Supplemental Material

Data and code availability

Data for all three experiments and custom code for the computational models are available at https://github.com/ycleong/AdviceTaking.

Bayesian Learning model

We adapted the ideal observer model previously described in Waskom, Frank, and Wagner (2016) to model how participants tracked an advisor's expertise. Each advisor is assumed to have a level of accuracy, a, which determines the probability that he will give accurate advice, y, on any given trial:

$$p(y = 1|a) = a \quad \text{and} \quad p(y = 0|a) = 1 - a \tag{1}$$

The model does not assume a fixed a, but instead allows it to change from trial to trial. Under this assumption, recent outcomes have a greater effect on the estimate of a than outcomes further in the past. Specifically, we assumed changes in a to be a Markov Decision Process, where a on trial $t + 1$ is determined by:

$$p(a_{t+1}|\, a_t, v_t) = \beta(a_t, e^{v_t}) \tag{2}$$

That is, a on trial $t + 1$ is drawn from a Beta distribution with mean a_t and variance e^{v_t}. Following Behrens et al. (2007), we used a reparameterization of the Beta distribution such that the first parameter denotes the mean of the distribution while the second parameter determines the width of the distribution. The parameter v scales with the expected rate of change of a: A large value of v implies a wide distribution, implying that a changes greatly from trial-to-trial, while a small value of v implies a narrow distribution, implying that a changes slightly from trial-to-trial. As mentioned in the main text, v functions as an "optimal learning rate" estimated using Bayesian inference.

The above description constitutes a *generative model* of how an advisor's accuracy relate to observed outcomes, i.e. $(p(y_{1:t}|a_{1:t}, v_{1:t})$. Using Bayes rule, we can invert this model to compute the posterior probability of an advisor's accuracy given a series of outcomes. Specifically, we can compute the posterior probabilities over parameters a and v *for any history* of observed outcomes from trial 1 to t, $y_{1:t}$ as

$$p(a_{1:t}, v_{1:t}|y_{1:t}) \propto p(a_1)p(v_1)\prod_{j=1}^{t} p(y_j|a_j)p(a_j|a_{j-1}, v_j) \tag{3}$$

where $p(a_1)$ and $p(v_1)$ denote the prior on a and v respectively, while $p(y_j|a_j)p(a_j|a_{j-1}, v_j)$ computes the likelihood of observing outcome y_j given accuracy a_j, weighted by the probability of observing a_j based on the transition function of a_j from a_{j-1}.

To get the posterior distribution of a and v at a *particular* trial i, $p(a_i, v_i|y_{1:i})$, we integrate over the history of a up to trial i:

$$p(a_i, v_i|y_{1:i}) \propto \int p(a_1)p(v_1)\prod_{j=1}^{i} p(y_j|a_j)p(a_j|a_{j-1}, v_j)da_{1:i-1} \tag{4}$$

How can we compute this integral on each trial in a tractable manner? We note that the posterior distribution of a and v on the previous trial, $i-1$, can be written as:

$$p(a_{i-1}, v_{i-1}|y_{1:i-1}) \propto \int p(a_1)p(v_1)\prod_{j=1}^{i-1} p(y_j|a_j)p(a_j|a_{j-i}, v_j)da_{1:i-2} \tag{5}$$

and $p(a_{i-1}, v_{i-1}|y_{1:i-1})$ can be substituted into equation 4, such that:

$$p(a_i, v_i|y_{1:i}) \propto p(y_i|a_i) \int p(a_{i-1}, v_{i-1}|y_{1:i-1})p(a_i|a_{i-1}, v_i)da_{i-1} \tag{6}$$

$p(y_i|a_i)$ is given by equation 1 and $p(a_i|a_{i-1}, v_i)$ is given by equation 2. As such, only the joint distribution of a and v, i.e. $p(a_i, v_i|y_{1:i})$ needs to be saved from trial to trial.

To obtain the distribution of a at trial i, we marginalize over v,

$$p(a_i) = \int p(a_i, v_i) dv_i \tag{7}$$

We then compute the mean estimate of the advisor's accuracy on trial i, or $\hat{a}_i$, as

$$\hat{a}_i = \int a_i \, p(a_i) \, da_i \tag{8}$$

The prior on a, that is the initial belief about the advisor prior before seeing any outcomes, i.e. $p(a_1)$ was defined by a Beta distribution,

$$a_1 \sim Beta(\alpha, \beta) \tag{9}$$

in which α and β were shape parameters that were fit to participants' bets (details in main text).

Confirmation Bias model

The confirmation bias model differs from the Bayesian Learning model in how it relates the likelihood of an outcome to the accuracy of an advisor (Eq. 1). In the Bayesian Learning model, the likelihood of an outcome is independent of the current estimate of the advisor's accuracy. In the Confirmation Bias model, the likelihood function of an unexpected outcome is the weighted combination of the likelihood of the observed outcome and that of the most likely outcome. We denote this new likelihood function as $p_b(y|a)$ to distinguish it from the Eq. 1. Specifically, if the current estimate of the advisor's accuracy is above 0.5, and the advisor gives inaccurate advice (Fig. S1 top row), the likelihood function of this unexpected outcome is given as,

$$p_b(y = 0|a, b) = b(p(y = 1|a)) + (1 - b)(p(y = 0|a)) \tag{10}$$

In contrast, if the current estimate of the advisor's accuracy is less than 0.5, and the advisor gives accurate advice (Fig. S1 bottom row), the likelihood function is then given as,

$$p_b(y = 1|a, b) = b(p(y = 0|a)) + (1 - b)(p(y = 1|a)) \tag{11}$$

b is a bias term that weights influence of expectations in computing the likelihood function. b changes from trial-to-trial and scales linearly with participants' current estimates of the advisor's accuracy:

$$b = | \hat{a} - 0.5 | \tag{12}$$

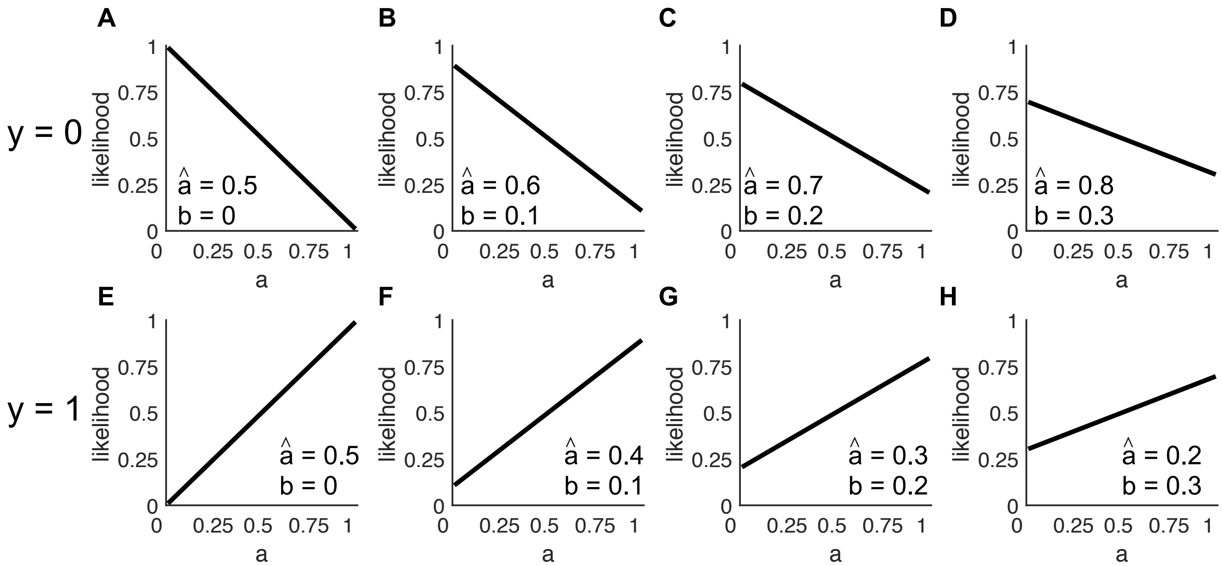

Figure S1. Likelihood function ($p_b(y|a,b)$) of observing inaccurate (top row) and accurate advice (bottom row) used by the Confirmation Bias model. A. Likelihood of observing inaccurate advice from an advisor expected to be at chance. The downward slope indicates that it is highly likely for an inaccurate advisor (e.g., $a = 0$) to give inaccurate advice and highly unlikely for an accurate advisor (e.g., $a = 1$) to give inaccurate advice. This is also the 'unbiased' likelihood function for observing inaccurate advice used by the Bayesian Learning model. **B-D.** Likelihood of observing inaccurate advice from an advisor expected to be accurate. As the estimate of the advisor's accuracy increases from 0.6 to 0.8, the slope of the likelihood function decreases. That is, there is greater likelihood that an accurate advisor (e.g., $a = 1$) gives inaccurate advice, and as such, inaccurate advice is *weaker* evidence for an inaccurate advisor. **E.** Likelihood of observing accurate advice from an advisor expected to be at chance. This is also the 'unbiased' likelihood function for observing accurate advice used by the Bayesian Learning model. **F-H**. Likelihood of observing accurate advice from an advisor expected to be inaccurate. As the estimate of the advisor's accuracy decreases from 0.4-0.2, the model assigns greater likelihood to an inaccurate advisor giving accurate advice, such that observing accurate advice is *weaker* evidence for an accurate advisor. $\hat{a}$: current estimate of the advisor's accuracy. b: bias term that determines the degree of confirmation bias.

Trial-by-trial comparison of Bayesian Learning model and Confirmation Bias model

To understand how the Bayesian Learning model and Confirmation Bias model differ, we simulated the two models learning about an advisor with chance accuracy. For both models, we assumed an optimistic prior defined by $a \sim \beta(5,3)$ ($\hat{a}_1 = 0.63$, Fig. S2A). We then plot the posterior belief distribution about the advisor's accuracy after having observed the outcome on the previous trial (y_{t-1}; Fig. S2B-J). After observing 9 outcomes, the Bayesian Learning model (correctly) estimates the advisor's accuracy to be 0.5. In contrast, the Confirmation Bias model remains optimistic about the advisor's accuracy.

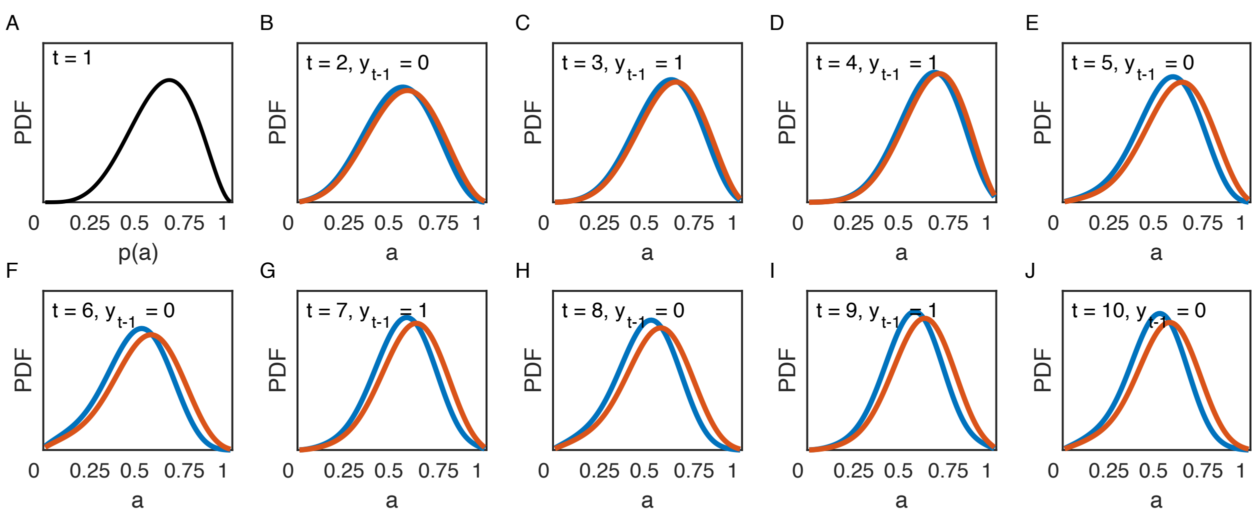

Figure S2. Trial-by-trial comparison of the posterior beliefs of the Bayesian Learning and Confirmation Bias models over the course of 10 trials. A. Prior belief about the advisor's accuracy at trial 1. For both models, we assumed an optimistic prior defined by $a \sim \beta(5,3)$ with $\hat{a}_1 = 0.63$. **B-J**. Posterior beliefs on subsequent trials. This advisor has chance accuracy. **J.** After observing 9 outcomes, the Bayesian learning model (correctly) estimates the advisor's accuracy to be 0.5. In contrast, the Confirmation Bias model remains optimistic about the advisor's accuracy. t = trial number, y_{t-1} = outcome observed on the previous trial.

Parameter Recovery Simulations

To examine if the two models were identifiable, we fit the models to data simulated with known values of α, β, and τ and assessed if we were able to accurately recover the true values of the model parameters. We ran 1000 simulations of the Confirmation Bias model performing the task with $\alpha = 3$, $\beta = 2$ and $\tau = 9$. We then fit the Confirmation Bias model to the simulated dataset and plot the recovered parameter values in Figure R6A. Estimates of τ were close to the true value of 9. Estimates of α and β were highly correlated, with a fair amount of variance around the true parameter values. However, we note that the estimates of $\hat{a}_1$, which depend on the ratio of α and β, were tightly defined around the true value of 0.6 with an interquartile range of 0.07. As statistical inference was performed over $\hat{a}_1$, the uncertainty over α and β do not confound our claims about participants' priors. Similar results were obtained when the Bayesian Learning model was fitted to a corresponding dataset simulated with the same parameter values, though with greater parameter uncertainty (Fig. 6B).

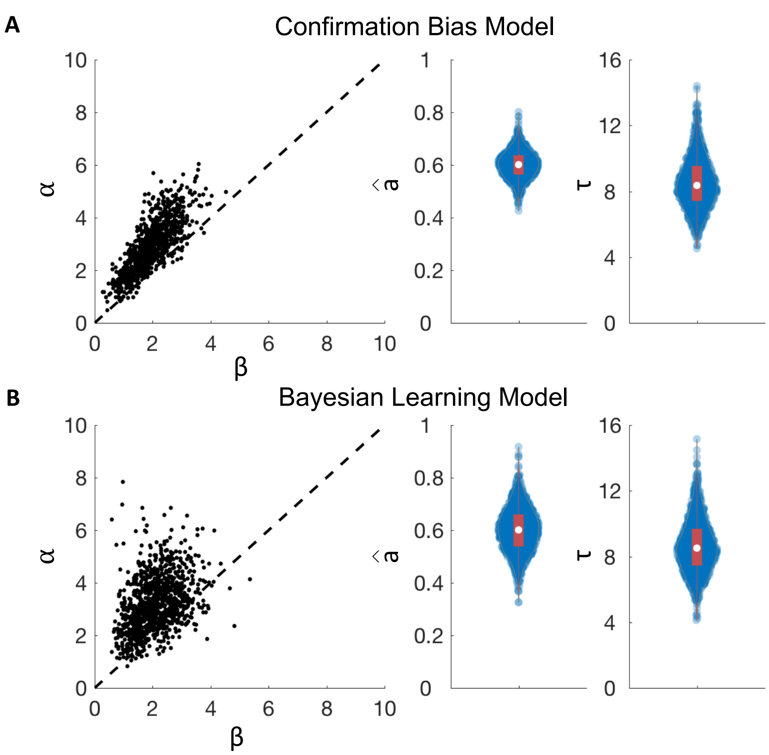

Figure S3. Parameter estimates when the models were fitted to simulated data (1000 iterations) with parameters $\alpha = 3, \beta = 2, \tau = 9$. A. Confirmation Bias Model. B. Bayesian Learning Model. Left. Each dot represents the best-fit values of α and β from 1 iteration of the simulation. Best-fit values of α and β are highly correlated. **Middle and Right**. Violin plots showing the distribution of best-fit $\hat{a}_1$and τ values. The median estimates (white dot) are close to the true values of 0.6 and 9 respectively. Red box indicates interquartile range.

Additional Comparison Models

Win-stay-lose-shift (WSLS) model. This model conditions the next bet on the previous bet. In particular, the probability of making the same bet after a correct bet ($P(stay|win)$) and the probability of making the opposite bet after an incorrect bet ($P(shift|lose)$ are free parameters estimated by fitting the model to each participant's data.

Reinforcement-learning model. This model learns an "expected value" ($E(V)$), for each advisor using a temporal difference learning rule (Sutton & Barto, 1998). Correct predictions are assigned a value of +1 and incorrect predictions are assigned a value of -1. On each trial t, the model computes a prediction error (δ_t), by taking the difference between the expected value of the advisor and whether the advisor was correct on that trial (R_t):

$$\delta_t = R_t - E(V_t)$$

The prediction error is used to update the expected value, scaled by a learning rate (η) that is fit to each participant's data:

$$E(V_{t+1}) = E(V_t) + \eta\delta_t$$

The model then makes bets based on the expected value of an advisor according to the same logistic function choice rule used by the other two models (see Main Text).

Null Model. This model fits a constant $p(bet = FOR)$ to each participant. The Null model is equivalent to an intercept-only model commonly used as a baseline model for regression analyses. If a participant has a general tendency to bet for or against an advisor regardless of an advisor's accuracy, this model will capture such behavior. The "Null" model serves as a baseline model for comparison with the other models.

Model Comparison. The corrected average likelihood per trial of the Confirmation Bias model was credibly higher than that of all the other models (Table S1), indicating that the Confirmation Bias model provided the best account of participants' data in the Advisor Evaluation phase. The WSLS model was the second best-fitting model. When we examined the best-fit parameters of the WSLS model, the average best-fit value of $P(stay|win)$ was 0.85 (SE = 0.03), indicating that participants were more likely to make the same bet after a correct bet. The average best-fit value of $P(shift|lose)$ was 0.37 (SE = 003), indicating that even after an incorrect bet, participants were still more likely to stay with the same bet than they were to switch to the opposite bet ($P(stay|lose)$ = 1 - $P(shift|lose)$ = 0.63). In other words, the model was implementing a "win-stay-lose-**stay**" strategy. This pattern of behavior is indicative of belief perseverance, and provides additional evidence that participants exhibited confirmation bias in how they learned about advisors' expertise. Unlike the Confirmation Bias model however, the WSLS model does not accumulate information over multiple trials, but instead relies only on the outcome of the previous trial to make the next bet. It would be interesting to investigate if a WSLS model that accumulates information over successive trials (e.g., see Worthy & Todd Maddox (2014) for a WSLS-RL hybrid) would approximate the behavior of the Confirmation Bias model, though this is beyond the scope of the current paper.

Model	Avg Lik	μ_{CB-X}	Effect Size
Confirmation Bias	0.65 [0.60, 0.70]	-	-
Bayesian Learning	0.62 [0.58, 0.66]	0.027 [0.006, 0.046]	0.612 [0.158, 1.056]
Win-stay-lose-shift	0.63 [0.59, 0.67]	0.022 [0.005, 0.039]	0.522 [0.095, 0.953]
TD-Learning	0.60 [0.57, 0.64]	0.044 [0.023, 0.064]	0.964 [0.473, 1.474]
Null	0.51 [0.50, 0.51]	0.143 [0.095, 0.193]	1.232 [0.691, 1.786]

Table S1. Additional model comparison results. Avg Lik = corrected average likelihood per trial averaged across participants. μ_{CB-X} = mean posterior estimate of the difference in corrected average likelihood per trial between the Confirmation Bias model and each comparison model in the first column. Also reported are the effect sizes when comparing μ_{CB-X} to 0. Square brackets denote the 95% HDI of the corresponding posterior distribution.

Robust Bayesian Estimation

Bayesian estimation was performed following the procedure described in Kruschke (2013). The statistic of interest was assumed to be described by a t-distribution, with posterior estimates of the mean (μ), standard deviation (σ), normality (ν) estimated from the data using Markov chain Monte Carlo (MCMC) with non-committal priors. Specifically, the prior on μ is a very broad normal distribution, with mean set to the mean of the sample, and standard deviation set to 1000 times the standard deviation of the sample. These settings allow the prior to be appropriately scaled to the scale of the data, while having minimal influence on the estimation. The prior on σ is a uniform distribution from a low value of one thousandth of the standard deviation of the sample, to a high value of one thousand times the standard deviation of the sample. The prior on ν is exponentially distributed, and balances nearly normal distributions (ν > 30) with heavy tailed distributions (ν < 30). MCMC was performed with a chain length of 90000, no thinning, and 1000 burn-in samples. We defined the 95% highest density interval (HDI) of the posterior distribution of μ as the 95% credible interval of μ. We refer to a "credible effect" whenever the 95% HDI does not include the comparison value. For example, if the 95% HDI of the posterior distribution of the mean within-participant difference in corrected average likelihood per trial (μ_{CB-BL}) does not include 0, we can conclude that that the corrected average likelihood per trials of the two models are credibly different. Effect size was defined as $\frac{\mu-0}{\sigma}$.

Fitting the models to the first 12 time periods

Our models were fit to data from all time periods. If the models do not provide a good characterization of participants' learning (i.e. learning too fast or too slow), the estimates of the priors could be unduly biased by the later trials. To investigate this possibility, we refit the models to only the first 12 trials for each advisor (out of 36 in Experiments 1, 28 in Experiment 2 and 20 in Experiment 3). The mean estimate of participants' priors were very similar when fitting to all data and when fitting to only the first 12 trials of each advisor. Fitting the model to only the first 12 trials of each advisor did not change any of the conclusions we made in our paper.

	Confirmation Bias Model		**Bayesian Learning Model**	
	$\hat{a}$ all data	**$\hat{a}$ first 12 trials**	**$\hat{a}$ all data**	**$\hat{a}$ first 12 trials**
Expt 1	0.62 [0.55, 0.70]	0.60 [0.55, 0.65]	0.59 [0.56, 0.65]	0.60 [0.56, 0.65]
Expt 2				
4-star	0.72 [0.70, 0.80]	0.72 [0.70, 0.79]	0.73 [0.71, 0.79]	0.75 [0.73, 0.79]
3-star	0.69 [0.64, 0.76]	0.67 [0.62, 0.72]	0.67 [0.63, 0.71]	0.69 [0.65, 0.73]
2-star	0.45 [0.36, 0.53]	0.45 [0.37, 0.51]	0.44 [0.36, 0.49]	0.43 [0.35, 0.48]
1-star	0.43 [0.36, 0.48]	0.40[0.34, 0.48]	0.41 [0.33, 0.47]	0.39 [0.33, 0.46]
Expt 3 Gen 1	0.58 [0.55, 0.61]	0.58 [0.55, 0.60]	0.59 [0.55, 0.61]	0.59 [0.56, 0.61]
Expt 3 Gen 2				
75% Advisor	0.78 [0.76, 0.79]	0.76 [0.74, 0.78]	0.77 [0.76, 0.78]	0.77 [0.75, 0.78]
50% Advisor	0.63 [0.60.0.66]	0.63 [0.60, 0.66]	0.62 [0.60. 0.65]	0.63 [0.61, 0.66]
25% Advisor	0.35 [0.32. 0.38	0.33 [0.30, 0.36]	0.33 [0.30, 0.36]	0.32 [0.29, 0.35]

Table S2. Comparison of the mean estimates of participants' priors about advisors' expertise, $\hat{a}$, when fitting the model to all data or to only the first 12 trials of each advisor. Square brackets indicate 95% credible intervals.

Analysis of participants' choice strategies

The logistic choice function interpolates between maximization (i.e. if $p(bet = FOR) = 1, \hat{a} > .5$, $p(bet = FOR) = 0$ otherwise) and random choice depending on the value of the gain parameter τ. For large values of τ (> 80), the choice rule approaches maximization (Fig. S4A), and for small values of τ (< 0.1), the choice rule approaches random choice (Fig. 4B). For moderate values of τ, the choice rule assumes "soft" maximization, such that $p(bet = FOR)$ scales with $\hat{a}$ (Fig. 4C). For $\tau = 4.5$, the choice rule comes reasonably close to exact probability matching (i.e. $p(bet = FOR) = \hat{a}$ for all values of $\hat{a}$, Fig. 4D). The best-fit values of τ for the Confirmation Bias and Bayesian Learning models in Experiment 1 are plotted in Figure S5. For most participants, the best-fit value of τ was well above 4.5, regardless of which model was fitted. This was also true for the participants in Experiments 2 and 3 (not shown). In particular, the range of values of τ suggests that most participants were following a soft-maximization rule whereby the probability of betting for an advisor increased with higher values of $\hat{a}$, with the probability being greater than $\hat{a}$ when $\hat{a}$ is above 0.5, and less than $\hat{a}$ when $\hat{a}$ less than 0.5 (Fig. R4C).

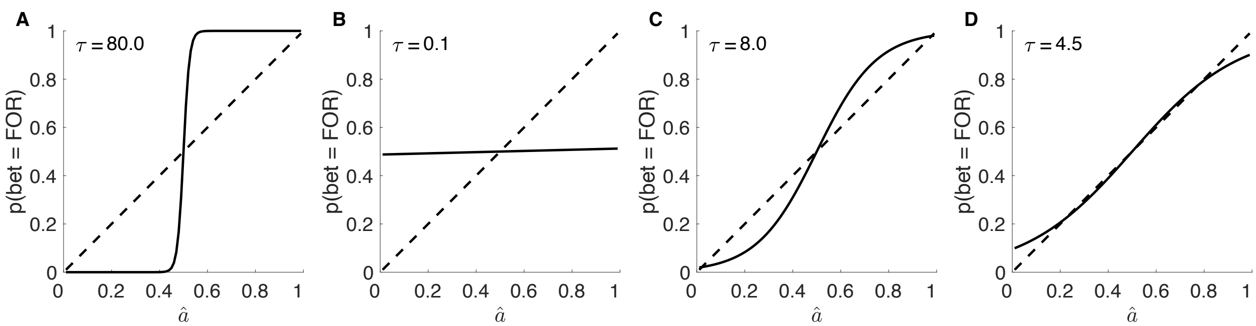

Figure S4. Choice rule at different values of τ. Dotted line indicates exact probability matching, p(bet = FOR) = $\hat{a}$. **A.** τ = 80, the choice rule approximates a maximization rule, and assumes that p(bet = FOR) = 1 when $\hat{a} > 0.5$, and that p(bet = FOR) = 0 when $\hat{a} < 0.5$. **B.** τ = 0.1, the choice rule approximates random choice and assumes that participants are equally likely to bet for or against the advisor's prediction at all levels of $\hat{a}$. **C.** τ = 8.0, the average τ when fitting the Confirmation Bias Model to participants in Experiment 1. This is a soft maximization rule that assumes that the probability of betting for the advisor is greater than $\hat{a}$ when $\hat{a}$ is above 0.5, and less than $\hat{a}$ when $\hat{a}$ less than 0.5. **D.** τ = 4.5, the choice rule approximates exact probability matching for $\hat{a}$ = (0.2, 0.8).

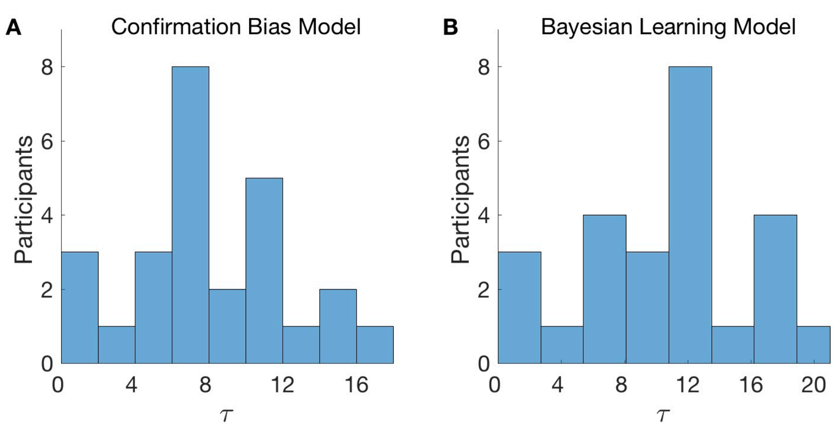

Figure S5. Histograms showing distribution of best-fit τ values in Experiment 1 for A. The Confirmation Bias model and B. the Bayesian Learning model. The average best-fit τ for the two models were 7.97 (SE = 0.84) and 11.23 (SE = 1.13) respectively.

Experiment 3 Modeling Results

Below we report the modeling results from Experiment 3:

Parameter	MAP Estimate Confirmation Bias Model	Bayesian Learning Model
α	2.88 (0.29)	3.09 (0.25)
β	2.06 (0.22)	2.13 (0.17)
τ	6.24 (0.84)	7.78 (1.01)
$\hat{a}_1$	0.58 [0.55, 0.61]	0.59 [0.56, 0.61]
	Model Comparison	
AIC	66.2 (2.2)	70.4 (2.0)
# of best-fit participants	59	27
Corrected Avg Lik per Trial	0.58 [0.55, 0.61]	0.56 [0.54, 0.58]
$\mu_{CB\text{-}BL}$	0.006 [0.002, 0.010]	
Effect Size	0.448 [0.141, 0.757]	

Table S3. Experiment 3 Generation 1 model-fitting results. We fit the two models to find the MAP estimates of the model parameters. For each participant, we fit the parameters that define their initial beliefs about the advisor (α and β), and the logistic function gain parameter τ. The table shows the average estimates of the model parameters. The mean estimate of $\hat{a}_1$ was optimistic for both the Confirmation Bias model and the Bayesian Learning model. We compared the two models based on AIC and the corrected average likelihood per trial. For statistical inference, we used robust Bayesian estimation to estimate the mean difference in corrected average likelihood per trial of the two models ($\mu_{CB\text{-}BL}$). Also reported is the effect size when comparing $\mu_{CB\text{-}BL}$ to 0. Parentheses indicate standard error of the mean, while square brackets denote 95% HDI of the corresponding posterior distribution.

Parameter	MAP Estimate Confirmation Bias Model	 Bayesian Learning Model
75% Advisor α	4.07 (0.18)	4.72 (0.25)
75% Advisor β	1.57 (0.16)	1.74 (0.13)
75% Advisor $\hat{a}_1$	0.78 [0.76, 0.79]	0.77 [0.76, 0.78]
50% Advisor α	3.71 (0.21)	4.57 (0.30)
5□□□Advisor β	2.24 (0.19)	2.67 (0.14)
50% Advisor $\hat{a}_1$	0.63 [0.60, 0.66]	0.62 [0.60, 0.65]
25% Advisor α	2.19 (0.19)	2.11 (0.16)
2□□□□□□□□□□ β	3.95 (0.18)	4.25 (0.21)
25% Advisor $\hat{a}_1$	0.35 [0.32, 0.38]	0.33 [0.30, 0.36]
τ	8.56 (0.76)	10.96 (0.97)
	Model Comparison	
AIC	44.0 (2.2)	48.0 (2.0)
# of best-fit participants	76	16
Corrected Avg Lik per Trial	0.63 [0.61, 0.65]	0.60 [0.58, 0.63]
$\mu_{CB\text{-}BL}$	0.016 [0.011, 0.021]	
Effect Size	0.887 [0.564, 1.217]	

Table S4. Experiment 3 Generation 2 model-fitting results. We fit the two models to find the MAP estimates of the model parameters. For each participant, we fit the parameters that define their initial beliefs about each advisor (α and β), and the logistic function gain parameter τ. The table shows the average estimates of the model parameters. We compared the two models based on AIC and the corrected average likelihood per trial. For statistical inference, we used robust Bayesian estimation to estimate the mean difference in corrected average likelihood per trial of the two models ($\mu_{CB\text{-}BL}$). Also reported is the effect size when comparing $\mu_{CB\text{-}BL}$ to 0. Parentheses indicate standard error of the mean, while square brackets denote 95% HDI of the corresponding posterior distribution.

Supplemental References

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–1221. https://doi.org/10.1038/nn1954

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General*, *142*(2), 573–603. https://doi.org/10.1037/a0029146

Sutton, R. S., & Barto, A. G. (1998). *Introduction to Reinforcement Learning* (1st ed.). Cambridge, MA, USA: MIT Press.

Waskom, M. L., Frank, M. C., & Wagner, A. D. (2016). Adaptive Engagement of Cognitive Control in Context-Dependent Decision Making. *Cerebral Cortex*, bhv333. https://doi.org/10.1093/cercor/bhv333

Worthy, D. A., & Todd Maddox, W. (2014). A comparison model of reinforcement-learning and win-stay-lose-shift decision-making processes: A tribute to W.K. Estes. *Journal of Mathematical Psychology*, *59*, 41–49. https://doi.org/10.1016/j.jmp.2013.10.001