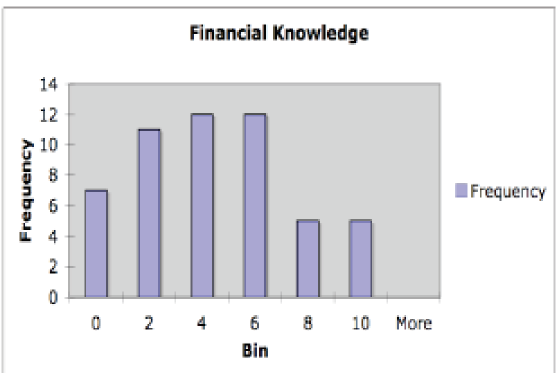
SI Appendix.

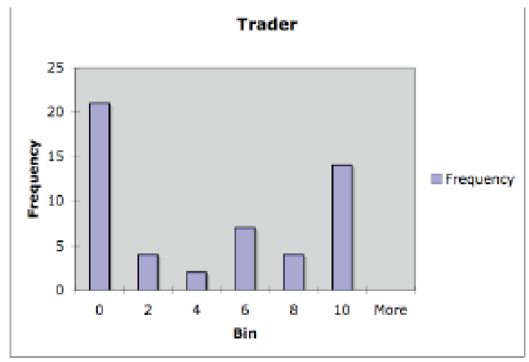
A. Questionnaire Data. Subjects answered three questions related to their investment knowledge and experience:

1. “I am educated in economics and finance” (Financial Knowledge).
2. “I make my own investment decisions” (Own Investments).
3. “I trade (or have traded) stocks bonds or commodities” (Trader).

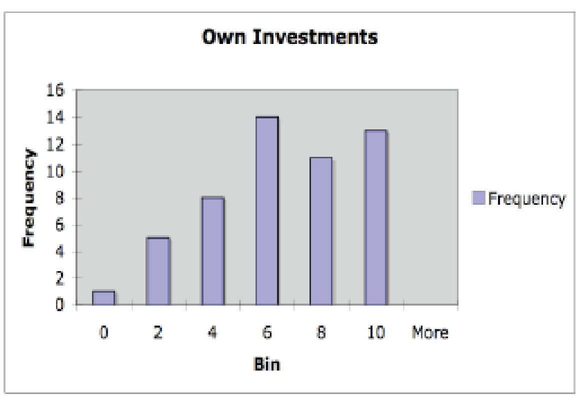
The answers were on a scale of 0-10 with 0 being “Strongly Disagree” and 10 being “Strongly Agree”. Below are histograms of the answers to these questions (n = 52, two reports missing).



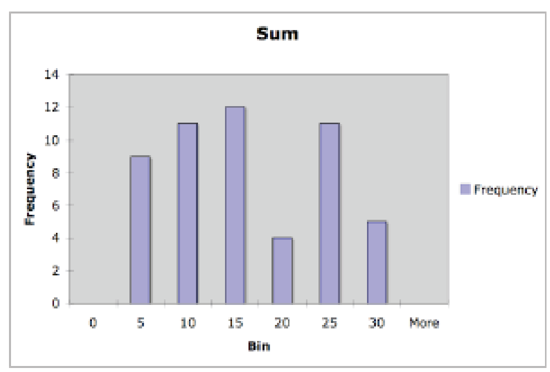
Bin	Frequency
0	7
2	11
4	12
6	12
8	5
10	5
More	0



Bin	Frequency
0	21
2	4
4	2
6	7
8	4
10	14
More	0



Bin	Frequency
0	1
2	5
4	8
6	14
8	11
10	13
More	0



Bin	Frequency
0	0
5	9
10	11
15	12
20	4
25	11
30	5
More	0

We also looked at the association between the scores on the three questions and overall performance in the experiment. Due to a record-keeping error 13 of the questionnaires are not unambiguously associable with a particular subject, leaving 41 subjects. For this cohort, regressions of the score (% gain in the experiment) versus the three question scores as well as the sum of the scores, gave only one significant regressor, the score on the "do you handle your own investments" (INVEST) question. The regression results were:

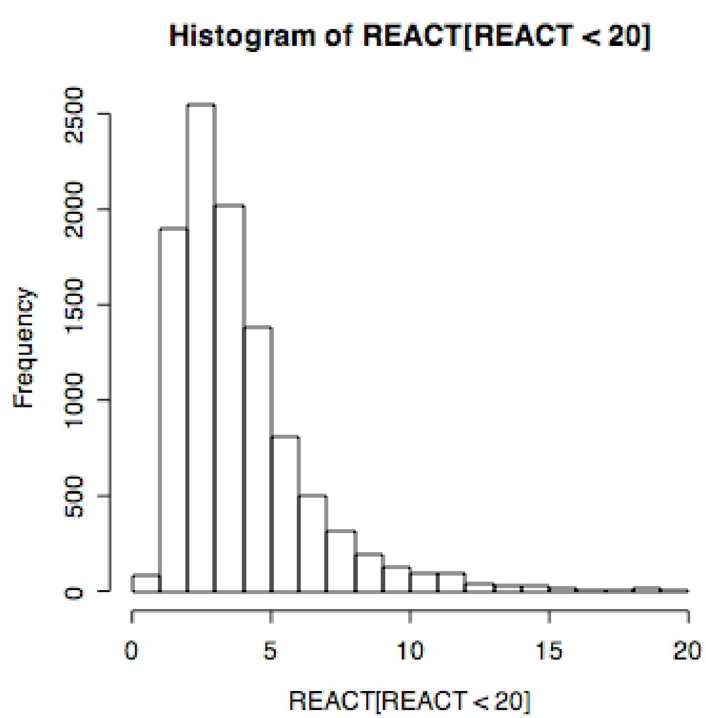
Coefficient	**Estimate**	**t-value**	P(>|t|)
Intercept	.066	.637	.5278
INVEST	.029	1.842	.0732

B. Reaction Time Data

1. Summary statistics for all reaction times:

Min	1st Qu.	Median	Mean	3rd Qu.	Max.
0.797	2.234	3.266	4.058	4.812	123.000

2. Histogram for all reaction times < 20 secs (66 events out of 10260 were > 20 secs.).



3. Below we show some single variable regressions of reaction time with variables of interest (Again REACT is the time interval in seconds between the reveal of a market snippet and the next bet submission. NEGMKT is the absolute value of the market return when it is negative, POSMKT is the absolute value of the market return when it is positive, and BETNUM is the bet-number in a subject's sequence. Note: the regressions were made for the market return positive and negative as appropriate i.e. the POSMKT regression is only over reaction time following positive market returns). While some of the regressors reach significance (N is large) only one, BETNUM, actually explains much variance.

Regression	*Estimate*	*t-value*	*$p(>\lvert t \rvert)$*	*Adj. R-Squared*
REACT~NEGMKT	*2.681*	*2.498*	*.0125*	*.001*
REACT~POSMKT	*3.245*	*2.956*	*.003*	*.001*
REACT~FICTIVE	*6.093*	*4.332*	*1.50e-05*	*.003*
*REACT~BET*NEGMKT*	*9.113*	*4.942*	*8.01e-07*	*.005*
*REACT~BET*POSMKT*	*-1.330*	*-.712*	*.476*	*.000*
REACT~BETNUM	*-.016*	*-26.02*	*<2e-16*	*.062*

C. Activation Data

Activation Tables. All tables are for activations with p < 1e-05 (uncorrected) and cluster size ≥ 5. Coordinates are Talairach. Note: Only the activation table for Fig. 5 had deactivations.

Fig. 4. Activation Table

Fictive Error

cluster	cluster size	x	y	z	t-value	Structure
1	260	-24	-68	56	8.27	Posterior Parietal Cortex L
2	184	16	64	60	8.00	Posterior Parietal Cortex R
3	178	-20	8	-4	7.75	Ventral/Dorsal Striatum L
4	194	16	12	-4	7.46	Ventral/Dorsal Striatum R
5	170	-32	52	-4	7.20	Dorsolateral Prefrontal Cortex L
6	34	16	52	44	6.53	Superior Frontal Gyrus R
7	162	-28	-4	52	6.39	Middle Frontal Gyrus L
8	35	28	-92	12	6.05	Middle Occipital Gyrus R
9	14	56	32	-4	5.74	Inferior Frontal Gyrus R
10	29	28	0	56	5.63	Middle Frontal Gyrus R
11	8	-8	20	48	5.51	Medial Cingulate L
12	17	-28	-80	28	5.38	Cuneus L
13	9	-36	40	36	5.37	Middle Frontal Gyrus L
14	11	-8	-12	-8	5.35	Ventral Striatum L
15	11	-48	20	4	5.34	Inferior Frontal Gyrus L
16	8	40	56	-4	5.28	Middle Frontal Gyrus R
17	20	-12	60	-4	5.25	Medial Frontal Gyrus
18	10	56	-28	-8	5.08	Middle Temporal Gyrus R

PosMktNL

cluster	cluster size	x	y	z	t-value	Structure
1	598	24	-96	12	8.45	Middle Occipital Gyrus
2	8	-32	56	-12	5.5	Superior Frontal Gyrus

Fig. 5. Activation Table.

TD (return - bet) (Activations)

cluster	cluster size	x	y	z	t-value	Structure
1	114	16	12	-12	8.09	Ventral Striatum R
2	199	60	-56	28	7.77	Supramarginal Gyrus R
3	114	-16	8	-12	7.53	Vetral Striatum L
4	283	-48	-64	48	7.37	Inferior Parietal Lobe L
5	286	-44	44	-8	7.10	Dorsolateral Prefrontal Cortex L
6	214	-32	16	52	6.95	Middle Frontal Gyrus L
7	59	40	24	36	6.28	Middle Frontal Gyrus R
8	74	20	-28	72	6.24	Precentral Gyrus R
9	18	16	40	48	5.49	Superior Frontal Gyrus R
10	9	-16	40	-12	5.42	Middle Frontal Gyrus L
11	8	60	-72	-4	5.31	Inferior Temporal Gyrus R
12	6	36	-84	36	5.26	Precuneus R
13	8	40	40	36	4.97	Superior Frontal Gyrus R

TD (return - bet) (Deactivations)

cluster	cluster size	x	y	z	t-value	Structure
1	7	0	-36	-4	5.61	corpus collosum
2	7	-26	-60	4	5.48	superior colliculus
3	6	4	16	12	5.28	middle occipital gyrus white matter

Fictive Error

cluster	cluster size	x	y	z	t-value	Structure
1	69	-20	-68	56	7.22	Superior Parietal Lobule L
2	41	20	-64	60	6.82	Superior Parietal Lobule R
3	57	-20	-80	-8	6.67	Lingula Gyrus L
4	51	-20	-100	8	6.52	Cuneus R
5	24	-8	8	4	5.89	Caudate Head L
6	30	-32	-8	56	5.78	Precentral Gyrus L
7	18	8	12	4	5.39	Caudate Head R
8	5	0	-12	12	4.86	Thalamus L

Fig. 6. Activation Table

TD (Q-learning)

cluster	cluster size	x	y	z	t-value	Structure
1	746	12	8	-8	10.06	Ventral Striatum R (Note extends to L)
2	400	-28	-72	56	7.88	Superior Parietal Lobule L
3	212	20	-100	8	7.75	Cuneus R
4	248	20	-60	52	6.79	Superior Parietal Lobule R
5	37	-56	8	32	6.17	Middle Frontal Gyrus L
6	18	28	-8	52	5.91	Middle Frontal Gyrus R
7	56	-20	-16	60	5.58	Middle frontal Gyrus L
8	25	-4	-4	72	5.58	Superior Frontal Gyrus L
9	5	20	-20	68	5.33	Precentral Gyrus R
10	9	-28	16	44	5.30	Middle Frontal Gyrus L
11	6	-12	32	44	5.07	Medial Frontal Gyrus L
12	9	0	-32	20	4.93	Corpus Callosum/Cingulate Gyrus
13	5	24	-24	20	4.82	white matter

Fictive Error

cluster	cluster size	x	y	z	t-value	Structure
1	49	16	-64	60	7.09	Superior Parietal Lobule R
2	72	-32	-4	56	6.01	Middle Frontal Gyrus L
3	33	-24	-68	56	5.95	Superior Parietal Lobule L
4	21	52	-24	-8	5.7	Middle Temporal Gyrus R
5	35	-12	12	4	5.68	Caudate Head L
6	23	-52	-56	44	5.66	Inferior Parietal Lobule L
7	6	-12	32	56	5.43	Superior Frontal Gyrus L
8	11	56	32	-4	5.33	Inferior Frontal Gyrus R
9	13	16	52	44	5.16	Superior Frontal Gyrus R
10	5	16	12	0	4.82	Putamen/Caudate Head R

D. Regressor Descriptions

Description of Regressors. The following regressors were present in all models and were entered as the SPM canonical hemodynamic response functions time-locked to the events listed.

1) Market Type Screen: when Screen comes on saying if market id “Live” (L) or “Not Live” (NL).
2) Clear Screen 1: when Market Type Screen goes off.
3) Initial Reveal NL: when initial segment of market history is displayed in the NL condition.

4) First Reveal NL: first market reveal in NL condition after initial reveal.
5) Reveal NL: market reveals 2- 19 in NL condition.
6) Final Reveal NL: 20th market reveal in the NL condition.
7) ClearScreen2NL: when the screen went blank after the final market reveal in the NL condition.
8) Initial Reveal L: when the initial segment of market history is revealed in the L condition.
9) First Reveal L: first market reveal after the initial reveal.
10) Reveal L: market reveals 2- 19 in the L condition.
11) Final Reveal L: 20th market reveal in the L condition.
12) ClearScreen2L: when the screen went blank after the final market reveal in the L condition.
13) Keypress: all keypresses except the actual submit. Keypresses that were closer than 2 seconds apart were collapsed into the first keypress.

A regressor for the actual choice submission, and the 'bar on' (bar going from gray to red) was no explicitly included since they were temporally so close to the event 'reveal'.

Additionally, each subject's head-motion data (6 parameters) were entered as covariates.

A regressor PosMktNL was also included all models. It was constructed by modulating (multiplication pointwise in time) the RevealNL regressor by the function r_t^+ (the value of the market return in the NL condition when the return was positive, 0 otherwise).

For Fig. 4, the regressor FictiveError was added to the basic model. It was constructed by modulating (multiplication pointwise in time) the RevealL regressor by f_t^+ (the fictive error, see main text).

For Fig. 5, two regressors were added to the basic model. The regressor TD was entered into the model. It was constructed by modulating (multiplication pointwise in time) the RevealL regressor by TD (the basic TD error; see text). The second regressor was the FictiveError (described above) orthogonalized with respect to TD. This operation was described in the main text.

For Fig. 6, two regressors were added to the basic model. The regressor QTD was entered into the model. It was constructed by modulating (multiplication pointwise in time) the RevealL regressor by QTD (the TD error derived from the Q-learning model; see text). The second regressor was the FictiveError (described above) orthogonalized with respect to QTD. This operation was described in detail in the main text.

D. Details of 'Endogenous Supervisor' Calculation

The fictive error can be connected to concepts recently developed in machine learning. Rosenstein and Barto outline a theory combining reinforcement and supervised learning (36). Consider a child learning to throw a ball at a target (36). Whether the ball goes to the left or right of the target, and by how much, constitutes an "evaluation signal" received by the child. On the other hand coach (a "supervisor") watching the throw may contribute error information in the form of explicit instruction as to what went wrong. These two forms learning are combined for the actor's error signal in an actor-critic structure. We interpret the fictive error as the report of an 'endogenous supervisor'. More formally (following (36) closely), if we denote the actor's policy by $\pi^a(\theta)$ where θ is a vector of parameters, then the actor 's policy update is given by

$$\theta \leftarrow \theta + k\Delta^{RL} + (1-k)\Delta^{SL} .$$

Here $k \in [0,1]$ is a parameter that measures the relative weight of the reinforcement versus supervised aspects of learning. The reinforcement update term is

$$\Delta^{RL} = \alpha\ \delta\ \nabla_\theta \pi^a$$

where

$$\delta = r + \gamma V(S_{t+1}) - V(S_t)$$

is the familiar TD error, and α is the learning rate.

The supervisory contribution to the total error is chosen to minimize

$$E = \frac{1}{2}(\pi^S - \pi^A)^2 .$$

Steepest descent dictates that the update is given by

$$\Delta^{SL} = -\alpha \nabla_\theta E = (a^S - a^A)\nabla_\theta \pi^A ,$$

where a^S and a^A are the actions given by the supervisor and actor respectively.
This is close to our fictive error, but not quite. Intuitively, errors in policy should not count as much when the market return is small. Thus we modify the error to include an importance weight $|r|$, the absolute value of the market return:

$$E = \frac{1}{2}|r|(\pi^S - \pi^A)^2 .$$

Now the same calculation as above yields:

$$\begin{aligned}\Delta^{SL} &= |r|(a^S - a^A)\nabla_\theta \pi^A \\ &= r^+(1 - a^A)\nabla_\theta \pi^A - r^- a^A \nabla_\theta \pi^A\end{aligned}$$

The first term on the bottom right is the fictive error of this paper, and the second term is identical to loss, since ‘shorting’ (negative bets) was not allowed.