	0.15	0.17	0.19	0.21	0.25	**0.3**	0.35	0.4	0.45	0.5
7	0.21	0.23	0.26	0.3	0.35	0.4	0.45	0.5	0.55	0.6
8	0.28	0.31	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7
9	**0.36**	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8
10	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9

Bayesian agents therefore use the probabilities in the left-most column and invest for all signals s_A such that $P(\bar{x}|s_A) > \frac{c}{\bar{x}} \approx 0.294$. Clearly, this inequality is only satisfied for $s_A = 9$ and $s_A = 10$, implying that the risk neutral Bayesian agent will follow the threshold strategy $s^{Bayes} = 9$.

To find the symmetric equilibrium investment strategy for naive investors, recall from the main text that the symmetric pure strategy equilibrium is given by s^{SN1} should be such that $\widehat{P}(\bar{x}|s^{SN1}; q_{inv}^{SN1}) > \frac{c}{\bar{x}}$ and $\widehat{P}(\bar{x}|s^{SN1} - 1; q_{inv}^{SN1}) < \frac{c}{\bar{x}}$ if $s^{SN1} > 1$. Since one can read the beliefs of a naive investor, $\widehat{P}(\bar{x}|s_A; q_{inv}^{SN1}) = P(\bar{x}|s_A; s_B \geq s^{SN1})$, directly off Table 5, one can see that the symmetric equilibrium in pure strategies is given by $s^{SN1} = 6$. The logic is as follows: given that $\frac{c}{\bar{x}} = \frac{1}{3.4} \approx 0.294$, moving down the $s^{SN1} = 6$ column in Table 5 shows that the Type A individual who holds these beliefs will invest for $s_A = 6$, but not for $s_A = 5$.

The information contained in Table 5 can be depicted graphically as in Figure 9. This visualisation of the data provides a second illustration of the equilibrium threshold strategies followed by the Bayesian and naive investor. In Figure 9, the Bayesian investor's beliefs are represented by the lowest (blue) curve.[43] The ratio $\frac{c}{\bar{x}} \approx 0.294$ is drawn as a horizontal (light blue) line. The Bayesian agent invests for signals which yield a belief about the probability of success that exceeds the ratio $\frac{c}{\bar{x}}$ (i.e. when the dark blue curve is above the light blue horizontal line). Clearly this is the case for signals $s_A = 9$ and $s_A = 10$.

For the naive investor, each of the curves in Figure 9 depict her beliefs as a function of the

[43]This lowest (blue) curve depicts the information contained in the left-most column of Table 5.

threshold strategy being followed by the individuals generating her feedback. For example, the top-most curve reflects the naive investors beliefs when all other agents are following a threshold strategy, where they only invest after a signal of 10, but not for lower signals. Looking at the $s^{SN1} = 6$ curve shows that the individual will invest for $s_A = 6$, but not for $s_A = 5$ (since the first belief is above the horizontal line, while the second is below).

Figure 9: Selection Neglect Agent's Perceived Success Probabilities

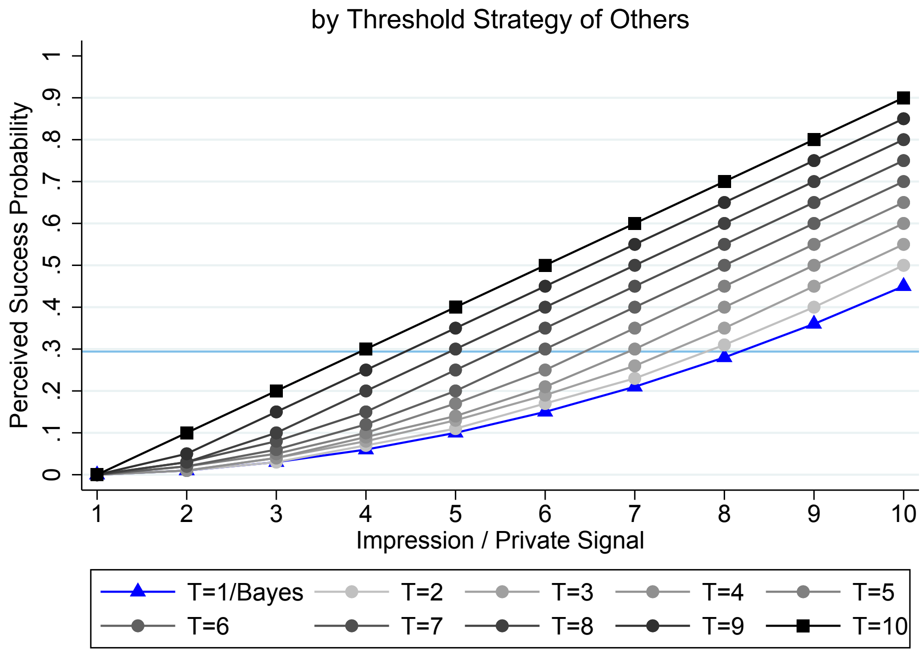


Figure 9 also provides an illustration of how individuals can exert an externality on others through the feedback generated. As the threshold strategy followed by other agents moves upwards, the naive agent's beliefs become increasingly distorted. This leads to poorer decision making by the naive agent. For example, if other agents are following the rational threshold strategy, $s^{Bayes} = 9$, the naive agent will hold highly distorted beliefs (i.e., refer to the distance between the T=1/ Bayes curve and the T=9 curve).

A.2.2 Calculating the equilibrium in scenario 2

As discussed in the main text above, in scenario 2, for all $s_A \geq s^{SN2}$, $\widehat{P}(\bar{x}|s_A; q_{inv}^{SN2}) = P(\bar{x}|s_A)$. The feedback received when there are perfectly correlated signals symmetric play is shown in Table 6. However, allowing for some trembling would imply that every column of the table is equivalent to the left-most column which contains the Bayesian success probabilities, conditional on s_A. Therefore, in this scenario risk neutral naive investor chooses $s^{SN2} = s^{Bayes}$.

Table 6: Expected success fractions for observed projects

	$s^{SN2} =$									
	1	2	3	4	5	6	7	8	9	10
$s_A =$										
1	0									
2	0.01	0.01								
3	0.03	0.03	0.03							
4	0.06	0.06	0.06	0.06						
5	0.1	0.1	0.1	0.1	0.1					
6	0.15	0.15	0.15	0.15	0.15	0.15				
7	0.21	0.21	0.21	0.21	0.21	0.21	0.21			
8	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28		
9	**0.36**	0.36	0.36	0.36	0.36	0.36	0.36	0.36	**0.36**	
10	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45

A.2.3 Calculating the equilibrium in scenario 3

Recognizing that the omniscient players invest iff a project is successful gives $\mathbb{E}_{s_B,s_C}[q^R(s_A,s_B,s_C)] = P(\bar{x}|s_A)$, which implies that:

$$
\begin{aligned}
\widehat{P}(\bar{x}|s_A;q_{inv}^{SN3}) &= \frac{\sum_{s_B}[(1-\lambda)\cdot P(\bar{x}|s_A;s_B)\cdot 1_{s_B\geq s^{SN3}} + \lambda\cdot P(\bar{x}|s_A)]}{\sum_{s_B}[(1-\lambda)\cdot 1_{s_B\geq s^{SN3}} + \lambda\cdot P(\bar{x}|s_A)]} \\
&= \frac{(1-\lambda)\cdot \widehat{P}(\bar{x}|s_A;q_{inv}^{SN1}=q_{inv}^{SN3})\cdot(11-s^{SN3}) + \lambda\cdot P(\bar{x}|s_A)\cdot 10}{(1-\lambda)\cdot(11-s^{SN3}) + \lambda\cdot P(\bar{x}|s_A)\cdot 10} \\
&= \frac{(1-\lambda)[(11-\widetilde{s^{SN3}})\cdot(s_A-11)+\frac{1}{2}(10\cdot 11-\widetilde{s^{SN3}}(\widetilde{s^{SN3}}-1))] + \lambda\cdot\frac{s_A^2-s_A}{2}}{(1-\lambda)\cdot 10\cdot(11-s^{SN3}) + \lambda\cdot\frac{s_A^2-s_A}{2}}
\end{aligned}
\tag{5}
$$

where

$$
\widetilde{s^{SN3}} := \begin{cases} s^{SN3} & \text{if } s_A + s^{SN3} \geq 12 \\ 12 - s_A & \text{if } s_A + s^{SN3} < 12 \end{cases}
$$

Using the parameterization relevant for our experimental design (i.e. $\lambda = \frac{1}{2}$) yields the following table of beliefs of a type A naive investor as a function of the threshold followed by other non-omniscient players, s^{SN3}, and her own signal, s_A.

Table 7: Expected success fractions for observed projects

	$s^{SN3} =$									
	1	2	3	4	5	6	7	8	9	10
$s_A =$										
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.02	0.02	0.02	0.03	0.03	0.04	0.05	0.06	0.10	0.18
3	0.06	0.06	0.07	0.08	0.10	0.11	0.14	0.18	0.26	0.38
4	0.11	0.13	0.14	0.16	0.18	0.21	0.26	0.33	0.42	0.56
5	0.18	0.20	0.22	0.25	**0.29**	**0.33**	0.40	0.48	0.57	0.70
6	0.26	0.29	0.32	0.35	0.40	0.46	0.53	0.60	0.69	0.80
7	0.35	0.38	0.42	0.46	0.52	0.58	0.64	0.71	0.78	0.87
8	0.44	0.47	0.52	0.57	0.63	0.68	0.74	0.79	0.85	0.92
9	0.53	0.57	0.62	0.67	0.72	0.77	0.82	0.86	0.91	0.96
10	0.62	0.67	0.71	0.76	0.80	0.84	0.88	0.92	0.95	0.98

For the Bayesian agent, the information in Table 7 is irrelevant. Instead she relies on her knowledge of the DGP and, as above, uses the probabilities $P(\bar{x}|s_A)$ to guide her decision making (the left-most column of Table 5 contains these probabilities). She therefore again follows the threshold strategy $s^{Bayes} = 9$.

There is no pure strategy symmetric equilibrium for risk neutral naive investors. Rather, there is exists a symmetric equilibrium in mixed strategies in which players mix between $s^{SN3} = 5$ and $s^{SN3} = 6$, such that all players are indifferent between these two strategies.[44] This condition is satisfied when players choose $s^{SN3} = 5$ with probability $\mu = 0.8$.

[44]This is the case when agents are indifferent between investing and not investing after the signal $s_A = 5$ (i.e. when the perceived probability of success is 0.294118). The equilibrium mixture probability is not obtained by simply weighting the two relevant probabilities reported in Table 7, but rather by calculating the equilibrium, allowing for mixing.

A.2.4 Learning dynamics

The above construction can also be used to analyze the learning dynamics (in a similar fashion as what is considered in the online Appendix of Jehiel (2018)).

If we consider the dynamics in which at every round t, players use the feedback generated in round $t-1$ (as considered in Jehiel (2018), then one gets a dynamics of thresholds given by (assuming there is a continuum of investment data in each round) s^{SN1k} where $s^{SN1(k+1)}$ is defined so that for

- $s_A = s^{SN1(k+1)}$, $\widehat{P}((\overline{x} \mid s_A, s_B \geq s^{SN1k}) > \frac{c}{\overline{x}}$ and
- $s_A = s^{SN1(k+1)} - 1$, $\widehat{P}((\overline{x} \mid s_A, s_B \geq s^{SN1k}) < \frac{c}{\overline{x}}$.

This defines a sequence of thresholds parameterized by the first threshold s^{SN11}.

One can check that no matter what s^{SN11} is chosen, this sequence converges to the equilibrium threshold $s^{SN1} = 6$ in just a few rounds. For example, starting from $s^{SN11} = 10$, we get (from inspecting Table 4 or Figure 7), $s^{SN12} = 4, s^{SN13} = 7, s^{SN1k} = 6$ for $k \geq 4$. Starting from $s^{SN11} = 1$, we get $s^{SN12} = 9, s^{SN13} = 5, s^{SN14} = 7, s^{SN1k} = 6$ for $k \geq 5$.

In our experiment, the feedback did not consist only of the projects implemented in the last round but of all projects implemented in earlier rounds. This would lead to modify the dynamics of thresholds s^{SN1k} as follows.[45]

Define

$$\widehat{P}_t(\overline{x} \mid s_A, s^{SN1k}, k < t) = \frac{\sum_{k<t} \widehat{P}((\overline{x} \mid s_A, s_B \geq s^{SN1k}) P(s_B \geq s^{SN1k})}{\sum_{k<t} P(s_B \geq s^{SN1k})}$$

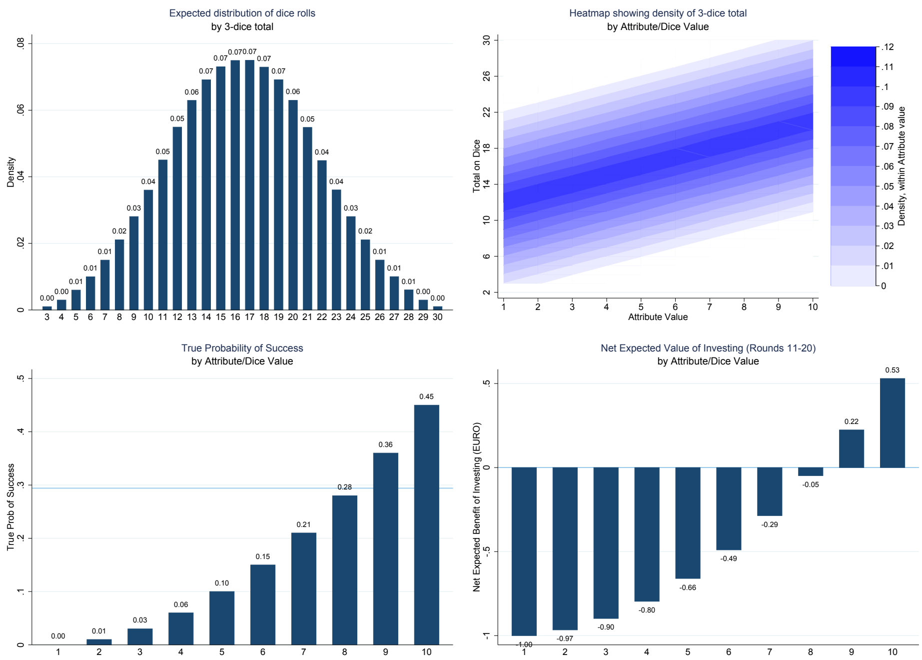
This induces a a value of s^{SN1t} defined so that for $s_A = s^{SN1t}$, $\widehat{P}_t(\overline{x} \mid s_A, s^{SN1k}, k < t) > \frac{c}{\overline{x}}$ and for $s_A = s^{SN1t} - 1$, $\widehat{P}_t(\overline{x} \mid s_A, s^{SN1k}, k < t) < \frac{c}{\overline{x}}$.

One can check again that such a modified dynamics would lead to quick convergence no matter how s^{SN11} is chosen. For example, with $s^{SN11} = 10$, we would have $s^{SN12} = 4, s^{SN1k} = 6$ for $k \geq 3$ (with a convergence one round before it occurs in the previous dynamics).

[45]The various rounds should be weighted by the corresponding mass of implemented projects, hence the terms $P(s_B \geq s^{SN1k})$.

Appendix B: The data generating process (DGP)

Figure 10: Overview of the DGP and Decision Problem



The data generating process is summarized in Figure 10.[46] The top left panel shows the unconditional distribution over the total when rolling three fair ten-sided dice. The top-right panel uses a heatmap to show the conditional distributions of the 3-dice totals, conditional on observing one of the dice values. Since the investor does not really care about the the full distribution of 3-dice totals, but rather cares about how this distribution maps onto the binary success / failure variable, the bottom-left panel displays the probability of success after observing each dice roll from 1 to 10. Given the parameters that we chose in our experiment in the last 10 rounds (i.e. the cost of investing, €1, and value of a successful investment, €3.40), the bottom-right panel then translates this into the net expected value of investing after observing each dice value from 1 to 10. This last panel illustrates that for the rational, risk neutral investor, it is only attractive to invest after observing a value of 9 or 10, although a value of 8 is marginal. A risk averse investor should be even more cautious about investing.

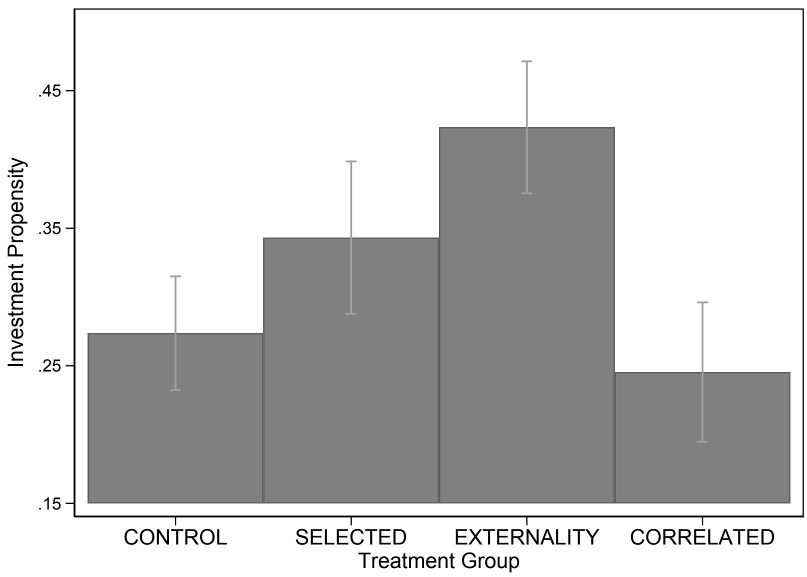
[46]The figures in the top two panels are generated by simulating 10 million projects and are therefore (fairly precise) approximations of the true distribution. The bottom two panels reflect the precise distribution.

Appendix C: Additional results for Experiment 1

Appendix C.1: Robustness checks

As a robustness check to our main results from Experiment 1[47], we also replicate our analysis for the subset of individuals who are "well behaved" in the sense that they follow a threshold strategy in each of the last five rounds. Restricting attention to this subsample rules out individuals who deviated from following a threshold strategy at least once (e.g. within a given round, invested for an attribute value of s, but did not invest for s' where $s < s'$). This restriction is quite conservative as applies not only to individuals who, (i) made a mistake, or (ii) lacked understanding, but also to those who (iii) understood fully, but simply believed that a lower attribute had a higher probability of leading to a successful project in at least one instance. It is therefore reassuring that the vast majority of participants (i.e. 77%) satisfy this restriction. We refer to this subsample as the "Restricted Sample", both in the regression tables in the main text, and in the figures below. Restricting attention to this subsample does not change any of the results presented above. An additional benefit of focusing on "well behaved" participants, who follow a threshold strategy, is that we can also present some results relating to the threshold strategy chosen. However, since an individual's threshold strategy is mechanically related to her investment propensity, these results do not yield novel insights into behavior.

Figure 11: Investment fraction in last five rounds, by treatment (Restricted Sample)



[47]The main results don't exclude any participants.

Figure 12: Propensity to invest across rounds, by treatment (Restricted Sample)

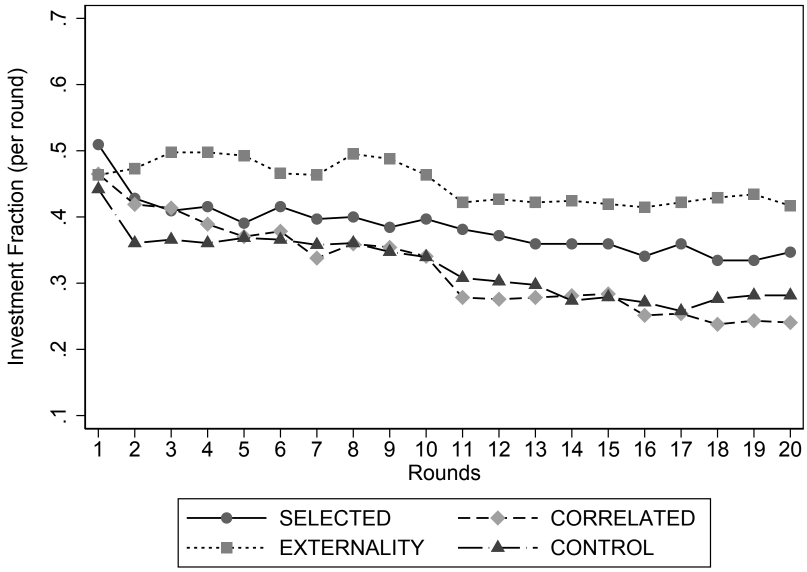
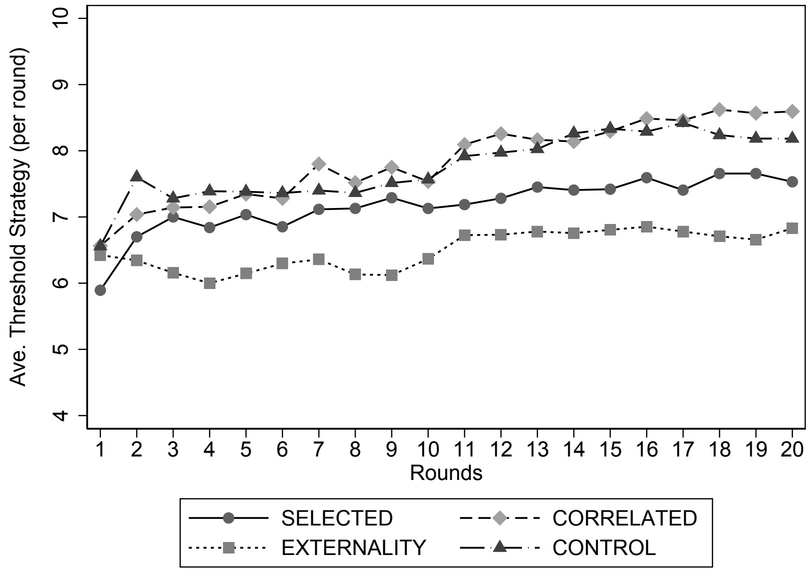


Figure 11 and 12 reproduce figures 3 and 4 respectively, for the restricted sample. The observed pattern of behavior is very similar in the restricted sample to that observed in the full sample. This is supported by the regression results reported in tables 2 and 3, showing similar results for the restricted and full sample. Together, these results suggest that our treatment effects are not driven by individuals who failed to understand the game, or chose their strategy randomly.

Figure 13 shows the average threshold strategy followed across the twenty rounds for the restricted sample. We only present this figure for the restricted sample, as one needs to make further assumptions to assign a threshold to individuals who don't follow a proper threshold strategy. However, in our experimental design, the threshold strategy is the mechanical inverse of the investment propensity for individuals who follow a threshold strategy. Figure 13 uses the average of all individuals in the restricted sample for the last five rounds, but in the first fifteen rounds, additionally, for each round, excludes individuals who didn't follow a threshold strategy in that particular round. This explains why the figure 13 is not a perfect reflected image of figure 12 for the first fifteen rounds.

Figure 13: Average threshold strategy across rounds, by treatment (Restricted Sample)



The manifestation of the treatment effect

The evidence presented in the main text for Experiment 1 focuses on differences in the mean propensity to invest across all signals / attribute values that an individual observes before choosing to invest in a project. To complement this, one may look at the full investment strategies across treatments. For Naive Extrapolators, investment threshold should be shifted downwards in the Selected and Externality treatments — implying that the treatment effect should be concentrated at intermediate attribute values (e.g. an investor who would invest for all attribute values of 8 or higher in the Control treatment, might, in the Selected treatment, instead invest for all attribute values 7 or higher, and in the Externality treatment, invest for all attribute values 6 or higher).[48]

[48]That is, in all treatments, low attribute values should be below an individual's threshold, and high attribute values be above their threshold. Therefore, individuals should never invest for these low attribute values, and always invest for high attribute values. A shift downwards in their threshold would increase the propensity to invest at intermediate attribute values.

Figure 14: Investment by attribute value, between treatments (R16-20)

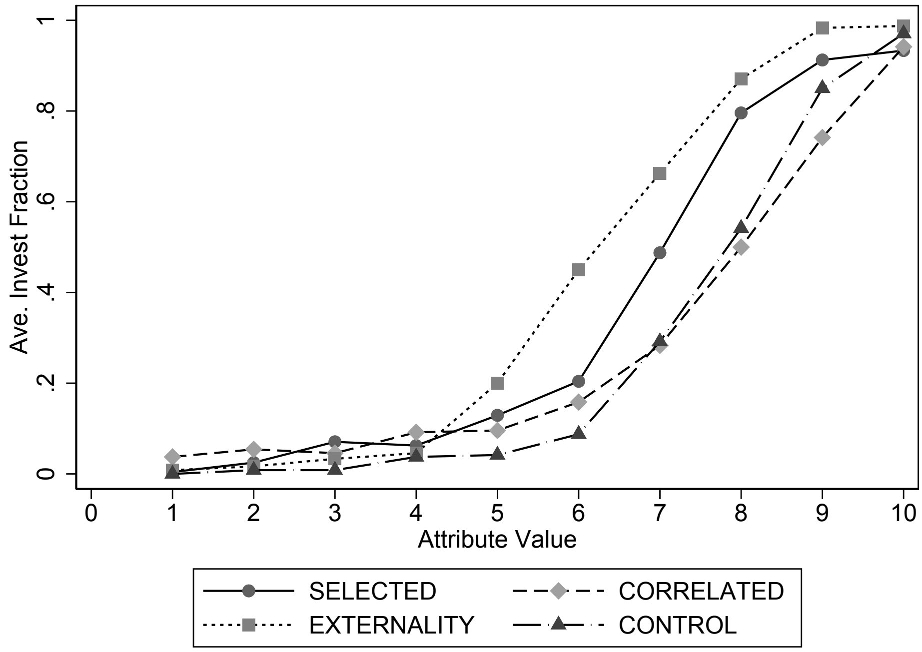


Figure 14 shows that our data is consistent with this, with the treatment differences arising at intermediate attribute values. The figure displays the average participant's propensity to invest in the last five rounds, conditional on each attribute value. Participants appear to have understood the task well as we observe very low investment rates for attributes 1 to 4. Similarly, we observe very high investment rates for attributes 9 and 10. The differences in investment between treatments occurs predominantly between attributes 6 and 8.

Figure 15 shows the average propensity to invest for each attribute value for the restricted sample. It is very similar to figure 14, with the exception that individuals in the restricted sample seem even better at not investing for attribute values 1 to 4 (as one might expect). Figure 16 provides some evidence on heterogeneity in the threshold strategy that participants followed in each of the treatments, but reporting the distribution of thresholds observed in round 20. In the figure, participants are grouped according to the attribute value, s^*, such that they invested for all $s \geq s^*$ in round 20. Therefore, participants with a threshold of $s^* = 11$ did not invest for any attribute value. The figure shows that the treatment shifted the threshold distribution to the left in the SELECTED and EXTERNALITY treatments, relative to the CONTROL and CORRELATED treatments.

Figure 15: Investment by attribute value, between treatments (R16-20, Rest. sample)

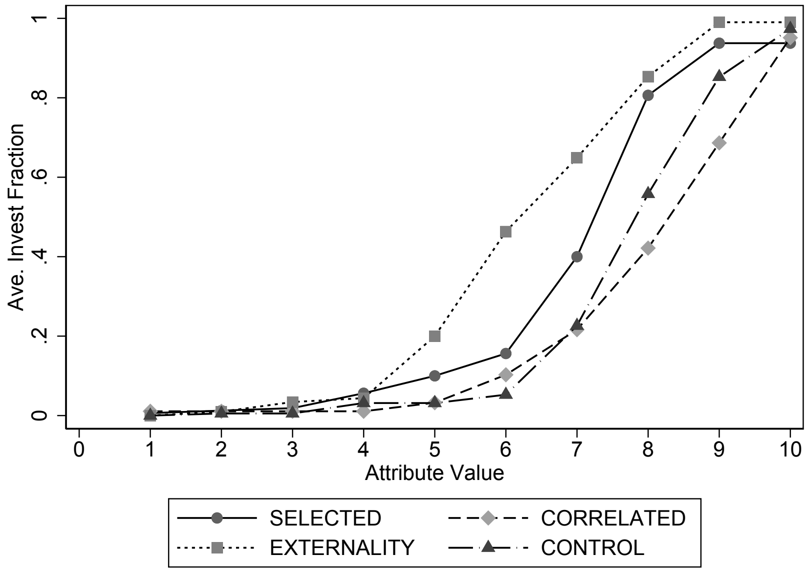
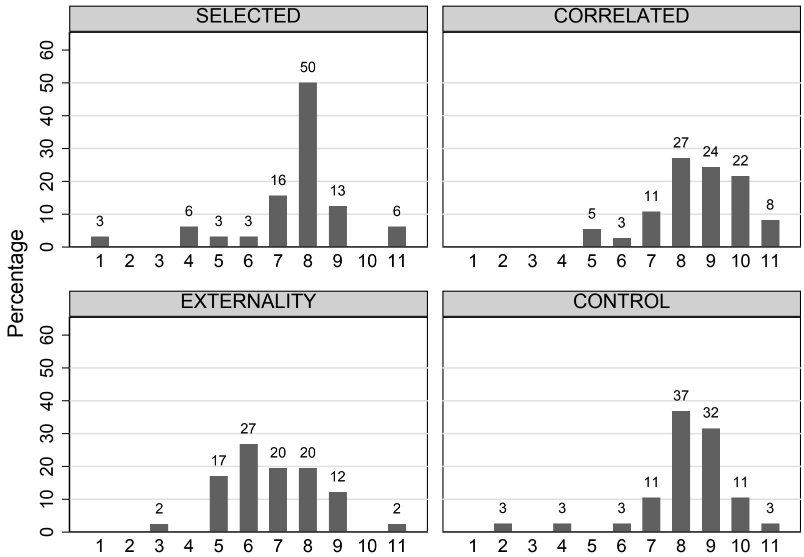


Figure 16: Distribution of threshold strategies in round 20 (Restricted sample)



Appendix C.2: Additional tables and figures from Experiment 1

Table 8: Propensity to invest in SELECTED and EXTERNALITY relative to CORRELATED

	Full Sample		Restricted Sample
	All	R16-20	R16-20
	(1a)	(1b)	(2a)
SELECTED	0.04	0.07*	0.10**
	(0.03)	(0.04)	(0.04)
EXTERNALITY	0.08***	0.13***	0.18***
	(0.03)	(0.03)	(0.04)
Constant	0.37***	0.30***	0.25***
	(0.02)	(0.03)	(0.03)
Observations	144	144	110
Adjusted R^2	0.039	0.091	0.184

Notes: (i) OLS regressions include one observation per individual, (ii) The dependent variable is an individual's average investment propensity, either over all rounds, or over rounds 16-20. (iii) The comparison treatment in the regression is the CORRELATED treatment, (iv) Standard errors are clustered at the interaction group level. This means that in the SELECTED and CORRELATED treatments, there are three individuals per cluster, and in the EXTERNALITY treatment there are two individuals per cluster. The standard errors are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Propensity to invest by treatment (observations for each round)

	Full Sample		Full Sample		Restricted Sample
	All	All	R16-20	R16-20	R16-20
	(1a)	(1b)	(2a)	(2b)	(2c)
SELECTED	0.07***	0.07***	0.08***	0.08***	0.07**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
EXTERNALITY	0.11***	0.11***	0.14***	0.14***	0.15***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
CORRELATED	0.03	0.03	0.01	0.01	-0.03
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
Constant	0.34***	0.45***	0.28***	0.28***	0.27***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Round FEs		Y		Y	Y
Observations	3840	3840	960	960	740
Adjusted R^2	0.052	0.115	0.106	0.103	0.167

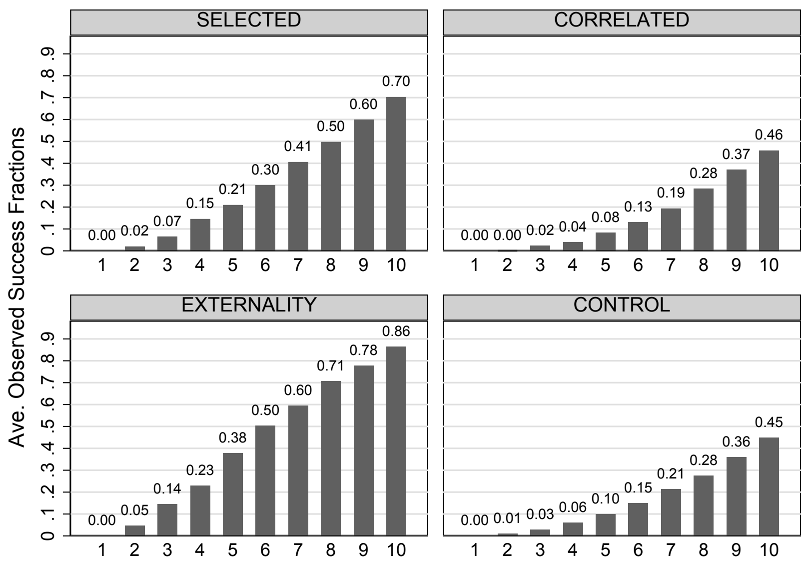
Notes: (i) OLS regressions include one observation per individual per round, (ii) The dependent variable is an individual's investment propensity in that round, (iii) In columns (1b), (2b) and (2c) we include round fixed effects (FEs). (iv) In column (2c), we restrict the sample to only those individuals who followed a pure threshold strategy in each of the last five rounds, (v) Standard errors are clustered at the interaction group level. This means that in the SELECTED and CORRELATED treatments all decisions by the relevant group of 3 individuals are included in a cluster, in the EXTERNALITY treatment there are two individuals per cluster, and in the CONTROL treatment each individual is a cluster, since their feedback is not influenced by their group members' choices. The standard errors are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Propensity to invest in SELECTED and EXTERNALITY (observations for each round)

	Full Sample		Full Sample		Restricted Sample
	All	All	R16-20	R16-20	R16-20
	(1a)	(1b)	(2a)	(2b)	(2c)
EXTERNALITY	0.04	0.04	0.06**	0.06**	0.08**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Constant	0.41***	0.50***	0.36***	0.36***	0.34***
	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)
Round FEs		Y		Y	Y
Observations	1920	1920	480	480	365
Adjusted R^2	0.013	0.053	0.036	0.028	0.046

Notes: (i) OLS regressions include one observation per individual per round, (ii) The dependent variable is an individual's investment propensity in that round, (iii) In columns (1b), (2b) and (2c) we include round fixed effects (FEs). (iv) In column (2c), we restrict the sample to only those individuals who followed a pure threshold strategy in each of the last five rounds, (v) Standard errors are clustered at the interaction group level. This means that in the SELECTED treatment, all decisions by the relevant group of 3 individuals are included in a cluster, while in the EXTERNALITY treatment there are two individuals per cluster. The standard errors are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 17: Average database observed by participants in round 20 (Experiment 1)



Appendix D: Additional material from Experiment 2

Appendix D.1: Description of treatments in Experiment 2

Experiment 2 consists of nine treatment conditions. They can be divided into four sets of treatments:

*The **3-Dice NoPastData** and **3-Dice PastData** treatments.*

The **3-Dice NoPastData** serves to populate the databases for other treatments. In this treatment, participants do not have access to any past data and must base their investment decisions in Part A solely on the information that they have about the data generating process. Additionally, the treatment will allow us to test whether additional (superfluous) information can impair decision quality.[49]

The **3-Dice PastData** treatment is identical to the 3-Dice NoPastData treatment, with the following two exceptions. First, the participants in the 3-Dice NoPastData treatment base their investment decisions on the realization of the **red** dice, while those in the 3-Dice PastData treatment base their investment decisions on the realization of the **purple** dice.[50] Second, in Part A of 3-Dice PastData, when participants make their investment decision they are provided with information about the outcomes of the projects that participants in the 3-Dice NoPastData treatment chose to invest in. Specifically, for each of the 100 participants who participate in the 3-Dice NoPastData treatment, the computer will randomly generate 50 projects. It will then use the decisions indicated by their strategy method investment choice in Part A to evaluate whether they invest or do not invest in each of the 50 projects. For each of the 100 participants, the computer will collect together the projects that they invested in.[51] It will then organize these projects into a database according to the **purple** dice value of the project. Investors in the 3-Dice PastData treatment will be able to

[49]Note, one implication of participants not having access to past data from other participants in Part A of the 3-Dice NoPastData treatment is that in Part C, the objective probabilities corresponding to the Part A beliefs of a "naive extrapolator" are not well defined (since the naive extrapolator has no data to extrapolate from in Part A). Therefore, to keep this treatment as similar as possible to the other treatments, we populate the Part C probabilities in the following way: We will use the beliefs that a "naive extrapolator" would hold in expectation if they were observing the past projects from individuals who all followed the decision rule: invest iff the project has a dice value of 6 or greater. These probabilities correspond to those reported in the fifth column from the right in Table 5 (i.e., the column headed "6").

[50]This is purely a difference in labelling since the projects are symmetric in the three dice. However, the reason that we use the different dice colors in the two treatments is because we wish to make it clear to participants that when they observe past projects that others invested in, the dice that the past investors observed is different to the one that they themselves observe.

[51]Specifically, participants are told that they will observe the outcomes of projects invested in by 100 past participants. They are told that " ... the computer has generated 50 projects for each of these 100 individuals. This implies that the computer generated a total of 5000 projects. For each of these 5000 projects, the computer rolled three dice," and "If it was invested in, then it is added to the database of past projects that you can observe. If it was not invested in, then you do not observe the outcome of that project and it is not added to the database."

observe the percent of these past projects that were successful, conditional on each **purple** dice value between 1 and 10.

*The **2-Dice NoPastData** and **2-Dice PastData** treatments.*

The second pair of treatments test the boundaries of naive extrapolation. Specifically, here we reduce the number of dice that define a project from 3 to 2, which reduces the complexity of the data generating process. The objective is to investigate whether individuals are more prone to extrapolate naively from past data when they face a decision problem that depends on a complex data generating process in comparison to when it is relatively simpler.

The **2-Dice NoPastData** treatment is the same as the 3-Dice NoPastData treatment, with the exception that a project in Part A is defined by rolling two ten-sided dice instead of three. In the 2-dice treatments, participants are told that a project is successful if the sum of the two dice is at least 16.

The **2-Dice PastData** treatment is the same as the 3-Dice PastData treatment, with the exception that the past data that participants now observe is collected from the 2-Dice NoPastData treatment. The probabilities in Part B and Part C are adjusted accordingly to fit the 2-dice case.[52]

*The **3-Dice PartialDGP**, **3-Dice Cue** and **3-Dice ExtraInfo** treatments.*

The third set of treatments aim to further explore the underlying mechanisms generating naive extrapolation. Each of the three treatments is very similar to the 3-Dice PastData, but one particular feature of the decision environment is varied.

The **3-Dice PartialDGP** is identical to the 3-Dice PastData treatment, with the exception that participants are not provided with a complete description of the data generating process. Here, while participants are told that the sum of the three ten-sided dice defines the success of the project, they are not told the exact threshold for success. Specifically, they are told that there is some value x and if the sum of the three dice is equal to at least x then the project is successful. Therefore, they are provided with sufficient information to know that success is monotonically increasing in the value of any individual dice realization (in particular, the one that they observe). This treatment allows us to investigate the influence of an investor not fully knowing the true underlying model that is generating the data.

The **3-Dice Cue** explores the limits of naive extrapolation. It is the same as the 3-Dice PastData treatment, except the participants read a series of leading questions before making their investment choice. The questions themselves do not reveal any additional information, but rather guide the participant through a series of logical steps that they may have followed themselves just by thinking hard about the problem (without requiring additional

[52]As in the 3-dice treatments, the probabilities in Part B correspond to the Bayesian probabilities of success for a project in Part A. The probabilities in Part C are also be calculated in the same way as for the 3-dice treatments.

knowledge). We can then ask whether this specific type of "cue" or "nudge" can induce the investor to view the problem through a different lens, helping them to see the logical error present when engaging in naive extrapolation.

The **3-Dice ExtraInfo** treatment is the same as the 3-Dice PastData treatment, except that participants are provided with a second table that summarizes the information about the number of projects that past investors chose to invest in. Similarly to the success percentages, these are organized according to the values of the **purple** dice. Therefore, the table shows the number of past investments that were invested in, conditional on each **purple** dice value between 1 and 10.

*The **3-Dice17 NoPastData** and **3-Dice17 PastData** treatments.*

The final two treatments in Experiment 2, the **3-Dice17 NoPastData** and **3-Dice17 PastData** treatments, are very similar to the core 3-Dice NoPastData and 3-Dice PastData treatments. The only difference is that the threshold for a project to be successful is reduced from 22 to 17. This means that the Bayesian probabilities of success are shifted upwards—conditional on any specific observed dice roll realization, a project is more likely to be successful with the lower threshold. These treatments aim to achieve two objectives. First, they provide an additional robustness check for our main result by assessing whether individuals still display selection neglect when the (Bayesian and Naive) probabilities associated with each dice roll are substantially higher. Under these two treatments, a risk-neutral Bayesian would invest for more than half of the possible dice values (in our other treatments, a risk-neutral Bayesian would invest for fewer than half the dice values). This implies that if there is a pull towards the middle of the choice list (e.g., a presentation effect) in these two treatments, this bias would now result in under-investment (in contrast to in our other treatments, where it would result in over-investment).

Second, in our other treatments, we observe over-investment even in the absence of additional information about past investments of others (i.e., when there is only information about the DGP).[53] These additional treatments with a lower success threshold allow us to shed light on why this is the case. Specifically, one potential explanation for the over-investment in the 3-Dice NoPastData treatment is that some individuals find inferring the probabilities of success from a description of the data-generating process to be a complex task. This may result in cognitive uncertainty, which leads to estimates that are distorted towards the cognitive default (see, e.g., Enke and Graeber, 2023). In the absence of information about past investments from others, this cognitive uncertainty may induce participants to shift their beliefs about success from the Bayesian probabilities implied by the

[53]As noted in the main text, these two treatments, 3-Dice17 NoPastData and
textsc3-Dice17 PastData, were not part of the initial seven pre-registered treatments in Experiment 2. We added these final two treatments in response to helpful suggestions that we received from the editor and referees. Therefore, we were already aware of the over-investment (relative to Bayes) in the 3-Dice NoPastData treatment when designing the 3-Dice17 NoPastData and
textsc3-Dice17 PastData treatments.

DGP towards 50% (which is a plausible candidate for the cognitive default). If this is the case, then by shifting the success threshold in these new treatments, the Bayesian probabilities are shifted upwards such that these Bayesian probabilities are closer to the cognitive default. This should reduce over-investment in the absence of past investment information from others. Crucially, this will then allow us to assess whether receiving this additional past information from others' investments leads to an increase in investment away from Bayes' and the cognitive default, making investors worse off.

Table 11 provides a summary of the key features of the treatments.

Table 11: Summary of the treatments in Experiment 2

Treatment Name	Treatment Variation	No. of Dice	Data on Past Projects?	Which Past Projects?	Part B	Part C
3-Dice NoPastData		3	No		Bayes	NaiveSim
3-Dice PastData	Add past project data	3	Yes	3-Dice NoPastData	Bayes	Naive
2-Dice NoPastData	2 dice instead of 3	2	No		Bayes	NaiveSim
2-Dice PastData	Add past project data	2	Yes	2-Dice NoPastData	Bayes	Naive
3-Dice PartialDGP	Partial info about DGP (success $\geq x$)	3	Yes	3-Dice NoPastData	Bayes	Naive
3-Dice Cue	Socratic method	3	Yes	3-Dice NoPastData	Bayes	Naive
3-Dice ExtraInfo	Add info. on no. of "invested in" projects	3	Yes	3-Dice NoPastData	Bayes	Naive
3-Dice17 NoPastData	Uses a success threshold of 17	3	No		Bayes	NaiveSim
3-Dice17 PastData	Add past project data	3	Yes	3-Dice17 NoPastData	Bayes	Naive

Notes: (i) "Bayes" refers to the correct Bayesian probabilities associated with the relevant **Part A** investment decision, (ii) When the participants have access to past project data in **Part A**, then "Naive" refers to using probabilities in **Part C** that are equal to the empirical success fractions observed in the past data in **Part A**, (iii) When the participants do not have access to past project data in **Part A**, then "NaiveSim" refers to simulated empirical success fractions that would be obtained (in expectation) if the participant were observing the past projects that a risk neutral Bayesian subject had invested in.

Appendix D.2: Pre-registered hypotheses for Experiment 2

The following hypotheses were pre-registered in the AEA Registry (AEARCTR-0010536) after the completion of Experiment 1, but prior to the data collection for Experiment 2. The description of these hypotheses relies on the definitions that are developed in Section 5.2.1 of the main text.

D.2.1 Is the average participant partially naive?

In Experiment 1, we presented evidence that the average participant displayed some degree of selection neglect, implying that this average participant deviated from Bayesian behavior and was classified as being partially naive. In Experiment 2, we aim to replicate this finding in a far simpler setting. While Experiment 2 retains the core elements of Experiment 1, we have introduced several important adjustments relative to Experiment 1. First, participants only make a single investment decision. Second, we have reduced the salience of the information about past projects by representing it in a simple table rather than graphically (thereby reducing potential experimenter demand effects). Third, we have substantially simplified the instructions, tranformed the task into an individual decision-making task rather than a dynamic interactive game, and added additional understanding checks to ensure that participants fully understand the setting. Therefore, if we replicate the finding in this setting, it would provide strong support that it is a robust result.

Hypothesis B. 1. ***(Naivety of the average individual)*** *The average participant in the 3-DICE PASTDATA treatment is partially naive.*

We test this hypothesis by evaluating whether the $\mu^n_{3D.PD} < 1$ for the average participant in the 3-DICE PASTDATA treatment.

D.2.2 Does complexity of the DGP influence the prevalence of naive extrapolation?

In Part A of Experiment 2 (and in Experiment 1), participants typically have two sources of information that they can draw on in order to infer the probabilities that an investment will succeed conditional on each possible dice value. First, they can draw on their knowledge of the data generating process. Second, they can draw on the empirical data that they observe from past participants. A key question here is how individuals decide which source of information to rely on. One possibility is that some participants will follow the data irrespective of other contextual features of the choice environment. However, an alternative possibility is that it depends on the difficulty (costliness) of processing the different sources of information. An important dimension here is the complexity of the data generating process. In scenarios where it is relatively simpler for individuals to process the information that they have about the data generating process, they may be more inclined to take it into account.

In our 2-Dice PastData treatment we reduce the complexity of the data generating process by reducing the number of dice that define a project from three to two. We then ask whether this reduction in complexity induces the average participant in the 2-Dice PastData treatment to move closer to using Bayesian probabilities in their Part A investment decisions in comparison to participants in the 3-Dice PastData treatment.

Hypothesis B. 2. ***(Complexity of the DGP)*** *Participants become more naive as the complexity of the data generating process increases—the average participant in the 2-Dice PastData treatment is closer to Bayesian than the average participant in the 3-Dice PastData treatment.*

We test this hypothesis by assessing whether $\mu^n_{2D.PD} > \mu^n_{3D.PD}$.

D.2.3 Does a reduction in DGP information exacerbate naive extrapolation?

In the real world, it is often the case that individuals do not have full information about the true underlying data generating process and have to draw on partial DGP information along with past data. It is plausible that in such scenarios individuals will rely more on the other data source at their disposal, namely the past data from previous participants. We therefore ask whether participants in our 3-Dice PartialDGP treatment are more inclined to the follow the past data than those in the 3-Dice PastData treatment.

Hypothesis B. 3. ***(Partial information about the DGP)*** *Participants become more naive when there is a reduction in the availability of information about the data generating process—the average participant in the 3-Dice PartialDGP treatment is further from Bayesian than the average participant in the 3-Dice PastData treatment.*

We will test this hypothesis by testing whether $\mu^n_{3D.PDGP} < \mu^n_{3D.PD}$.

D.2.4 Can cues shift the way individuals think about the problem?

Some forms of cognitive mistakes may emanate from individuals approaching a particular problem in the wrong way—i.e., they may *represent* the problem in the wrong way.[54] Many statistical biases take this form. Here, we are focused on selection neglect, which involves individuals treating a conditional distribution as if it representative of the unconditional distribution that they are interested in. The next hypothesis explores whether simply asking individuals a series of leading questions can induce them to approach the problem differently.[55] Importantly, the questions themselves contain no additional information, but simply aim to shift how an individual thinks about the problem.

[54]For example, imagine a school student that has learnt two algorithms for solving particular types of algebra problems—each algorithm is appropriate for solving one class of problems, but these classes are mutually exclusive. When faced with a new problem, he focuses on certain features of the problem and tries to assess which is the appropriate algorithm to use. If he represents the features of the problem incorrectly, he may adopt the wrong algorithm.

[55]This is reminiscent of the approach taken by Socrates in Plato's "Meno".

Hypothesis B. 4. ***(Socratic Method)*** *Participants become less naive when they are prompted (through means of a series of questions) to think about the investment problem in a different way—the average participant in the 3-DICE CUE treatment is closer to Bayesian than the average participant in the 3-DICE PASTDATA treatment.*

We will test this hypothesis by assessing whether $\mu^n_{3D.CUE} > \mu^n_{3D.PD}$.

D.2.5 Does additional information about the past data help alleviate naivety?

As noted above, a fully rational Bayesian agent does not need to draw on the past data at all in this experiment (except in the 3-DICE PARTIALDGP treatment), since they know the data generating process. However, if individuals do draw on the past data, a question of interest is whether having additional information about this past data that they are observing influences how they process it. Specifically, in the 3-DICE EXTRAINFO treatment, in addition to receiving information about the percentage of past projects that were successful for each dice value, participants also learn about the number of projects that were invested in for each value of the purple dice. Under the assumption that all the past participants followed the same threshold strategy, this information would be sufficient to reveal what that threshold was.

We ask whether being provided access to this additional information about the past projects helps participants to reduce their naive extrapolation. This could operate in different ways. First, participants could draw on the additional information directly and try to learn from the threshold followed by past participants (however, since the past participants had less information than them, this doesn't seem like an attractive strategy). Second, the additional information could help to alert them to the selection present in the data and therefore induce them to think more carefully about the decision problem.

Hypothesis B. 5. ***(Additional information about the past data)*** *Participants become less naive when they are provided with additional information about the past data—the average participant in the 3-DICE EXTRAINFO treatment is closer to Bayesian than the average participant in the 3-DICE PASTDATA treatment.*

We will test this hypothesis by assessing whether $\mu^n_{3D.EI} > \mu^n_{3D.PD}$.

D.2.6 Is being exposed to past data harmful?

Here we ask whether participants are better off with less information. Specifically, we compare a scenario in which a participant is provided with access to past data to a scenario in which she is not provided with access to past data, i.e. 3-DICE PASTDATA and 3-DICE NOPASTDATA. Since Part C of these two treatments are not comparable (in 3-DICE NOPASTDATA, the Naive Lottery probabilities are not well defined because the participants don't observe past data), we cannot use μ^n as our metric for comparison. However, we can construct a simpler measure that just uses Part A and Part B and reflects the distance between the Part A investment decision and the individual's behavior when they hold Bayesian beliefs. Specifically, define $\sigma_i^n = 1 - \frac{\tilde{n}_i - n_i^b}{10}$. Again, the closer this metric is to 1, the closer is investment behavior in Part A and Part B, implying that the individual's Part A behavior is closer to Bayesian (in the sense of holding Bayesian beliefs).

Hypothesis B. 6. ***(Harmful information)*** *Participants become less Bayesian when they are exposed to past data—the average participant in the 3-Dice NoPastData treatment is closer to Bayesian than the average participant in the 3-Dice PastData treatment.*

We test this hypothesis by assessing whether $\sigma^n_{3D.NPD} > \sigma^n_{3D.PD}$.

Appendix D.3: Results for pre-registered hypotheses for Experiment 2

Table 12 summarizes the results from all six of the hypotheses discussed above.[56] The left-most column shows that we observe a substantial amount of naivety in the 3-Dice PastData treatment. The average participant has a μ^n that is less than 0, which is significantly below the Bayesian benchmark of $\mu^n = 1$ ($p < 0.01$) and close to the naive benchmark of $\mu^n = 0$. This suggests that on average individuals are engaging in investment behavior in the investment game in Part A that is more similar to that in the Naive Lottery than in the Bayesian Lottery. This conclusion is consistent with the picture that emerges from a visual inspection of the cdf of investment behavior in the Investment Game (Part A) in comparison to the Bayesian Lottery (Part B) and the Naive Lottery (Part C). The left-hand panel of Figure 6 depicts the cdfs for the 3-Dice PastData treatment. The figure suggests that at the aggregate level, behavior in the Investment Game more closely resembles that in the Naive Lottery task in comparison to that in the Bayesian Lottery task. Taken together, these results provide strong support for the finding from Experiment 1 that individuals tend to make investment choices in the Investment Game as if they hold naive probabilities as their beliefs, rather than Bayesian probabilities. Since we find this result in a new participant pool with a simplified version of the game and a within-individual benchmark for Bayesian and Naive beliefs, this lends support to the conclusion that naive extrapolation from data is a robust finding.

Result B. 1. *(Hypothesis B.1) The average participant in the 3-Dice PastData treatment is (at least) partially naive. Observed behavior in the Investment Game is more similar to that in the Naive Lottery than that in the Bayesian Lottery.*

Table 12: Hypothesis tests from Experiment 2

	Hyp B.1	Hyp B.2	Hyp B.3	Hyp B.4	Hyp B.5	Hyp B.6
Alternative Hypothesis (H_a):	$\mu^n_{3D.PD} < 1$	$\mu^n_{3D.PD} < \mu^n_{2D.PD}$	$\mu^n_{3D.PD} > \mu^n_{3D.PDGP}$	$\mu^n_{3D.PD} < \mu^n_{3D.CUE}$	$\mu^n_{3D.PD} < \mu^n_{3D.EI}$	$\sigma^n_{3D.PD} < \sigma^n_{3D.NPD}$
Mean (3D.PD)	-0.09	-0.09	-0.09	-0.09	-0.09	0.83
Std err. (3D.PD)	0.07	0.07	0.07	0.07	0.07	0.01
Mean (Comparison Group)	1.00	0.27	0.14	-0.11	-0.02	0.82
Std err. (Comparison Group)		0.06	0.12	0.07	0.11	0.02
Diff.	1.09	0.36	0.23	-0.03	0.07	-0.01
T-Statistic	14.61	3.15	1.68	-0.23	0.50	-0.60
p-value	<0.01	<0.01	0.95	0.59	0.31	0.73
N	202	303	299	301	303	302
H_0 Rejected	Y	Y	N	N	N	N

Notes: (i) The table contains the tests of the pre-registered hypotheses from Experiment 2, (ii) The column headers denote the alternative hypothesis (H_a), where the null hypothesis (H_0) always involves a test that the difference equals 0, (iii) Hypothesis 1 involves a one-sample test and Hypothesis 6 involves a different outcome variable.).

[56] Table 14 in the Appendices replicates this analysis for the restricted subset of individuals who did not violate stochastic dominance by investing more often in the Bayesian Lottery than in the Naive Lottery. The results remain very similar.

The second column of Table 12 shows that individuals in the 2-Dice PastData moved closer to behaving as if they held Bayesian beliefs in the Investment Game in comparison to individuals in the 3-Dice PastData treatment. This is indicated by the shift in the average μ^n from -0.09 to 0.27, implying a statistically significant difference of 0.36 ($p < 0.01$). While we do observe evidence that reducing the complexity of the data generating process results in a significant reduction in naivety, it is important to also note that the μ^n observed in the 2-Dice PastData treatment is still far away from the Bayesian benchmark of $\mu^n = 1$, indicating that there remains a large amount of naive investment behavior even in this setting with a simpler DGP.

Result B. 2. *(Hypothesis B.2) Participants become less naive as the complexity of the data generating process is reduced—the average participant in the 2-Dice PastData treatment is closer to Bayesian than the average participant in the 3-Dice PastData treatment.*

Columns 3, 4, 5, and 6 of Table 12 show that we fail to reject the null hypothesis for each of the remaining four hypotheses. This suggests that in our setting, none of these four channels influence the propensity of individuals to invest in the Investment Game. These results suggest several insights. First, individuals over-invest in our setting even without access to past data. This can be seen in the right-most column of Table 12 (Hyp. B.6), which shows that participants in the 3-Dice NoPastData treatment are not closer to Bayesian than those in the 3-Dice PastData treatment.[57] This indicates that participants are not very good at inferring the Bayesian probabilities from their knowledge of the data-generating process in this context, which is also a feature of many real-world scenarios where individuals may try to learn from observing past outcomes of others. Second, the treatments that try to vary the difficulty of drawing inference directly from the 3-dice DGP information did not shift investment behavior (Hyp B.3, B.4, and B.5). This indicates that in contrast to the reduction of complexity present when moving to the simpler 2-dice setting, the average participant appears to have a cognitive constraint that leads them to form a noisy mental model of the mapping from the DGP to the success probabilities in the 3-dice setting. Third, importantly, even with this cognitive constraint in accurately extracting Bayesian probabilities from the DGP information, individuals who are not naive should still move closer to acting like a Bayesian when they receive information about the success frequencies of projects that were invested in. Upon seeing the past data, if an individual recognizes and appreciates the selection process that generates the success frequencies that she observes, she should discount the observed success frequencies and therefore invest less in the Investment Game than in the Naive Lottery. Taken together, these results suggest that naivety is quite a robust phenomenon that is difficult to alleviate in scenarios where the true DGP is complex, and it is non-trivial to extract the relevant Bayesian probabilities from knowledge about the DGP.

[57]It can also be seen in Table 16, which summarizes the propensity to invest in the Investment Game in the two NoPastData treatments as well as the two corresponding PastData treatments.

Unfortunately, this is the case for many real-world scenarios of interest.

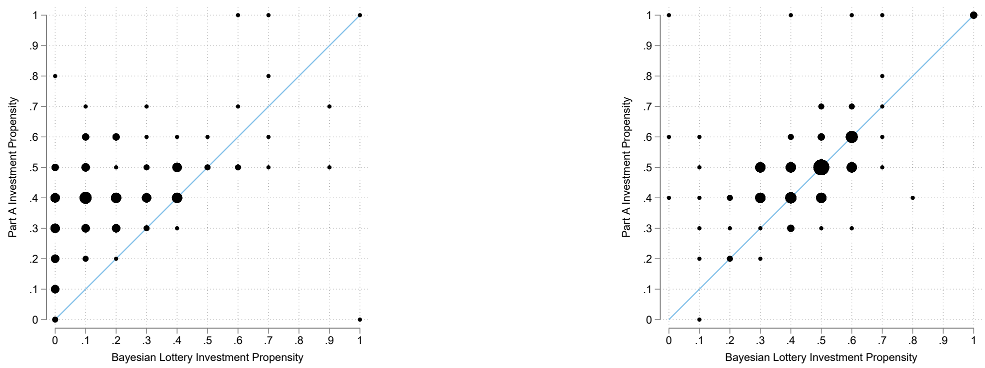
Result B. 3. *(Hypotheses B.3, B.4, B.5, B.6) When the DGP is complex, participants do not become less naive when provided with additional information or cues about the DGP, nor do they become more naive when there is a partial reduction in the availability of information about the DGP. Instead, the majority of individuals appear to be willing to follow the past data naively.*

Appendix D.4: Examining overinvestment in the absence of past data

One of the features of the decision context used in the majority of our treatments is that a risk-neutral Bayesian should invest in fewer than 50% of the investment opportunities that she faces. This means that biases other than the naivety that we are studying here, such as a tendency to shade towards a cognitive default of investing in 50% of the available investment opportunities, could lead to over-investment (Enke and Graeber, 2023). As discussed in the main text, this is not an issue for the aggregate-level between-treatment comparisons in Experiment 1 and Experiment 2 (since the decision context is held constant across treatments).

For the individual-level analyses in Experiment 2, a simple bias toward investing in 50% of the available investment opportunities would also not present an issue for comparing Part A, B, and C and, therefore, would not influence our main results or classification exercises (since, there again, the decision context is held constant across parts). However, if such a bias is related to the complexity of the decision context, then it is plausible that it could exert a stronger influence on investment decisions in Part A relative to Part B and C. For example, this might occur if cognitive uncertainty is larger in Part A relative to Part B and C. This does not seem implausible. Consider an individual in the 3-Dice NoPastData treatment. In Part A, this individual only receives a description of the data-generating process (DGP) and must infer the probability of success associated with each of the dice values. Since this may be a challenging exercise for at least a subset of participants, they may shade their assessments toward holding beliefs that the probability of success is closer to 50%. Since all of the Bayesian probabilities are below 50%, this shading towards 50% could inflate the propensity to invest. Provided this occurs more in Part A (where cognitive uncertainty is plausibly higher) than in Part B (where cognitive uncertainty is plausibly lower), this mechanism could lead to over-investment in Part A.

Figure 18: Individual investment propensities (comparing two 3D NoPastData treatments)



(a) **3D NoPastData**

(b) **3D17 NoPastData**

Consistent with this idea, we do see over-investment (relative to Bayes) even in the absence of past data. For example, in Figure 4, we see over-investment in the first round of Experiment 1. Similarly, in Figure 18a, we see that there is more investment in the investment game than in the Bayesian lottery for the majority of participants in the 3-Dice NoPastData treatment. This raises the following question: Could (part of) the overinvestment observed in the PastData treatments be driven by a mechanism other than naively learning from the past data? We argue that even the naivety that we detect in our individual-level analysis is robust to this concern for the following two reasons.

First, even if individuals form success probability beliefs that are inflated towards 50% when they only receive a description of the DGP (as in the 3-Dice NoPastData treatment), there is still scope for learning from the past data and adjusting their belief downward if they are not naive and factor in the presence of selection. Consider an individual who, upon receiving a description of the DGP, forms the belief that the probability of success after observing a dice value of 7 is 50%. Now, she observes selected past data with an empirical success frequency of 40% for a dice value of 7. If the individual takes into account the positive selection in the data, she should adjust her belief downwards to end up well below 40%. Even if she observed an empirical success frequency of 50%, she should adjust her belief downwards if she realizes that the data is positively selected.[58]

Second, to directly examine this question, we added two additional treatments with a lower success threshold, 3-Dice17 NoPastData and 3-Dice17 PastData. The rationale for adding these two treatments is that they shift the Bayesian probabilities of success upwards. This implies that a bias that distorts beliefs towards 50% will no longer result in overinvestment relative to Bayes. Figure 18b shows that the lower success threshold is effective in shifting beliefs closer to the Bayesian benchmark in the absence of past data (i.e., in 3-Dice17 NoPastData). Therefore, since we do observe individuals overinvesting when past data is added in the 3-Dice17 PastData treatment (see Figure 8a), this indicates that naivety is causing individuals to shift their beliefs and invest too much.

This can also be seen clearly in Table 13. Columns (*a) and (*b) show the distributions of individual investment behavior with and without past data by comparing choices in the investment game and the Bayesian lottery. The first two columns (1*) report the results for a project success threshold of 22, while the last two columns (2*) do so for a success threshold of 17. In column (1a), we see that without past data, consistent with the discussion above, 78% of individuals invest more in the investment game than in the Bayesian lottery. When past data is added in column (1b), 82% of individuals invest more in the investment game. Importantly, the vast majority of these individuals are investing precisely the same amount in

[58]An analogous hypothetical example is the following. Suppose that there is a university competition where students from a particular class all complete a quiz. Thereafter, each student chooses whether to submit their quiz and enter the competition or not. A spectator is asked to estimate the average quiz score of all students in the class—she guesses that the average is 70 points out of 100. Now, suppose the spectator is told that the average score of those who entered the competition is 70 points and is asked whether she would like to adjust her guess. Even though the information coincides with the spectator's guess, if she takes into consideration the likely positive selection into the competition, she should lower her guess.

the investment game and Naive lottery in 3-Dice PastData (see Figure 7b). Therefore, even though the fraction of individuals overinvesting is similar with and without past data when the success threshold is 22, the fact that the majority of individuals invest *exactly* the same amount in the investment game and Naive lottery when they have past data suggests that they do believe that empirical frequencies observed in the past data are the true probabilities.

Nevertheless, these results do not cleanly parse the naivety explanation from the above-mentioned cognitive uncertainty explanation. Columns (2a) and (2b), however, do successfully achieve this. Column (2a) shows that with the lower threshold for success, 42% of individuals invest more in the investment game than in the Bayesian lottery without past data. When past data is added in column (2b), the classification is nearly identical to that in column (1b)—now, 82% of individuals invest more in the investment game. Furthermore, Figure 8b shows that the majority of them are behaving in exactly the same way in the investment game and the Naive lottery, which suggests that they hold the same beliefs in both. This shows that naivety is playing an important role in driving the results that we observe.[59]

Table 13: Classification of individuals relative to Bayes.

	3-Dice		**3-Dice17**	
	NoPastData (1a)	**PastData** (1b)	**NoPastData** (2a)	**PastData** (2b)
Less than Bayes ($\tilde{n}_i < n_i^{BL}$)	8	7	22	7
Exactly Bayes ($\tilde{n}_i = n_i^{BL}$)	14	11	36	11
More than Bayes ($n_i^{BL} < \tilde{n}_i$)	78	82	42	82
Observations (N)	100	202	100	100

Notes: (i) The table provides a classification of individuals into discrete types according to their investment behavior in Part A, B and C of Experiment 2, (ii) Each of the columns reports the distribution of types within a particular treatment in percentage points, (iii) The first two columns report the distributions for the 3-Dice NoPastData and 3-Dice PastData treatments, while the last two columns report the same information for the 3-Dice17 NoPastData and 3-Dice17 PastData treatments, (iv) When comparing this table to Table 4, it is worth taking into consideration the following. First, since there is no naive benchmark in the NoPastData treatments, this table only compares the choices in the investment game to the Bayesian benchmark, resulting in only three types. This allows a comparison between the NoPastData and PastData treatments, which is not possible using the more fine-grained classification employed in Table 4. Second, this simplified classification pools together several categories from Table 4 and categories such as "Violates Stochastic Dominance" are not identified here. This means that the category labels do not have exactly the same meaning between tables—for example, "Exactly Bayes" here is not identical to the category labelled "Exactly Bayes" in Table 4. The precise definitions are denoted by the comparisons in parentheses next to the category labels.

[59]It is important to note that we are not claiming that cognitive uncertainty is playing no role. Rather, it does seem to be an important factor for understanding behavior in the NoPastData treatments. However, the results here illustrate that it is not explaining the coincidence of behavior between the Naive lottery and the investment game in the PastData treatments. In other words, the results robustly show that naivety is playing an important role in inflating beliefs and generating overinvestment.

Appendix D.5: Additional tables and figures from Experiment 2

Table 14: Hypothesis tests from Experiment 2 (Restricted Sample)

	Hyp B.1	Hyp B.2	Hyp B.3	Hyp B.4	Hyp B.5	Hyp B.6
Alternative Hypothesis (H_a):	$\mu^n_{3D.PD} < 1$	$\mu^n_{3D.PD} < \mu^n_{2D.PD}$	$\mu^n_{3D.PD} > \mu^n_{3D.PDGP}$	$\mu^n_{3D.PD} < \mu^n_{3D.CUE}$	$\mu^n_{3D.PD} < \mu^n_{3D.EI}$	$\sigma^n_{3D.PD} < \sigma^n_{3D.NPD}$
Mean (3D.PD)	-0.13	-0.13	-0.13	-0.13	-0.13	0.81
Std err. (3D.PD)	0.06	0.06	0.06	0.06	0.06	0.01
Mean (Comparison Group)	1.00	0.20	0.08	-0.16	-0.07	0.79
Std err. (Comparison Group)		0.06	0.11	0.07	0.11	0.02
Diff.	1.13	0.33	0.21	-0.03	0.06	-0.02
T-Statistic	17.77	3.37	1.76	-0.29	0.48	-1.20
p-value	<0.01	<0.01	0.96	0.61	0.31	0.88
N	187	280	279	283	283	281
H_0 Rejected	Y	Y	N	N	N	N

Notes: (i) The table contains the tests of the pre-registered hypotheses from Experiment 2, (ii) The column headers denote the alternative hypothesis (H_a), where the null hypothesis (H_0) always involves a test that the difference equals 0, (iii) Hypothesis 1 involves a one-sample test and Hypothesis 6 involves a different outcome variable.), (iv) The sample is restricted to the set of individuals who did not violate stochastic dominance by investing more often in the Bayesian Lottery (Part B) than the Naive Lottery (Part C).

Table 15: Average investment propensity across 3-Dice22 PD treatments (Investment Game)

	(1)		(2)		(3)		(4)		(5)		(6)		(7)	
	3DicePastData		**3DiceExtraInfo**		**3DicePartialDGP**		**3DiceCUE**		**3DPD:3DEI**		**3DPD:3DPDGP**		**3DPD:3DCUE**	
	mean	sd	mean	sd	mean	sd	mean	sd	diff.	t-stat.	diff.	t-stat.	diff.	t-stat.
Investment Propensity	0.401	(0.211)	0.395	(0.219)	0.434	(0.221)	0.387	(0.192)	-0.006	(-0.225)	0.033	(1.229)	-0.014	(-0.580)
Observations	202		101		97		99		303		299		301	

Notes: (i) The first eight columns [labelled (1) to (4)] report the mean and standard deviation of participants' propensity to invest across the ten possible dice values in the investment game (Part A), (ii) The final six columns [labelled (5) to (7)] report the mean comparison of each of the treatments considered in (2) to (4) against the 3DicePastData treatment in (1); t-statistics are reported in parentheses, (iii) Symbols: * for $p < 0.1$, ** for $p < 0.05$, *** for $p < 0.01$.

Table 16: Average investment propensity across 3-Dice22 and 2-Dice Baseline Treatments

	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
	3D PastData		**3D NoPastData**		**2D PastData**		**2D NoPastData**		**3DPD:3DNPD**		**3DNPD:2DNPD**		**3DPD:2DPD**		**2DPD:2DNPD**	
	mean	sd	mean	sd	mean	sd	mean	sd	diff.	t-stat.	diff.	t-stat.	diff.	t-stat.	diff.	t-stat.
Investment Propensity	0.401	(0.211)	0.425	(0.190)	0.351	(0.152)	0.333	(0.203)	0.024	(0.995)	0.092***	(3.310)	-0.050**	(-2.337)	0.018	(0.731)
Observations	202		100		101		100		302		200		303		201	

Notes: (i) The first eight columns [labelled (1) to (4)] report the mean and standard deviation of participants' propensity to invest across the ten possible dice values in the investment game (Part A), (ii) The final eight columns [labelled (5) to (8)] report mean comparisons of the treatments considered, with the relevant comparison described in the column header; t-statistics are reported in parentheses, (iii) Symbols: * for $p < 0.1$, ** for $p < 0.05$, *** for $p < 0.01$.

Table 17: Average investment propensity in 3-Dice17 PastData and 3-Dice17 NoPastData

	(1)		(2)		(3)	
	3D17 PastData		**3D17 NoPastData**		**3D17PD:3D17NPD**	
	mean	sd	mean	sd	diff.	t-stat.
Investment Propensity	0.602	(0.175)	0.507	(0.184)	-0.095***	(-3.741)
Observations	100		100		200	

Notes: (i) The first four columns [labelled (1) and (2)] report the mean and standard deviation of participants' propensity to invest across the ten possible dice values in the investment game (Part A), (ii) The final two columns [labelled (3)] report the comparison of the means the two treatments considered; t-statistics are reported in parentheses, (iii) Symbols: * for $p < 0.1$, ** for $p < 0.05$, *** for $p < 0.01$.

Table 18: Comparison of investment game and lottery games

	Investment Game	Bayes Lottery	Naive Lottery	Diff. (Bayes:Invest)	Diff. (Naive:Invest)
3-Dice 22 Past Data (Pooled)	0.40	0.22	0.37	0.18***	0.03***
	(0.21)	(0.24)	(0.21)	(22.70)	(4.46)
N	499	499	499		
2-Dice Past Data	0.35	0.22	0.35	0.13***	0.00
	(0.15)	(0.18)	(0.18)	(9.80)	(0.28)
N	101	101	101		
3-Dice 17 Past Data	0.60	0.45	0.60	0.15***	0.00
	(0.17)	(0.19)	(0.19)	(10.41)	(0.12)
N	100	100	100		

Notes: (i) This table contains comparisons of the rate of investment in Part A, B and C of Experiment 2, (ii) The 3-Dice 22 sample consists of individuals in the four 3-Dice22 treatments where participants observed past data, pooled together, (iii) The statistics in parentheses below the group means denote standard deviations, while those below differences denote t-statistics, (iv) Symbols: * for $p < 0.1$, ** for $p < 0.05$, *** for $p < 0.01$.

Table 19: Comparison of investment game and lottery games (Restricted Sample)

	Investment Game	Bayes Lottery	Naive Lottery	Diff. (Bayes:Invest)	Diff. (Naive:Invest)
3-Dice 22 Past Data (Pooled)	0.39	0.20	0.37	0.19***	0.02***
	(0.21)	(0.22)	(0.20)	(25.33)	(3.52)
N	471	471	471		
2-Dice Past Data	0.34	0.20	0.35	0.14***	-0.01
	(0.15)	(0.17)	(0.18)	(10.21)	(-1.03)
N	93	93	93		
3-Dice 17 Past Data (Pooled)	0.59	0.43	0.61	0.16***	-0.02
	(0.16)	(0.17)	(0.17)	(11.93)	(-1.55)
N	94	94	94		

Notes: (i) This table contains comparisons of the rate of investment in Part A, B and C of Experiment 2, (ii) The 3-Dice 22 sample consists of individuals in the four 3-Dice22 treatments where there was past data, pooled together, (iii) The statistics in parentheses below the group means denote standard deviations, while those below differences denote t-statistics, (iv) The sample restriction removes individuals who violated stochastic dominance between Part B and Part C, (iv) Symbols: * for $p < 0.1$, ** for $p < 0.05$, *** for $p < 0.01$.

Table 20: Classification of individuals into discrete types (percentages).

	All 4 PD22 Treatments	**3-Dice PD**	**2-Dice PD**
	(1)	(2)	(3)
Less than Bayes ($\tilde{n}_i < n_i^{BL} < n_i^{NL}$)	3.6	3.5	5.9
Exactly Bayes ($\tilde{n}_i = n_i^{BL} < n_i^{NL}$)	3.8	3.0	5.9
Bayes = Naive = Invest ($\tilde{n}_i = n_i^{BL} = n_i^{NL}$)	6.8	5.9	9.9
Partially Naive ($n_i^{BL} < \tilde{n}_i < n_i^{NL}$)	6.6	8.4	12.9
Exactly Naive ($n_i^{BL} < \tilde{n}_i = n_i^{NL}$)	46.5	47.0	37.6
Extremely Naive ($n_i^{BL} < n_i^{NL} < \tilde{n}_i$)	27.1	24.8	19.8
Violates Stochastic Dominance ($n_i^{BL} > n_i^{NL}$)	5.6	7.4	7.9
Observations (N)	499	202	101

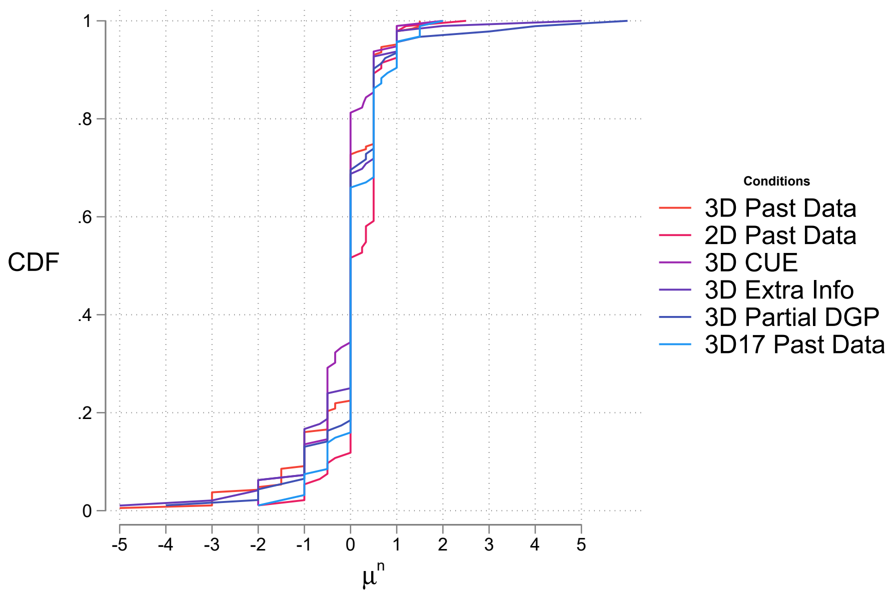
Notes: (i) The table provides a classification of individuals into discrete types according to their investment behavior in Part A, B and C of Experiment 2, (ii) The columns report the distribution of types in percentage points, (iii) Column (1) pools together all individuals in the 4 PastData treatments with a threshold of 22 (i.e., 3-Dice PastData, 3-Dice PartialDGP, 3-Dice CUE, and 3-Dice ExtraInfo).

Table 21: Correlates of μ^n

	3-Dice PastData Only	All 4 PD22 Treatments
	(1)	(2)
Female = 1	-0.349***	-0.223**
	(0.129)	(0.087)
At Least Some College = 1	-0.014	-0.053
	(0.174)	(0.110)
Age	0.079**	0.041**
	(0.032)	(0.021)
Age^2	-0.001**	-0.000*
	(0.000)	(0.000)
Zero Understanding Mistakes = 1	0.061	-0.001
	(0.134)	(0.090)
Time Spent on Part A Decision Screen (seconds)	0.003*	0.001
	(0.002)	(0.001)
Constant	-1.795***	-0.918**
	(0.689)	(0.453)
Treatment FEs	N	Y
Observations	186	468
Adjusted R^2	0.047	0.014

Notes: (i) OLS regressions include one observation per individual, (ii) The dependent variable is μ^n, (iii) The variable "At Least Some College" is a binary variable that takes a value of one if the individual reported an education level higher than high school, (iv) The regressions exclude individuals who violated stochastic dominance by investing more often in Bayesian Lottery than in the Naive Lottery (15 individuals or 7% in column 1 and 28 individuals or 6% in column 2), as well as individuals who chose not to report their gender as male or female (1 individual in column 1 and 3 individuals in column 2), (v) Standard errors are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 19: CDF of μ^n across treatments (Restricted sample)



Appendix E: Instructions for the SELECTED treatment

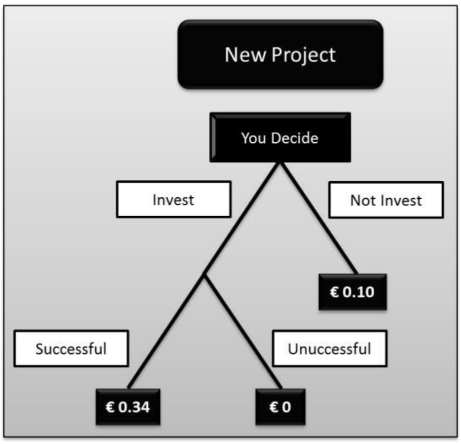
In Game I, you are going to have the opportunity to make a series of investment decisions. For each decision, you will be presented with a "project". Every project will turn out to either be a *successful project* or an *unsuccessful project*. Each project will have three "attributes". These attributes will be related to whether the project will be successful or not. You will be able to observe one of the three attributes for every project you face. (This will always be the same attribute).

In order to help you learn about whether a project you face will be a *successful project* or an *unsuccessful project,* you will be placed in a group with two other participants and you will be able to learn from the success or failure of the past investments of your group members. In particular, you will be able to observe whether projects similar to the one that you are currently considering were successful or not. Projects are "similar" when the attribute you observe is exactly the same between the projects. (We will provide you with more details on the exact information you will receive below).

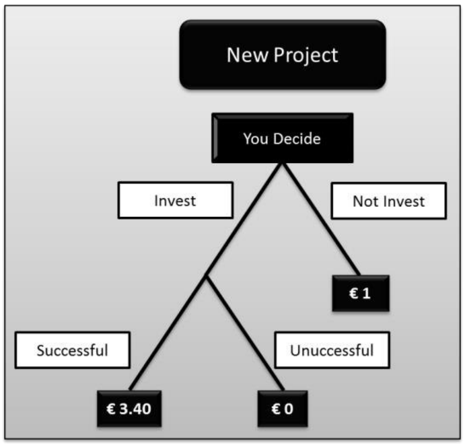
There will be two phases in the Investment Game. In Phase 1 (Low Stakes), each time you face a new project, you will be given an amount of money, €0.10. This is exactly the amount that it costs to make the investment. You can then choose whether to INVEST this €0.10; or NOT INVEST and keep the €0.10 to be paid at the end of the experiment. If you INVEST and the project is *successful*, you will be paid a high prize, which has value €0.34. If you INVEST and the project is *unsuccessful*, you will receive nothing and you will lose the €0.10 that you invested.

Phase 2 (High Stakes) is exactly the same, except all the amounts are multiplied by ten. This explanation is summarized in the following diagram.

Phase 1: Low Stakes



Phase 2: High Stakes



In **Phase 1 (Low Stakes)**, you will face low cost, low prize investment opportunities. In this phase, the cost of investment will be **€0.10** and the prize when the investment is successful will be **€0.34**. This phase will allow you to learn how the investment game works, and also to learn about which projects are likely to be *successful projects* and which projects are likely to be *unsuccessful projects*.

In **Phase 2 (High Stakes)**, the game is exactly the same. In particular, the chances that a project will be *successful* or *unsuccessful* are exactly the same as in Phase 1. The only difference is that the cost of investment will now be **€1.00** and the prize when the investment is successful will be **€3.40**.

Phase 1 (Low Stakes) investment costs and returns

Decision	Successful Project	Unsuccessful Project
Invest	Receive €0.34	Receive €0
Do Not Invest	Receive €0.10	

Phase 2 (High Stakes) investment costs and returns

Decision	Successful Project	Unsuccessful Project
Invest	Receive €3.40	Receive €0
Do Not Invest	Receive €1.00	

Learning about the project:

In order to help you learn about whether a project will be *successful* or *unsuccessful,* you will be able to observe whether similar projects that were invested in by your group members in the past were successful. The text below provides you with details explaining: (i) what is meant by "similar" projects; and (ii) which past projects you will observe.

Who is in your group?

As mentioned above, you will be randomly assigned into a group with two other participants in this experiment (three in total, including yourself – called Group Members **A**, **B** and **C**). You will stay in this group for both Phase 1 and Phase 2 of the experiment.

Which past projects will you observe?

In order to learn about which projects are likely to be successful, you will be able to observe the success of projects that were invested in by the other members of your group. You will also observe one of the attributes of these past projects.

What are project attributes?

Every project has three attributes, called **a**, **b** and **c**. For every project, each of these three attributes takes an integer[1] value between 1 and 10. The attribute values are determined by the computer rolling three ten-sided dice – i.e. Attribute **a** is equal to the number shown on the purple dice, called Dice **a**; Attribute **b** is equal to the number shown on the red dice, called Dice **b**; and Attribute **c** is equal to the number shown on the green dice, Dice **c**. All three dice are fair dice (i.e. each dice has an equal chance of showing every number between 1 and 10).

[1] An integer is a whole number, so each attribute takes one of the following ten values: 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10.

Who observes which attributes?

Each of the three group members (including you) will observe only one of the three attributes – and each group member will observe a different attribute from the other two.

More specifically, Group Member **A** will always observe Attribute **a**, Group Member **B** will always observe Attribute **b**; and Group Member **C** will always observe Attribute **c**.

When is a project successful?

The success of a project is determined by adding up the three attribute values (i.e. adding up the numbers on the three dice). If this total is equal to 22 or more (i.e. 22 to 30) then the project is *successful*; if the total is equal to 21 or less (i.e. 3 to 21) then the project is *unsuccessful*.

Example: What information does Group Member C observe?

*Information about current project that Group Member **C** is considering*

If, for example, you are Group Member **C**, then when you are considering a new project, you will always be able to observe Attribute **c** for this new project before you decide whether to invest in it.

*Information about past projects that Group Members **A** and **B** invested in*

As Group Member **C**, you would also have access to information about the Attribute **c** value of all projects that other members of your group invested in in the past. You will also be told whether these past projects were successful or not. In other words, you will know what proportion of the projects that other members of your group invested in with a particular Attribute **c** value were successful.

An example

Let's consider a concrete example to clarify this. Consider the following hypothetical project:

Attribute a	Attribute b	Attribute c	Successful / Unsuccessful
10	5	7	**Successful**

The information above is a complete summary of the project - it contains all three attribute values and the information about its success.

Recall that the three attribute values are determined by throwing three fair ten-sided dice (Dice **a**, Dice **b**, and Dice **c**). Furthermore, notice that the project is "successful" because adding up the three attribute values = 10 + 5 + 7 = 22, which is between 22 and 30. Of course, as explained above, a participant in the experiment cannot observe such a complete description of projects.

Rather, assume that Group Member **C** is the one that has the opportunity to invest in this project. When she is deciding whether or not to invest, she will only observe the value of Attribute **c**:

Attribute a	Attribute b	Attribute c	Successful / Unsuccessful
		7	

In addition, Group Member **C** will also observe whether other past projects that (i) had a value of 7 for Attribute **c** *and* (ii) that either **A** or **B** invested in in the past; were successful or not. (You will see more detailed information about this below.)

*Now, what information will be revealed after **C** makes her decision?*

If she **decides <u>not</u> to invest**, then nobody, neither **A**, **B** nor **C** will receive any further information about this project.

If she **decides to invest**, then **after** the investment has been made, Group Member **A** will observe the following data about the project:

Attribute a	Attribute b	Attribute c	Successful / Unsuccessful
10			Successful

while Group Member **B** will observe:

Attribute a	Attribute b	Attribute c	Successful / Unsuccessful
	5		Successful

Notice that, while Group Members **A** and **B** observe whether the project that Group Member **C** invested in is successful or not, **C** herself will only find out at the end of the experiment. Group Members will never receive immediate feedback on the success of their own projects.

Summary

The following summarises some of the key information from above:

- The game is completely symmetric for the three players (i.e. there is no difference between being assigned to be Group Member **A**, **B** or **C** at the start of the game. The only difference between players is the information they receive about the other two group members' past investments).
- Each of the three participants in your group will observe the value of one of the three attributes for: (i) all the projects they face, as well as (ii) all projects that other members of their group have previously invested in. Each will observe a different attribute.
- You will only observe the outcome of projects that your group members invested in.
- You won't observe the outcome of the projects that you invest in until the end of the experiment.
- The relationship between an attribute value and the chances of the corresponding project's success will stay the same throughout the experiment.
- The information on past projects accumulated in Phase 1 is carried over to Phase 2.

Collecting and summarising the information about past projects

Since the other members of your group may invest in many projects, the computer will organize the information about these past projects for you.

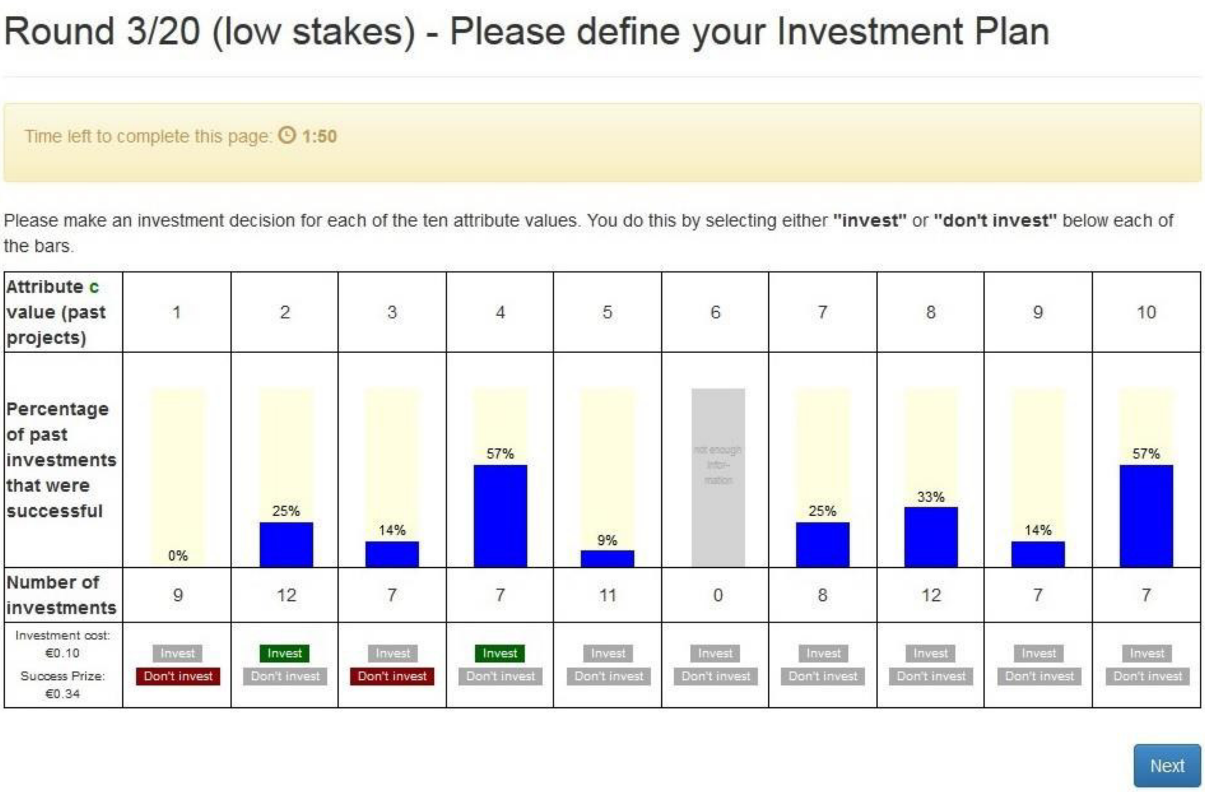
Instead of showing you the individual data from each of the projects that other members of your group have invested in, the information for all past projects invested in by other members of your group will be summarized in the following way. If you are Group Member **C**, then the computer will collect together all of these projects which have the same Attribute **c** value and tell you:

(i) The number of times other members of your group (**A** and **B**) invested in a project with that value of Attribute **c**; and

(ii) The proportion of projects invested in by other members of your group (**A** and **B**) with the same Attribute **c** value that were successful.

Specifically, Group Member **C** will see a screen that looks similar to the following – except with different information on success rates (other participants will see a similar screen, corresponding to their own attribute):

Figure 1: The Information Screen of Group Member C about Past Projects



Note: In the example in Figure 1, Group Member **A** and **B** have invested in 7 projects with an Attribute **c** value of 4. Of these 7 projects, 57% were successful. Group Member **C** sees this information.

Timeline for decision making

The text above has described the information that you will receive when making a decision about investing in a **single new project**. However, in the experiment, there will be 10 rounds of investment decisions in Phase 1 and 10 rounds of investment decisions in Phase 2. This makes 20 rounds of investment decisions in total. In each of these rounds of investment decision, you will be asked to state whether you would like to INVEST or NOT INVEST in **50 projects**.

In each round, the way you will do this is by telling the computer whether you would like to INVEST or NOT INVEST in projects with each of the 10 possible Attribute values (i.e. you make ten decisions and are providing the computer with an "Investment Plan" for how to act on your behalf – see the decision buttons in the bottom row of Figure 1). In each round, the computer will then see 50 randomly selected new projects, and it will follow your "Investment Plan" to decide whether to INVEST or NOT INVEST in each of these projects on your behalf. So, you make 10 decisions, and the computer uses these decisions to decide whether to INVEST or NOT INVEST in 50 projects on your behalf.

Providing the computer with an "Investment* Plan*" for investing on your behalf

More specifically, we will ask you to report whether you would like to INVEST or NOT INVEST in **projects with each possible Attribute value between 1 and 10**. Therefore, during each round of decisions, we will ask you to make 10 investment decisions[2] – one for each possible Attribute value between 1 and 10. The other members of your group will also report whether they would like to INVEST or NOT INVEST for each Attribute value between 1 and 10 for the Attribute they observe (Attribute **a** for Group Member **A**; **b** for **B**; **c** for **C**).

Computer acts according to your "Investment Plan"

In each round of decisions, after every member of the group has made their 10 decisions, each group member will face 50 randomly selected new projects. Given the choices that you have made, the computer will look at the relevant Attribute value of each of these 50 projects you face and then INVEST or NOT INVEST, according to the "Investment Plan" you gave it (i.e. if you draw a project which has an Attribute value of 6 and you said that you would like to invest in projects with an Attribute value of 6, then the computer will invest in this project on your behalf). You can change your "Investment Plan" for the next round of decisions.

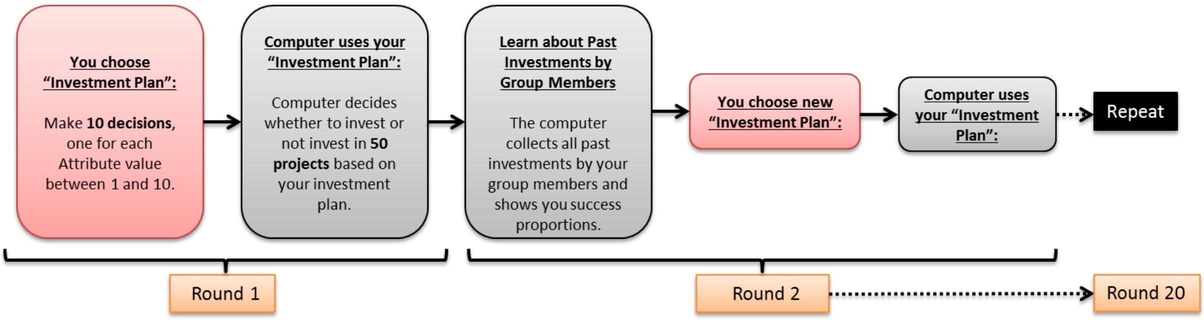
The "Investment Plan": An example:

If this is a little bit complicated, it might be simpler to think about the "Investment Plan" that you give the computer in the following way. If you are Group Member **C**, and you select INVEST for an Attribute **c** value equal to 6, then you are telling the computer: "In this round, every time you see a project with an Attribute **c** value of 6, please choose INVEST on my behalf". The computer will then go through 50 randomly chosen new projects and act on your behalf as you have instructed it.

[2] i.e. we will ask if you would like to INVEST or NOT INVEST when the Attribute value is 1; when the Attribute value is 2; when the Attribute value is 3; …; when the Attribute value is 10.

Updating the information you observe about past projects:

Once these 50 project decisions have been made, your database of information regarding the past success rates of projects invested in by other Group Members will be updated with their new investments, and you will be given the opportunity to revise your “Investment Plan” which will determine your investment strategy for the next 50 projects.



How your payment will be calculated

As described above, in every round of investment decisions, you will make 10 decisions. These 10 investment decisions will provide the computer with an “Investment Plan” for how to act on your behalf when faced with the next 50 randomly drawn new projects. **In every round, one of these 50 projects will be randomly selected to be the one that affects your payment in the experiment.** For this project selected for payment, the computer will act as you have instructed it to and either INVEST or NOT INVEST. If your “Investment Plan” prescribes that the computer invest for this particular project then the contribution to your payment will depend on whether the project is *successful* or *unsuccessful*, as described above.

- Therefore, the **maximum** you can earn in Game I is: 10 * €0.34 + 10 * €3.40 = **€37.40** if you always invest <u>and</u> every project you invest in is successful.
- The **minimum** you can earn in Game I is **€0** if you always invest <u>and</u> every project you invest in is unsuccessful.
- If you never invest in Game I, you would earn: 10 * €0.10 + 10 * €1 = **€11**.

You will not learn about the outcomes of your own investments during the experiment. At the end of the experiment, you will be informed about the investment decisions that you made that are relevant for your payment. In each of the two phases, for the ten projects chosen to contribute to your payment, you will learn: the number of projects that you chose to invest in, the number that were successful, and how they contributed to your final payment.

We will now proceed to Game I. Before we do, if you have any questions at this moment, please raise your hand. The experimenter will come to you.