

Experience and advice consequences shape information sharing strategies

INDEX

Supplementary Note 1: INSTRUCTIONS	2
Supplementary Note 2: ADDITIONAL ANALYSES	4
Supplementary Note 3: CHOICE PROBABILITY, EXPLOITATION AND TRIAL PROGRESSION	10
Supplementary Note 4: DIFFERENCES BETWEEN NO COST AND REPUTATION CONDITIONS (EXPERIMENT 2)	11
Supplementary Note 5: FULL SAMPLE RE-ANALYSIS	12
Supplementary Note 6: COMPUTATIONAL MODELING	14
Supplementary Note 7: MODEL PARAMETERS	18
Supplementary Note 8: PRE-REGISTRATION MATCHING TABLE	21
Supplementary references	24

Supplementary Note 1: INSTRUCTIONS

A demo of the experiment is available at <http://socialdecisionlab.net/resources.html>

Instructions Experiment 1:

1. *Welcome to the fishing game. In this game you're going to go on three fishing vacations.
In each one you have to choose which lake you want to fish from. The aim of the task is to catch as many fish as possible. For every 50 fish you catch, you will get a £0.15 bonus, with the maximum of £1.8 for 600 fish. *No fish was harmed in the preparation of this experiment.*
2. *Each of the fishing vacations will be 20 days long. On each vacation day you need to choose from one of two lakes to fish in.*
3. *After you choose a lake, you will have the opportunity to send your choice as an advice to players that will participate in a follow-up experiment. If you choose to publish your advice, it may be displayed to <i>all</i> participants in a follow-up experiment, involving the same lakes/vacations. Accurate advice can greatly accelerate learning, leading to greater bonus for other players. However, inaccurate advice can hinder learning, resulting in lower bonuses for other players.*
4. *After you choose, you'll see how many fish you caught that day (between 0-10). Remember, every 50 fish you catch is worth £0.15. At the end of the three vacations you will be asked to fill 2 questionnaires. * Sometimes it takes a couple of seconds to the next page to load, please do not refresh the page as it will send you back to the beginning of the experiment. Good luck!*
5. *Remember: your advice may be published to all players in the follow-up experiment. These participants will see your advice about the recommended lake. If you choose not to publish your advice, your choice will remain confidential. Accurate advice can greatly accelerate learning, leading to greater bonus for other players. However, inaccurate advice can hinder learning, resulting in lower bonuses for other players.*

Instructions in Experiment 2:

Experiment 2 included the same instructions as Experiment 1, with the addition of another page which differed between experimental conditions.

Monetary consequence:

6. Each advice you give will cost you one fish, which will be deducted from your own catch and will reduce your final payment at the end of the experiment.
 If you give advice 50 times, you will receive £0.15 less bonus.

Reputation consequence:

6. Your advice for future players will be displayed along your **Adviser-Rating**. Every time you publish an advice, an algorithm is used to evaluate this advice using simulations, and to update your Adviser-Rating. If the algorithm indicates positive advice outcome, your Adviser-Rating increases. If it indicates negative advice outcome your Adviser-Rating decreases.
 If you choose not to publish your choice as advice, your Adviser-Rating stays the same.
 Your Adviser-Rating is initialized in each vacation to 0. Your advice will be published to future players with your Adviser-Rating and your participant number - your number is: #' + SubID + '.

Responsibility consequence:

6. In the follow-up experiment, participants will have the opportunity to use a "help-line" when they are stuck, and ask for advice. Only participants that use this help-line and ask for advice will be presented with your advice. Your advice will displayed along with your participant number - your number is: #' + SubID + '.

Participants using the help-line will be presented with your advice, and will follow it automatically.

No-cost consequence:

No additional information was provided.

Supplementary Note 2: ADDITIONAL ANALYSES

Choice behavior analysis: correct choice ratio

Ratio of correct choices ($P(\text{correct})$) in Study 1 was defined as the proportion of choices that maximized expected value, and modeled as pre-registered, i.e.

$$P(\text{correct}) \sim \text{Block difficulty} * \text{Trial number} + \text{PTM altruism score} + \text{SPIN fear score} + (\text{Block difficulty} + \text{Trial number} \mid \text{Participant})$$

where Block difficulty was a 3-level factor (easy/medium/hard). As shown in Fig. 2A, $P(\text{correct})$ was consistent with successful value learning: it increased along trial progression, plateaued above chance, and was coherently modulated by task difficulty (Block main effect: $\chi^2(2) = 102$, $p < 0.001$; Trial number main effect: $\chi^2(1) = 234$, $p < 0.001$; Block x Trial number interaction: $\chi^2(2) = 103$, $p < 0.001$). A summary of all relevant statistical effects and interactions can be found below in Table S1.

P(correct) ~ Block x Trial N + SPIN + PTM				P(advice) ~ Block x Trial N + SPIN + PTM		
	χ^2	DF	P	χ^2	DF	P
Block	102	2	<.001	28	2	<.001
Trial N	234	1	<.001	109	1	<.001
SPIN (scaled)	2	1	0.162	0.02	1	0.902
PTM (scaled)	0.33	1	0.566	4	1	0.046
Block x Trial N	103	2	<.001	59	2	<.001

Table S1. Result summary for choice and advice behavior. Type 2 Wald χ^2 test for main effects and interactions. SPIN (fear) and PTM (altruism) dimensions were centered (rescaled) for convergence purposes. χ^2 = Chi square; DF = degrees of freedom; P = p-value; SPIN = Social phobia Inventory - fear dimension; PTM = Prosocial Tendencies Measure - altruism dimension.

We then executed the same analysis for each of the consequence conditions in Study 2. As seen on Figure 3.A in the main manuscript, correct choice rates evolved with trial progression consistently with successful value learning in all experiments (Trial N main effect: $\chi^2(1) = 309$, $p < 0.001$), with advice consequence significantly impacting the learning speed (AM x Trial N interaction: $\chi^2(3) = 9$, $p = 0.026$), particularly on the Reputational condition.

P(correct) ~ AC x (Trial N + SPIN + PTM)				P(advice) ~ AC x (Trial N + SPIN + PTM)		
	χ^2	DF	P	χ^2	DF	P
Adv. consequence	3	3	0.458	56	3	<.001
Trial N	309	1	<.001	376	1	<.001
SPIN (scaled)	0.1	1	0.713	3	1	0.083
PTM (scaled)	0.4	1	0.512	1	1	0.302

AC x Trial N	9	3	0.026	129	3	<.001
AC x SPIN	1	3	0.821	1.6	3	0.657
AC x PTM	2	3	0.676	1.9	3	0.599

Table S2. Result summary for choice and advice behavior. Type 2 Wald χ^2 test for main effects and interactions. SPIN (fear) and PTM (altruism) dimensions were centered (rescaled) for convergence purposes. χ^2 = Chi square; *DF* = degrees of freedom; *P* = p-value; *SPIN* = Social phobia Inventory - fear dimension; *PTM* = Prosocial Tendencies Measure - altruism dimension; *AC* = advice consequence.

Advice tendencies and block difficulty (full model as preregistered)

	P(advice) ~ AC x Block x Trial N x (SPIN + PTM) + (Block + Trial N ID)		
	χ^2	DF	P
Adv. Consequence	37	3	< 0.001
Block	26	2	< 0.001
Trial N	59	1	< 0.001
PTM (scaled)	1	1	0.354
SPIN (scaled)	2	1	0.124
AC x Block	7	6	0.300
AC x Trial N	32	3	< 0.001
AC x PTM	2.6	3	0.455
AC x SPIN	1.6	3	0.653
Block x Trial N	20	2	< 0.001
Block x PTM	2.5	2	0.283
Block x SPIN	0.7	2	0.694
Trial N x PTM	0.04	1	0.837
Trial N x SPIN	0.46	1	0.497
AC x Block x Trial N	23	6	0.001
AC x Block x PTM	4	6	0.642
AC x Block x SPIN	5.8	6	0.451
AC x Trial N x PTM	0.14	3	0.987
AC x Trial N x SPIN	0.42	3	0.936
Block x Trial N x PTM	5.7	2	0.057
Block x Trial N x SPIN	4.9	2	0.086
AC x Block x Trial N x PTM	15	6	0.019
AC x Block x Trial N x SPIN	10	6	0.115

Table S3. Result summary for advice behavior taking into account block difficulty. Type 2 Wald χ^2 test for main effects and interactions. SPIN (fear) and PTM (altruism) dimensions were centered (rescaled) for convergence purposes. χ^2 = Chi square; *DF* = degrees of freedom; *P* = p-value; *SPIN* = Social phobia Inventory - fear dimension; *PTM* = Prosocial Tendencies Measure - altruism dimension; *AC* = advice consequences.

Advice and performance for all difficulty levels and conditions

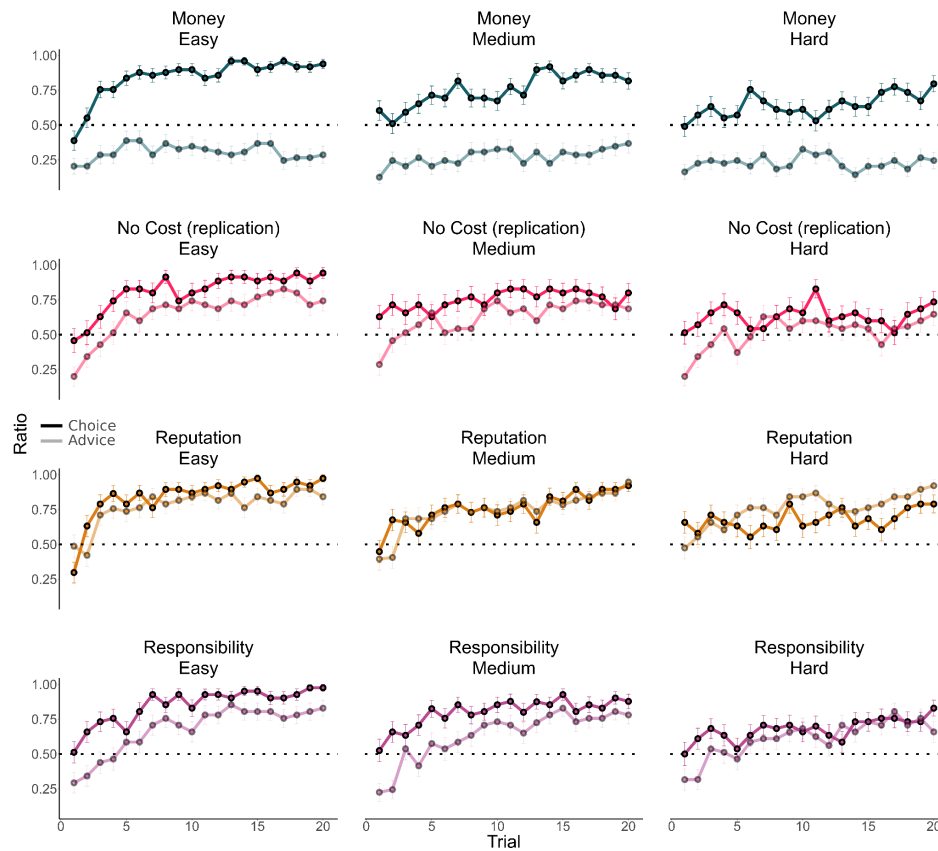


Figure S1. Correct (i.e. value-maximizing) choice and advising ratios per trial, for all difficulties and advice consequence conditions [Points: mean ratio per trial; bars: bootstrapped standard error of the mean per trial; $n = 210$]

Detection of over-advisers

The criterion for identifying over-advisers consisted of advice broadcasting in all trials. A tolerance threshold of 1 misclick was enabled, so that participants who broadcasted advice in every trial except for one would also be classified as over-advisers. Exceptionally, participants who gave advice in every trial except for the first one were not to be considered as over-advisers, as arguably said participants could consider that a single trial was enough learning experience to start broadcasting advice (no such cases were found). Proportions of over-advisers in Study 1 are presented below in figure S2.

Advice-giving ratio

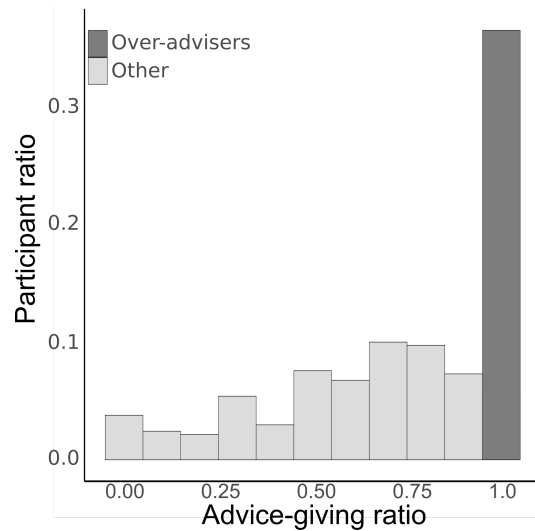


Figure S2. Binned advice-giving ratio against proportion of participants on each bin [Over-advisers: participants choosing to disclose advice on 100% of trials; n = 373]

Additionally, we conducted model-based clustering analyses on the sample of Study 1 to determine if the property of being an over-adviser would naturally emerge from a data-driven search for the modality of the distribution of advising rate. We found significant evidence in favor of a multimodal vs. unimodal distribution (ARC method, Excess mass = 0.04, $p < .001$; ¹). Fitting of a mixture model of multiple univariate normal distributions through model-based clustering (default priors; see ²) confirmed 9 components of unequal variance (clustering table: 25, 33, 29, 21, 54, 19, 30, 27 and 135 participants per component). The last of these components overlapped perfectly with the participants identified as over-advisers (n=135), and was the most far-removed component of all (Fig S3).

Clustering for advice-giving ratios

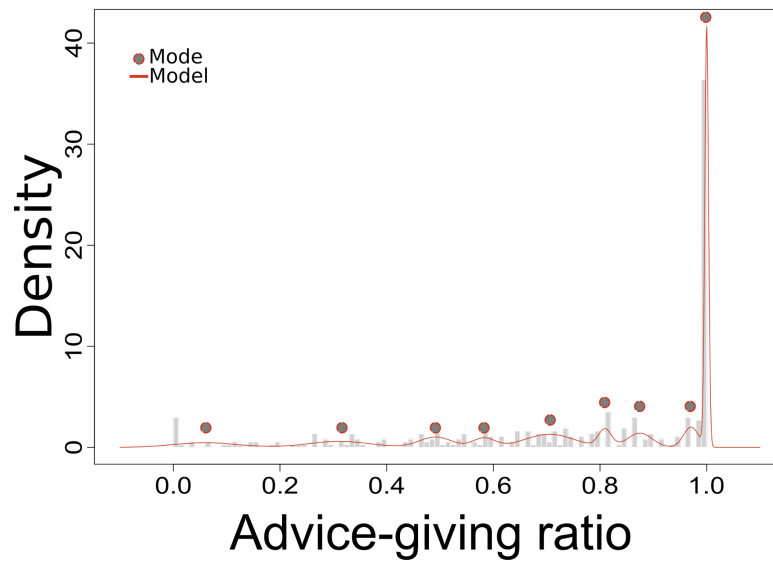


Figure S3. Clustering analysis for detection of modes revealed a mixture model of 9 distinct univariate components. The 9th component (rightmost mode) overlapped perfectly with overadvisers (n=135). [complete n = 373]

The same criteria were used to identify the proportion of over-advisers in Study 2 for each advice consequence condition (see Fig. S4 below)

Advice-giving ratio per cost condition

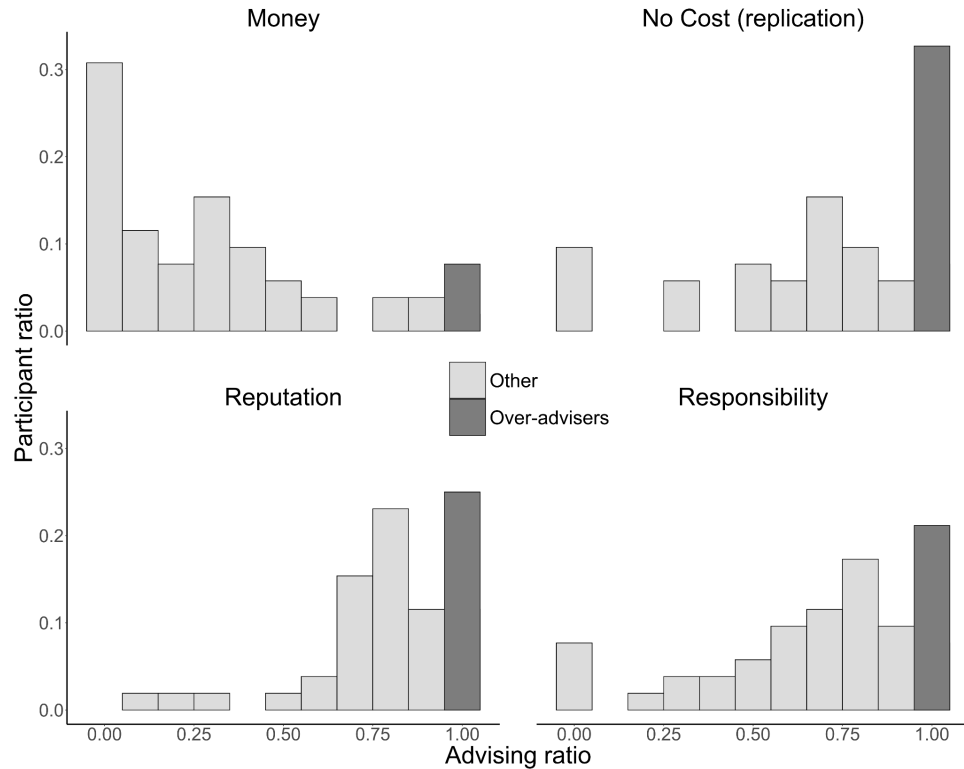


Figure S4. Binned advice-giving ratio against proportion of participants on each bin [Over-advisers: participants choosing to disclose advise on 100% of trials; ; n = 210]

Over-advising and individual psychosocial differences

A generalized linear regression considering the individual psychological characteristics SPIN (fear subscale; ³) and PTM (altruism subscale; ⁴2002) revealed a non-significant trend for a lower probability to be an over adviser at higher levels of social anxiety (coefficient = -0.203 (95% CI = -0.418, -0.010); Likelihood ratio Chi-square $\chi^2(1) = 3.49$, $p = 0.049$), and a non-significant trend favoring higher levels of over-advising at higher altruistic prosociality (coefficient = 0.17 (-0.048, 0.385); $\chi^2(1) = 2.30$, $p = 0.130$). Both trends are presented in Figure S5 below.

Over-advising probabilities & individual psychological differences

Probability of Over-advise

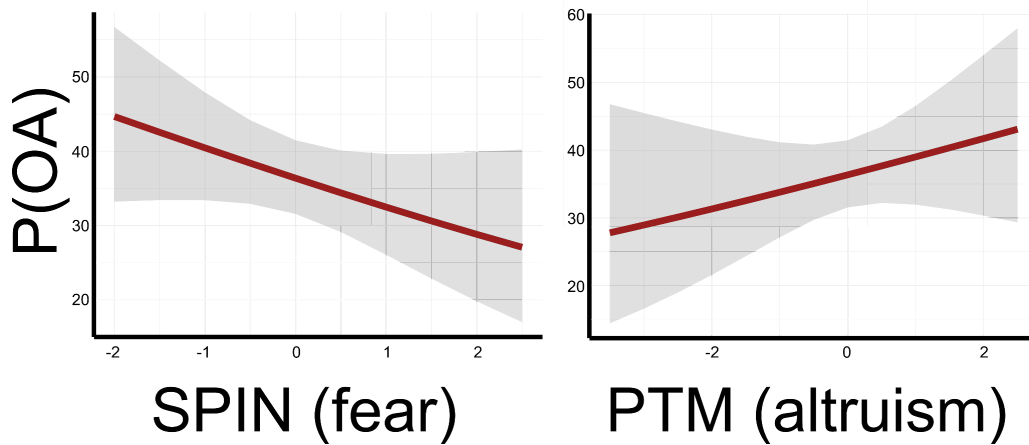


Figure S5. A. Predicted probabilities of being an over adviser (disclosing advice at 100% of trials) as a function of the fear dimension of the Social Phobia Inventory. **B.** Predicted probabilities of being an over adviser (disclosing advice at 100% of trials) as a function of the altruistic dimension of the Prosocial Tendencies Measure. *Shaded area is the 95% Confidence Interval of the fit. Inventory scores were rescaled (centered) for model convergence [n = 373].*

Propensity to give advice and prosociality

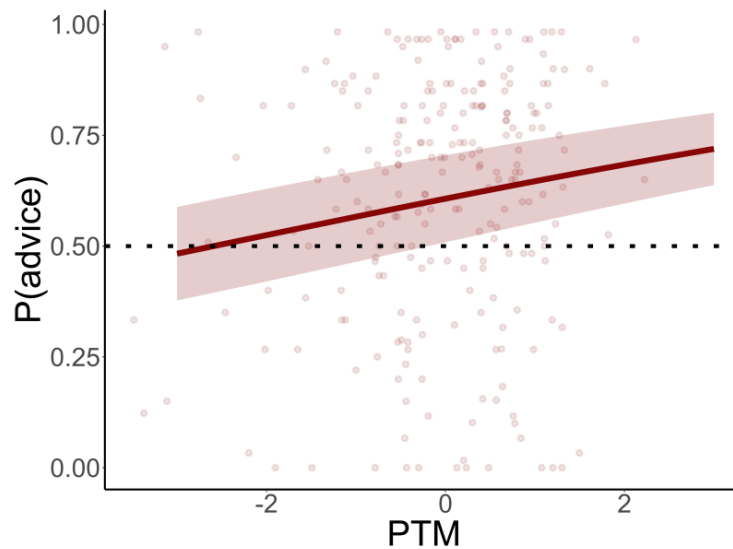


Figure S6. Prosocial tendencies (altruism) score predicts higher probability of disclosing advice. Dots are the average advice probability per participant. PTM scores are scaled. Line is mean advice scores across participants. Shaded area is the standard error of the mean [n = 373].

Propensity to give advice and social anxiety levels

In both Studies, we observed that more socially-anxious participants advised less on earlier trials (i.e. exploratory and uncertain, $P(\text{choice}) < 0.5$), and more on later trials (i.e. exploitative,

$P(\text{choice}) < 0.5$), possibly as a result of perceived social pressure for sharing useful information. We observed that the tendency to give advice increased with the progression from exploratory ($P(\text{choice}) < 0.5$) to exploitative ($P(\text{choice}) > 0.5$) choice behavior, and was favored by positive prediction errors. The effect of choice probability on advice was found to be sensitive to participants' self-perceived levels of social anxiety, as probability to give advice increased with SPIN (fear) score for exploitative trials, but decreased for exploratory trials both in Study 1 (Fig. S7 *left*) and Study 2 (Fig. S7 *right*).

Advice disclosure preference and individual psychological differences

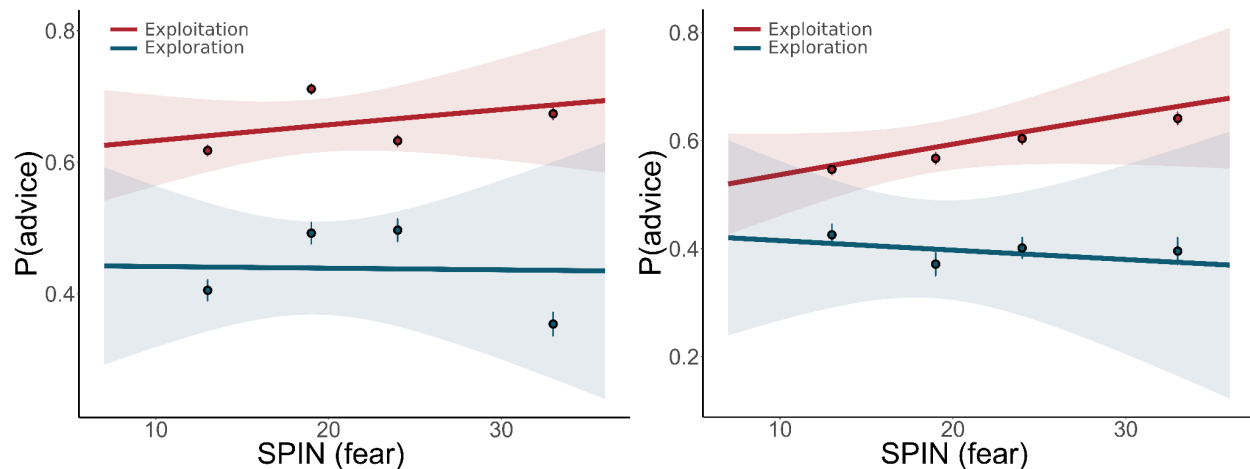


Figure S7. A. Ratio of advice disclosure as a function of the fear dimension of the Social Phobia Inventory for Study 1. **B.** Ratio of advice disclosure as a function of the fear dimension of the Social Phobia Inventory for Experiments 2 to 5 combined. *Dots represent the mean advice ratio for all participants at the bounds of SPIN inter-quantile range. Bars represent the bootstrapped Standard Error of the mean. Shaded area is the 95% Confidence Interval. Inventory scores were rescaled (centered) for model convergence. Exploration trials were defined as trials for which $P(\text{choice}) < 0.5$, while exploitation trials were defined as trials for which $P(\text{choice}) > 0.5$ (see Daw et al 2011). [n = 373]*

Supplementary Note 3: CHOICE PROBABILITY, EXPLOITATION AND TRIAL PROGRESSION

One of the main findings of the present work indicated that individuals with high SPIN tended to give more advice when $P(\text{choice})$ was high. We interpreted this phenomenon as a tendency of socially-anxious individuals to share more information when manifesting an exploitative rather than exploratory behaviour. We determined whether trials were primarily exploratory or exploitative based on a classic conceptualization in Reinforcement Learning literature⁵, where

by definition exploitation tendencies consist of choosing the reward-maximising option with a probability consistently higher than chance.

This interpretation, however, relied heavily on the assumption that in later trials $P(\text{choice})$ must have been increasingly higher. To verify this assumption, we fitted a regression that considered all participants across both experiments, and used trial number and SPIN to predict the value of $P(\text{choice})$. Results are presented in the Table S4 below:

P(choice) ~ Trial number + SPIN + Trial number:SPIN				
	Estimate (for intercept = 0.6)	χ^2	DF	P
Trial number	0.02	1628	1	<0.001
SPIN (scaled)	-0.003	1	1	0.243
Trial number x SPIN (scaled)	0.0001	7	1	0.008

Table S4. Impact of SPIN and trial progression on choice probability (as obtained through model fitting). Type 2 Wald χ^2 test for main effects and interactions. SPIN (fear) dimension was centered (rescaled) for convergence purposes. χ^2 = Chi square; DF = degrees of freedom; P = p-value; SPIN = Social phobia Inventory - fear dimension. $N = 583$.

Confirming our expectations, the regression showed that choice probability increased with trial progression, and while it was not affected by SPIN directly, SPIN did have an incremental effect as it interacted with trial number.

Supplementary Note 4: DIFFERENCES BETWEEN NO COST AND REPUTATION CONDITIONS (EXPERIMENT 2)

Our findings show that within people who do not over-advise, unmotivated advising appears to be close to responsible advising. We posit it is possible that the awareness or default responsibility that led said participants to not over-advise in the first place may have also influenced the degree of responsibility with which they decided to give advice every trial. At the same time, we did observe that the relation between PTM and advice calibration appear to reveal that these conditions are indeed different. Therefore, we decided to conduct a Bayes Factor analysis (BIC approximation to BF, as before) evaluating the strength of the general difference between the two conditions. Interestingly, while the difference is indeed significant ($\chi^2(1) = 6$, $p = 0.014$), the Bayes Factor indicated that in terms of evidence strength it was barely worth mentioning when all model predictors were considered ($BF < 0.001$). For reference, the strength of the evidence for a general difference between the No Cost and the Monetary conditions was $BF = 151259$. Overall, we interpret this to mean that despite significant finer-grained differences in the calibration of advice, driven by PTM scores, the general advising behavior between both conditions is demonstrably similar.

Supplementary Note 5: FULL SAMPLE RE-ANALYSIS

For theoretical and methodological reasons Over-advisers and the rest of the participants were analyzed separately. Here we replicate our entire analysis strategy over the consolidated sample, both for Study 1 and Study 2. We found that, despite the obvious noise added by collapsing regular and over- advisers together, all the main effects that were relevant to our hypotheses could be replicated in the full sample analyses.

Study 1

Choice behaviour & Advising

Advice-giving was still dependent on both the learning process and task difficulty, as expected. Easier blocks led to better performances and more advising, and we found no significant evidence that either SPIN (fear) behaviour. One key difference with our main analysis findings was that what previously was just a trend of PTM affecting advising, now became a statistically significant effect: higher prosociality led to higher advising tendencies in the composite sample (coefficient = 0.51; $\chi^2(1) = 4.7$, $p = 0.030$). A summary of all relevant statistical effects and interactions can be found below in Table S4.

	P(correct) ~ Block x Trial N + SPIN + PTM			P(advice) ~ Block x Trial N + SPIN + PTM		
	χ^2	DF	P	χ^2	DF	P
Block	183	2	<.001	0.6	2	0.752
Trial N	340	1	<.001	64	1	<.001
SPIN (scaled)	1	1	0.321	1.9	1	0.163
PTM (scaled)	0.07	1	0.795	4.7	1	0.030
Block x Trial N	126	2	<.001	55	2	<.001

Table S5. Result summary for choice and advice behavior. Type 2 Wald χ^2 test for main effects and interactions. SPIN (fear) and PTM (altruism) dimensions were centered (rescaled) for convergence purposes. χ^2 = Chi square; DF = degrees of freedom; P = p-value; SPIN = Social phobia Inventory - fear dimension; PTM = Prosocial Tendencies Measure - altruism dimension.

Advising and correct choice probability

Results in this section were identical to what had been previously observed: probability of broadcasting advice per trial was significantly predicted by the estimated correct choice probability given expected option value. Again, we did not find evidence for either social anxiety directly modulating advice disclosing tendencies, or of prosociality affecting advice as a function

of choice probability. We did find again a significant interaction between social anxiety and P(choice) that aligned with our prediction. Details are presented in Table S5.

	P(advice) ~ P(choice) + PE + SPIN + PTM + P(choice):SPIN + P(choice):PTM		
	χ^2	DF	P
P(choice)	1199	1	<.001
PE	18	1	<.001
SPIN (scaled)	3	1	0.099
PTM (scaled)	5	1	0.030
P(choice) x SPIN (scaled)	7	1	0.008
P(choice) x PTM (scaled)	2	1	0.163

Table S6. Result summary for the impact of expected-value-based choice probability (P(choice)) and advice probability (P(advice)). Type 2 Wald χ^2 test for main effects and interactions. SPIN (fear) and PTM (altruism) dimensions were centered (rescaled) for convergence purposes. χ^2 = Chi square; DF = degrees of freedom; P = p-value; SPIN = Social phobia Inventory - fear dimension; PTM = Prosocial Tendencies Measure - altruism dimension.

Study 2

Choice behaviour & Advising

In accordance with the main analyses, as predicted (H2.1), advice consequences modulated advice-giving tendencies without disturbing learning, with the monetary loss consequence still inhibiting advice and the Reputation consequence globally increasing advising across the board. A summary of all relevant statistical effects and interactions can be found below in Table S6.

	P(correct) ~ AC x (Trial N + SPIN + PTM)			P(advice) ~ AC x (Trial N + SPIN + PTM)		
	χ^2	DF	P	χ^2	DF	P
Adv. consequence	4	3	0.314	52	3	<.001
Trial N	421	1	<.001	381	1	<.001
SPIN (scaled)	0.3	1	0.606	0.05	1	0.829
PTM (scaled)	0.2	1	0.680	3	1	0.095
AC x Trial N	12	3	0.007	130	3	<.001
AC x SPIN	0.3	3	0.953	0.3	3	0.952
AC x PTM	0.2	3	0.971	0.3	3	0.952

Table S7. Result summary for choice and advice behavior. Type 2 Wald χ^2 test for main effects and interactions. SPIN (fear) and PTM (altruism) dimensions were centered (rescaled) for convergence purposes. χ^2 = Chi square; DF = degrees of freedom; P = p-value; SPIN = Social phobia Inventory - fear dimension; PTM = Prosocial Tendencies Measure - altruism dimension; AC = advice consequence.

Advising and correct choice probability

One more time, results largely coincided with what was presented in the main analyses. Namely, we still observed an impact of advice conditions on advising, but not to the extent of

extinguishing the effect of choice probability on advice. We also replicated the double and triple interactions involving psychosocial characteristics. A summary of all relevant statistical effects and interactions can be found below in Table S7.

	P(advice) ~ P(choice) x AC x (PTM altruism score + SPIN fear score)		
	χ^2	DF	P
P(choice)	680	1	<.001
AC	50	3	<.001
SPIN (scaled)	0	1	0.871
PTM (scaled)	3	1	0.064
P(choice) x AC	79	3	<.001
P(choice) x SPIN (scaled)	31	1	<.001
P(choice) x PTM (scaled)	4	1	0.036
AC x SPIN (scaled)	0.5	3	0.924
AC x PTM (scaled)	0.6	3	0.895
P(choice) x AC x SPIN (scaled)	27	3	<.001
P(choice) x AC x PTM (scaled)	21	3	<.001

Table S8. Result summary for the impact of expected-value-based choice probability (P(choice)) and advice probability (P(advice)) across different types of advice consequences. Type 2 Wald χ^2 test for main effects and interactions. SPIN (fear) and PTM (altruism) dimensions were centered (rescaled) for convergence purposes. χ^2 = Chi square; DF = degrees of freedom; P = p-value; SPIN = Social phobia Inventory - fear dimension; PTM = Prosocial Tendencies Measure - altruism dimension.

Supplementary Note 6: COMPUTATIONAL MODELING

We operationalized our Reinforcement Learning (RL) models through STAN⁶, which facilitates Bayesian statistical inference because it allows for the specification of complex statistical models together with their corresponding posterior inferences. Our RL model quantified how individuals updated the perceived value (Q-value) of different options based on choice outcomes (rewards). Thus, it was structured around two critical parameters: the learning rate (α) and the inverse temperature (β). The learning rate delineated the speed at which each participant individual's q-values were updated in light of new information, whereas the inverse temperature quantified the extent to which these outcomes influenced decision-making, acting as a proxy for choice stochasticity.

We used hierarchical model fitting, as described in detail in the methods section. Details about group parameters' priors and the way individual parameters were derived is detailed in the main text as well, and full scripts can be found in our repository. Bayesian inferences were conducted using Markov Chain Monte Carlo (MCMC) sampling, specifically employing the No-U-Turn Sampler (NUTS), an extension of Hamiltonian Monte Carlo (HMC). The model ran for 2000

iterations across 4 chains, leveraging parallel computing with 7 cores to enhance computational efficiency. We checked the Rhat for any high values that could indicate problems with convergence and found that none exceeded 1.01. Full convergence and parameters are available in the CSV files accompanying the analysis code (<https://github.com/hrl-team/AdvisingRL>). For evaluation of individual level parameters we used the group level and random effects recovered from the model fitting procedure, and used the mean of the posterior as point estimate for further analyses.

To evaluate how adequate the fitting procedure was in regard to our sample, we implemented a tried and tested approach^{8,9}, that consists of fitting our model to each individual participant's behaviour, recover the value of each individual's parameter, and then use said model and parameters to generate new simulated behaviour at the individual level. Crucially, this produces simulations that are comparable to specific participant behaviour. Results of this computation are shown in Figure S8.

Simulations vs. participant behavior.

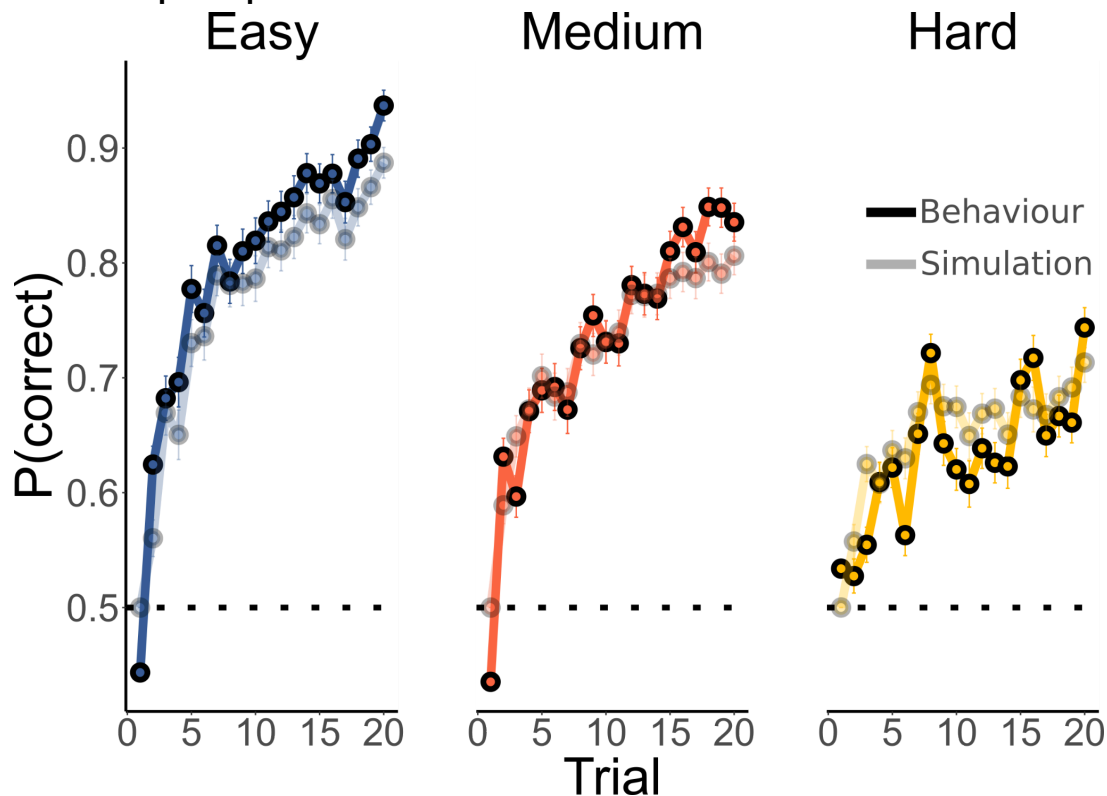


Figure S8. Mean performance per block difficulty on a trial basis, for both participants and simulations. (Bars are standard error of the mean; $P(\text{correct})$ is correct choice ratio, where correct choice is the choice that maximizes value; $N = 373$, simulated agents $N = 373$).

Visual inspection of the results shows a simulated performance that closely mirrors actual behaviour. To confirm this similarity, we computed the Bayes Factor for an effect of a difference between real and simulated behaviour. We did so by using the Bayesian Information Criterion approximation to the Bayes Factor approach¹⁰, fitting two identical regression models that

differed only on a “Data Type” categorical factor (levels: real behaviour vs. simulation). Overall, results greatly favored the model without the “data type” factor ($P(\text{correct}) \sim \text{Difficulty}$), regardless of whether we considered it as a main effect ($P(\text{correct}) \sim \text{Difficulty} + \text{Data Type}$; $\text{BF} = 0.0024$), an interaction ($P(\text{correct}) \sim \text{Difficulty} + \text{Data Type}:\text{Difficulty}$; $\text{BF} = \text{BF} < 0.0001$), or both ($P(\text{correct}) \sim \text{Difficulty} + \text{Data Type} + \text{Data Type}:\text{Difficulty}$; $\text{BF} < 0.0001$).

We finished our robustness tests by completing the implementation of the “simulation plus parameter recovery” approach^{11,12}. Following the successful use of our model to simulate participant behaviour above, we generated artificial agents with predefined random assortments of parameter values. Then, we presented these agents with our default task exactly as it was presented to human participants (i.e. three blocks of increasing difficulty, 20 trials per block), and refitted our model on their behaviour to evaluate if parameters could be successfully recovered. The model of our task relied fundamentally on two parameters (i.e. learning rate [α] and inverse choice temperature [β]). Thus, we picked 9 random values for each of these parameters, and conceived one simulated agent for every possible parameter combination (for a total of $N = 81$). As shown in Figure S9.A, trial-by-trial simulated performance was coherent with task design (i.e. learning and decreasing performance with increased difficulty were observed). Figure S9.B shows that the correlation between the randomly-generated parameters and the recovered parameters was acceptable both for learning rates (Pearson's product-moment correlation $t = 85$, $p < 0.001$, $R = 0.772$ (95% confidence interval: 0.760, 0.783) and inverse temperatures ($t = 92$, $p < 0.001$, $R = 0.800$ (0.788, 0.808)).

Parameter recovery

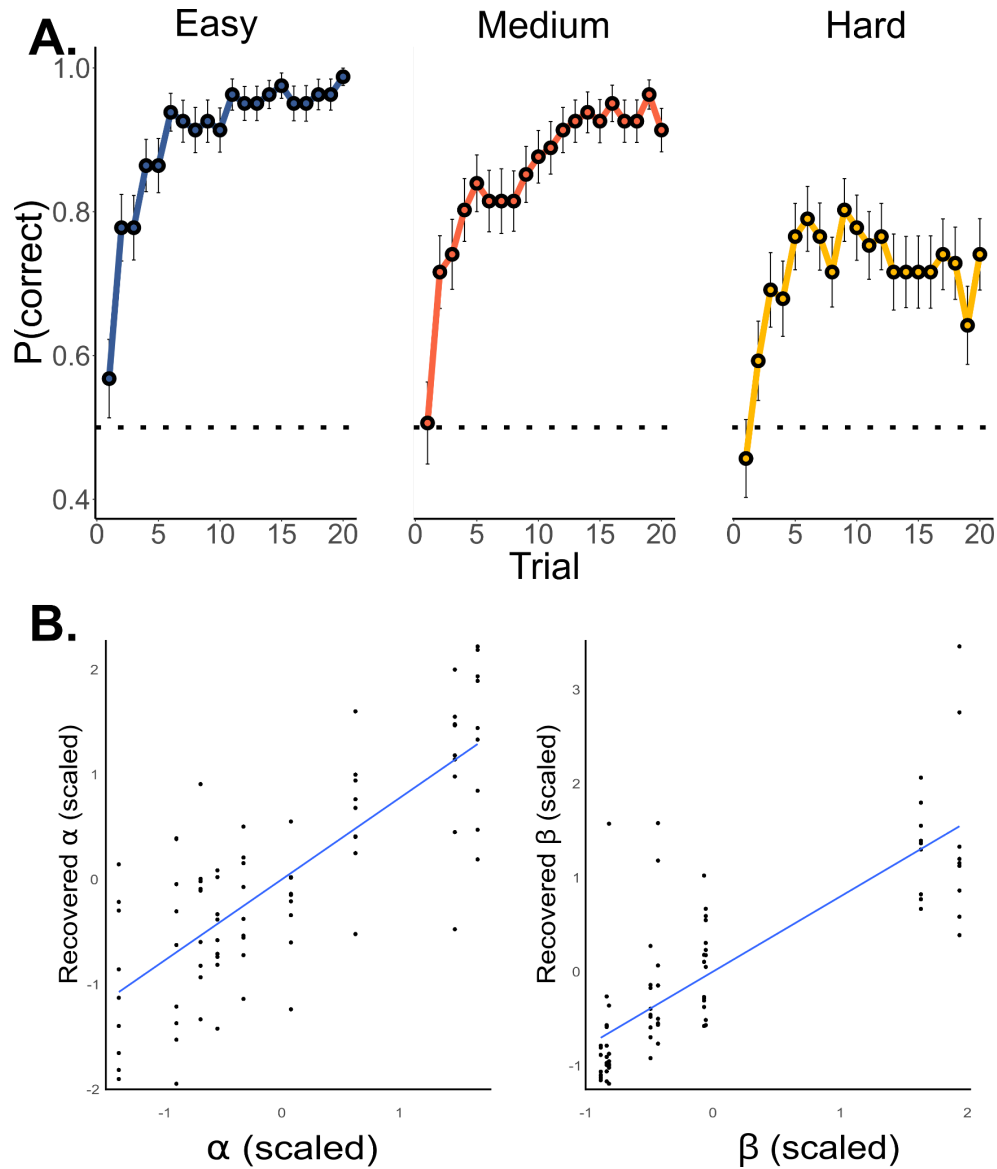


Figure S9. A. Per trial performance of simulated agents. **B.** Correlations of original and refitted learning rates and inverse temperatures (scaled). (*Bars are standard deviation; $P(\text{correct})$ is correct choice ratio, where correct choice is the choice that maximizes value; simulated agents $N = 81$*)

Following the obtention of the recovered parameters, we conducted an Intraclass Correlation Coefficient analysis (ICC), for both learning rate and inverse temperature, and observed significant moderate-high to high reliability¹³. Results of this procedure are outlined in Table S8 below.

ICC Learning rate (α)							
	ICC	F	DF1	DF2	CI (lower)	CI (higher)	P
Single raters (absolute)	0.7	5.8	80	81	0.58	0.80	<.001

Single random raters	0.7	5.8	80	80	0.58	0.80	<.001
Single fixed raters	0.7	5.8	80	80	0.57	0.80	<.001
Average raters (absolute)	0.83	5.8	80	81	0.73	0.89	<.001
Average random raters	0.83	5.8	80	80	0.73	0.89	<.001
Average fixed raters	0.83	5.8	80	80	0.73	0.89	<.001
ICC Inverse temperature (2)							
Single raters (absolute)	0.65	4.7	80	81	0.51	0.76	<.001
Single random raters	0.65	4.9	80	80	0.51	0.76	<.001
Single fixed raters	0.66	4.9	80	80	0.52	0.77	<.001
Average raters (absolute)	0.79	4.7	80	81	0.67	0.86	<.001
Average random raters	0.79	4.9	80	80	0.67	0.86	<.001
Average fixed raters	0.80	4.9	80	80	0.68	0.87	<.001

Table S9. Intraclass correlation coefficient analysis for recovered vs randomly-generated parameters. *Single raters absolute: single raters using an absolute agreement definition. Each participant is rated by a different, randomly selected rater. Single random raters: single raters under the assumption that raters are randomly selected. Reflects consistency of ratings, not just agreement. Single fixed raters: raters are fixed, consistency is measured within this fixed set of raters. Average raters (absolute): the average of multiple raters using an absolute agreement definition. Average random raters: average of multiple raters assuming that raters are randomly selected. Average fixed raters: average of multiple raters with the assumption that the set of raters is fixed. ICC: intraclass correlation coefficient value. DF1: degrees of freedom for numerator. DF2: degrees of freedom for denominator. CI: 95% confidence interval bounds. P: p-value.*

Overall, these results confirmed that for the purpose of estimating choice probability, which was the key element of interest here given its role as a proxy for the accumulation of knowledge on each option, our model provided reliable estimations.

Supplementary Note 7: MODEL PARAMETERS

Experiment 1:

Individual level parameters (full tables including convergence data available at <https://github.com/hrl-team/AdvisingRL>)

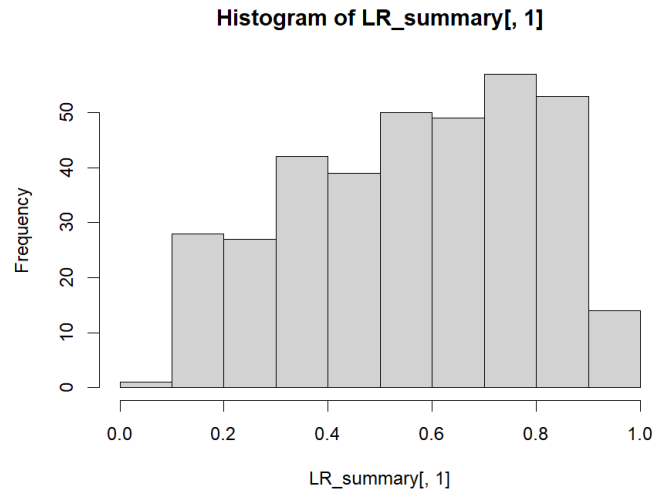


Figure S10. Learning rate parameter histogram. Mean $\pm$ std 0.56 ± 0.23 .

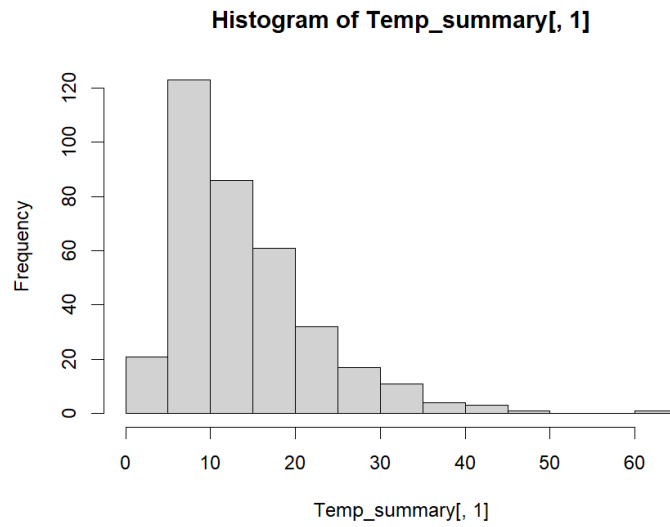


Figure S11. Inverse Temperature parameter histogram. Mean $\pm$ std 13.96 ± 8.36

Experiment 2

Individual level parameters (full tables including convergence data available at <https://github.com/hrl-team/AdvisingRL>)

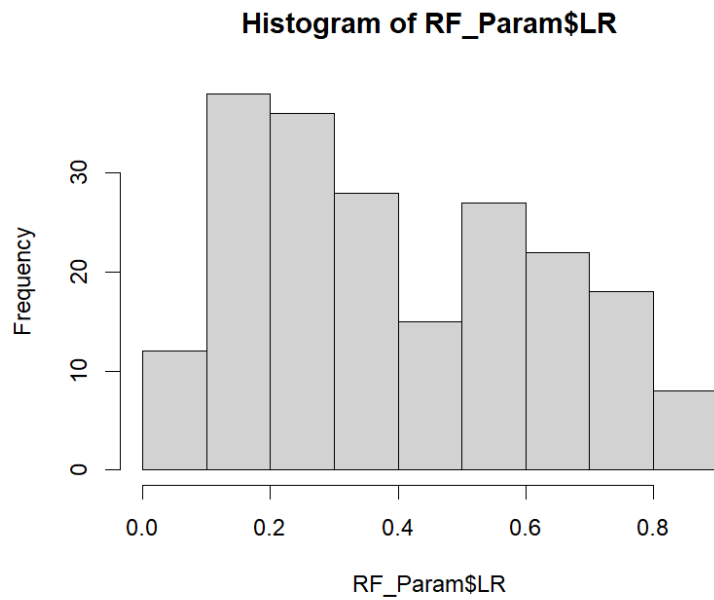


Figure S12. Learning rate parameter histogram. Mean $\pm$ std 0.39 ± 0.23 .

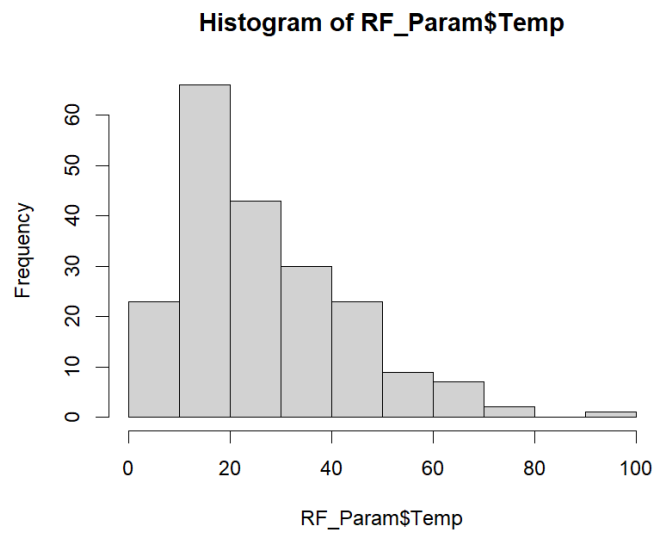


Figure S13. Inverse temperature parameter histogram. Mean $\pm$ std 26.46 ± 16.14 .

Supplementary Note 8: PRE-REGISTRATION MATCHING TABLE

Pre-registered hypothesis	As presented in the manuscript	Results
Study 1		
H1: Learners are more likely to give advice when the difference between options are more distinguishable	“Our baseline hypothesis was that likelihood to give advice would increase with expertise. Namely, we expected advisers to be more likely to give advice the easier the task was (H1.1)...”	See Results, Study 1
H2: Learners are more likely to give advice later in a learning task, after learning the options' values. H2a: An interaction between trial number and block difficulty will indicate that people learn faster when the task is easier.	“... and later in the experiment, i.e. after learning options' values (H1.2). We therefore expected to detect a significant interaction effect between trial number and block difficulty on advising (H1.2a) ”	See Results, Study 1
H3: Advice giving is related to the probability of choice, defined as the computed probability of choosing an option, which relates to the relative learned value of an option compared with all other options. H3a: Advice giving will not be related to prediction error, i.e. short term surprise outcome.	“On the basis of these predictions, we hypothesised that advising would be correlated to the probability of choice (H1.3), as estimated from a classic Q-learning model...”	See Results, Study 1 NOTE: Hypothesis H1.3 random slope component was modified from its original pre-registered version for convergence (<i>this is also disclosed in the manuscript</i>)
H4: Social anxiety levels will be negatively related to advice giving. H4a: A positive interaction	“ Finally, we expected participants' proneness to social anxiety and prosocial tendency to affect their	See Results, Study 1

between social anxiety and probability of choice will indicate that once participants with social anxiety learn about the expected value they are more likely to give advice (social pressure for advice giving	tendency to give advice (H1.4 and H1.5 respectively)”	
H5: Prosocial tendencies will be positively related to advice giving.	“ Finally, we expected participants’ proneness to social anxiety and prosocial tendency to affect their tendency to give advice (H1.4 and H1.5 respectively)”	See Results, Study 1
Study 2		
H1: The introduction of cost will reduce the number of over-advisers. [exploratory: it is possible that certain types of cost reduce over-advising more than others]	“Our main motivation consisted of comprehending how the introduction of various consequences to the act of advising would influence advising tendencies among participants. Specifically, we anticipated that the inclusion of advice consequences would significantly diminish the rate of over-advising, compared to the No Cost condition (H2.1)...”	See Results, Study 2 For simplicity of interpretation and to improve convergence, we collapsed task difficulty and reduced the random slope, namely $P(\text{Advice}) \sim AC \times (\text{Trial } N + \text{SPIN fear score} + \text{PTM altruism score}) + (1 \text{Participant})$ but the original pre-registered model that takes into account the different task difficulty levels and more complex slope is also available (see Supplementary Materials - Table S3 & Figure S1) [<i>this is also disclosed in the manuscript</i>]
H2: Prosocial individuals will have a higher tendency to over-advise. [exploratory: socially-anxious people may have a higher tendency to over-advise too, potentially because they may feel this is required of them.]	“...and that prosocial levels and social anxiety levels would again condition over-advising behaviour (H2.2). ”	See Results, Study 2

H3: Learners are more likely to give advice when the difference between options are more distinguishable. H3a: This link will be modulated by cost conditions, and potentially by prosociality and social anxiety.	“For the majority subgroup (i.e. those who did not overadvise), we expected advising to increase when the better lakes were easier to identify (i.e., on easier blocks H2.3)...”	See Results, Study 2
H4: Learners are more likely to give advice later in a learning task, after learning the options’ values. H4a: The tendency to give advice later will disappear in conditions where cost of advice is considered steep, but not where advice is still seen as affordable	“...as well as in later stages of the learning process (H2.4). These relations, in turn, would be modulated (but not extinguished) by the newly-introduced advising consequences: advising tendencies would change to reduce costs and maximise benefits of sharing information, but would remain dependent on accrued knowledge.”	See Results, Study 2
H5: Advice giving is related to the probability of choice, defined as the computed probability of choosing an option, which relates to the relative learned value of an option compared with all other options. H5a: Advice-giving will be independent of choice probability where cost of advice is considered steep, but not where advice is still seen as affordable. Prediction error will remain independent of advice-giving regardless of cost	“On the computational level, we predicted that the introduction of advice consequences would decrease the correlation between the probability of correct choice (as a proxy for accumulated knowledge on choice outcome) and the probability of advice (H2.5), as tested in Study 1.”	See Results, Study 2 Hypotheses 2.5-7 were analysed in the model: $P(advice) \sim P(choice) \times AC \times (PTM \text{ altruism score} + SPIN \text{ fear score}) + (1 Participant)$ where the random slope component was again modified from its original pre-registered version (which included probabilities and prediction error) for convergence issues [This is also disclosed in the manuscript]
H6: Social anxiety levels will be negatively related to advice giving. H6a: A positive interaction	“Lastly, at the individual differences level, we anticipated a replication of the effects of social anxiety	See Results, Study 2

between social anxiety and probability of choice will indicate that once participants with social anxiety learn about the expected value they are more likely to give advice (social pressure for advice giving). H6b: Socially-anxious participants will be more vulnerable to reputational costs	observed in Study 1, with higher social anxiety predicting higher advising probability for exploitative trials (H2.6)."	Hypotheses 2.5-7 were analysed in the model: $P(\text{advice}) \sim P(\text{choice}) \times AC \times (PTM \text{ altruism score} + SPIN \text{ fear score}) + (1 \text{Participant})$ where the random slope component was again modified from its original pre-registered version (which included probabilities and prediction error) for convergence issues [This is also disclosed in the manuscript]
H7: Prosocial tendencies will be positively related to advice giving. H7a: Prosocial participants will be more willing to pay the cost of advice	"We also predicted that higher prosociality would result in a greater tendency to give advice, particularly in the Monetary cost, as we expected these participants to privilege social exchange over cost (H2.7)"	See Results, Study 2 Hypotheses 2.5-7 were analysed in the model: $P(\text{advice}) \sim P(\text{choice}) \times AC \times (PTM \text{ altruism score} + SPIN \text{ fear score}) + (1 \text{Participant})$ where the random slope component was again modified from its original pre-registered version (which included probabilities and prediction error) for convergence issues [This is also disclosed in the manuscript]

Supplementary references

1. Ameijeiras-Alonso, J. & Crujeiras, R. M. Multimode: An R package for mode assessment.

arXiv preprint arXiv (2018).

2. Scrucca, L., Fop, M., Murphy, T. B. & Raftery, A. E. mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models. *R J.* **8**, 289–317 (2016).
3. Connor, K. M. *et al.* Psychometric properties of the Social Phobia Inventory (SPIN). New self-rating scale. *Br. J. Psychiatry* **176**, 379–386 (2000).
4. Carlo, G. & Randall, B. A. The development of a measure of prosocial behaviors for late adolescents. *J. Youth Adolesc.* doi:10.1023/A:1014033032440.
5. Beeler JA, Daw N, Frazier CRM, Zhuang X, Tonic dopamine modulates exploitation of reward learning, *Front. Behav. Neurosci.*, Sec. Motivation and Reward, Volume 4 - 2010 | <https://doi.org/10.3389/fnbeh.2010.00170>
6. Carpenter, B. *et al.* Stan: A Probabilistic Programming Language. *J. Stat. Softw.* **76**, (2017).
7. Golub, G. H. & Van Loan, C. F. *Matrix Computations*. (JHU Press, 1996).
8. Bavard S., Rustichini A., Palminteri S. Two sides of the same coin: Beneficial and detrimental consequences of range adaptation in human reinforcement learning. *Sci. Adv.* **7**, eabe0340 (2021). DOI:10.1126/sciadv.abe0340
9. Anlló, H., Bavard, S., Benmarrakchi, F. *et al.* Comparing experience- and description-based economic preferences across 11 countries. *Nat Hum Behav* **8**, 1554–1567 (2024). DOI:10.1038/s41562-024-01894-9
10. Wagenmakers, E. J. A practical solution to the pervasive problems of p values. *Psychonomic bulletin & review*, **14**(5), 779-804 (2007).
11. Palminteri, S., Wyart, V., & Koehlin, E. (2017). The Importance of Falsification in Computational Cognitive Modeling. *Trends in cognitive sciences*, **21**(6), 425–433. <https://doi.org/10.1016/j.tics.2017.03.011>

12. Wilson R.C., Collins A.G.E. (2019) Ten simple rules for the computational modeling of behavioral data eLife 8:e49547, <https://doi.org/10.7554/eLife.49547>
13. McGraw, K. O., & Wong, S. P. Forming inferences about some intraclass correlation coefficients. Psychological Methods, 1(1), 30–46. (1996)
<https://doi.org/10.1037/1082-989X.1.1.30>