



Flexible use of confidence to guide advice requests

Nomi Carlebach^{*}, Nick Yeung

Department of Experimental Psychology, University of Oxford, Oxford, UK

ARTICLE INFO

Keywords:

Confidence
Metacognition
Decision making
Advice
Cognitive control

ABSTRACT

Subjective confidence plays an important role in guiding behavior, for example, people typically commit to decisions immediately if high in confidence and seek additional information if not. The present study examines whether people are flexible in their use of confidence, such that the mapping between confidence and behavior is not fixed but can instead vary depending on the specific context. To investigate this proposal, we tested the hypothesis that the seemingly natural relationship between low confidence and requesting advice varies according to whether people know, or do not know, the quality of the advice. Participants made an initial perceptual judgement and then chose between re-sampling evidence or receiving advice from a virtual advisor, before committing to a final decision. The results indicated that, when objective information about advisor reliability was not available, participants selected advice more often when their confidence was high rather than when it was low. This pattern reflects the use of confidence as a feedback proxy to learn about advisor quality: Participants were able to learn about the reliability of advice even in the absence of feedback and subsequently requested more advice from better advisors. In contrast, when participants had prior knowledge about the reliability of advisors, they requested advice more often when their confidence was low, reflecting the use of confidence as a self-monitoring tool signaling that help should be solicited. These findings indicate that people use confidence in a way that is context-dependent and directed towards achieving their current goals.

1. Introduction

People are faced with choices almost continuously during an ordinary day. These choices range from the brief and inconsequential, such as what socks to wear, to those involving some deliberation, such as what films we want to watch, to those with lifelong impact, such as what career path to pursue. Often, many different considerations and uncertainties are involved in reaching a decision, making the best course of action unclear. This uncertainty is reflected in a varying sense of confidence that accompanies the decisions we reach. Typically, people can evaluate their own performance reliably, reporting levels of confidence that correlate with objective accuracy (Baranski & Petrusic, 1994; Koriat & Goldsmith, 1996), and detecting errors they have recently made (Rabbitt, 1966). Researchers have suggested that this ability to evaluate the quality of our decisions plays a crucial role in adaptive behavior, and is particularly important given that many decisions we reach are not accompanied by reliable and immediate external feedback (Meyniel, Sigman, & Mainen, 2015; Yeung & Summerfield, 2012, 2014).

Recent research has documented the adaptive influence of subjective confidence on subsequent behavior. In perceptual decision making, for

example, confidence has been found to modulate serial dependence, so that past decisions that were made with higher confidence more strongly bias current decisions (Samaha, Switzky, & Postle, 2019), and lead to more liberal response thresholds (Desender, Boldt, Verguts, & Donner, 2019). Furthermore, confidence has been shown to guide high-level strategic choices: research investigating metamemory has demonstrated that learners self-regulate their study behavior to enhance memory, for instance by allocating more time to studying items they are less confident they will remember later (Metcalfe & Finn, 2008; for a review, see Bjork, Dunlosky, & Kornell, 2013); similarly, people employ cognitive offloading strategies, such as setting external reminders, more often when unconfident that they will succeed without external help (Boldt & Gilbert, 2019; Gilbert, 2015; Hu, Luo, & Fleming, 2019); confidence guides decisions regarding whether to choose the option which is currently believed to have the best outcome, or explore new alternatives in the hope of finding an even better option (Boldt, Blundell, & De Martino, 2019); and people choose to engage more with tasks in which they experience higher confidence in their performance (Carlebach & Yeung, 2020; Rouault, Dayan, & Fleming, 2019). Of particular relevance to the present study, two recent studies demonstrated that the

^{*} Corresponding author at: Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford OX3 7LF, UK.

E-mail address: nomi.carlebach@bdi.ox.ac.uk (N. Carlebach).

amount of information people gather before committing to a choice depends on their subjective confidence, with people seeking out more information when they are initially less confident (Desender, Boldt, & Yeung, 2018; Desender, Murphy, Boldt, Verguts, & Yeung, 2019).

Taken together, these studies provide substantial evidence that confidence plays a functional role in a variety of adaptive behaviors. Importantly, these previous studies have all revealed a consistent mapping between confidence in a decision and subsequent behavior when performing a specific task, so that low confidence leads to one course of action and high confidence to another. For example, when performing an information sampling task, low confidence leads to gathering additional information, whereas high confidence leads to committing to a choice. However, it remains unclear whether this mapping is necessarily fixed, or could exhibit some degree of flexibility and context-sensitivity. Here, we explore the hypothesis that people's adaptive use of confidence can vary flexibly according to the immediate context and goal at hand. This framing is consistent with theories that view explicit, conscious representations—such as feelings of confidence—as serving crucially to guide behavior in a flexible manner to support evaluation, planning, and voluntary action (Dehaene & Changeux, 2011).

To investigate this proposal, we studied the relationship between confidence and advice seeking. Previous social decision-making studies have shown that people utilize advice more when a task is difficult, and they are uncertain about their ability to make an accurate decision on their own (Gino & Moore, 2007; Morgan, Rendell, Ehn, Hoppitt, & Laland, 2012; Pescetelli, Hauperich, & Yeung, 2021). In these situations, confidence is used as a self-monitoring tool, signaling to a decision maker that they are unlikely to succeed and should seek the help of others. We propose that this seemingly natural relationship between low confidence and advice requests is not fixed but rather context-dependent. Specifically, we predicted that when people want to learn about advice quality, they will request more advice, not less, as their confidence increases.

The importance of taking advisor quality into account when deciding whether to request or follow advice is intuitive. Accordingly, studies have shown that people are sensitive to the quality of advice based on the reputation of the advisor, relative to their own level of knowledge (Morgan et al., 2012; Yaniv, 2004; Yaniv & Kleinberger, 2000). However, we often do not have any previous knowledge about the reliability of information provided by other people. In such situations, when feedback is available, people can integrate the reliability of advisors' responses over time, updating their trust in advisors through associative learning processes (Behrens, Hunt, Woolrich, & Rushworth, 2008). Importantly, it has recently been demonstrated that people can learn about advisor reliability even when feedback is not available, by comparing advice to their own beliefs when they are certain they know the correct answer (Pescetelli & Yeung, 2021): If we are certain that a given decision is correct, we can equally be certain that anyone who disagrees with us is wrong, and accordingly reduce our trust and reliance on their advice now and in the future. Conversely, if we reach a decision but with low confidence, we should still reduce our trust in advisors who disagree, but to a lesser extent.

Here, we explore a novel implication and prediction of this agreement-in-confidence hypothesis. The implication of this hypothesis is that our best opportunity to learn about the reliability of an advisor comes when we are ourselves sure of what we think about a particular decision. The counterintuitive prediction that follows is that if people want to learn about advisor reliability in the absence of external feedback, they should strategically sample advice when they are highly confident in their own decisions, breaking away from the typical relationship between low confidence and advice requests. In this way, we extend Pescetelli and Yeung's (2021) earlier findings about advisor trust and influence, to test a novel prediction about how people's subjective sense of confidence will guide their choices about whether and when to seek advice.

To illustrate this prediction, consider the situation of being unsure

which of two movies to watch, and therefore seeking advice from online reviews—a straightforward example of subjective confidence guiding information seeking. Suppose you find reviews for both movies written by the same film critic, who had a strong preference for one of the two movies. However, you have never heard of this critic before, and cannot be sure you should trust their opinion. Suppose, though, that on the website you notice that they have also reviewed a movie you saw last week and loved. Now you might be motivated to read the review, even though you are confident in your opinion about this movie, so as to learn about the reviewer: If you find that the critic loved the movie too, you might infer that the critic can be trusted, and therefore follow their recommendation for tonight's film choice.

The example above demonstrates that confidence can be used flexibly as part of a coherent overall strategy, both as a self-monitoring tool (indicating when you need advice), and as a proxy for feedback when objective feedback is lacking (as a yardstick against which to evaluate the reliability of advice). In the present study, we tested the prediction that people can utilize confidence in such a flexible way. We expected the relationship between confidence and advice requests to vary depending on the way in which confidence is currently being utilized. Specifically, we hypothesized that increasing advice requests as confidence decreases reflects the use of confidence as a self-monitoring tool, signaling that help should be requested. In contrast, more frequent advice requests when confidence is highest reflects an attempt to learn about advisor quality, by using confidence as an internal feedback proxy.

We tested our predictions in three experiments using a Judge-Advisor System (JAS) approach. In its most basic form, the JAS consists of a decision maker (the judge), and one or more advisors who communicate their advice to the judge (Bonaccio & Dalal, 2006; Sniezek & Buckley, 1989, 1995). In the present study, participants (i.e., acting as 'judge') performed a perceptual task in which they reported which of two boxes contained more dots, and how confident they were in their response. Following each initial judgement, they chose between receiving advice from a virtual advisor or seeing the stimulus again. After receipt of the additional information chosen, they reported their final decision and confidence on which box contained more dots. In Experiment 1, we tested the novel prediction that, when advisor reliability is initially unknown and no objective feedback is provided, participants will ask for advice when they themselves are most confident in a decision (and can therefore learn if advisors are providing accurate advice). Experiments 2 and 3 manipulated the availability of explicit information about advisor reliability (rendering learning about advisors unnecessary) and of external feedback on task performance (so that the use of confidence as an internal feedback proxy is redundant). Our data reveal that the relationship between confidence and advice requests is not fixed, but instead varies depending on the availability of such information. We conclude that confidence is used strategically, both as a self-monitoring tool, and as a proxy for feedback, depending on the task context and current goal.

2. Experiment 1

Experiment 1 tested our main hypothesis, that people use confidence flexibly depending on their immediate goal. To test this hypothesis, we used a perceptual decision-making task in which, after an initial decision, participants chose between seeing the same stimulus again for a longer period or receiving advice from a virtual advisor. They were paired with a different advisor in each block but were not told how reliable the advice would be, and were not given trial-by-trial feedback with which to evaluate the quality of the advice (or their own decisions). Given these conditions, we predicted a U-shaped relationship between confidence and advice requests, reflecting the use of confidence for two different purposes. Specifically, we predicted that people would seek advice frequently when confidence was very low, to improve their accuracy by using the help of others, but critically also when confidence

was very high, enabling them to learn about advisor quality. Key predictions and analyses were preregistered at <https://aspredicted.org/26py4.pdf>.

2.1. Methods

2.1.1. Participants

Twenty-six participants took part in the experiment (17 females, ages 18–35 years, $M = 23.5$, $SD = 5.5$). All participants had normal or corrected-to-normal vision. They provided written informed consent and were compensated with course credit or payment for their participation. One participant who did not use enough of a range of confidence values for us to analyse their advice seeking behavior as a function of confidence (i.e., whose data could not be divided into quantiles as defined below) was removed from the data analyses. All procedures were approved by the University of Oxford's Medical Sciences Interdisciplinary Research Ethics Committee.

2.1.2. Task and procedure

Participants performed a series of trials in which they first reported which of two boxes contained more dots and how confident they were in their response, and then chose to sample one of two additional sources of information before giving a final response and confidence judgement. Fig. 1 depicts an example trial in the main experimental blocks. Trials began with a fixation cross displayed at the center of the screen for 800 ms in total. The fixation cross flashed briefly 400 ms after it appeared, indicating the perceptual stimulus will be presented shortly after. The stimulus, two boxes containing different numbers of dots, appeared to the left and right of the fixation cross for 150 ms. Participants reported which of the two boxes contained more dots, and how confident they were in their response, on a scale ranging from 50% (guess), to 100% (sure). After the initial response was registered, images representing the advisor and the dot stimulus appeared above and below the fixation cross. Participants chose between viewing the stimulus again for a longer period (making it easier to judge which box contained more dots) versus receiving advice from a virtual advisor, by selecting the appropriate image with the mouse cursor. If participants chose to view the stimulus again, they were presented with a fixation cross for 800 ms, followed by the same dot stimulus for an extended period of 300 ms. If they chose advice, a fixation cross appeared for 300 ms, after which two empty boxes were presented to the right and left with an arrow pointing at one of the boxes to indicate the advisor's recommended response, displayed for 1500 ms. Advice trials were longer than review-stimulus trials to counteract a bias displayed towards advice selections in a pilot study. Following the additional information (stimulus re-viewing or advice), participants reported their final decision and associated confidence, with their initial response marked on the confidence scale for reference. The next trial then began (with no feedback provided to

indicate the box that objectively contained more dots).

The experiment consisted of six blocks of 60 trials each. Participants encountered a new advisor during each block. Advisor accuracy varied across the 6 blocks, at 50%, 60%, 70%, 80%, 90%, and 100% correct, with the order of advisor accuracy randomised across blocks for each participant (and, across participants, randomly assigned to the 6 different advisor faces). Participants were instructed that advisors would be of different quality but received no information about their accuracies. To encourage participants to choose according to their preference, the instructions indicated that 'It is up to you to decide on each trial if receiving advice or viewing the stimulus again for a longer duration is more useful to you'. Notably, since the relative value of the two options depended not only on objective factors such as advisor accuracy and task difficulty, but also on internal factors such as participants' level of attention during the initial presentation and their willingness to change their minds given contradictory evidence, it is very difficult to define an optimal level of advice seeking for this task. The locations in which the advisor and review-stimulus options were presented during information choice (i.e., above vs. below the fixation) were assigned randomly for each block, with each information type appearing in each location an equal number of times throughout the experiment.

To induce a wide range of confidence judgements (that would allow us to study the relationship between different confidence levels and information choice), each block contained an equal number of trials in three difficulty levels (easy, medium, hard), with trial order determined randomly. Trial difficulty was controlled by manipulating the difference in dot numbers between the boxes, with the numbers in each box determined by the equation $nDots = 200 \pm d * l$, where d represents the number of dots added and subtracted to the correct and incorrect boxes correspondingly, and $l = [3, 1, 1/3]$, for the easy, medium, and hard difficulties correspondingly. To achieve similar accuracy levels across participants, the medium difficulty condition was staircased to ~70% accuracy using a 2-down 1-up procedure titrating the d parameter value throughout the experiment. The easy and difficult trials were defined relative to the staircased medium difficulty by multiplying or dividing d by 3, as captured by the l parameter. The side of the box with more dots varied randomly across trials, and the locations of the dots within each box were sampled randomly on every trial. To encourage task engagement, participants received a bonus payment based on final-response accuracy during the main experimental blocks.

Before beginning the main experiment, participants viewed instructions explaining the task and performed three practice blocks. The first two blocks consisted only of the dot judgement task and contained 42 trials each. During these two training blocks, participants received feedback on their performance in the form of a tone after error trials, as well as feedback on their overall accuracy at the end of each block (based on participants' final responses). The third practice block

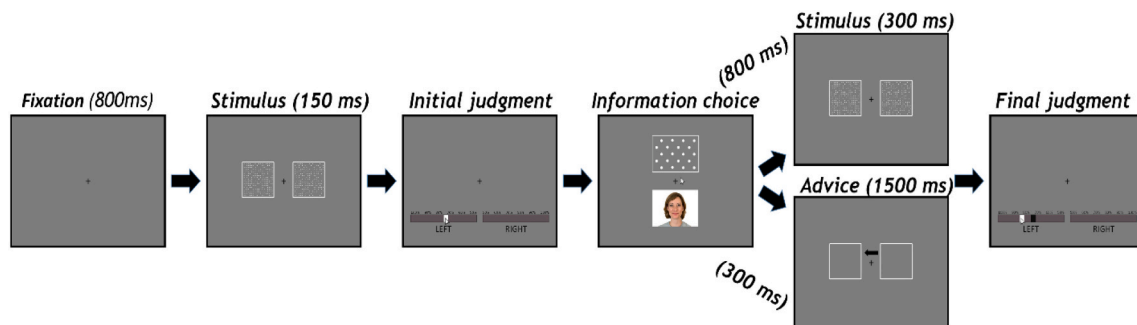


Fig. 1. Experimental Design. Participants made a speeded judgement about which of two squares contained more dots, and reported their confidence in this decision. Following the initial judgement, participants chose between viewing the stimulus again for a longer period, or receiving advice from a virtual advisor by selecting the corresponding image. After receiving the additional information they selected, participants reported their final decision and confidence on which box contained more dots.

introduced the information choice component, with participants performing 36 trials of the full task. The practice advisor had an accuracy of 80%. As in the main experiment, participants did not receive information about advisor accuracy or feedback following each trial of this block, but received feedback on their overall final response accuracy at the end of the block.

After completing the main experiment, participants completed the Concern over Mistakes, Personal Standards and Doubts about Actions subscales from the Multi-dimensional perfectionism scale (Frost, Marten, Lahart, & Rosenblate, 1990). These surveys were included in this experiment, and in Experiment 2, to allow exploratory analysis of how perfectionism might relate to confidence and information choices. However, since no clear relationships were found, we do not report any analyses on the questionnaire data.

2.1.3. Stimuli and apparatus

All stimuli were created and presented in Matlab using the Psychtoolbox3 extension (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). Stimuli were presented on a 20-in. CRT screen with grey background (127, 127, 127), a refresh rate of 60 Hz and 1600 × 1200 resolution. Dot stimuli comprised two white rectangular frames in which white dots were presented in random locations in a 20 × 20 grid. Advisors were represented by a picture of a white female face with a neutral expression selected from the NimStim database (Tottenham et al., 2009). The experiment took place in a dimly lit room, with participants sitting at a comfortable viewing distance from the screen.

2.1.4. Statistical analyses

Confidence Quantiles. To study the relationship between subjective confidence and behavior, trials were grouped according to confidence ratings. For each participant, N quantiles were computed on initial confidence ratings (with N depending on the specific analysis), and trials grouped into bins according to trial confidence so that $Q(n-1) < \text{conf}(t) \leq Q(n)$. If certain confidence ratings were used frequently, resulting in bins with shared cut off values or containing fewer than 20 trials, new quantiles were computed above and below these values. For example, if $Q(n) = Q(n-1)$, $n-1$ quantiles were calculated on $\text{conf} \leq Q(n)$, and the remaining quantiles were calculated on $\text{conf} > Q(n)$. If the last quantiles had shared cut off values, or the cutoff value was 100 (the highest confidence rating), $N-1$ quantiles were calculated on $\text{conf} < Q(n)$, so that the highest confidence rating was binned in the last quantile.

2.2. Results

2.2.1. Accuracy and confidence judgements

A first set of analyses confirmed that participants' basic task performance was as expected. Their accuracy varied according to trial difficulty for both their initial responses and final decisions (for initial responses: $M = 93\%, 71\%, 59\%$, for easy, medium, and hard trials, respectively; $F(2, 48) = 625.25, p < .001, \eta^2 = 0.96$; for final decisions: $M = 95\%, 76\%, 62\%$, $F(2, 48) = 443.73, p < .001, \eta^2 = 0.95$). Participants' subjective confidence ratings reflected these differences in objective accuracy across conditions, albeit indicating relative underconfidence in easy trials and overconfidence in difficult trials (for initial responses: $M = 74, 68, 67$, for easy, medium, and hard trials, respectively, $F(2, 48) = 133.77, p < .001, \eta^2 = 0.85$; for final decisions: $M = 79, 71, 69$, $F(2, 48) = 152.02, p < .001, \eta^2 = 0.86$). Participants made effective use of additional available information (i.e., receive advice or view stimulus again), so that their mean accuracy was higher for their final decisions than their initial responses ($M = 77\%$ vs. 74% , respectively; $t(24) = 6.14, p < .001$). Again, this difference in accuracy was mirrored in participants' confidence ratings, which were higher for final decisions than initial responses ($M = 73$ vs. 70 , respectively; $t(24) = 7.06, p < .001$).

2.2.2. Learning about advisor quality

Across blocks, participants encountered advisors whose accuracy varied from 50% to 100%, but no explicit trial-by-trial feedback was provided against which to evaluate the quality of this advice. Participants nevertheless learned about advisor quality when deciding whether to ask for advice: As demonstrated in Fig. 2, the proportion of advice requested throughout the experiment depended on advisor accuracy. A repeated-measures ANOVA on the proportion of advice requests as a function of advisor accuracy showed a significant effect ($F(5, 120) = 2.80, p = .02, \eta^2 = 0.82$), with a significant within-subject linear trend ($F(1, 24) = 6.87, p = .015, \eta^2 = 0.71$), indicating that advice requests increased monotonically as a function of advisor quality. These results provide further evidence that people can learn about the quality of advice even when objective feedback is lacking (Pescetelli & Yeung, 2021). As described above, our hypothesis was that this learning would be driven by participants seeking advice when they were already sure about the correct answer, so that they could evaluate advisors against these firmly held beliefs, rather than only seeking advice when low in confidence as seen in previous research.

2.2.3. Confidence effects on information choices

Overall, participants chose to receive advice on 0.39 of the trials on average, with large individual differences in preferences for advice or viewing the stimulus again (proportion advice choice: range = 0–0.91, $SD = 0.26$). Both advice and viewing the stimulus again increased participants' accuracy on their final response compared to their initial response (from 74% to 77% and 72% to 74%, respectively), indicating that both types of information were of value ($t(23) = 3.97, p < .001$ and $t(24) = 4.23, p < .001$, for advice and view-stimulus trials, respectively; note that one participant never requested advice).

To answer our main question—whether people use confidence strategically depending on their immediate goal—we studied the relationship between confidence and information choices. For each participant, we divided trials into six bins according to initial-confidence sextiles, and then computed the proportion of trials in which advice was selected for each bin. We predicted a U-shaped relationship between confidence and advice selections, whereby people would choose to receive advice, rather than view the stimulus again, more often both when they were initially low in confidence (reflecting an attempt to use advice in order to improve accuracy), and when they were initially high in confidence (reflecting an attempt to learn about advisor quality). A preregistered repeated-measures ANOVA revealed a significant effect of confidence on information choices ($F(5, 120) = 5.34, p = .012, \eta^2 = 0.18$), and a

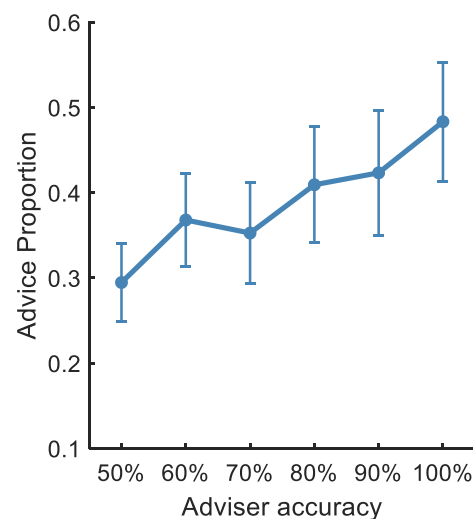


Fig. 2. Results of Experiment 1. Mean proportion of advice selections according to advisor accuracy. Error bars represent standard errors of the mean (SEM).

significant quadratic trend ($F(1, 24) = 5.82, p = .024, \eta^2 = 0.195$), both as predicted. Surprisingly, however, we found that people requested advice more often when confidence was high, but not low—the significant quadratic trend reflected a levelling off, rather than increase, in advice choices at the lowest levels of confidence (see Fig. 3A). Pairwise comparisons of confidence sextile bins with Bonferroni correction revealed significant differences in the frequency of advice selections between trials with the highest confidence ratings (Q6) and medium confidence trials (Q5 and Q4, $ps < 0.05$), but no significant differences for the lowest confidence trials (Q1). Thus, while these results support the hypothesis that people will sometimes seek advice when highly confident—specifically, when they can use confidence as a feedback proxy to learn about advisor reliability—they do not support our initial prediction of a U-shaped relationship between confidence and advice requests.

Since confidence varied reliably as a function of trial difficulty, an alternative interpretation of the above results is that information choices were driven by objective trial difficulty, rather than subjective confidence. This seems unlikely given the brief stimulus presentation time and the plausible influence of fluctuations in internal states, such as attention, on participants' experience of difficulty (as represented here by confidence). Nonetheless, we performed further analysis that supported the conclusion that this advice-seeking behavior is driven by confidence: Confidence ratings overlapped substantially across difficulty conditions (Fig. 3B) and, when we applied logistic regression to predict each participant's trial-wise advice seeking choices (vs. review the stimulus) with predictors for their initial confidence and objective task difficulty, we observed a consistently positive beta for confidence

across participants ($t(24) = 2.16, p = .04$), while no consistent effects emerged for trial difficulty ($t(24) = -0.6, p = .57$, Fig. 3C).

2.2.4. Evidence for use of confidence for learning

We predicted the observed increase in advice requests at the highest levels of confidence based on the hypothesis that participants were using their own judgements as a yardstick against which to judge (and learn about) advisor quality. If so, we might expect this relationship between confidence and advice seeking to be most pronounced at the beginning of each block, when participants did not have any information about the current advisor's accuracy. As the block progressed, we would expect fewer advice choices when confidence was high, since participants should have gained knowledge about advisor quality and would no longer needed to use high confidence trials to learn about advice. Consistent with these predictions, an exploratory two-way repeated-measures ANOVA on the proportion of advice choices, with time-in-block (first half, second half) and confidence (Q1-Q6) as factors, revealed significant main effects for time-in-block ($F(1, 24) = 19.95, p < .001, \eta_p^2 = 0.45$) and confidence ($F(5, 120) = 5.02, p = .016, \eta_p^2 = 0.17$), and, importantly, a reliable interaction effect between these factors ($F(5, 120) = 3.95, p = .01, \eta_p^2 = 0.14$): As shown in Fig. 3D, advice choices varied more strongly as a function of confidence in the first half of blocks than the second. Follow-up one-way ANOVAs on advice choices for each block half separately revealed that confidence reliably predicted advice requests during the first time period ($F(5, 120) = 7.04, p = .003, \eta^2 = 0.23$) but not during the second time period ($F(5, 24) = 1.93, p = .15, \eta^2 = 0.07$).

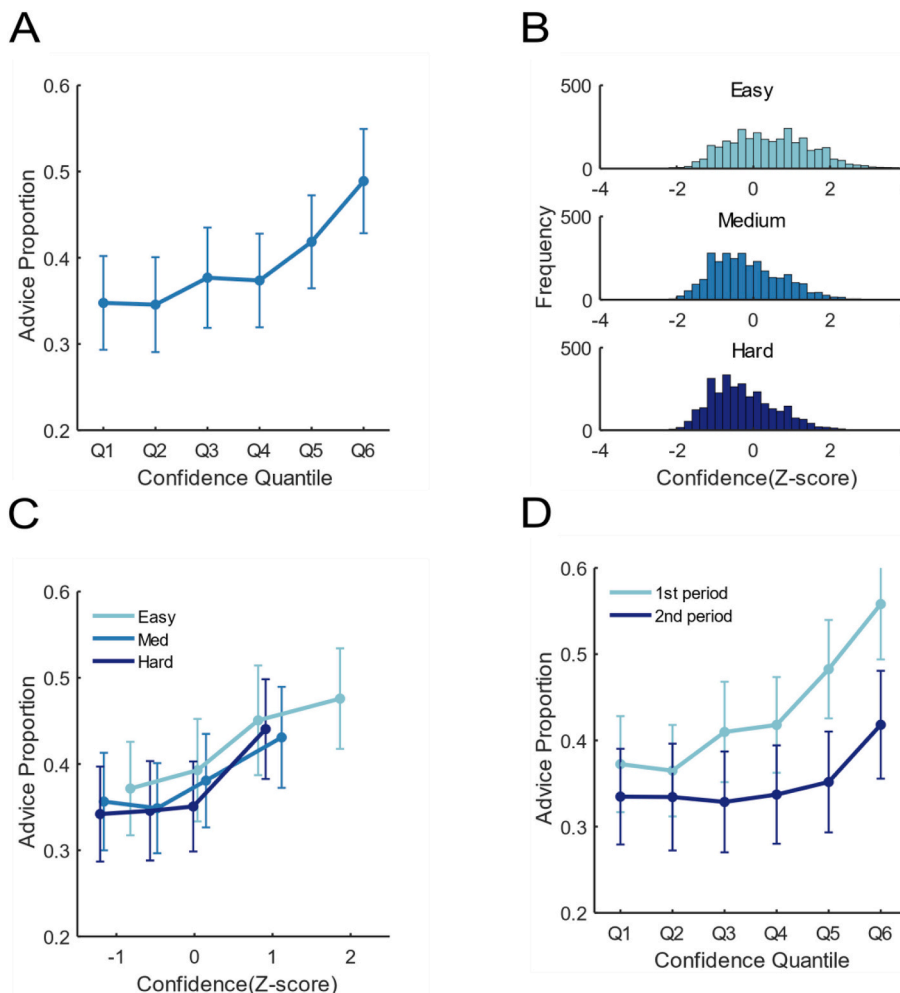


Fig. 3. Results of Experiment 1: (A) Mean proportion of advice selections according to confidence bins across the entire experiment. (B) Normalized initial confidence ratings, divided according to trial difficulty. Confidence ratings were first normalized within each participant, across all trials regardless of difficulty. Then, normalized confidence ratings were aggregated across all participants and grouped according to trial difficulty. (C) Mean proportion of advice selections according to normalized initial confidence and trial difficulty. For each participant, trials within each difficulty were divided into four bins according to confidence ratings. (D) Mean proportion of advice selections according to confidence and time in block. For each participant, trials were divided into two time periods (first half vs. second half of block), and then into six bins according to confidence ratings, within each time period. Error bars represent SEMs.

2.3. Discussion

The results of Experiment 1 provided some support for our main hypothesis, but also some surprising results. As predicted, we found that participants were able to learn about the accuracy of their advisors even without objective feedback, and that they sought advice frequently when they were most confident in their own initial decisions, particularly early in each block as they were learning about the quality of advice on offer. These findings are consistent with our hypothesis that confidence is used strategically depending on peoples' immediate goal: Advice-seeking is not solely associated with situations of high uncertainty, as explored in previous research, but can also be seen in situations of high confidence when an important goal is to learn about the current task context. Surprisingly, however, we did not find that participants were more likely to request advice when they were unconfident in their initial response, such that we did not observe the U-shaped relationship we had originally predicted between confidence and advice seeking: Instead, participants' preferences for receiving advice vs. reviewing the stimulus were relatively stable from medium to low levels of confidence. A possible interpretation of these findings is that, although people can strategically select a mapping between confidence and behavior that matches their current goals (here, requesting more advice when highly confident in order to learn about advisor reliability), they might find it difficult to maintain multiple mappings active at the same time or to switch between mappings at a rapid rate. However, participants showed individual differences in strategies and degree of flexibility (see Fig. S3 in the Supplementary Material) and, as we will see below, the effect varied across experiments, with some evidence that participants are more likely to request advice when confidence is low in Experiments 2 and 3, so we do not discuss this further for now.

Although we interpret high-confidence advice requests to reflect an attempt to learn about advisor quality, several alternative explanations exist. For example, one possibility is that viewing the stimulus again when confidence is high holds very low informational value, thereby leading people to request advice instead. Alternatively, high confidence advice requests might reflect an attempt to receive social approval when participants know they are correct. To address these explanations and others, Experiment 2 built on the main observation here of advice seeking at high levels of confidence, to compare conditions in which participants received prior information about advice quality, rendering learning unnecessary, to conditions in which they did not receive prior information, and thus had to learn about the quality of advice. We predicted that participants would show evidence of using confidence in a strategic fashion, such that the relationship between confidence and advice-seeking would vary across these conditions. Since conditions were identical in all other aspects, such results would further favor our interpretation—that advice requests when confidence is high reflect an attempt to learn about advice quality—above alternative interpretations.

3. Experiment 2

Experiment 2 tested whether confidence is used flexibly depending on the availability of external information, which we varied in a between-participants design. The "Explicit information absent" condition was a replication of Experiment 1 in which participants were not told about advisor accuracy. Thus, for this condition we again predicted that participants would seek advice when highly confident, using these trials to learn about advisor quality via the agreement-in-confidence heuristic. In contrast, in the "Explicit information available" condition, participants were notified of the current advisor's average accuracy at the start of each block. Given that learning about advisors was unnecessary, we did not expect to see a higher frequency of advice selections when confidence was high. We did not have a specific prediction regarding the relationship between confidence and advice seeking in this case: It could be that people might request more advice when

confidence is low in order to improve their immediate accuracy, but it seemed equally plausible that they might not use confidence to guide advice requests in this condition, and that advice requests would depend only on advisor quality, which was known. Key predictions and analyses were preregistered at <https://aspredicted.org/xr9tt.pdf>.

3.1. Methods

3.1.1. Participants

Fifty-one participants took part in Experiment 2. One participant whose accuracy on the initial response was more than two standard deviations below the mean accuracy was removed, resulting in a final sample of 50, with 25 participants in each condition (35 females, ages 18–33 years, $M = 21.7$, $SD = 4.2$). All participants had normal or corrected-to-normal vision. They provided written informed consent and were compensated with course credit or payment for their participation. All procedures were approved by the University of Oxford's Medical Sciences Interdivisional Research Ethics Committee.

3.1.2. Task and procedure

Experiment 2 used a mixed design, with explicit information about advisor accuracy (absent, available) varying between participants. For each participant, advisor accuracy varied randomly between blocks across 5 levels, from 50%–90%, and within each block participants performed the perceptual task across three levels of task difficulty. The perceptual task and information choice were identical to Experiment 1. Thus, the 25 participants assigned to the information absent condition provided a replication of Experiment 1. The remaining 25 participants were assigned to the information available condition in which, at the beginning of every block, a message on the screen informed them of the advisor's accuracy in the current block (e.g., "Advisor accuracy on this block: 70%"). This experiment did not include a block with 100% accurate advice, to avoid participants in the information available condition always requesting advice from the 100% accurate advisor. The end of block feedback differed slightly from Experiment 1, with participants receiving feedback on their average initial-response accuracy rate, in addition to average final-response accuracy.

3.1.3. Stimuli and apparatus

The stimuli and apparatus were identical to Experiment 1, except for advisor stimuli: In Experiment 2, advisors were represented by a picture of a smiling white female face selected from the Chicago Face database (Ma, Correll, & Wittenbrink, 2015).

3.2. Results

3.2.1. Accuracy and confidence judgements

The two groups were well-matched in the accuracy of their initial responses, but participants who received information about advisor accuracy benefited more from the additional information on each trial (i.e., from advice or reviewing the stimulus), as shown in a two-way mixed ANOVA on mean accuracy that revealed a reliable effect for response (initial vs. final; $F(1, 48) = 108.27$, $p < .001$, $\eta_p^2 = 0.69$) and a reliable interaction between response and group ($F(1, 48) = 6.38$, $p < .02$, $\eta_p^2 = 0.12$). For participants provided with explicit information about advisor reliability, accuracy improved from 74% to 79% from initial response to final decision ($t(24) = 8.68$, $p < .001$). For participants without this information, the corresponding values were 75% and 78% ($t(24) = 5.90$, $p < .001$). A corresponding ANOVA on confidence ratings revealed a significant effect for response ($F(1, 48) = 51.08$, $p > .001$, $\eta_p^2 = 0.52$), but not for group or the interaction between response and group ($F_s < 1$): Participants in both experimental groups reported higher confidence in their final responses compared to the initial responses ($M = 74$ vs. 71 , and $M = 73$ vs. 71 , for participants in the Explicit information available vs. absent groups, respectively).

3.2.2. Effects of advisor quality

Participants in both groups showed sensitivity to advice quality in their choices to receive advice vs. review the stimulus, but this sensitivity was unsurprisingly much stronger in the group with Explicit information available (Fig. 4A). Thus, a two-way mixed measures ANOVA on the proportion of advice requests, with factors for advisor accuracy and group showed reliable effects for advisor accuracy ($F(4, 192) = 42.46, p < .001, \eta_p^2 = 0.47$) and for the interaction effect between advisor accuracy and group ($F(4, 192) = 12.46, p < .001, \eta_p^2 = 0.21$). Separate ANOVAs for each group revealed that the proportion of advice requests was reliably predicted by advisor accuracy even when external information was unavailable ($F(4, 96) = 8.38, p < .001, \eta^2 = 0.26$), with a reliable linear contrast ($F(1, 24) = 21.73, p < .001, \eta^2 = 0.48$), replicating the results of Experiment 1. A very strong effect of advisor accuracy was apparent for participants who were told about advisor accuracy before each block ($F(4, 96) = 36.52, p < .001, \eta^2 = 0.60$), again with a significant linear trend ($F(1, 24) = 69.73, p < .001, \eta^2 = 0.74$).

Fig. 4B and C plot ten-trial running averaged proportions of advice requests according to advisor accuracy, separately for each group, and demonstrate the differences in learning according to information availability. For participants without explicit information about advisor accuracy (Fig. 4B), differences in advice requests across advisors developed throughout the block, primarily in terms of asking for less advice from low quality advisors as the block progressed. In contrast, when participants were given explicit information about advisor accuracy (Fig. 4C), the proportion of advice requests varied as a function of advisor quality stably throughout the block. Thus, patterns of advisor choice reflected advisor quality for both groups of participants, but visual inspection suggests that only the group without external information about advisor quality showed evidence of learning. Of interest, then, was whether confidence influenced advisor choice differently across the two groups, reflecting their different needs and goals.

3.2.3. Effects of confidence on information choices depend on explicit information availability

Experiment 1 showed that participants selected more advice when confidence was high, presumably in order to learn about advice quality. In the current experiment, we predicted that participants would use confidence differently depending on the availability of information about advisor accuracy. Specifically, we predicted that when no explicit information about advice quality was available (i.e., replicating the conditions of Experiment 1), participants would again request more

advice when confidence was high, allowing them to learn about the advisors. In contrast, we did not expect to see this pattern for participants in the explicit information condition, since learning about advisor accuracy was unnecessary. Consistent with this prediction, a preregistered two-way mixed ANOVA on the proportion of advice requests with factors for group (Explicit information available vs. absent) and confidence (Q1-Q6, see Fig. 5A) revealed a significant main effect for confidence ($F(5, 240) = 3.54, p = .02, \eta_p^2 = 0.07$), but not for information availability ($F < 1$) and, crucially, a significant interaction between confidence and group ($F(5, 240) = 6.73, p < .001, \eta_p^2 = 0.12$). Thus, confidence influenced information sampling differently depending on the availability of external information about advice.

Separate ANOVAs for the two groups revealed that for participants for whom explicit information was not available, their confidence predicted their information choices reliably ($F(5, 120) = 7.51, p < .001, \eta^2 = 0.24$), with a reliable quadratic trend found in the data ($F(1, 24) = 22.57, p < .001, \eta^2 = 0.49$). Replicating the key results of Experiment 1, the proportion of advice requests was highest for the highest confidence bin (Q6), with Bonferroni-corrected pairwise comparisons revealing reliable differences in the frequency of advice requests between high confidence trials (Q6) and medium confidence trials (Q2, Q3, Q4 and Q5, $ps < 0.05$). Interestingly, the data revealed a numerical increase in advice requests for lowest confidence (Q1) compared to Q2 ($t(24) = 2.75$), but this difference was not reliable when correcting for multiple comparisons ($p = .17$). In contrast, for participants who were provided with information about advisor quality before each block, confidence reliably predicted advice requests ($F(5, 120) = 3.05, p = .046, \eta^2 = 0.11$), with a tendency for a decreasing proportion of advice requests as initial confidence increased, although the linear trend was not significant ($F(1, 24) = 4.17, p = .052$). Pairwise comparisons did not reveal any reliable differences after Bonferroni correction. Additional analyses demonstrating that advice requests were driven by confidence, above and beyond trial difficulty, can be found in the Supplementary Material.

3.2.4. Effects of time-in-block depend on need to learn

Replicating the results of Experiment 1 again, for the participant group who were not told about advisor accuracy, advice seeking at high levels of confidence was more strongly apparent in the first half vs. second half of each block (Fig. 5B). Thus, for this group, a preregistered two-way repeated measures ANOVA on the proportion of advice requests, revealed reliable main effects of confidence (Q1-Q6: $F(5, 120) = 7.49, p < .001, \eta_p^2 = 0.24$), and time-in-block (first vs. second half of

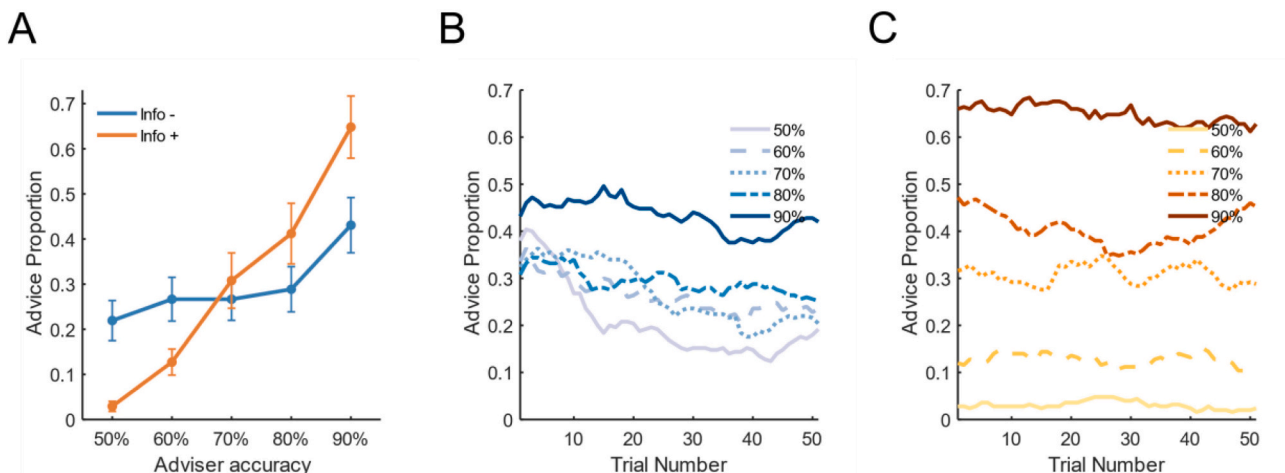


Fig. 4. Results of Experiment 2: (A) Mean proportion of advice selections according to advisor accuracy and explicit information availability. Participants who did not receive explicit information about advisor accuracies are depicted in blue, and participants who did receive explicit information about advisors are depicted in orange. Error bars indicate SEMs. (B) Ten trial running average showing the proportion of advice requests according to advisor accuracy, across blocks and participants when explicit information about advisors was not available, and (C) when explicit information about advisor accuracy was available. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

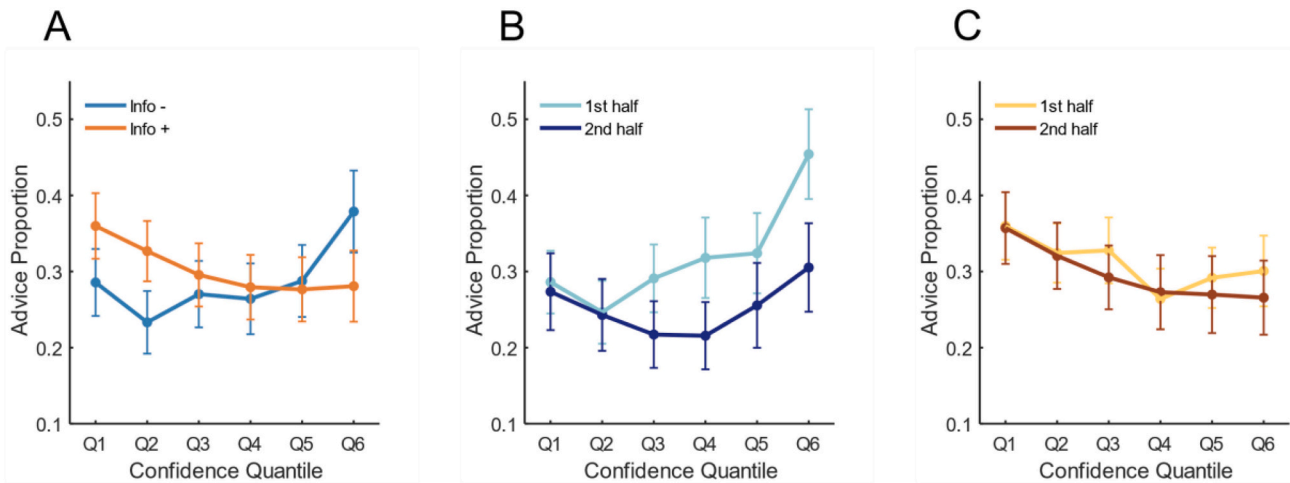


Fig. 5. Results of Experiment 2. (A) Mean proportion of advice choices according to explicit information availability and confidence. Participants who did not receive explicit information about advisor accuracies are depicted in blue, and participants who did receive explicit information about advisors are depicted in orange. (B, C) Mean proportion of advice choices according to information availability, confidence, and time in block. For each participant, trials were divided into two time periods (first half vs. second half of block), and then into six bins according to confidence ratings, within each time period. (B) Participants who did not receive explicit information about advisor accuracy. (C) Participants who did receive explicit information about advisor accuracy. Error bars represent SEMs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

blocks: $F(1, 24) = 6.85, p = .015, \eta_p^2 = 0.22$) and, importantly, a significant interaction between these factors ($F(5, 120) = 3.72, p = .014, \eta_p^2 = 0.13$). Confidence reliably predicted advice requests during the first time period ($F(5, 120) = 8.64, p < .001, \eta^2 = 0.26$), where the dominant trend was for participants to seek advice most often when they were highly confident in their initial response. In the second half of blocks, this trend was reduced, and there was a more pronounced tendency for increased advice seeking also at the lowest levels of confidence, but these effects were not statistically reliable ($F(5, 120) = 2.63, p = .061, \eta^2 = 0.10$). In contrast, advice-seeking behavior was more stable from first to second halves of blocks for the group of participants told about advisor accuracy: Their advice requests were reliably predicted by confidence ($F(5, 120) = 3.29, p = .04, \eta_p^2 = 0.12$), but with no significant effect of time-in-block ($F(1, 24) = 1.95, p = .5, \eta_p^2 = 0.08$) and no interaction between time-in-block and confidence on their tendency to seek advice ($F < 1$, see Fig. 5C).

3.3. Discussion

Experiment 2 investigated whether, when choosing between advice and re-sampling the evidence, people use their confidence differently depending on the availability of explicit information about advisor reliability. Replicating the results from Experiment 1, we found that when explicit information about advisor quality was unavailable, participants requested advice more frequently when they were highly confident about their initial response. In contrast, when external information about advisor quality was available, we see a trend towards participants requesting advice more often when they were lowest in confidence. Our results further demonstrated that the relationship between confidence and information choices varied from the first to second half of blocks only when explicit information about advisors was not available. Collectively, these results support our interpretation that people use their confidence as an internal feedback cue, allowing them to learn how good advice is by comparing the advice they received with their own beliefs when they are highly confident. When participants do not want to learn about advice quality, either because explicit information is available or because they have already learned earlier in the block, they do not use confidence as a feedback proxy for evaluating advisors.

Taken together, Experiments 1 and 2 provide evidence for our two key hypotheses: that confidence guides adaptive information seeking in

a flexible manner, and that confidence can function as an internal feedback cue, used to evaluate the quality of advice when explicit information about advisor accuracy is not available. In Experiment 3, we aimed to provide further support for these claims by investigating how people use confidence when explicit information about advisor accuracy is not available but external feedback is provided. If, as we hypothesize, confidence was previously used as an internal feedback proxy, we would expect that its effects would change when external feedback is provided, rendering this proxy redundant.

4. Experiment 3

Experiment 3 used the same perceptual decision task and JAS approach as Experiments 1 and 2, but half of the participants now received trial-by-trial objective feedback, while the other half of the participants did not (replicating the conditions of the previous experiments). Neither group was given explicit information about advisor reliability. We expected that only participants without access to external feedback would request advice more frequently on high confidence trials, using confidence as a feedback proxy in order to learn about advisor quality. We did not expect this relationship between confidence and advice requests when participants received external feedback, since feedback would provide information about advice accuracy, regardless of subjective confidence on a given trial. Key predictions and analyses were preregistered at <https://aspredicted.org/2wq7f.pdf>.

4.1. Methods

4.1.1. Participants

Eighty-eight participants were tested in order to achieve a final sample of 40 participants in each experimental group (32 females, ages 19–42 years, $M = 27.9, SD = 5.9$). A larger sample was recruited compared to Experiments 1 and 2, because we were collecting data online and expected key effects might be weaker given lower compliance and attention levels for online participants than for individual, in-person testing as used in previous experiments. Four participants were excluded because their data could not be divided into appropriate confidence bins, and four additional participants were excluded because their accuracy was >2 standard deviations below the average accuracy rate. Participants were recruited via Prolific Academic (<https://www.prolific.co>) and were compensated for participation. They received online

instructions and agreed to take part before beginning the experiment. All procedures were approved by the Oxford University Medical Sciences Interdivisional Research Ethics Committee.

4.1.2. Stimuli and apparatus

Experiment 3 was an online study. All stimuli were created and presented using JavaScript and CSS. Stimulus sizes were defined in pixels, and so varied according to the specific screen used by each participant.

4.1.3. Task and procedure

Participants performed a perceptual judgement task, similar to the task used for Experiments 1 and 2. Several small changes were made to the procedure to accommodate the change to an online platform, but these did not alter the substance of the task. Each trial began with a black fixation cross at the center of a grey screen, with two white framed boxes appearing to the left and right of the fixation. After 500 ms the fixation flashed, indicating the dot stimuli will appear briefly. The dots appeared after an additional 300 ms and remained on the screen for 150 ms. Participants had to report which box contained more dots and how confident they were in their decision, in a single response given on a scale ranging from 100% left to 100% right. If participants selected 50%, the message “please choose one side or the other” appeared in red under the confidence scale, and they had to update their response and select one side before they could proceed. After recording their initial response, participants chose between viewing the stimulus again or receiving advice from a virtual advisor. Two boxes representing the information types appeared above and below the center of the screen, and participants selected which information to receive by pressing the corresponding box with the left mouse button. If participants chose to receive advice, the box representing the stimulus choice disappeared, and a text box emerged from the advisor box, pointing in the direction of the advisor’s response, with the text “I think it was on the RIGHT / LEFT” appearing within. After 1000 ms the confidence scale appeared below the advice, with the advice remaining on the screen until participants submitted their final response. If participants chose to view the stimulus again, they were presented once more with a fixation for 800 ms (flashing after 500 ms), followed by the same dot stimulus for an extended period of 300 ms. The confidence scale appeared immediately after the dot stimulus disappeared.

Participants were divided into the two experimental groups randomly (feedback available vs. feedback absent). Participants in the feedback available group received feedback after every trial, in the form of the dot stimulus appearing again in the correct box for 300 ms after they submitted their final decision (the incorrect box remained empty). Participants in the feedback absent group viewed empty boxes, without the dot stimulus, for the same duration. The next trial began after a 300 ms inter-trial interval. All participants received feedback on their average accuracy at the end of each block. Participants completed five blocks of 60 trials, with advisor accuracy ranging from 60% - 100% accurate, randomly ordered across blocks. As in the previous experiments, trials were in three varying levels of difficulty, with the medium difficulty staircased using a 2-down 1-up procedure.

Before beginning the main experimental blocks, participants received online instructions and training. Training comprised two blocks of 54 trials of only the dot judgement task (with feedback) followed by a 12-trial practice block of the entire task including information choice and final decision (with feedback only for the Feedback Available group).

Participants were asked to complete the experiment in a quiet and comfortable environment. To encourage task engagement, they were informed they could earn additional bonus payment depending on their accuracy levels.

4.2. Results

4.2.1. Accuracy and confidence judgements

Analysis of mean accuracy, via a two-way mixed measures ANOVA with factors for group (Feedback Absent vs. Available) and response (initial, final), revealed the usual benefit of additional information in the main effect of response ($F(1, 78) = 259.94, p < .001, \eta_p^2 = 0.77$), but no main effect of group ($F < 1$), nor a reliable interaction ($F(1, 78) = 1.57, p = .21, \eta_p^2 = 0.02$): Accuracy improved in final decisions relative to initial responses in both the Feedback Absent ($M = 81\%$ vs. 76%) and Feedback Available ($M = 83\%$ vs. 76%) groups. Participants’ subjective confidence ratings mirrored these effects (Feedback Absent: $M = 78$ vs. 72 , for final decisions and initial responses, respectively; Feedback Available: $M = 81$ vs. 74 , $SDs = 10$), with a two-way mixed measures ANOVA revealing a reliable effect for response ($F(1, 78) = 140.01, p < .001, \eta_p^2 = 0.64$), but not group ($F(1, 78) = 1.07, p = .30, \eta_p^2 = 0.14$), and no significant interaction ($F < 1$).

4.2.2. Effects of advisor quality

Participants in both groups showed sensitivity to advisor accuracy in terms of their proportion of advice requests (see Fig. 6A). A two-way mixed measures ANOVA on the proportion of advice requests with factors for advisor accuracy and feedback availability showed reliable effects for advisor ($F(4, 312) = 10.08, p < .001, \eta_p^2 = 0.11$) and for the interaction effect between advisor accuracy and feedback availability ($F(4, 312) = 4.93, p < .005, \eta_p^2 = 0.06$). Follow-up ANOVAs were performed to examine the influence of advisor within each group separately. For participants in the no-feedback condition, the influence of advisor accuracy on advice requests approached significance ($F(4, 156) = 2.47, p = .058, \eta^2 = 0.06$) with a significant linear trend ($F(1, 39) = 5.77, p = .02, \eta^2 = 0.13$), indicating that participants tended to request more advice from better advisors, even when feedback was not available. When feedback was provided, the proportion of advice requests was reliably predicted by advisor quality ($F(4, 156) = 11.25, p < .001, \eta^2 = 0.22$, with a significant linear trend, $F(1, 39) = 31.21, p < .001, \eta^2 = 0.44$).

4.2.3. Effects of confidence on information choices depend on feedback availability

As in the previous experiments, participants displayed large individual differences in their preferences for selecting advice vs. viewing the stimulus again. On average, participants in the group not provided with trial-by-trial feedback chose to receive advice on 0.30 of the trials (range = 0–0.82, $SD = 0.25$), and participants in the group given feedback chose to receive advice on 0.37 of trials (range = 0–0.88, $SD = 0.28$). Fig. 6B shows the relationship between confidence and advice requests within each feedback condition. Consistent with our key prediction, a preregistered two-way mixed ANOVA on the proportion of advice selections with factors for group (Feedback Absent vs. Available) and confidence (Q1–Q6) revealed a significant interaction between these two factors ($F(5, 390) = 3.36, p = .036, \eta_p^2 = 0.04$), demonstrating that confidence influenced advice-seeking differently according to the availability of feedback. No significant main effects were found for group ($F(1, 78) = 1.56, p = .22, \eta_p^2 = 0.02$) or confidence ($F(5, 390) = 2.20, p = .11, \eta_p^2 = 0.03$).

To study the relationship between confidence and advice requests within each feedback condition, follow-up one-way ANOVAs were performed for each group separately. For participants in the Feedback Absent group, confidence reliably predicted advice requests ($F(5, 195) = 3.48, p = .035, \eta^2 = 0.08$), with participants requesting more advice when confidence was lowest (Q1), and highest (Q6), compared to medium confidence trials. This was captured by a significant quadratic trend ($F(1, 39) = 19.03, p < .001, \eta^2 = 0.33$), although no pairwise comparisons between confidence bins were significant after Bonferroni correction. In contrast, for participants receiving explicit trial-by-trial feedback, confidence did not reliably predict advice requests ($F(5,$

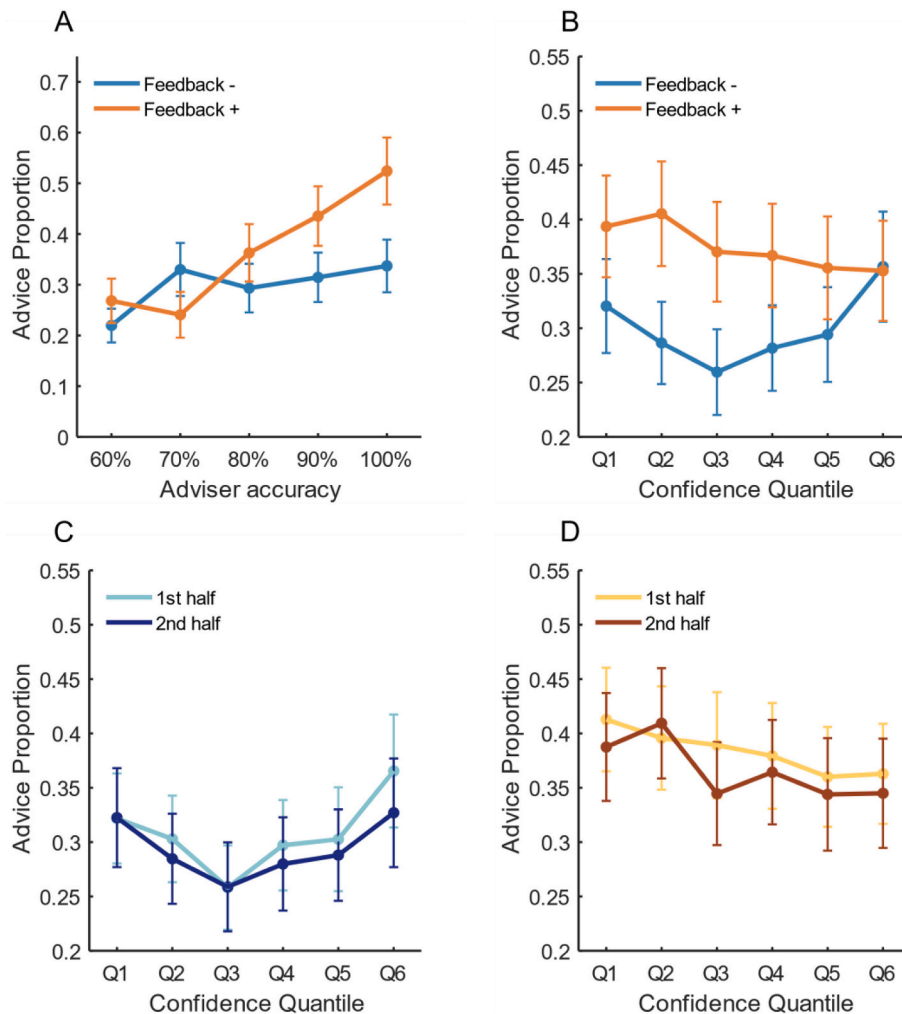


Fig. 6. Results of Experiment 3. Average proportion of advice selections according to: (A) Feedback availability and advisor accuracy. (B) Feedback availability and confidence bin. (C, D) Feedback availability, confidence, and time in block. For each participant, trials were divided into two time periods (first half vs. second half of block), and then into six bins according to confidence ratings, within each time period. (C) Participants who did not receive trial-by-trial feedback, and (D) Participants who did receive trial-by-trial feedback. Error bars indicate standard errors of the mean (SEM).

195) = 1.82, $p = .17$, $\eta^2 = 0.04$), but, numerically, the rate of advice requests was lowest on trials where participants reported highest confidence in their initial response.

4.2.4. Confidence and advice-seeking over time

In contrast to the previous experiments, here we observed only weak and inconsistent changes in advice seeking behavior from the first to second half of blocks (Fig. 6C and D). We performed a preregistered two-way repeated measures ANOVAs on the proportion of advice requests for each group separately, as a function of time-in-block and confidence bin (after dividing trials into sextiles according to initial confidence). For the Feedback Absent group, the tendency for high-confidence advice seeking was only slightly more pronounced in the first half of blocks than in the second, such that we observed no reliable main effect of time-in-block ($F(1,39) = 1.26$, $p = .27$, $\eta_p^2 = 0.03$), and no significant interaction between time-in-block and confidence ($F < 1$). Likewise, in the Feedback Available group, we did not find any reliable effects for time-in-block ($F(1, 39) = 1.46$, $p = .23$, $\eta_p^2 = 0.04$), confidence ($F(5, 195) = 1.91$, $p = .09$, $\eta_p^2 = 0.05$), or their interaction ($F < 1$).

4.3. Discussion

Experiment 3 confirmed our prediction that the availability of objective feedback would modulate how people use confidence when deciding whether to request advice or re-sample stimulus information. When external feedback was unavailable, we again observed that advice-seeking varied with participants' confidence in their initial

responses, with the highest rate of requesting advice when confidence was maximal. These results lend further support to our suggestion that, when external information is not available, confidence is used as a feedback proxy, allowing assessment of advisor accuracy. This group of participants also showed a tendency to seek advice at higher rates when very low in confidence, overall showing the quadratic relationship between confidence and advice-seeking that we originally predicted in Experiment 1. In contrast, participants in the group given trial-by-trial feedback showed only a weak relationship between confidence and advice-seeking, and had numerically the lowest rate of advice seeking when they were most confident in their initial response.

Although the results were in line with our critical predictions, in some details, performance of participants in the Feedback Absent group of this experiment differed from that in the comparable conditions of the previous experiments: There were weaker learning effects in terms of distinguishing between advisors of differing accuracy and, perhaps correspondingly, these participants' use of confidence was not reliably modulated as blocks progressed. These discrepancies might be attributed to Experiment 3 being conducted on-line, such that it drew on a different (perhaps less engaged) participant sample than the in-lab studies in Experiments 1 and 2.

5. General discussion

Confidence reflects a subjective estimate of the likelihood that a decision reached was a good one, and as such can play an important role in guiding adaptive behavior. Recent studies have demonstrated the

involvement of confidence in a variety of behaviors, such as learning (Guggenmos, Wilbertz, Hebart, & Sterzer, 2016; Hainguerlot, Vergnaud, & de Gardelle, 2018) and information seeking (Desender et al., 2018; Desender, Murphy, et al., 2019). These studies have typically revealed consistent mappings between confidence and behavior, whereby low confidence leads to one course of action, and high confidence to another. In the present study, we tested the hypothesis that the relationship between confidence and behavior will vary strategically according to the specific goal and context, based on the notion that explicit representations of confidence and uncertainty serve crucially to guide behavior in a flexible, goal-directed manner (Dehaene & Changeux, 2011; Nelson & Narens, 1990).

We tested this hypothesis in three experiments examining the relationship between confidence and advice seeking. In our judge-advisor system (JAS) design, participants chose between receiving advice and re-sampling the evidence after each initial decision. We found that, in most cases, participants' information choices were predicted by their previous confidence judgements, and, critically, that this relationship varied depending on the availability of information about advisor reliability. When participants did not have any external information about advisor accuracy, they selected advice most often when they were most confident in their initial responses. In contrast, when they had access to objective information about advisor accuracy, participants tended to request advice most often when their confidence was lowest. These patterns of behavior appeared in our data consistently, regardless of the type of information available about advisor quality: whether it was provided explicitly, learnable via external feedback, or learnt through exposure earlier in the block. Thus, whereas low-confidence advice requests reflect the use of confidence as a self-monitoring system—with low confidence acting as a cue that participants are unlikely to succeed in the task—high-confidence advice requests reflect the use of confidence as a feedback proxy, allowing participants to learn about advisor reliability by comparing advice to the expected answer when they are certain they know the correct answer. Supporting the suggestion that confidence can be used as a feedback proxy (Pescetelli & Yeung, 2021), our participants were able to learn about the quality of advice, and requested more advice from better advisors, even in the absence of external feedback. Together, these results extend findings from previous studies, which have demonstrated that confidence can be used as a learning signal during perceptual tasks (Guggenmos et al., 2016), to show that confidence can also be valuable as a feedback signal when learning about social information.

The present findings complement previous research into advice taking using the Judge Advisor System. Like many other actions, the decision whether to seek advice can be viewed in terms of the costs and benefits associated with each line of action (Van Swol, Paik, & Prah, 2018). Often, soliciting the help of external advisors requires the investment of time and money, particularly when advice is given by professionals or experts. Typically, the main benefit of advice is to improve performance and gains (for example when requesting advice from a banker about investments). Interestingly, there have been suggestions that advice might provide benefits in addition to improved performance, such as joint responsibility (El Zein, Bahrami, & Hertwig, 2019; Harvey & Fischer, 1997). Here we demonstrate that advice seeking can also provide knowledge about advisor reliability. By requesting advice when cost is low (e.g., when we already know the answer), people can learn about the reliability and accuracy of others and build up their trust in an advisor. This information can be important when deciding whether to invest in advice in the future. Future studies might examine whether these effects replicate in contexts other than perceptual tasks, for example when people are faced with more complex value-based choices that do not have an evident ground truth.

Our finding that confidence is used to learn about the reliability of others resonates with the conclusions of recent studies investigating the formation of global estimates of self-performance (SPEs; Boldt, Schiffer, Waszak, & Yeung, 2019; Rouault et al., 2019). In these studies, the

researchers demonstrated that local confidence (confidence in a single decision) is aggregated and used for building a global self-performance estimate over time. These estimates affect neural preparation for stimulus processing (Boldt, Schiffer, et al., 2019), and, importantly, allow people to decide which tasks to pursue next, based on their beliefs of how likely they are to succeed (Carlebach & Yeung, 2020; Rouault et al., 2019). Here, we demonstrate that local confidence is used not only to form a global estimate of self-performance, but also global estimates about the performance of others. This ability to monitor and evaluate one's own performance, and compare it to the performance of others, is particularly important in social contexts, when decisions must be made as a group, or when deciding whether to seek or use other people's advice. Thus, our results suggest that confidence acts as a signal that allows not only the comparison and prioritisation of different tasks (Carlebach & Yeung, 2020; de Gardelle, Le Corre, & Mamassian, 2016; de Gardelle & Mamassian, 2014), but also the comparison of different sources of information. Future studies might extend these findings to consider different types and sources of information beyond advisors.

Previous studies exploring advice taking have demonstrated that people use advice more often when a task is difficult (Gino & Moore, 2007), and they are unconfident (Gino, Brooks, & Schweitzer, 2012; Morgan et al., 2012; Pescetelli et al., 2021). Based on these results, we initially predicted a U-shaped relationship between confidence and advice requests when external information about advisor accuracy was not available. However, whereas participants consistently requested advice more often when confidence was high, they did not show a consistent effect for low confidence across the three experiments. Thus, although our initially-predicted quadratic relationship between confidence and advice-seeking was significant in all three experiments, only in Experiments 2 and 3 did we observe a clear trend towards frequent advice requests at the lowest levels of initial confidence. Several methodological differences between the experiments might explain this inconsistency, such as the testing methodology (on-line vs. in-lab testing), different advisor accuracies, and differences in the end-of-block feedback. Regardless, it is striking that, whereas this intuitive relationship between low confidence and advice seeking was only inconsistently observed across experiments, the critical pattern of participants seeking advice when highly confident in their own initial responses was replicated in all experiments.

In a broader context, although we investigated the use of confidence for two specific purposes—as a cue that a task is hard and external assistance should be used, and as a feedback signal for learning—our main interest was not in these specific uses. Rather, the key idea that we address is people's ability to use confidence strategically, as demonstrated by the flexibility in the relationship between confidence and behavior. Importantly, this flexibility can be understood as a coherent overall strategy to maximize current performance, by simultaneously guiding information seeking about the world (via advice received) and by guiding learning about the reliability of that information (via evaluation of advice received against one's own judgements). The flexible use of confidence in our study is perhaps most evident when considering the results of all three experiments together. Results from Experiments 1 and 2 showed that, when no external information about advice was available, the relationship between confidence and advice requests changed as time progressed, with people requesting less advice when confidence was high once they had learned about the quality of advice. In Experiment 3, where participants did not learn about advice quality as effectively, participants continued requesting advice when confidence was high throughout the entire experiment. Taken together, these results show that participants changed their use of confidence strategically, in a way that advanced their current goals. These results correspond well with findings from metamemory research demonstrating that the relationship between metacognitive judgements of learning and study time allocation varies strategically depending on available study time and learners' expertise (Metcalfe, 2002, 2009).

Our findings are somewhat surprising given previous evidence that

the confidence signal is integrated in the brain automatically when making decisions (Lebreton, Abitbol, Daunizeau, & Pessiglione, 2015). Thus, it might be expected that consistent associations are learned between confidence and behavior. For example, a learned association might be that when confidence is low external help is sought to improve performance, while when confidence is high advice is discarded. Our results demonstrate that the relationship between confidence and behavior is more flexible than would be expected based on such associative mechanisms. This prompts a question for future research of the cognitive and neural mechanisms that enable people to learn different uses of confidence and apply them to situations appropriately, particularly when situations differ only in people's internal state or level of knowledge, as in the experiments we describe here.

In conclusion, the results of the current investigation show that people use their confidence both as a self-monitoring tool and as a feedback proxy allowing them to learn about the reliability of others. Our findings fit in with the concept of subjective confidence specifically, and metacognitive cues in general, acting as a tool that can be used to monitor behavior and adjust performance accordingly (Nelson & Narens, 1990). As such, our subjective sense of confidence is a powerful tool that we carry with us and use flexibly according to our current goals and needs.

Data availability

Raw data are available via Open Science Framework: <https://osf.io/vbnux/>.

CRediT authorship contribution statement

Nomi Carlebach: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Nick Yeung:** Supervision, Conceptualization, Methodology, Writing – review & editing.

Data availability

All raw data can be freely accessed at <https://osf.io/njmez/>

Acknowledgements

We thank Niccolo Pescetelli for his help with Matlab scripts and Matt Jaquiere for his help with JavaScript. Nomi Carlebach was supported by a DPhil studentship from the Department of Experimental Psychology, University of Oxford, and grants from the British Friends of the Hebrew University, Funds for Women Graduates, B'nai B'rith Leo Baeck (London) Lodge and the Anglo-Israel Association. Nick Yeung was supported by the UK Ministry of Defence through the Defence Science and Technology Laboratory (DSTLX 1000128890), under the Bilateral Academic Research Initiative programme.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cognition.2022.105264>.

References

- Baranski, J. V., & Petrusic, W. M. (1994). The calibration and resolution of confidence in perceptual judgments. *Perception & Psychophysics*, 55(4), 412–428. <https://doi.org/10.3758/BF03205299>
- Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature*, 456(7219), 245–249. <https://doi.org/10.1038/nature07538>
- Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual review of psychology*, 64, 417–444.
- Boldt, A., Blundell, C., & De Martino, B. (2019). Confidence modulates exploration and exploitation in value-based learning. *Neuroscience of Consciousness*, 2019(1), 1–12. <https://doi.org/10.1093/nc/niz004>
- Boldt, A., & Gilbert, S. J. (2019). Confidence guides spontaneous cognitive offloading. *Cognitive Research: Principles and Implications*, 4(1). <https://doi.org/10.1186/s41235-019-0195-y>
- Boldt, A., Schiffer, A.-M., Waszak, F., & Yeung, N. (2019). Confidence predictions affect performance confidence and neural preparation in perceptual decision making. *Scientific Reports*, 9(1), 4031. <https://doi.org/10.1038/s41598-019-40681-9>
- Bonaccio, S., & Dalal, R. S. (2006). 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. <https://doi.org/10.1016/j.obhdp.2006.07.001>
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9176952>.
- Carlebach, N., & Yeung, N. (2020). Subjective confidence acts as an internal cost-benefit factor when choosing between tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 46(7), 729–748. <https://doi.org/10.1037/xhp0000747>
- Dehaene, S., & Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, 70(2), 200–227. <https://doi.org/10.1016/j.neuron.2011.03.018>
- Desender, K., Boldt, A., Verguts, T., & Donner, T. H. (2019). Confidence predicts speed-accuracy tradeoff for subsequent decisions. *eLife*, 8, 1–25. <https://doi.org/10.7554/eLife.43499>
- Desender, K., Boldt, A., & Yeung, N. (2018). Subjective confidence predicts information seeking in decision making. *Psychological Science*. <https://doi.org/10.1177/0956797617744771>
- Desender, K., Murphy, P., Boldt, A., Verguts, T., & Yeung, N. (2019). A postdecisional neural marker of confidence predicts information-seeking in decision-making. *Journal of Neuroscience*, 39(17), 3309–3319. <https://doi.org/10.1523/JNEUROSCI.2620-18.2019>
- El Zein, M., Bahrami, B., & Hertwig, R. (2019). Shared responsibility in collective decisions. *Nature Human Behaviour*, 3(6), 554–559. <https://doi.org/10.1038/s41562-019-0596-4>
- Frost, R. O., Marten, P., Lahart, C., & Rosenblate, R. (1990). The dimensions of perfectionism. *Cognitive Therapy and Research*, 14(5), 449–468. <https://doi.org/10.1007/BF01172967>
- de Gardelle, V., Le Corre, F., & Mamassian, P. (2016). Confidence as a common currency between vision and audition. *PLoS One*, 11(1), Article e0147901. <https://doi.org/10.1371/journal.pone.0147901>
- de Gardelle, V., & Mamassian, P. (2014). Does confidence use a common currency across two visual tasks? *Psychological Science*, 25(6), 1286–1288. <https://doi.org/10.1177/0956797614528956>
- Gilbert, S. J. (2015). Strategic use of reminders: Influence of both domain-general and task-specific metacognitive confidence, independent of objective memory ability. *Consciousness and Cognition*, 33, 245–260. <https://doi.org/10.1016/j.concog.2015.01.006>
- Gino, F., Brooks, A. W., & Schweitzer, M. E. (2012). Anxiety, advice, and the ability to discern: Feeling anxious motivates individuals to seek and use advice. *Journal of Personality and Social Psychology*, 102(3), 497–512. <https://doi.org/10.1037/a0026413>
- Gino, F., & Moore, D. A. (2007). Effects of task difficulty on use of advice. *Journal of Behavioral Decision Making*, 20(1), 21–35. <https://doi.org/10.1002/bdm.539>
- Guggenmos, M., Wilbertz, G., Hebart, M. N., & Sterzer, P. (2016). Mesolimbic confidence signals guide perceptual learning in the absence of external feedback. *eLife*, 5, 1–19. <https://doi.org/10.7554/eLife.13388>
- Hainguerlot, M., Vergnaud, J.-C., & de Gardelle, V. (2018). Metacognitive ability predicts learning cue-stimulus associations in the absence of external feedback. *Scientific Reports*, 8(1), 5602. <https://doi.org/10.1038/s41598-018-23936-9>
- Harvey, N., & Fischer, I. (1997). Taking advice accepting help, improving judgment. *Organizational Behavior and Human Decision Processes*, 70(2), 117–133. Retrieved from <https://pdfs.semanticscholar.org/2987/5f69c4e39bee1c415c6e9fb7b1f02bc19b.pdf>.
- Hu, X., Luo, L., & Fleming, S. M. (2019). A role for metamemory in cognitive offloading. *Cognition*, 193(August 2018). <https://doi.org/10.1016/j.cognition.2019.104012>
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, 36(14), 1–16.
- Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, 103(3), 490–517. <https://doi.org/10.1037/0033-295X.103.3.490>
- Lebreton, M., Abitbol, R., Daunizeau, J., & Pessiglione, M. (2015). Automatic integration of confidence in the brain valuation signal. *Nature Neuroscience*, 18(8), 1159–1167. <https://doi.org/10.1038/nn.4064>
- Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, 47(4), 1122–1135. <https://doi.org/10.3758/s13428-014-0532-5>
- Metcalfe, J. (2002). Is study time allocated selectively to a region of proximal learning? *Journal of Experimental Psychology: General*, 131(3), 349–363. <https://doi.org/10.1037/0096-3445.131.3.349>
- Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science*, 18(3), 159–163. <https://doi.org/10.1111/j.1467-8721.2009.01628.x>
- Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, 15(1), 174–179. <https://doi.org/10.3758/PBR.15.1.174>

- Meyniel, F., Sigman, M., & Mainen, Z. F. (2015). Confidence as Bayesian probability: From neural origins to behavior. *Neuron*, 88(1), 78–92. <https://doi.org/10.1016/j.neuron.2015.09.039>
- Morgan, T. J. H., Rendell, L. E., Ehn, M., Hoppitt, W., & Laland, K. N. (2012). The evolutionary basis of human social learning. *Proceedings of the Royal Society B: Biological Sciences*, 279(1729), 653–662. <https://doi.org/10.1098/rspb.2011.1172>
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation*, 26, 125–173. [https://doi.org/10.1016/S0079-7421\(08\)60053-5](https://doi.org/10.1016/S0079-7421(08)60053-5)
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437–442. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9176953>.
- Pescetelli, N., Hauperich, A. K., & Yeung, N. (2021). Confidence, advice seeking and changes of mind in decision making. *Cognition*, 215(June). <https://doi.org/10.1016/j.cognition.2021.104810>
- Pescetelli, N., & Yeung, N. (2021). The role of decision confidence in advice-taking and trust formation. *Journal of Experimental Psychology: General*, 150(3), 507–526. <https://doi.org/10.1037/xge0000960>
- Rabbitt, P. M. (1966). Errors and error correction in choice-response tasks. *Experimental Psychology*, 71(2), 264–272. <https://doi.org/10.1037/h0022853>
- Rouault, M., Dayan, P., & Fleming, S. M. (2019). Forming global estimates of self-performance from local confidence. *Nature Communications*, 10(1), 1–11. <https://doi.org/10.1038/s41467-019-09075-3>
- Samaha, J., Switzky, M., & Postle, B. R. (2019). Confidence boosts serial dependence in orientation estimation. *Journal of Vision*, 19(4), 1–13. <https://doi.org/10.1167/19.4.25>
- Snizek, J. A., & Buckley, T. (1989). *Social influence in the advisor-judge relationship. Annual meeting of the judgment and decision making society, Atlanta, Georgia.*
- Snizek, J. A., & Buckley, T. (1995). Cueing and cognitive conflict in judge-advisor decision making. *Organizational Behavior and Human Decision Processes*, 62, 159–174. <https://doi.org/10.1006/obhd.1995.1040>
- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, 168(3), 242–249. <https://doi.org/10.1016/j.psychres.2008.05.006>
- Van Swol, L. M., Paik, J. E., & Pahl, A. (2018). In E. L. MacGeorge, & L. M. Van Swol (Eds.), *1. Advice recipients*. <https://doi.org/10.1093/oxfordhb/9780190630188.013.2>
- Yaniv, I. (2004). Receiving other people's advice: Influence and benefit. *Organizational Behavior and Human Decision Processes*, 93(1), 1–13. <https://doi.org/10.1016/j.obhdp.2003.08.002>
- Yaniv, I., & Kleinberger, E. (2000). Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational Behavior and Human Decision Processes*, 83(2), 260–281. <https://doi.org/10.1006/obhd.2000.2909>
- Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: Confidence and error monitoring. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 367(1594), 1310–1321. <https://doi.org/10.1098/rstb.2011.0416>
- Yeung, N., & Summerfield, C. (2014). Shared mechanisms for confidence judgements and error detection in human decision making. In S. M. Fleming, & C. D. Frith (Eds.), 9783642451. *The cognitive neuroscience of metacognition* (pp. 147–167). https://doi.org/10.1007/978-3-642-45190-4_7