Pescetelli, N., Hauperich, A.-K., & Yeung, N. (2021). **Confidence, advice seeking and changes of mind in decision making.** *Cognition, 215*, Article 104810. https://doi.org/10.1016/j.cognition.2021.104810

Confidence, advice seeking and changes of mind in decision making

Niccolò Pescetelli[1,2,*], Anna-Katharina Hauperich[1], and Nick Yeung[1]

[1]Department of Experimental Psychology, University of Oxford, Oxford, UK
[2]Max Planck Institute for Human Development, Berlin, Germany
[*]Corresponding author: Niccolò Pescetelli, niccolo.pescetelli[at]gmail.com

Abstract

Humans and other animals rely on social learning strategies to guide their behavior, especially when the task is difficult and individual learning might be costly or ineffective. Recent models of individual and group decision-making suggest that subjective confidence judgments are a prime candidate in guiding the way people seek and integrate information from social sources. The present study investigates the way people choose and use advice as a function of the confidence in their decisions, using a perceptual decision task to carefully control the quality of participants' decisions and the advice provided. The results show that reported confidence guides the search for new information in accordance with probabilistic normative models. Moreover, large inter-individual differences were found, which strongly correlated with more traditional measures of metacognition. However, the extent to which participants used the advice they received deviated from what would be expected under a Bayesian update of confidence, and instead was characterised by heuristic-like strategies of categorically ignoring vs. accepting advice provided, again with substantial individual differences apparent in the relative dominance of these strategies.

keywords: metacognition; advice taking; opinion change; social learning

Introduction

We often look for other people's opinions to improve our own judgments. Seeking advice can be costly in terms of time, energy, or opportunities forgone, yet people and organisations can be willing to pay large amounts of money to hire professional advisors – such as consultants and third-party organisations – to obtain a professional and impartial view. The pervasive reliance on learning from others' advice, despite these costs, speaks to the utility of social learning: Although solitary learning performs well in simple tasks, it becomes increasingly costly and unreliable in more difficult ones. Here, social learning achieves better performance and increased use of social information has been observed in both human and non-human animals [Toyokawa et al., 2019, Kendal et al., 2004, Morgan et al., 2012, Barkoczi and Galesic, 2016, Wisdom et al., 2013, Rendell et al., 2010]. Intuitively, looking for a second opinion can be important when outcomes of our decisions are consequential but we struggle to make a good decision alone due to lack of time, information, expertise or confidence.

Internal uncertainty monitoring is crucial to adaptive advice seeking. For example, seeking costly advice might be less beneficial when we are already certain of our decisions than when we are unsure. Notice that in this example, accurately knowing when to be uncertain and when to be sure (being *calibrated* [Fleming et al., 2014]) will influence how efficiently we look for advice. Understanding how internal metacognitive processes of uncertainty monitoring are related to social learning and advice seeking should therefore shed light on the mechanisms at the interface between subjective confidence judgments and overt behaviour [Kendal et al., 2018]. Moreover, understanding this link might provide useful when investigating metacognition in situations where obtaining a verbal confidence report might not be possible, as in non-human animals and babies.

In social and organisational psychology, a large literature has investigated the conditions under which people seek and use advice [Sniezek and Buckley, 1995, Yaniv and Kleinberger, 2000, Soll and Larrick, 2009, Harvey and Fischer, 1997]. This line of work commonly uses a "judge-advisor system" in which participants make judgments (e.g., estimating quantities, providing forecasts, or answering general knowledge questions) that are informed by advice from various sources. This approach has provided a rich corpus of findings on advice-taking behaviour and strategies. In general, people seem to give advice less weight than their own judgments ('egocentric discounting', [Yaniv and Kleinberger, 2000, Yaniv and Milyavsky, 2007]), dependent on the difficulty of the task [Gino and Moore, 2007], their own initial confidence [See et al., 2011] and the confidence and expertise of the advisor [Sniezek and Van Swol, 2001]. Advice also seems to carry more weight when paid for than when provided for free [Gino, 2008]. There are differing theoretical perspectives on how people could and should integrate advice into their judgements and decisions, contrasting normative probabilistic perspectives (e.g.,[Robalo and Sayag, 2018]) with more heuristic approaches [Soll and Larrick, 2009].

Extending this work, recent developments in the cognitive sciences, fostered by growing interest in the role of metacognition in decision making [Yeung and Summerfield, 2012], have made a connection between metacognitive and decision processes in social contexts [Fleming and Lau, 2014, Pleskac and Busemeyer, 2010, Sorkin et al., 2001, Bahrami et al., 2010]. Formal models grounded in probability theory and cognitive science allow us to mathematically describe opinion integration among individuals using subjective estimations of confidence [Bahrami et al., 2010, Sorkin et al., 2001]. Meanwhile, the use of well-characterised psychophysical decision tasks has allowed precise control over the information available to decision makers, both directly from sensory evidence and from their advisors, to reveal subtle features of how people use advice from and learn about social sources of information [Bang et al., 2017, Mahmoodi et al., 2015, Pescetelli and Yeung, 2020, 2021]. This work has focused on categorical decision tasks that require commitment to a course of action (e.g., indicating whether or not a target stimulus was presented, or which of two stimuli was larger on some perceptual dimension), rather than tasks requiring estimating or forecasting values on a continuous scale (e.g., guessing the date of historical events, the total value of coins in a pictured jar, or the distance between US cities).

The current study extended this use of carefully-controlled psychophysical decision tasks in a judge-advisor paradigm, to precisely characterise the impact of confidence on two key aspects of social decision making: whether a decision maker chooses to seek advice in the first place, and then how they subsequently integrate any advice received to update their decisions. Participants performed perceptual judgments in which they made binary choices (about which of two boxes contained more dots) and rated their confidence after each choice. In some blocks, participants could ask for advice from a virtual advisor after each decision, and were able to revise their decision (and associated confidence) in light of the advice received, which could either agree or disagree with their initial choice. Our basic expectation was that participants would ask for advice less often when they were more confident in their initial decisions [Swol and Sniezek, 2005, Sniezek and Van Swol, 2001, Tost et al., 2012, Gibbons et al., 2003]. Based on prior research, we expected subjective confidence rather than objective task difficulty would be the primary determinant of advice-seeking choices [Desender et al., 2018, 2019]. However, going beyond simply showing that people look for advice more when uncertain, we were specifically interested in how reliably they do so. Thus, we assessed the novel questions of whether participants' metacognitive ability predicted how efficiently they would ask for advice, and how consistent is the relationship between confidence and advice seeking.

In other blocks, participants were given advice freely regardless of their own decision and associated confidence, and again were invited to revise their initial decisions in light of this advice. Our basic expectations were that advice would have less influence when participants were more confident in their initial decisions, as has previously been shown in a variety of tasks including judge advisor systems, jury decision tasks and estimation tasks [Swol and Sniezek, 2005, Sniezek and Van Swol, 2001, Tost et al., 2012, See et al., 2011, Park et al., 2017, Fleming et al., 2018], and that advice would have more impact when asked and paid for than when provided freely [Gino, 2008]. Going beyond replicating these prior results, we were specifically interested in how participants would integrate advice when revising their beliefs and decisions: We investigated whether participants updated their beliefs, as reflected in their decisions and associated confidence ratings, in accordance with a normative Bayesian perspective that treats subjective confidence as a probabilistic estimate of decision accuracy (i.e., confidence as a readout of p(correct); [Aitchison et al., 2015, Meyniel et al., 2015, Park et al., 2017, Pouget et al., 2016], although see [Maniscalco et al., 2021]), and assessed how consistently this behavior was observed across participants.

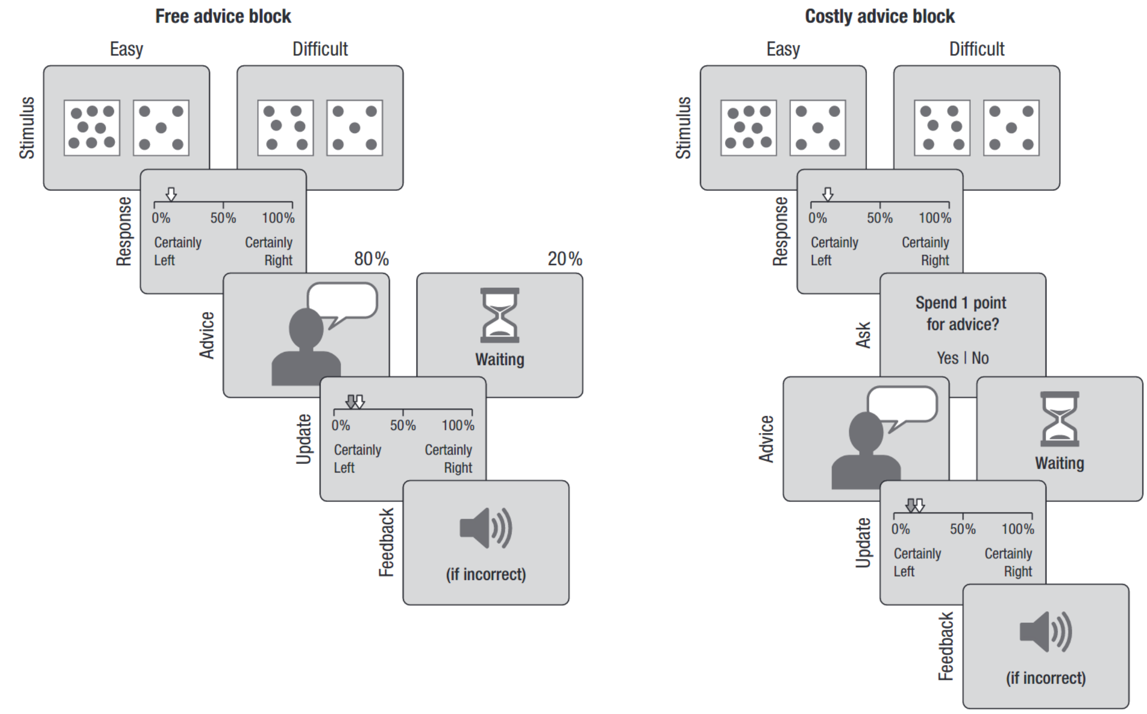

Figure 1: Experimental design. Participants made binary perceptual judgments and responded on a 100-point response scale, indicating the most likely option and their confidence. On some trials they received binary advice (agree vs. disagree) after which they could revisit their initial response. Feedback was provided at the end of each trial with a text message and auditory tone. Two difficulty conditions were alternated within-participant and across blocks, by manipulating the difference between dots in the two stimuli. Advice cost was orthogonally manipulated by providing participants with advice on 80% of free trials at no cost, or for 1 coin on Costly trials. When advice was not presented on Free trials (20%) or waived on Costly trials participants waited and skipped the revision stage.

Methods

Participants. Participants ($N = 24$, 9 females, mean age = 21.20, SD = 2.24) were recruited from the local Oxford community. All participants provided written informed consent. The study was approved by the University of Oxford's Research Ethics Committee.

Procedure. The task comprised 480 experimental trials, divided into eight blocks, with alternating conditions across blocks as described below. Each trial started with a dot-count perceptual judgment task [Boldt and Yeung, 2015, Pescetelli and Yeung, 2021]. Participants had to determine which of two boxes, presented briefly on the left and right of a central fixation cross, contained more dots (Figure 1). On each trial, one box contained $D = 200 + d$ dots, while the other contained $D = 200 - d$ dots, arrayed randomly across a 20x20 grid within each box. Task difficulty was manipulated by varying the d parameter according to two parallel interleaved staircase procedures for easy vs. difficult trials, which continued for the duration of the task. Specifically, easy and difficult trials were defined by a 3-down-1-up and a 2-down-1-up staircase procedures, respectively (step-down = 4, step-up = 3). Convergence accuracy was around 75% for easy trials and around 65% for difficult trials. Due to the fixed area of the stimuli, numerosity and density of dots perfectly correlate. We thus remain agnostic about which perceptual mechanisms lead to correct discrimination. Trial order was pseudo-randomised at the beginning of the experiment.

Participants entered their response and their confidence simultaneously on a continuous confidence scale, ranging from "certainly left" to "certainly right", that was interrupted at the centre so that participants had to commit to one or other decision on every trial (2-alternative forced-choice task). For analysis, confidence values were quantified as ranging from 1 (minimal confidence) to 50 (certain that the chosen side contains more dots). Participants entered their response by clicking with the mouse along the confidence scale and confirming their response with space bar. As soon as participants confirmed their initial response, advice could either be given to participants or not, as described below. Regardless of whether or not advice was provided, a second response was then prompted, on the same confidence scale as above, with a reminder of the original judgment kept on the scale as a shaded mark. Auditory feedback (150 ms, 400 Hz tone) was provided at the end of the trial in case of an incorrect decision, which related to the final judgment. Participants were told that they would score points on each trial according to their accuracy, +5 if correct and -5 if incorrect, and were given feedback on their current total at the end of each block.

We manipulated advice availability across blocks. In Free Advice blocks, participants received advice on a pre-determined 80% of trials. Advice was given in the form of a binary judgment ("I think it was on the [LEFT,RIGHT]"), presented in a pre-recorded native English female voice over noise-cancelling headphones. Accompanying the spoken advice was the picture of a smiling female character (NimStim database, [Tottenham et al., 2009]). Advice accuracy was fixed at 10% above the participant's expected accuracy, namely $\sim$ 85% on easy trials and $\sim$ 75% on difficult trials, to ensure that the advice was useful overall. Advice content on each trial (Left vs. Right) was pre-determined and hence was not dependent on participants' initial choice or accuracy, and thus agreed or disagreed according to the independent probabilities of participants and advisors being correct or incorrect across trials. On the remaining 20% of trials in Free Advice blocks, the advice phase was skipped and participants' initial decision was taken as final. These trials were included to encourage participants to register meaningful answers in their initial (pre-advice) decisions. In Costly Advice blocks (CA condition), after their initial decision, participants could choose between committing immediately to this decision or instead to pay a small cost (1 point) to receive advice (with properties as above) and have the opportunity to revise their judgment accordingly. The cost of advice was chosen so that the expected payoff was the same on average for skipping advice versus choosing and then following it (which would lead to a 10% improvement in accuracy, but at a 10% cost in terms of the difference between correct and incorrect answers). Trial time in Costly Advice trials was equalised to avoid participants skipping advice to shorten the duration of the study.

Prior to the main experimental blocks, participants completed three short practice blocks of 10 trials each, which successively introduced the perceptual decision task, advice, and then the option to receive vs. skip advice. After that, the 8 experimental blocks alternated between Free and Costly advice blocks. Each experimental session took approximately 1 hour. Based on total points accumulated in the task, the participant with the highest score was rewarded with a £10 shopping voucher.

Analysis. Unless otherwise specified, we report two-tail within-participant T-test statistics to show differences in averages across conditions. We use Pearson product-moment correlations in individual difference correlations between our key measures of interest. We use mixed-effect regression models (*lme4* package in R) to estimate the effect of our independent variables on final confidence, advice requests and influence (Tables S1-3). We defined random effects to take into account non-independence within participants. All p-values for regression models are estimated using the *lmerTest* package in R. Regression models were compared with a log-likelihood ratio test using the *compare* function in Matlab.

Results

Task performance and advice. Accuracy of initial decisions averaged 76% on easy trials and 66% on difficult trials, a reliable difference, ($t(23) = 16.93, p < .001$), indicating that the staircase procedure worked broadly as designed. Participants were correspondingly more confident in their initial decisions on easy trials than on difficult trials, although the difference was small (M=18 vs. M=16, respectively, on the 50-point confidence scale; $t(23) = 6.45, p < .001$). Further indicating that confidence tracked performance, participants were more confident in initial decisions when they were correct than when they were incorrect (M=18 vs. M=14, respectively; $t(23) = 8.25, p < .001$). When given the choice to receive or skip advice in Costly Advice blocks, participants opted to receive advice on 25% of trials on average. Advice overall led to improved performance, with participants being more accurate in their final decisions, after advice, than in their initial decisions before it (M=76% vs. M=71%, respectively; $t(23) = 6.18, p < .001$), and with participants who chose to receive advice more often being more accurate in their final decisions (Pearson's $r(22) = .76, p < .001$). Altogether, participants' basic task performance was as expected: Accuracy varied as a function of task difficulty, confidence varied as a function of accuracy, and participants made use of advice to improve the quality of their decisions. Of interest, therefore, are the details of how they sought and used advice across conditions and in relation to their expressed decision confidence.

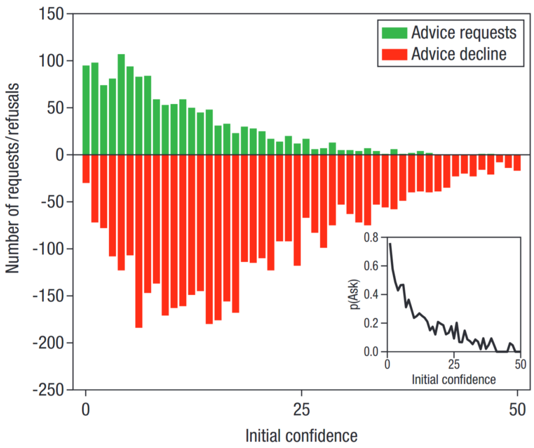

Figure 2: Number of requests and declines of advice as a function of initial confidence (pooled data across all participants). It can be seen that the probability of asking for advice declines with greater confidence, as indicated by the inset panel. Negative values on the y-axis indicate number of times the advice was declined.

Confidence and advice seeking. Advice should be more useful when participants experience low confidence because greater judgment improvements should be expected when private information is uncertain. Plotting the pooled number of advice requests vs. waivers as a function of participants' initial confidence (Figure 2) revealed a pattern consistent with this reasoning: The probability of asking for advice $p(Ask)$ decayed from about 80% for the lowest confidence levels to 0% for the highest levels. To quantify this effect, and distinguish it from a direct effect of task difficulty on advice seeking, a logistic regression was run for each participant to predict the trial-wise choice to seek vs. waive advice in Costly Advice blocks using predictors of objective difficulty and subjectively-rated confidence. Second-order statistics performed on the resulting beta coefficients revealed that confidence (mean $\beta = 0.17 \pm 0.11, t(23) = 7.31, p < .001, d = 1.49$) but not difficulty (mean $\beta = 4.10 \pm 20.50, t(23) = 0.98, p > .3, d = 0.2$) significantly predicted advice requests. The same pattern was evident in a second logistic regression analysis in which task difficulty was coded as a continuous variable according to the particular dot difference present on each trial, which was staircased throughout the experiment for both difficulty conditions (confidence: $t(23) = 7.52, p < .001, d = 1.53$; dot-difference: $t < 1$). Thus, participants' seeking of advice from social sources appeared to be guided by internal monitoring of decision confidence rather than objective task difficulty. Correspondingly, participants were more likely to ask for advice when they were initially incorrect (associated with lower confidence) than when initially correct (associated with higher confidence), $t(23) = 6.05, p < .001$.

Confidence was nevertheless an imperfect predictor of advice seeking, both between and within individuals. Figure 3 plots participants' likelihood of asking for advice as a function of their initial confidence (binned into quartiles per participant). Most participants sought less advice when confident in their initial decisions, but across individuals there was considerable variability both in the overall likelihood of seeking advice (M=25%, range = 1 - 79%, which did not correlate with participants' average confidence, Pearson's $r(22) = .15, n.s.$) and in the consistency with which confidence predicted advice requests.

In principle, changing the cost of advice would have resulted in more conservative strategic requests, or – using a signal detection theory terminology [Macmillan and Creelman, 2005] – in a change of the criterion used. To analyse the consistency of advice requests independently from the criterion used, we adopted an A_{ROC} approach, as is commonly used in metacognition research to quantify the consistency of the relationship between confidence and objective accuracy [Fleming and Lau, 2014]. Here, the probabilities of seeking vs. declining advice, conditional on initial confidence, are plotted against each other, and the area under the resulting Receiver Operating Characteristic curve (Az_{ask}) is taken as a measure of consistency in advice requests, with Az=0.5 indicating no systematic relationship between confidence and advice seeking and Az=1 indicating perfect consistency in how confidence related to choices to seek vs. waive advice. Thus calculated, Az_{ask} values ranged

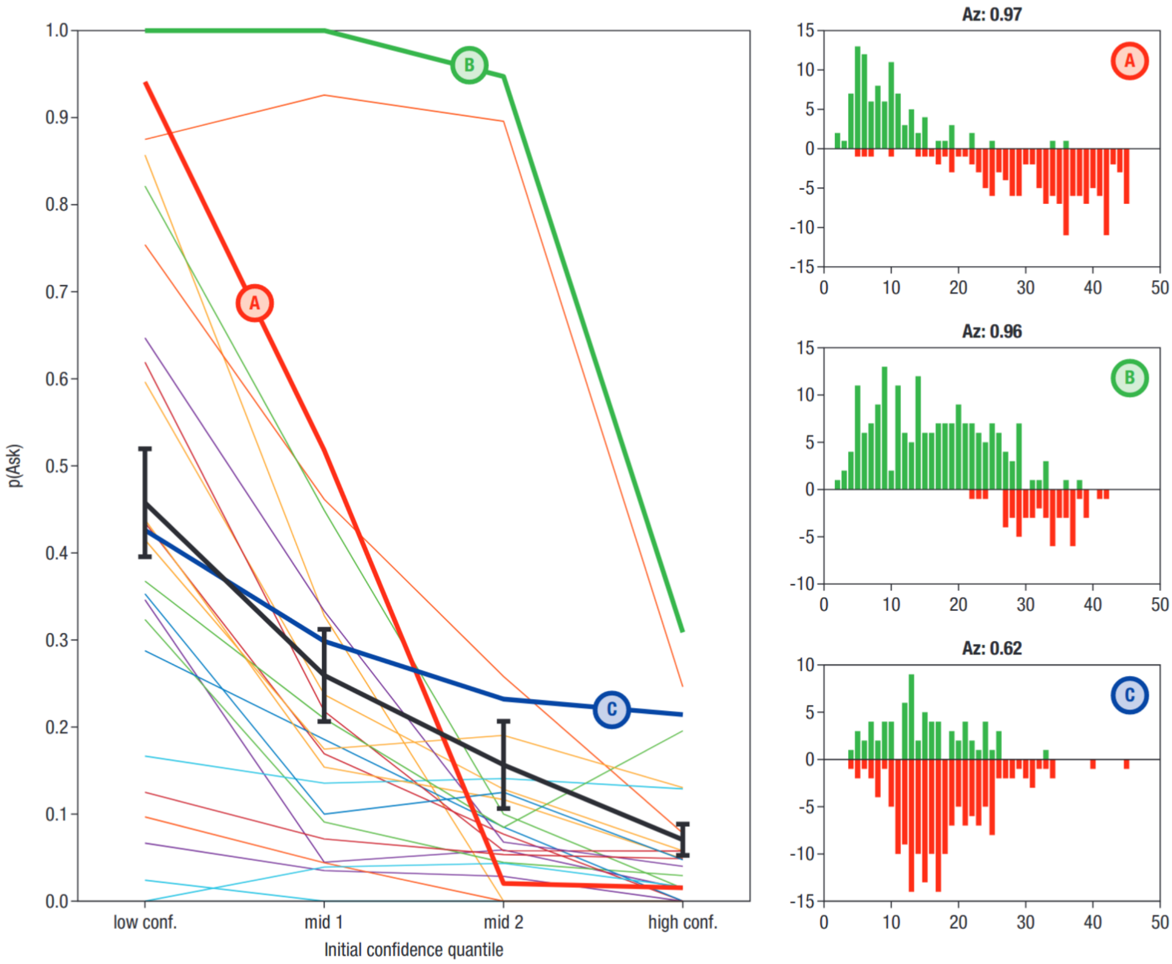

Figure 3: Probability of advice request in costly advice blocks for each individual participant. The solid black line indicates averages across individuals. Error bars represent s.e.m. The three inset panels show the detailed pattern of advice seeking behavior as a function of confidence for sample participants, two showing a consistent relationship between confidence and advice seeking (A and B inset panels) and one showing an inconsistent relationship (C inset panel). Negative values on the y-axis indicate number of times the advice was declined.

from 0.41 to 0.97 across participants (M=0.77, Figure S4): Some participants were very consistent in their advice-seeking choices, seeking advice at lower levels of confidence and declining advice at higher levels of confidence, even if the confidence criterion determining this choice was idiosyncratic (compare A and B inset panels in Figure 3). Other participants showed a much less consistent relationship between initial confidence and advice choice (C inset panel in Figure 3). Expected value differences between asking and waiving advice in different conditions seemed to predict the probability of asking for advice as well as explain some of the inter-individual variation observed in Az_{ask} (Supplementary Methods 1.1-1.2). However, as we did not manipulate cost in Costly blocks, we limited our modelling analysis to a normative account of advice use and participants' deviations from it.

We compared the consistency of participants' advice seeking choices (Az_{ask}) with the calibration of their confidence judgements; that is, the correlation of these judgements with objective accuracy [Roediger III et al., 2012, Henmon, 1911, Fleming and Lau, 2014, Koriat, 2012]. Measures of calibration – or the ability to accurately represent the probability of being correct with a confidence report – are often used as an indicator of the sensitivity of internal metacognitive processes, which monitor the perceptual uncertainty of an agent. We calculated each participant's metacognitive calibration using a corresponding ROC measure, Az_{conf}, based on plots of the probabilities of correct and incorrect responses conditional on rated confidence (Figure S5). We found a consistent correlation between participants' Az_{conf} and Az_{ask} scores, even when calculated for separate trial blocks (from Free Advice and Costly Advice blocks, respectively), Pearson's $r(22) = .67, p < .001$, providing further evidence for the relationship between subjective confidence and advice seeking consistency.

However, we also observed meaningful differences between confidence and advice-seeking. Az_{conf} was reliably higher on easy than difficult trials, $t(23) = 4.28, p < .001$, as is typically observed, whereas Az_{ask} tended to be slightly higher in difficult trials, $t(22) = 2.30, p < .05$ (this t-test excluded one participant who asked for advice very rarely, and not at all in one condition). Thus, the consistency of advice seeking did not depend on task difficulty, even though the reliability of confidence judgments as a predictor of objective accuracy did show this dependence. Dissociations of this kind are perhaps relevant to studies in non-verbal participants such as animals and infants, where confidence is typically inferred from behavioral proxies including information seeking [Call and Carpenter, 2001, Goupil et al., 2016], opting-out of a choice [Kiani and Shadlen, 2009], willingness to pay [Kepecs and Mainen, 2012] or willingness to wait for reward [Kepecs et al., 2008]. In our dataset, advice seeking was only a modestly valid proxy for confidence: We calculated, separately for each participant, the probability that declining advice was associated with high confidence (i.e., above median) and that seeking advice was associated with low confidence (i.e., below median). Both values were consistently above chance, but showed considerable variability across individuals (p(HighConfidence|DeclineAdvice): M=0.59, $t(23) = 3.41, p < .01$, range 0.46 - 1.0; p(LowConfidence|SeekAdvice): M=0.78, $t(23) = 7.45, p < .001$, range 0.2-1.0). This variability reflects two features of the data shown in Figure 3: large individual differences in the overall likelihood of asking for advice, and (variable) inconsistency in the relationship between confidence and advice-seeking.

Confidence and advice use. Participants rated their confidence both before and after receiving advice, enabling us to assess the impact of advice in terms of how participants changed their mind or their confidence from initial to final decision. To illustrate key trends in the data, Figure 4 plots final decision confidence as a function of initial confidence, pooled across all advice trials from all participants (data separated per participant are shown in Figure S2). We divided trials according to trial-varying consensus (advice agreed vs. disagreed with the participant's initial decision) and advice condition (Free vs. Costly). Negative values on the y-axis indicate that the participant changed their mind about which box contained more dots after receiving advice.

On average, participants changed their mind less than half the time they received advice disagreeing with their initial decision (M=32% in Free Advice blocks, where advice was delivered irrespective of participants' initial confidence, <50%: $t(23) = 4.33, p < .001$). Thus, participants tended to weight their own initial decision more heavily than the advice they received, consistent with previous evidence of 'egocentric discounting' [Yaniv and Kleinberger, 2000, Yaniv and Milyavsky, 2007] and 'naïve realism' [Liberman et al., 2012, Ross and Ward, 1995]. Indeed, many data points in Figure 4 fall exactly along the line y = x that represents no-change from pre- to post-advice phase; i.e., ignoring the advice received. However, use of advice varied considerably across participants, with changes of mind following disagreeing advice varying from 0% to 72% across participants. Higher rates of changes in mind were seen in participants who asked for advice more often in Costly Advice blocks (Pearson's $r(22) = .66, p < .001$), indicating stable individual differences in how advice was valued despite careful control of its objective utility. Participants were more likely to change their mind when their initial decision was incorrect than when they were correct (M=35% vs. 25%, $t(23) = 3.96, p < .001$), likely because they were more confident in decisions that were objectively correct. Indeed, changes of mind decreased monotonically as a function of confidence when trials were binned into quartiles according to participants' initial confidence ratings, from 54% to 34% to 20% to 11% across confidence quartiles (all successive pairwise contrasts $t(23) > 3.4, p < .01$).

Graded changes in confidence were also observed: On average, confidence increased in agreement trials (data points above the red line) and decreased in disagreement trials (data points below the red line) as would be expected. The impact of advice was broadly similar whether it was free or costly; the main difference in the plots is the sparser sampling of trials with high initial confidence in the Costly Advice condition, reflecting participants' tendency to decline advice in these trials. We modeled how advice is used–to evaluate whether people approximate Bayesian belief updating–independently from the costs and benefits associated with choosing vs. waiving advice (because we did not vary the cost of advice across Costly Advice blocks, and the benefits were equivalent across Free and Costly Advice blocks because the advice source was always the same)(Supplementary Methods 1.1). We found that the detailed patterns of belief update following advice are not those predicted by a straightforward Bayesian account in which confidence is treated as a probabilistic estimate of the probability of bring correct that is updated as new evidence arrives [Aitchison et al., 2015, Meyniel et al., 2015, Pouget et al., 2016]. On this account, the impact of advice should be maximal when initial confidence is low (see Supplementary Methods 1.1 and Figure S6). However, if anything, we see a larger average impact of advice as initial confidence increases, as shown in Figure 5 which plots the mean confidence change (from first to second decisions) that was observed in agreement and disagreement trials as a function of initial confidence quantile, trial difficulty and condition (Free vs. Costly). Particularly for disagreement trials, we find a larger average influence of advice as initial confidence increased. A further detail of the results that is inconsistent with a normative Bayesian account is that costly advice was more influential than free advice, despite carrying equivalent informational value (because it always came from the same source of fixed reliability).

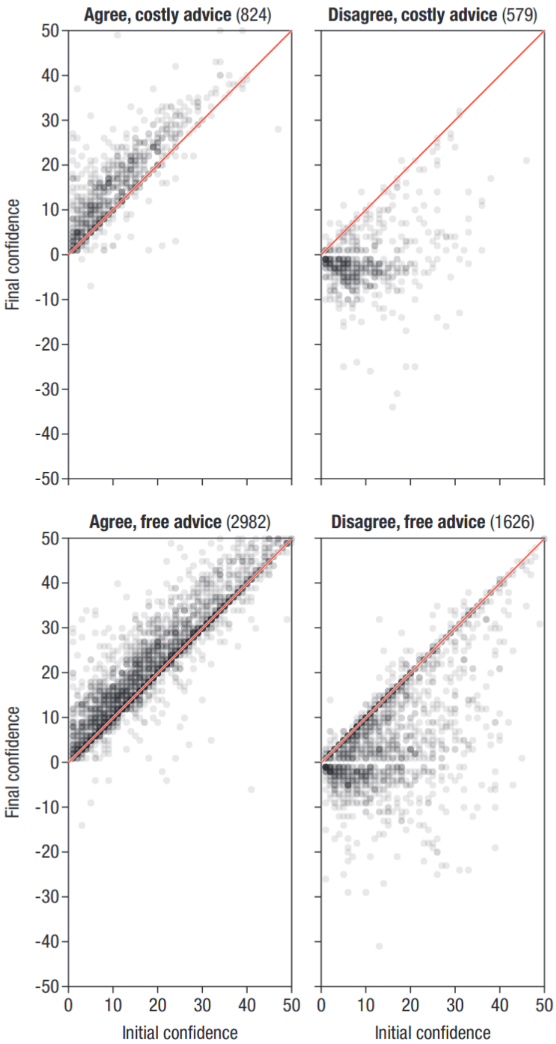

Figure 4: Post-advice confidence as a function of pre-advice confidence, divided by consensus (agree *vs.* disagree) and condition (free vs. costly). Numbers in parenthesis indicate the number of data points plotted.

To analyze these results, we ran a mixed-effects linear regression on final confidence with predictors or initial confidence, consensus (agree vs. disagree), condition (free vs. costly), difficulty (easy vs. difficult), and whether advice was requested on the previous trial. All variables except for confidence were declared as categorical. The analysis has the benefit of clustering trial-level data according to participant identity and thus identifies significant effects over and above inter-individual differences in response measure (intercept as the only random effect). The full model was run, namely all main effects and all interaction terms. For the sake of space, we report only significant effects in Table 1 and refer to Table S1 for the full model, including 95% confidence intervals. It can be seen that both initial confidence and advisor agreement are strong predictors of final confidence ($\beta > 0.72, SE < 0.48$), with a positive interaction between them ($\beta = .21, SE = .02$). This interaction indicated that on average the agreement effect, namely greater final confidence after agreement compared to disagreement, increases for larger initial confidence. Importantly however, the interaction is negatively modulated by condition, indicating that when advice came with a personal cost this initial confidence by agreement interaction was stronger compared to when the advice was freely available ($\beta = 0.27, SE = .07$). Individual participants' plots are shown in Supplementary Figures S1-S2.

As shown in Figure 5, in agreement trials, greater confidence increases took place after low initial confidence judgments and smaller confidence increases took place after high initial confidence judgments, as expected by normative Bayesian models (Supplementary Methods 1.1, Figure S6). Notice however the possible ceiling effect – at highest levels of confidence, it is not possible to increase confidence any further. In disagreement on the contrary, larger confidence decreases were observed when participants started with higher levels of confidence than when they started with lower levels of confidence. This result is at odds with the Bayesian prediction that larger updates should be observed when initial confidence (i.e., p(correct)) is low. Instead, most of the datapoints seem to fall within three distinct responses to receiving disagreeing advice [Soll and Larrick, 2009]: ignoring it completely (in Figure 4, points on the line y = x), keeping with the initial decision but with slightly reduced confidence, or a change of mind that is associated with minimal confidence in the final decision (points close to the line y=0 in Figure 4). This latter response explains the apparent larger impact of disagreeing advice at higher levels of initial confidence: As initial confidence increases, a larger shift on the response scale is needed to register a change of mind (with an approximately fixed, low level of confidence).

Once again we observed marked individual differences in how people treated advice, which were manifest here in terms of the relative prevalence of these categorically different responses to disagreement (see Supplemental Figure S3): Some participants always ignored the advice, some always showed graded changes in confidence but rarely changed their mind, and others frequently changed their minds but with minimal confidence in their final decision. Most participants showed a mix of these responses. None showed a clear Bayesian-like pattern whereby disagreeing advice had its largest impact when initial confidence was low.

The average correlation coefficient across participants between the empirical confidence change observed and the confidence change predicted by a Bayesian confidence update model was significantly larger in costly advice (M=0.75, SD=0.13) than free advice (M=0.53, SD=0.18) ($t(22) = 6.529, p < .001, d = 1.414$), likely due to the larger number of advice refusals observed in high confidence costly trials. Thus, even controlling for initial confidence, participants seemed to use advice more when they paid for it. This effect could be interpreted as a form of sunk cost fallacy. However, it actually made people more rational because advisors were by design more accurate on average than participants and participants tended to under-use free advice [Yaniv and Kleinberger, 2000].

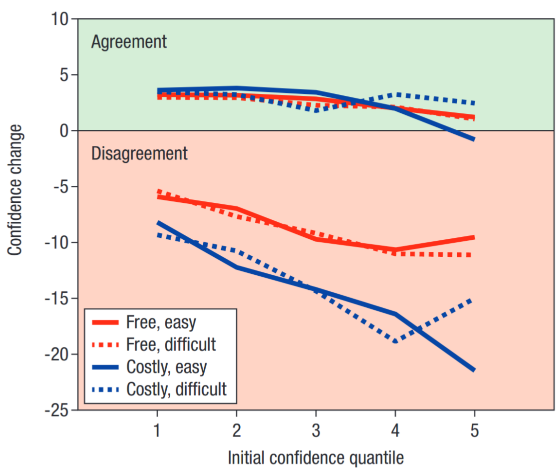

Figure 5: Confidence change as a function of initial confidence quantile, divided by condition and difficulty. Some participants displayed empty cells, which were thus not included in the averages plotted here. For single data points trends (not averaged), please refer to the mixed effects analysis (Tables 1 and S1) which is less affected by the problem of missing cells.

	Estimate	SE	tStat	DF	p
$Intercept$	-4.23	0.53	-7.91	5960	<.001***
$conf_{init}$	0.72	0.02	30.08	5960	<.001***
$agree$	7.78	0.53	14.43	5960	<.001***
$conf_{init} : agree$	0.21	0.02	7.73	5960	<.001***
$conf_{init} : cost$	-0.28	0.06	-4.59	5960	<.001***
$conf_{init} : ask$	-0.88	0.51	-1.70	5960	.08
$conf_{init} : agree : cost$	0.27	0.07	3.50	5960	<.001***
$conf_{init} : diff : ask$	1.24	0.57	2.15	5960	0.03
$conf_{init} : cost : diff : ask$	-1.51	0.59	-2.54	5960	0.01

Table 1: Significant effects of a linear mixed-effects model run on participants final confidence. Main effects include: initial confidence ('conf$_{init}$'), agreement ('agree', reference: disagreement), advice cost ('cost', reference: free advice), task difficulty ('diff', reference: easy), and whether the participant asked for advice in the previous trial ('ask', reference: participant did not ask for advice). The full model is reported in Supplementary Table 1.

Reciprocity. Past research shows that prior agreement with an advisor affects social influence of the advice beyond accuracy [Pescetelli and Yeung, 2021] and advisors who are more susceptible to advice themselves are also more influential [Mahmoodi et al., 2018]. Furthermore, people show a strong equality bias, namely they tend to weigh each other's opinion equally regardless of differences in their reliability, notwithstanding explicit performance feedback and monetary incentives [Mahmoodi et al., 2015]. We thus tested whether influence was predicted by reciprocity, defined as whether the advisor agreed or disagreed with the participant in the previous trial. We tested whether agreement in the previous trial ($t-1$) temporarily increased the advisor's influence on the following trial ($t0$). We ran a linear regression on influence and included fixed effects for initial confidence and agreement in the current trial $t0$, and included all interactions terms. Results are reported in Table S3. We found that agreement on the previous trial predicted larger influence on the current trial ($\beta = 1.81, SE = 0.62, t(df) = 2.89(4179), p = .003$), although to a much smaller degree than current agreement ($\beta = 9.44, SE = 0.63, t(df) = 14.76(4179), p < .001$). As expected, greater initial confidence on the current trial predicted lower influence ($\beta = -0.19, SE = 0.02, t(df) = -6.73(4179), p < .001$). Finally, we found a negative interaction between initial confidence in the current trial and prior trial agreement ($\beta = -0.11, SE = 0.03, t(df) = -3.22(4179), p = .001$), and a positive interaction between current trial initial confidence and agreement ($\beta = 0.16, SE = 0.03, t(df) = 4.54(4179), p < .001$). The difference in sign is attributed to the choice of reference (disagreement) as well as the larger effect of current agreement compared to prior agreement. Overall, these results show that reciprocity had a small but significant effect on advisor influence, evident as larger confidence changes in trials following agreement trials. On the contrary, reciprocity did not seem to affect the probability of asking for advice (not reported).

Discussion

Situations in which advice is freely provided or actively sought are common in our daily life, yet much remains unknown regarding the mechanisms governing how people search for and integrate new information from social others. In the current experiment, participants performed a series of binary-choice perceptual decisions. We systematically manipulated task difficulty and the cost of advice, and recorded participants' trial-wise confidence in their initial decisions, as three potentially critical determinants of advice seeking and advice use. Of interest was how participants' advice-seeking behaviour related to their subjectively-rated confidence, and conversely how their decisions and confidence were impacted by the advice provided.

We found that, when offered the choice to pay for advice, participants used this opportunity coherently with their initially expressed confidence, asking for advice more often when unsure than when sure [Gibbons et al., 2003, See et al., 2011, Tost et al., 2012]. The probability of asking for advice was not predicted by task difficulty, over and above this effect of initial confidence. Thus, what mattered for advice seeking was the perceived difficulty of trial, represented by a confidence judgment, rather than objective difficulty. Although objective difficulty is known to affect performance and should therefore affect advice seeking, trial-by-trial fluctuations in sensory or internal noise, attention and other contingent factors might have lessened the effect of trial difficulty. Variability in confidence reports on the contrary, precisely reflect these sources of variability and correspondingly emerge as a stronger predictor of advice seeking. Other factors, such as reciprocity, did not affect advice seeking even though they had a small effect on advice influence.

Notwithstanding this consistent relationship between confidence and advice-seeking, we observed striking variation across participants in their advice-seeking behavior. One dimension of variation was in the simple likelihood of asking for advice, which ranged widely from 1% to over 75% across participants. Advice in our task was helpful by design. Correspondingly, participants who asked for advice more often also showed greater final task accuracy. Nevertheless, even when provided with the opportunity to learn the reliability of their advisors, because objective feedback was provided, participants relied on the advice to different extents. This variation was not due to differences in performance (e.g., some participants needed the advice less than others) because task performance and the quality of advice were both carefully controlled. We have observed similarly high variability in information-seeking behavior in other contexts, in which participants could seek an external hint rather than advice ostensibly from another person [Desender et al., 2018, 2019], suggesting that the variation is not a purely social-learning phenomenon. The state and trait factors that determine these variations in advice seeking are an important issue for future research, and likely reflect a range of factors including sensitivity to costs and payoffs of advice as well as social factors such as agreement and reciprocity [El Zein et al., 2019, Mahmoodi et al., 2015] (Supplementary Methods 1.3 and Table S3). Our findings identify one notable source of this variability, in terms of the correlation we observed such that participants who chose advice more often (in Costly Advice blocks) were also more influenced by the advice they received (in Free Advice blocks). One interpretation of this correlation is that participants seek advice to the extent that they expect it could materially affect their decisions. In the Supplementary Material, we outline a simple computational model that incorporates likely costs and benefits into decisions whether or not to seek advice.

A second key dimension of variability across participants was in terms of the consistency of their advice requests and declines. Advice-seeking consistency as a function of confidence (or Az_{ask}) strongly correlated with an established measure of confidence calibration, Az_{conf}. Conceptually, this finding provides further evidence that advice seeking depends on subjective confidence and, more broadly, that metacognitive processes play a critical role in regulating social learning and decision making [Bang et al., 2017, Bahrami et al., 2010, Pescetelli and Yeung, 2021, Bonaccio and Dalal, 2006]. Moreover, it indicates that individual differences in confidence reports do not solely reflect idiosyncrasies in how people communicate their internal states [Navajas et al., 2017], but that these individual differences at least partly reflect meaningful variation in the internal states that govern decision-making strategies such as information seeking. In this way, methodologically, our findings support the assumption that advice-seeking behaviour can provide a valid external measure of metacognitive ability in participants who are unable to provide explicit verbal reports, such as animals and pre-verbal infants [Goupil et al., 2016, Kornell et al., 2007]. However, our results also suggest limitations in the validity of behavioral proxies for confidence: Even within our small sample of participants, all performing the task to the same overall level of accuracy, some individuals' advice-seeking was an almost perfectly consistent readout of their confidence, whereas for others the relationship was indistinguishable from chance.

Participants' use of advice in our paradigm replicated previous findings that people egocentrically discount advice [Yaniv and Kleinberger, 2000, Yaniv and Milyavsky, 2007, Liberman et al., 2012, Ross and Ward, 1995] and place greater weight on costly advice than advice that is freely provided [Gino, 2008]. Beyond this, we found that use of advice to update beliefs and decisions showed similar dependence on confidence and individual differences as did advice seeking. Confidence, but not task difficulty, was a significant predictor of advice use such that, once again, subjective estimations of uncertainty were a more reliable predictor of behaviour than objective task measures. This dependence on confidence was evident both in terms of overt changes of mind from pre- to post-advice, as well as in more subtle adjustments in confidence. In this regard, our findings replicate previous observations that people rely more on advice when they are initially unsure [Swol and Sniezek, 2005, Sniezek and Van Swol, 2001, Tost et al., 2012, See et al., 2011, Park et al., 2017, Fleming et al., 2018]. This pattern is consistent with a Bayesian interpretation of belief updating, which takes into account the relative reliability of one's initial opinion and the received advice [Park et al., 2017, De Martino et al., 2017]. However, patterns of graded change in confidence following advice deviated from what normative accounts would predict based on a probabilistic interpretation of confidence as a subjective estimate of p(correct) that can be updated following advice [Maniscalco et al., 2021]. Graded changes in confidence were also affected by task irrelevant factors such as reciprocity [Mahmoodi et al., 2018].

A Bayesian observer who treated confidence as monotonically related to their p(correct), and similarly treated advice as an imperfect guide to the correct answer, would be more influenced by advice when low in confidence in their initial response [Bahrami et al., 2010, Park et al., 2017]. This pattern was somewhat evident in the case of agreeing advice, although here it might be an artifact of ceiling effects in the confidence scale. In disagreement trials, on the contrary, participants showed evidence of being more influenced by advice the more confident they initially were, exhibiting larger average decreases in confidence as their initial confidence increased. In principle, a more complex Bayesian observer could demonstrate such a pattern if they did not treat advice as having a fixed reliability across all trials, but rather attributed the same level of confidence they have on a given trial to the advisors as well. That is, when participants found the trial impossible and guessed their answer, they could treat advice as similarly unreliable; when they felt there was useful evidence in the stimulus, advice might correspondingly be treated as more valid. This attribution error makes sense in a world in which individuals share similar perceptual cognitive systems and biases (e.g., where a difficult task is considered difficult by everybody). In this case, the influence of advice would, paradoxically, tend to increase with a participants' own initial confidence.

However, this more complex Bayesian inference model still does not capture the full pattern of results evident in our confidence-change plots (Figure 4), which exhibit categorically distinct responses to disagreement across trials and across participants and, hence, are not readily captured by any single computation. Instead, our participants' responses to disagreement indicate a mix of strategies, a conclusion that concurs with previous findings from estimation tasks [Soll and Larrick, 2009], where participants have been found to adopt distinct strategies of choosing (i.e., choosing either their initial estimate or the advisors') vs. averaging the two opinions. In our data, we likewise observed distinct tendencies to ignore advice completely on some trials vs. making use of it. When using advice, participants again showed a mixture of strategies: keeping with their initial decision but with reduced confidence vs. changing their mind in line with the advice, in which case they typically expressed minimal confidence in the final decision irrespective of their own initial confidence. The prevalence of each response varied substantially across participants. As noted above, the tendency to use advice when it was provided correlated strongly with participants' choices to seek advice when it was offered, suggesting widely differing perceptions of the value of advice, the value of updating beliefs and decisions, or both. The source of these individual differences– which, as noted above, are apparent despite our careful matching of decision accuracy and advice quality across participants–represents a potentially fruitful avenue for future research. We speculate that patterns of advice use are characteristic to individuals in the same way as are subjective confidence reports [Ais et al., 2016], and might reflect multiple aspects of an individual decision maker (e.g., their sensitivity to time, effort, risk and regret) and their social situation (e.g., their sensitivity to agreement and reciprocity), which would have interesting implications both theoretically and practically.

Conclusions

Integrating information coming from social others is essential to our daily life. People often receive advice and ask for it, particularly when they lack competencies, information or the ability to look for relevant evidence. The present work provides a systematic investigation of advice seeking and advice use in a binary decision task. We find that confidence plays a critical role in the way people seek and use advice, such that their confidence ratings predict their decisions to seek vs. decline advice, their metacognitive abilities predict the efficiency with they seek advice, and their advice-seeking behaviour is a reasonably reliable proxy for subjective confidence. Conversely, confidence ratings reveal key features of how people use advice to inform their decisions, which are better characterised as a mixture of heuristic strategies rather than Bayesian belief updating. Altogether our findings are consistent with the idea that people represent the certainty of their beliefs and decisions to guide adaptive information seeking, but nevertheless show a high degree of variability in their strategies for doing so and for updating their decisions and beliefs on the basis of the information uncovered.

Replicability

Data and code to reproduce analyses, tables, and figures can be found via Open Science Framework: Pescetelli, N., Yeung, N., & Hauperich, A.-K. (2020, October 4). Confidence, advice seeking and changes of mind in decision making. Retrieved from osf.io/z8vay.

Bibliography

References

Wataru Toyokawa, Andrew Whalen, and Kevin N. Laland. Social learning strategies regulate the wisdom and madness of interactive crowds. *Nature Human Behaviour*, 3(2):183–193, 2 2019. ISSN 2397-3374. doi: 10.1038/s41562-018-0518-x. URL `http://www.nature.com/articles/s41562-018-0518-x`.

R. L. Kendal, I. Coolen, and K. N. Laland. The role of conformity in foraging when personal and social information conflict. *Behavioral Ecology*, 15(2):269–277, 3 2004. ISSN 1465-7279. doi: 10.1093/beheco/arh008. URL `https://academic.oup.com/beheco/article-lookup/doi/10.1093/beheco/arh008`.

T J H Morgan, L E Rendell, M Ehn, W Hoppitt, and K N Laland. The evolutionary basis of human social learning. *Proceedings of Royal Society B*, 279 (July 2011):653–662, 2012. doi: 10.1098/rspb.2011.1172.

Daniel Barkoczi and Mirta Galesic. Social learning strategies modify the effect of network structure on group performance. *Nature Communications*, 7:13109, 10 2016. ISSN 2041-1723. doi: 10.1038/ncomms13109. URL `http://www.nature.com/doifinder/10.1038/ncomms13109`.

Thomas N. Wisdom, Xianfeng Song, and Robert L. Goldstone. Social Learning Strategies in Networked Groups. *Cognitive science*, 37:1383–1425, 2013. doi: 10.1111/cogs.12052.

L. Rendell, R. Boyd, D. Cownden, M. Enquist, K. Eriksson, M. W. Feldman, L. Fogarty, S. Ghirlanda, T. Lillicrap, and K. N. Laland. Why Copy Others? Insights from the Social Learning Strategies Tournament. *Science*, 328(5975): 208–213, 4 2010. ISSN 0036-8075. doi: 10.1126/science.1184719. URL `http://www.sciencemag.org/cgi/doi/10.1126/science.1184719`.

Stephen M Fleming, Brian Maniscalco, Yoshiaki Ko, Namema Amendi, Tony Ro, and Hakwan Lau. Action-Specific Disruption of Perceptual Confidence. *Psychological science*, 2014. ISSN 1467-9280. doi: 10.1177/0956797614557697.

Rachel L. Kendal, Neeltje J. Boogert, Luke Rendell, Kevin N. Laland, Mike Webster, and Patricia L. Jones. Social Learning Strategies: Bridge-Building between Fields. *Trends in Cognitive Sciences*, 22(7):651–665, 7 2018. ISSN 13646613. doi: 10.1016/j.tics.2018.04.003. URL `https://linkinghub.elsevier.com/retrieve/pii/S1364661318300949`.

Janet A. Sniezek and Timothy Buckley. Cueing and Cognitive Conflict in Judge-Advisor Decision Making. *Organizational Behavior and Human Decision Processes*, 62(2):159–174, 5 1995. ISSN 07495978. doi: 10.1006/obhd.1995.1040. URL `http://linkinghub.elsevier.com/retrieve/pii/S0749597885710400`.

Ilan Yaniv and Eli Kleinberger. Advice Taking in Decision Making: Egocentric Discounting and Reputation Formation. *Organizational behavior and human decision processes*, 83(2):260–281, 11 2000. ISSN 0749-5978. doi: 10.1006/obhd.2000.2909. URL `http://www.ncbi.nlm.nih.gov/pubmed/11056071`.

Jack B Soll and Richard P Larrick. Strategies for revising judgment: how (and how well) people use others' opinions. *Journal of experimental psychology: Learning, memory, and cognition*, 35(3): 780–805, 5 2009. ISSN 0278-7393. doi: 10.1037/a0015145. URL `http://www.ncbi.nlm.nih.gov/pubmed/19379049`.

Nigel Harvey and Ilan Fischer. Taking Advice: Accepting Help, Improving Judgment, and Sharing Responsibility. *Organizational Behavior and Human Decision Processes*, 70(2):117–133, 5 1997. ISSN 07495978. doi: 10.1006/obhd.1997.2697. URL `http://linkinghub.elsevier.com/retrieve/pii/S0749597897926972`.

Ilan Yaniv and Maxim Milyavsky. Using advice from multiple sources to revise and improve judgments. *Organizational Behavior and Human Decision . . .*, 103:104–120, 2007. doi: 10.1016/j.obhdp.2006.05.006. URL `http://www.sciencedirect.com/science/article/pii/S074959780600063X`.

Francesca Gino and Don A. Moore. Effects of task difficulty on use of advice. *Journal of Behavioral Decision Making*, 20(1):21–35, 1 2007. ISSN 08943257. doi: 10.1002/bdm.539. URL http://doi.wiley.com/10.1002/bdm.539.

Kelly E. See, Elizabeth W. Morrison, Naomi B. Rothman, and Jack B. Soll. The detrimental effects of power on confidence, advice taking, and accuracy. *Organizational Behavior and Human Decision Processes*, 116(2):272–285, 11 2011. ISSN 07495978. doi: 10.1016/j.obhdp.2011.07.006. URL http://linkinghub.elsevier.com/retrieve/pii/S0749597811000975.

Janet A. Sniezek and Lyn M. Van Swol. Trust, Confidence, and Expertise in a Judge-Advisor System. *Organizational behavior and human decision processes*, 84(2):288–307, 3 2001. ISSN 0749-5978. doi: 10.1006/obhd.2000.2926. URL http://www.ncbi.nlm.nih.gov/pubmed/11277673.

Francesca Gino. Do we listen to advice just because we paid for it? The impact of advice cost on its use. *Organizational Behavior and Human Decision Processes*, 107(2):234–245, 11 2008. ISSN 07495978. doi: 10.1016/j.obhdp.2008.03.001. URL https://linkinghub.elsevier.com/retrieve/pii/S0749597808000435.

Pedro Robalo and Rei Sayag. Paying is believing: The effect of costly information on Bayesian updating. *Journal of Economic Behavior & Organization*, 156:114–125, 12 2018. ISSN 01672681. doi: 10.1016/j.jebo.2018.09.016. URL https://linkinghub.elsevier.com/retrieve/pii/S0167268118302646.

Nick Yeung and Christopher Summerfield. Metacognition in human decision-making: confidence and error monitoring. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 367(1594):1310–21, 5 2012. ISSN 1471-2970. doi: 10.1098/rstb.2011.0416. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3318764&tool=pmcentrez&rendertype=abstract.

Stephen M. Fleming and Hakwan C. Lau. How to measure metacognition. *Frontiers in Human Neuroscience*, 8, 7 2014. ISSN 1662-5161. doi: 10.3389/fnhum.2014.00443. URL http://www.frontiersin.org/Human_Neuroscience/10.3389/fnhum.2014.00443/abstract.

Timothy J. Pleskac and Jerome R. Busemeyer. Two-stage dynamic signal detection: A theory of choice, decision time, and confidence. *Psychological Review*, 117(3):864–901, 2010. ISSN 1939-1471. doi: 10.1037/a0019737. URL http://doi.apa.org/getdoi.cfm?doi=10.1037/a0019737.

Robert D. Sorkin, Christopher J. Hays, and Ryan West. Signal-detection analysis of group decision making. *Psychological Review*, 108(1):183–203, 2001. ISSN 1939-1471. doi: 10.1037/0033-295X.108.1.183. URL http://doi.apa.org/getdoi.cfm?doi=10.1037/0033-295X.108.1.183.

Bahador Bahrami, Karsten Olsen, Peter E Latham, Andreas Roepstorff, Geraint Rees, and Chris D Frith. Optimally interacting minds. *Science (New York, N.Y.)*, 329(5995):1081–5, 8 2010. ISSN 1095-9203. doi: 10.1126/science.1185718. URL http://www.ncbi.nlm.nih.gov/pubmed/20798320.

Dan Bang, Laurence Aitchison, Rani Moran, Santiago Herce Castanon, Banafsheh Rafiee, Ali Mahmoodi, Jennifer Y. F. Lau, Peter E. Latham, Bahador Bahrami, and Christopher Summerfield. Confidence matching in group decision-making. *Nature Human Behaviour*, 1(0117):1–7, 2017. doi: 10.1038/s41562-017-0117.

Ali Mahmoodi, Dan Bang, Karsten Olsen, Yuanyuan Aimee Zhao, Zhenhao Shi, Kristina Broberg, Shervin Safavi, Shihui Han, Majid Nili Ahmadabadi, Chris D. Frith, Andreas Roepstorff, Geraint Rees, and Bahador Bahrami. Equality bias impairs collective decision-making across cultures. *Proceedings of the National Academy of Sciences*, page 201421692, 2015. ISSN 0027-8424. doi: 10.1073/pnas.1421692112. URL http://www.pnas.org/lookup/doi/10.1073/pnas.1421692112.

Niccol Pescetelli and Nick Yeung. The effects of recursive communication dynamics on belief updating. *Proceedings of the Royal Society B: Biological Sciences*, 287(1931):20200025, 7 2020. ISSN 0962-8452. doi: 10.1098/rspb.2020.0025. URL https://royalsocietypublishing.org/doi/10.1098/rspb.2020.0025.

Niccol Pescetelli and Nicholas Yeung. The role of decision confidence in advice-taking and trust formation. *Journal of Experimental Psychology: General*, 150(3):507–526, 3 2021. ISSN 1939-2222. doi: 10.1037/xge0000960. URL http://arxiv.org/abs/1809.10453http://doi.apa.org/getdoi.cfm?doi=10.1037/xge0000960.

Lyn M. Swol and Janet A. Sniezek. Factors affecting the acceptance of expert advice. *British journal of social psychology*, 44(3):443–461, 2005.

Leigh Plunkett Tost, Francesca Gino, and Richard P. Larrick. Power, competitiveness, and advice taking: Why the powerful dont listen. *Organizational Behavior and Human Decision Processes*, 117(1):53–65, 1 2012. ISSN 07495978. doi: 10.1016/j.obhdp.2011.10.001. URL http://linkinghub.elsevier.com/retrieve/pii/S0749597811001233.

A. M. Gibbons, J. A. Sniezek, and R. S. Dalal. Antecedents and consequences of unsolicited vs. explicitly solicited advice. In *Symposium in Honor of Janet Sniezek. Society for Judgment and Decision Making Annual Conference.*, volume 9, 2003.

Kobe Desender, Annika Boldt, and Nick Yeung. Subjective Confidence Predicts Information Seeking in Decision Making. *Psychological Science*, 29 (5):761–778, 5 2018. ISSN 0956-7976. doi: 10.1177/0956797617744771. URL http://journals.sagepub.com/doi/10.1177/0956797617744771.

Kobe Desender, Peter Murphy, Annika Boldt, Tom Verguts, and Nick Yeung. A Postdecisional Neural Marker of Confidence Predicts Information-Seeking in Decision-Making. *The Journal of Neuroscience*, 39(17):3309–3319, 4 2019. ISSN 0270-6474. doi: 10.1523/JNEUROSCI.2620-18.2019. URL http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.2620-18.2019.

Seongmin A. Park, Sidney Goïame, David A. O'Connor, and Jean-Claude Dreher. Integration of individual and social information for decision-making in groups of different sizes. *PLOS Biology*, 15(6):e2001958, 6 2017. ISSN 1545-7885. doi: 10.1371/journal.pbio.2001958. URL https://dx.plos.org/10.1371/journal.pbio.2001958.

Stephen M. Fleming, Elisabeth J. van der Putten, and Nathaniel D. Daw. Neural mediators of changes of mind about perceptual decisions. *Nature Neuroscience*, 21(4):617–624, 4 2018. ISSN 1097-6256. doi: 10.1038/s41593-018-0104-6. URL http://www.nature.com/articles/s41593-018-0104-6.

Laurence Aitchison, Dan Bang, Bahador Bahrami, and Peter Latham. Doubly Bayesian Analysis of Confidence in Perceptual Decision-Making. *PLoS Computational Biology*, 11(10):1, 2015.

Florent Meyniel, Mariano Sigman, and ZacharyF. Mainen. Confidence as Bayesian Probability: From Neural Origins to Behavior. *Neuron*, 88(1):78–92, 2015. ISSN 08966273. doi: 10.1016/j.neuron.2015.09.039. URL http://linkinghub.elsevier.com/retrieve/pii/S0896627315008284.

Alexandre Pouget, Jan Drugowitsch, and Adam Kepecs. Confidence and certainty: distinct probabilistic quantities for different goals. *Nature Neuroscience*, 19(3):366–374, 2 2016. ISSN 1097-6256. doi: 10.1038/nn.4240. URL http://www.nature.com/doifinder/10.1038/nn.4240.

Brian Maniscalco, Brian Odegaard, Piercesare Grimaldi, Seong Hah Cho, Michele A. Basso, Hakwan Lau, and Megan A. K. Peters. Tuned inhibition in perceptual decision-making circuits can explain seemingly suboptimal confidence behavior. *PLOS Computational Biology*, 17(3):e1008779, 3 2021. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1008779. URL https://dx.plos.org/10.1371/journal.pcbi.1008779.

Annika Boldt and Nick Yeung. Shared Neural Markers of Decision Confidence and Error Detection. *Journal of Neuroscience*, 35(8):3478–3484, 2 2015. ISSN 0270-6474. doi: 10.1523/JNEUROSCI.0797-14.2015. URL http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.0797-14.2015.

Nim Tottenham, James W Tanaka, Andrew C Leon, Thomas McCarry, Marcella Nurse, Todd A Hare, David J Marcus, Alissa Westerlund, B J Casey, and Charles Nelson. The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry research*, 168(3):242–9, 8 2009. ISSN 0165-1781. doi: 10.1016/j.psychres.2008.05.006. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3474329&tool=pmcentrez&rendertype=abstract.

Neil A. Macmillan and C. Douglas Creelman. *Detection Theory: a User's Guide.* Lawrence Erlbaum Associates, London, second edi edition, 2005. ISBN 0805842306.

Henry L. Roediger III, John H. Wixted, and K. Andrew Desoto. The Curious Complexity between Confidence and Accuracy in Reports from Memory. In *Memory and Law*, pages 84–117. Oxford University Press, 7 2012. doi: 10.1093/acprof:oso/9780199920754.003.0004. URL http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199920754.001.0001/acprof-9780199920754-chapter-4.

V. A. C. Henmon. The relation of the time of a judgment to its accuracy. *Psychological Review*, 18(3):186–201, 1911. ISSN 0033-295X. doi: 10.1037/h0074579. URL http://doi.apa.org/getdoi.cfm?doi=10.1037/h0074579.

Asher Koriat. When are two heads better than one and why? *Science (New York, N.Y.)*, 336(6079):360–2, 4 2012. ISSN 1095-9203. doi: 10.1126/science.1216549. URL http://www.sciencemag.org/cgi/doi/10.1126/science.1216549http://www.ncbi.nlm.nih.gov/pubmed/22517862.

Josep Call and Malinda Carpenter. Do apes and children know what they have seen? *Animal Cognition*, 3(4):207–220, 3 2001. ISSN 1435-9448. doi: 10.1007/s100710100078. URL http://link.springer.com/10.1007/s100710100078.

Louise Goupil, Margaux Romand-Monnier, and Sid Kouider. Infants ask for help when they know they dont know. *Proceedings of the National Academy of Sciences*, 113(13):3492–3496, 3 2016. ISSN 0027-8424. doi: 10.1073/pnas.1515129113. URL http://www.pnas.org/lookup/doi/10.1073/pnas.1515129113.

Roozbeh Kiani and Michael N Shadlen. Representation of confidence associated with a decision by neurons in the parietal cortex. *Science (New York, N.Y.)*, 324(5928):759–64, 5 2009. ISSN 1095-9203. doi: 10.1126/science.1169405. URL http://www.ncbi.nlm.nih.gov/pubmed/19423820http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2738936.

A. Kepecs and Z. F. Mainen. A computational framework for the study of confidence in humans and animals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1594):1322–1337, 4 2012. ISSN 0962-8436. doi: 10.1098/rstb.2012.0037. URL http://rstb.royalsocietypublishing.org/cgi/doi/10.1098/rstb.2012.0037.

Adam Kepecs, Naoshige Uchida, Hatim a Zariwala, and Zachary F Mainen. Neural correlates, computation and behavioural impact of decision confidence. *Nature*, 455(7210):227–31, 9 2008. ISSN 1476-4687. doi: 10.1038/nature07200. URL http://www.ncbi.nlm.nih.gov/pubmed/18690210.

Varda Liberman, Julia A. Minson, Christopher J. Bryan, and Lee Ross. Naïve realism and capturing the wisdom of dyads. *Journal of Experimental Social Psychology*, 48(2):507–512, 3 2012. ISSN 00221031. doi: 10.1016/j.jesp.2011.10.016. URL https://linkinghub.elsevier.com/retrieve/pii/S0022103111002599.

Lee Ross and Andrew Ward. Psychological Barriers to Dispute Resolution. pages 255–304. 1995. doi: 10.1016/S0065-2601(08)60407-4. URL https://linkinghub.elsevier.com/retrieve/pii/S0065260108604074.

Ali Mahmoodi, Bahador Bahrami, and Carsten Mehring. Reciprocity of social influence. *Nature Communications*, 9(1):2474, 12 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-04925-y. URL http://www.nature.com/articles/s41467-018-04925-y.

Marwa El Zein, Bahador Bahrami, and Ralph Hertwig. Shared responsibility in collective decisions. *Nature Human Behaviour*, 3(6):554–559, 6 2019. ISSN 2397-3374. doi: 10.1038/s41562-019-0596-4. URL http://www.nature.com/articles/s41562-019-0596-4.

Silvia Bonaccio and Reeshad S. Dalal. Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences. *Organizational Behavior and Human Decision Processes*, 101(2):127–151, 11 2006. ISSN 07495978. doi: 10.1016/j.obhdp.2006.07.001. URL http://linkinghub.elsevier.com/retrieve/pii/S0749597806000719.

Joaquin Navajas, Chandni Hindocha, Hebah Foda, Mehdi Keramati, Peter E. Latham, and Bahador Bahrami. The idiosyncratic nature of confidence. *Nature Human Behaviour*, 1(11):810–818, 11 2017. ISSN 2397-3374. doi: 10.1038/s41562-017-0215-1. URL http://www.nature.com/articles/s41562-017-0215-1.

Nate Kornell, Lisa K. Son, and Herbert S. Terrace. Transfer of Metacognitive Skills and Hint Seeking in Monkeys. *Psychological Science*, 18(1):64–71, 1 2007. ISSN 0956-7976. doi: 10.1111/j.1467-9280.2007.01850.x. URL http://journals.sagepub.com/doi/10.1111/j.1467-9280.2007.01850.x.

Benedetto De Martino, Sebastian Bobadilla-Suarez, Takao Nouguchi, Tali Sharot, and Bradley C. Love. Social Information Is Integrated into Value and Confidence Judgments According to Its Reliability. *The Journal of Neuroscience*, 37(25):6066–6074, 6 2017. ISSN 0270-6474. doi: 10.1523/JNEUROSCI.3880-16.2017. URL http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.3880-16.2017.

Joaqun Ais, Ariel Zylberberg, Pablo Barttfeld, and Mariano Sigman. Individual consistency in the accuracy and distribution of confidence judgments. *Cognition*, 146:377–386, 1 2016. ISSN 00100277. doi: 10.1016/j.cognition.2015.10.006. URL http://linkinghub.elsevier.com/retrieve/pii/S0010027715300846.

Supplemental Materials:

Confidence, advice seeking and changes of mind in decision making

Niccoló Pescetelli, Anna-Katharina Hauperich, Nick Yeung

1 Supplementary Methods

1.1 Optimal advice use: a Bayesian observer

In probabilistic terms, confidence judgments are conceived as a subjective estimation of the probability of being correct, given a certain decision d [Aitchison et al., 2015, Meyniel et al., 2015, Pouget et al., 2016] (although see [Maniscalco et al., 2021]). We thus modelled a Bayesian observer who uses the confidence reported by participants and the history of previous advisor's outcomes (correct vs. incorrect) to optimally update confidence according to Bayes' theorem. This modelling analysis is concerned with how advice is used–to evaluate whether people approximate Bayesian belief updating–and does not consider the costs and benefits associated with choosing vs. waiving advice. Subjective initial confidence judgments C_{pre} were transformed into 50 quantiles $\hat{C}_{pre}$ to normalise participants' confidence distributions and then transformed into a probability judgment using a linear mapping function: $p(Corr) = 0.5 + .01(\hat{C}_{pre})$. The function linearly transforms confidence judgments into a probability scale, with the minimum confidence judgment corresponding to 50% and the maximum confidence judgment corresponding to 1. A likelihood term was computed as $p(Adv|Corr) = \widehat{Acc}^{A}(1-\widehat{Acc})^{D}$, where Adv is the advice received on a given trial (advisor's agreement vs. disagreement), $\widehat{Acc}$ is the cumulative accuracy of the advisor, A is and indicator variable equals to 1 in agreement trials and 0 in disagreement trials, and $D = 1 - A$. In other words $p(Adv|Corr)$ assumes the value of the advisor's current accuracy rate on agreement trials and the advisor's current error rate on disagreement trials, capturing the idea that, when the participant is correct, agreement can happen only if the advisor too is correct and disagreement can happen only if the advisor makes a mistake. Posterior probability of being correct was calculated using the standard Bayes' formula:

$$p(Corr|Adv) = \frac{p(Corr)p(Adv|Corr)}{p(Corr)p(Adv|Corr) + p(Err)p(Adv|Err)} \tag{S1}$$

The posterior probability so obtained was transformed into a confidence scale using the inverse transformation applied to initial confidence judgments. Figure S6 shows the average confidence change that such Bayesian observer would have reported, given the sequence of trials experienced by our participants. It can be seen that greater confidence updates are experienced when initial confidence is in lower quintiles, compared to larger confidence values.

1.2 Expected value of asking for advice

The Bayesian model above was concerned with how advice is used–to evaluate whether people approximate Bayesian belief updating–and did not consider the costs and benefits associated with choosing vs. waiving advice. Experimentally, we included a cost of advice to encourage participants to ask for advice strategically rather than on all trials. However, this cost could (and indeed in our data, does) influence how participants seek and use advice. We here show how a Bayesian approach can be extended to consider the costs and benefits of advice.

Specifically, we computed the expected value (EV) difference between asking for vs. waiving advice. The expected value was calculated from the outcomes associated with being correct (+5) or incorrect (-5) and the cost of requesting advice (-1), weighted by the subjective probability of each outcome. Crucially, when waiving advice, the model uses its prior probability (i.e., from its initial confidence). In contrast, when requesting advice, the model uses the expected posterior probability.

$$EV_{diff} = EV_{ask} - EV_{waive} \tag{S2}$$

$$EV_{ask} = [(-6) * (1 - post(Corr))] + [(+4) * post(Corr)] \tag{S3}$$

$$EV_{waive} = [(-5) * (1 - prior(Corr))] + [(+5) * prior(Corr)] \tag{S4}$$

The model estimates the posterior probability from the expected posterior confidence in case of future agreement vs. in case of future disagreement, weighted by the observed past agreement rate.

$$post(Corr) = post(Corr|agree) * E[agree] + post(Corr|disagree) * (1 - E[agree]) \tag{S5}$$

where the posterior probability of being correct in case of agreement vs. disagreement with the advisor were computed from Bayes theorem:

$$post(Corr|agree) = \frac{prior(Corr) * E[AdvAcc]}{prior(Corr) * E[AdvAcc] + (1 - prior(Corr)) * (1 - E[AdvAcc])} \tag{S6}$$

$$post(Corr|disagree) = \frac{prior(Corr) * (1 - E[AdvAcc])}{prior(Corr) * (1 - E[AdvAcc]) + (1 - prior(Corr)) * E[AdvAcc]} \tag{S7}$$

We entered the expected value difference in the following binomial regression model as the only predictor on the binary dependent variable ask, representing whether on a given trial, the participant had asked for or waived advice (Table S2). We found a strong positive association of expected value difference, suggesting that the greater the expected value of asking for advice the more likely it was that participants requested advice. Furthermore, we find a positive correlation between individual differences in advice seeking (Az_{ask} in the main text) and the above regression model parameters (random intercepts ($r(22) = .54, p = .006$) and random slopes ($r(22) = .59, p = .002$) fitted to each participant).

Notice that any correlation between this model parameters and individual differences (such as Az_{ask}) cannot be attributed to cost as this was not manipulated within Costly advice blocks. We report this analysis here for completeness. However, because we did not manipulate cost, we cannot test the model fit or evaluate the degree to which participants incorporated advice costs into their information seeking behaviour in a rational manner. Nevertheless, we think this extended model usefully shows how the Bayesian modelling framework can be used to consider this aspect of peoples use of advice which is likely to be important in the real-world.

Going beyond this effect of cost on advice seeking, we could in principle have also extended the confidence update model to include effects of cost on advice use (which we observed empirically, with costly advice having greater influence even after controlling for participants initial confidence–see Figure 5). As indicated in the Discussion in the main text, the addition of such free parameter in the calculation of the Expected Value of seeking vs. waiving advice could capture some of the individual differences in advice use reported in the main text. However, our aim with the belief update part of the model was to compare participants behaviour with a rational model, which would not take cost into account, given that costly and free advice came from the same source.

2 Supplementary Figures

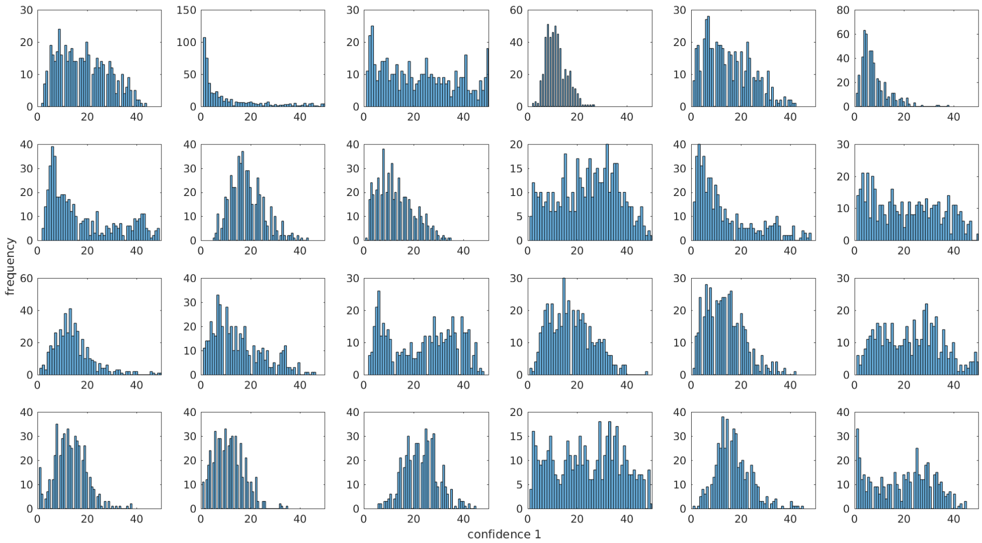

Figure S1: Individual confidence distributions of each participant, recorded in relation to participants' initial decisions.

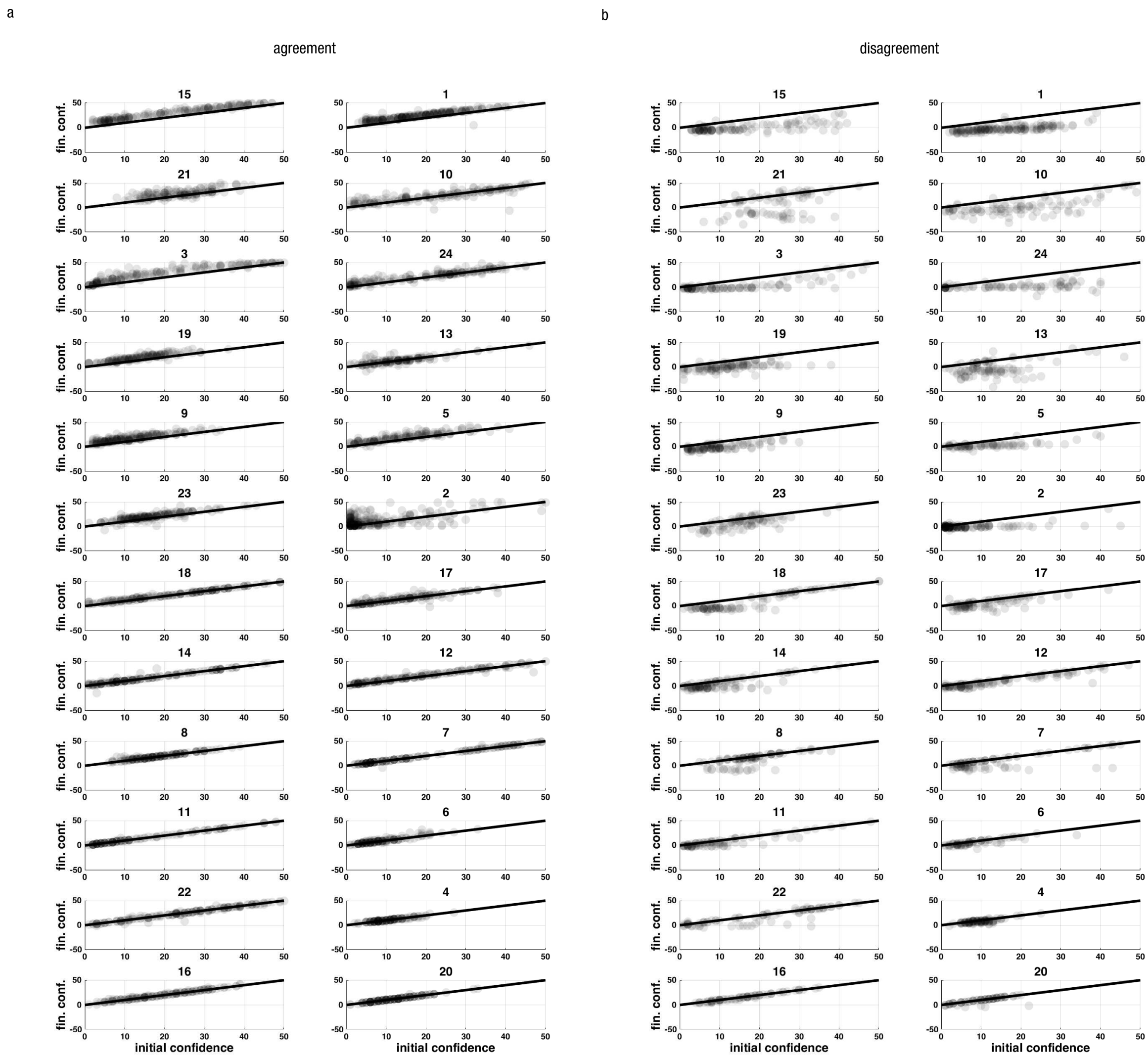

Figure S2: Scatter plot of the relation between post-advice confidence over initial confidence judgments for each participant tested, divided by consensus (agree vs disagree). The plots are sorted according to the average influence of advice on participant's belief, calculated as the difference between average confidence shift in agreement and average confidence shift in disagreement ($I = \delta_{agree} - \delta_{disagree}$), with top rows representing participants with the greatest average confidence change and bottom rows representing participants with the smallest average confidence change. Participants showing large average advice influence are displayed in the upper plots and the participants showing the smallest effects of advice in the lower plots.

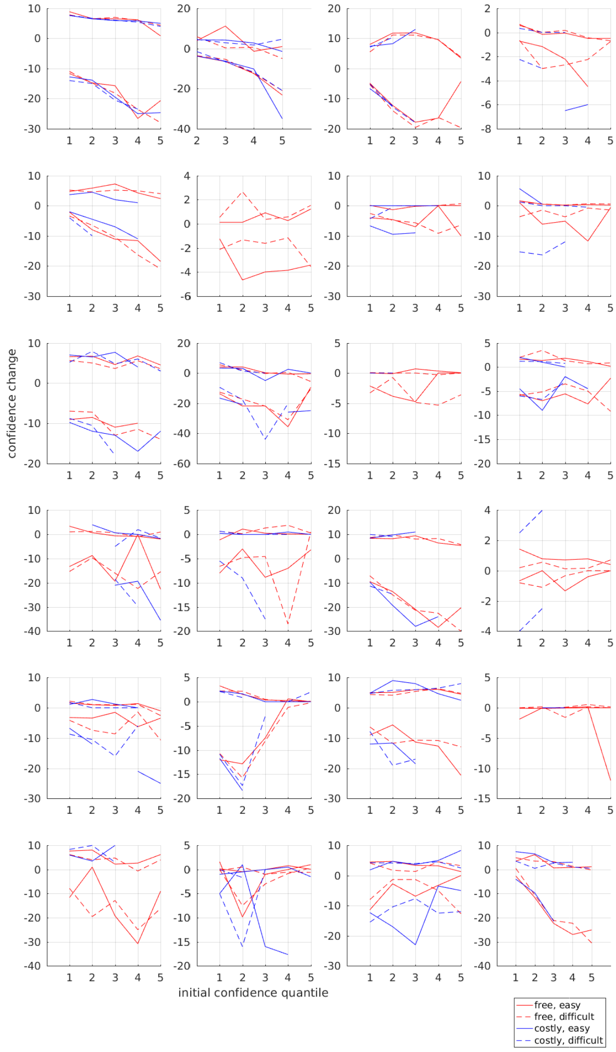

Figure S3: Confidence change observed for each participant as a function of initial confidence quintile, and divided for agreement and disagreement, condition and trial difficulty. Several participants show the unexpected pattern of larger confidence decreases following disagreement with the advisor.

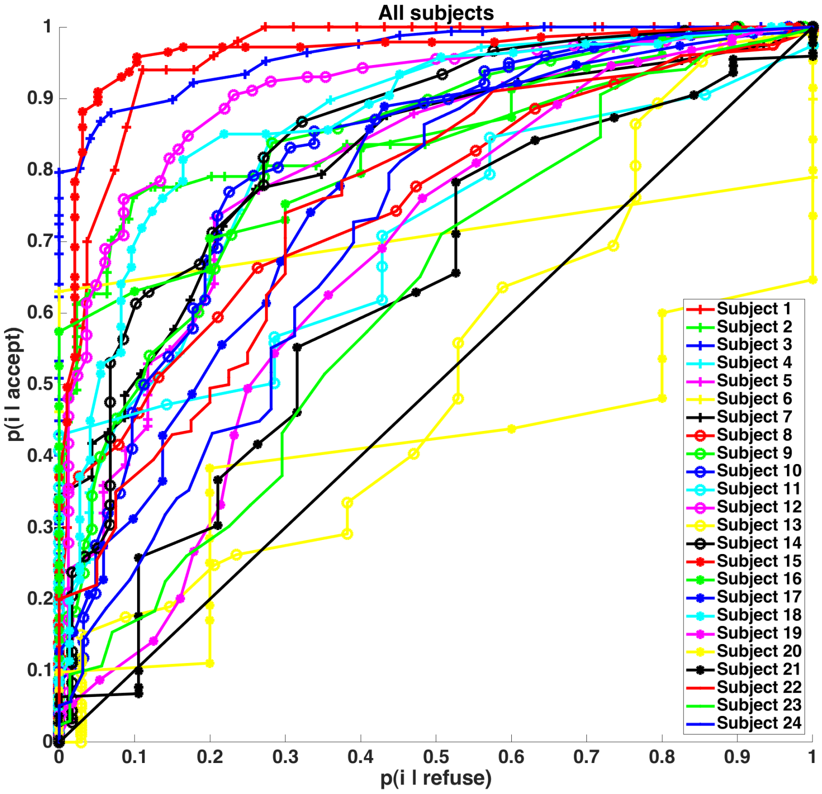

Figure S4: ROC curves used to calculate participants' Az_{ask} measure for advice seeking in costly blocks. The ROC curve is calculated for the probability of asking for *vs.* refusing advice, for each of the 50 levels of initial confidence. It is observed that great inter-individual differences exist in the consistency with which participants request vs. waive advice as a function of their initial confidence (range: [0.41, 0.96]).

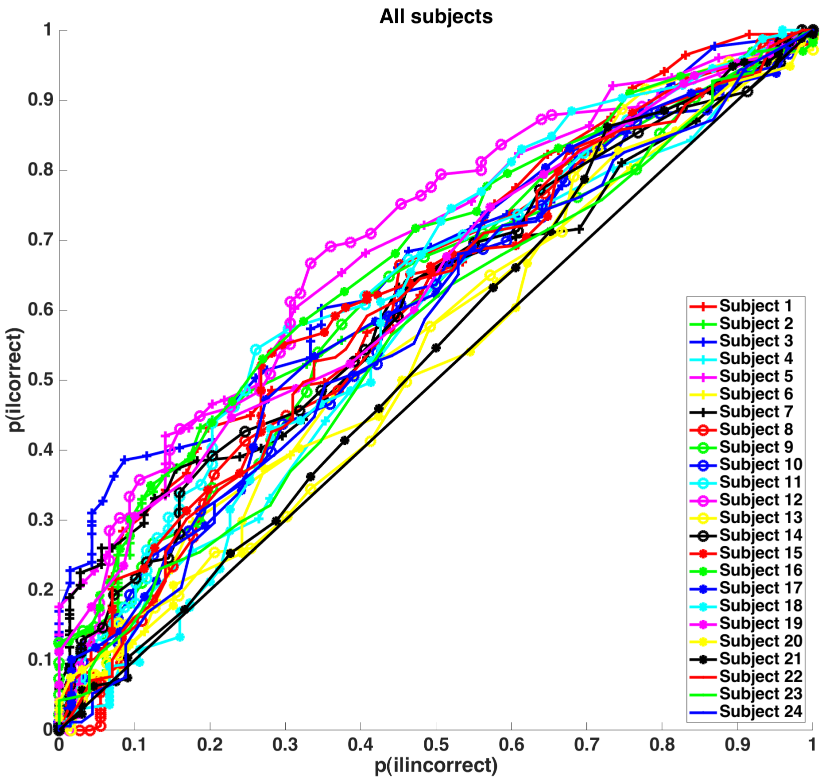

Figure S5: ROC curves used to calculate participants' Az measure for accuracy in all trials. Inter-individual differences are observed in the calibration of participants' confidence (range: [0.54, 0.69]).

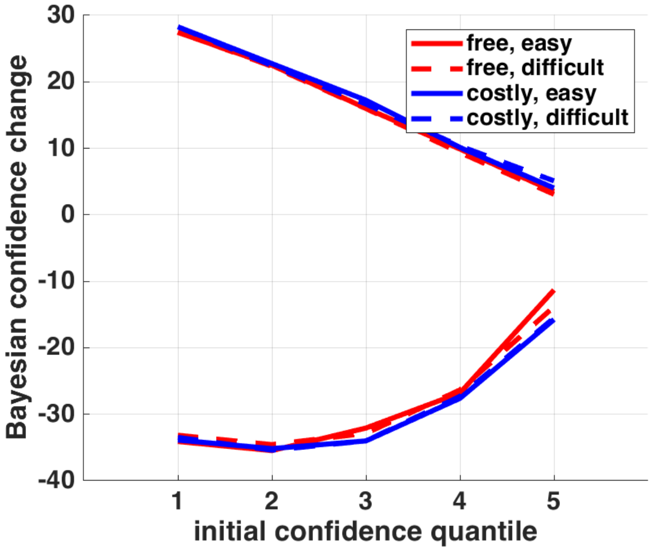

Figure S6: Confidence change pattern over initial confidence quintile, to be expected by a Bayesian observer update confidence based on a linear confidence-probability scaling.

3 Supplementary Tables

Source	Estimate	SE	t(df=5960)	p	95% CI
Intercept	-4.2388	0.5355	-7.9155	**2.9149e-15**	(-5.2885, -3.189)
conf_{init}	0.72983	0.024259	30.085	**3.0231e-185**	(0.68227, 0.77739)
agree	7.7867	0.53933	14.438	**1.8091e-46**	(6.7294, 8.844)
cost	-0.98748	0.91297	-1.0816	0.27947	(-2.7772, 0.80228)
diff	-0.012432	0.58563	-0.021228	0.98306	(-1.1605, 1.1356)
asked	4.0729	6.4601	0.63047	0.52841	(-8.5912, 16.737)
conf_{init}:agree	0.21546	0.027858	7.7342	**1.2135e-14**	(0.16085, 0.27008)
conf_{init}:cost	-0.28119	0.061158	-4.5977	**4.3595e-06**	(-0.40108, -0.1613)
agree:cost	1.231	1.1787	1.0443	0.29637	(-1.0797, 3.5417)
conf_{init}:diff	0.01273	0.031642	0.4023	0.68748	(-0.0493, 0.07476)
agree:diff	-0.21516	0.74123	-0.29027	0.77162	(-1.6682, 1.2379)
cost:diff	-0.83929	1.2699	-0.66089	0.50871	(-3.3288, 1.6503)
conf_{init}:asked	-0.88822	0.51992	-1.7084	*0.087621*	(-1.9075, 0.13102)
agree:asked	-1.8358	10.378	-0.1769	0.85959	(-22.18, 18.509)
cost:asked	-4.0527	6.6151	-0.61264	0.54013	(-17.021, 8.9152)
diff:asked	-8.2665	8.3133	-0.99438	0.32008	(-24.564, 8.0305)
conf_{init}:agree:cost	0.27809	0.079238	3.5096	**0.00045211**	(0.12276, 0.43343)
conf_{init}:agree:diff	-0.011443	0.038171	-0.29979	0.76435	(-0.086272, 0.063386)
conf_{init}:cost:diff	0.027971	0.094687	0.29541	0.76769	(-0.15765, 0.21359)
agree:cost:diff	1.1056	1.6711	0.66163	0.50823	(-2.1703, 4.3816)
conf_{init}:agree:asked	0.71448	0.56898	1.2557	0.20926	(-0.40092, 1.8299)
conf_{init}:cost:asked	0.7798	0.53018	1.4708	0.14139	(-0.25954, 1.8191)
agree_1:cost:asked	2.5616	10.533	0.2432	0.80786	(-18.087, 23.21)
conf_{init}:diff:asked	1.2403	0.57603	2.1531	**0.031349**	(0.11104, 2.3695)
agree_1:diff:asked	5.6149	13.569	0.4138	0.67904	(-20.985, 32.215)
cost_2:diff:asked	11.462	8.5295	1.3438	0.17905	(-5.2586, 28.183)
conf_{init}:agree:cost:diff	0.025014	0.1238	0.20204	0.83989	(-0.21769, 0.26772)
conf_{init}:agree:cost:asked	-0.54478	0.58315	-0.93421	0.35024	(-1.688, 0.5984)
conf_{init}:agree:diff:asked	-1.1504	0.9373	-1.2273	0.21975	(-2.9878, 0.68707)
conf_{init}:cost:diff:asked	-1.5145	0.59578	-2.5421	**0.011045**	(-2.6825, -0.34656)
agree_1:cost:diff:asked	-7.2052	13.796	-0.52227	0.6015	(-34.25, 19.84)
conf_{init}:agree:cost:diff:asked	1.2773	0.95625	1.3357	0.18169	(-0.59729, 3.1519)

Table 1: Full table of the linear mixed-effect model reported in the main text. Final confidence is modeled as a function of agreement (baseline: disagree), difficulty $diff$ (baseline: easy), advice cost (baseline: free), initial confidence $conf_{init}$, and whether the participant requested advice *asked* (baseline: advice waived). Full model: $conf_{final} \sim conf_{init} * agree * cost * diff * asked + (1|subjectID)$

Source	Estimate	SE	t(df=5960)	p	95% CI
(Intercept)	0.101	0.089	1.131(11512)	0.257	(-0.074 - 0.276)
EVdiff'	0.323	0.065	4.912(11512)	$<$**.001**	(0.194 - 0.452)

Table 2: The model above shows a positive effect of expected value difference on the probability of asking for advice. Formula: $ask \sim 1 + EVdiff + (1 + EVdiff|s)$. Expected value takes into account the cost of advice against the expected rewards after requesting the advice (Equations S2-7).

Source	Estimate	SE	t(df=4179)	p	95% CI
(Intercept)	-5.93	0.59	-10.04	**<.001**	(-7.09, -4.77)
agree_{t-1}	1.81	0.62	2.89	**0.003**	(0.58, 3.04)
agree_{t0}	9.44	0.63	14.76	**<.001**	(8.19, 10.70)
conf_{init}	-0.19	0.02	-6.73	**<.001**	(-0.25, -0.14)
agree_{t-1}:agree_{t0}	-1.15	0.80	-1.43	0.15	(-2.72, 0.41)
agree_{t-1}:conf_{init}	-0.11	0.03	-3.22	**0.001**	(-0.18, -0.04)
agree_{t0}:conf_{init}	0.16	0.03	4.54	**<.001**	(0.09, 0.22)
agree_{t-1}:agree_{t0}:conf_{init}	0.07	0.04	1.65	0.09	(-0.01, 0.15)

Table 3: We model social influence ($i.e.$change in confidence after social interaction) as a function of past and current agreement with the advisor (reference: disagree). Formula: $influence \sim agree_{t-1} * agree_{t0} * conf_{init} + (1|subject)$. Including past agreement as a predictor to model influence is a good measure of reciprocity, namely the degree to which participants listened to advisors who agreed with them in previous trials [Pescetelli and Yeung, 2021, Mahmoodi et al., 2015, 2018]. We compared this regression model with a more complex one including a predictor for accuracy. We found no main effect of accuracy and no interaction of accuracy with any other predictor. The model including accuracy was not superior to the model reported in this table ($\chi^2 = 12.39, \delta(d.f) = 8, p = .13$).