Dissociating Confidence Bias and Confidence Noise in Perceptual and Knowledge-Based Decisions

Matteo Lisi
Royal Holloway, University of London

Marc Pabst & Lottie Wood
University College London

Abstract

People often misjudge how reliable their decisions are, leading to confidence errors. These errors can be due to systematic distortions (confidence bias) or random variability (confidence noise). In this study we use a dual-decision method, in which participants are required to use confidence in a prior decision to inform expectations about subsequent choices, to examine the importance of these two sources of confidence error in both perceptual and knowledge-based tasks. Across a reanalysis of published data and two new studies, we find that perceptual tasks elicit under-confidence relative to a Bayesian ideal model, while knowledge-based tasks exhibit increased confidence noise. Additional conditions using calorie estimation tasks suggest that some domains blend perceptual and knowledge-based decision features. These findings provide novel insights into the computational structure of confidence, suggesting that different cognitive domains are subject to distinct metacognitive constraints.

Keywords: confidence; metacognition; uncertainty; perception; decision-making; dual-decision paradigm

Introduction

Confidence judgments guide behavior in uncertain situations, yet are often inaccurate. Two distinct sources of error have been identified: *confidence bias*, referring to systematic distortions in estimated uncertainty, and *confidence noise*, or trial-to-trial variability in internal confidence (Shekhar & Rahnev, 2022; Boundy-Singer et al., 2023).

These errors are typically studied using explicit confidence ratings, which are prone to response biases. To address this limitation, we adopt an implicit modeling approach using a dual-decision paradigm (Lisi et al., 2021), where confidence is inferred from how performance in one decision informs the next.

The present study builds on this framework to ask whether confidence errors differ across decision domains. We compare perceptual and knowledge-based tasks — including a calorie estimation condition that combines perceptual and knowledge elements — to test whether confidence bias and noise vary systematically across decision domains.

Methods

We analyzed data from two previously published studies (Lisi et al., 2021; Constant et al., 2023) and two new experiments using the same dual-decision paradigm, adapted to perceptual (dot numerosity) and knowledge-based (GDP comparison, food calories) tasks. The knowledge-based tasks followed the trivia-style design of Lund et al. (2023). GDP per capita values were sourced from the World Bank API (World Bank, 2024), and food stimuli with caloric annotations were taken from the Full4Health Image Collection (Charbonnier et al., 2016; Smeets, 2022).

In the dual-decision paradigm (Fig. 1), participants first made a binary choice (e.g., more dots or higher GDP), followed by a second decision on a new stimulus pair. Critically, the correct answer on the second decision appeared on the right if the first response was correct, and on the left if it was not. This mapping was deterministic and explicitly explained to participants, who were encouraged to use it to improve performance. Each correct response (in either decision) earned an entry into a £50 Amazon voucher lottery.

Task difficulty was adjusted using a 3-up-1-down staircase based on first-decision performance. In the dot numerosity task, difficulty was manipulated via the log-ratio of dot counts (always summing to 100, initially 60–40, then reduced in 2-dot steps). In the GDP and calorie tasks, difficulty followed a ranked sequence of unique stimulus pairs, with each pair shown only once per participant.

In Study 1 (S1), 21 participants completed 150 GDP trials. Study 2 (S2) was a within-participant design: 23 participants completed 250 trials each of the GDP and dot numerosity tasks in randomized order. A subset of 14 participants in S2 also completed 250 trials of the food-calories task.

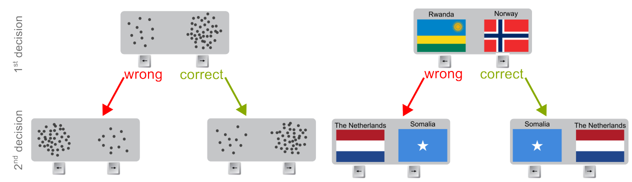

Figure 1: Schematic of the dual-decision task for the perceptual (left) and knowledge-based (right) conditions. On each trial, participants first chose the stimulus with the greater quantity — either more dots (perception) or higher GDP per capita (knowledge, shown with flags and names). This was followed by a second decision on a new stimulus pair. If the initial decision was correct, the correct option appeared on the right; otherwise, on the left.

Computational models

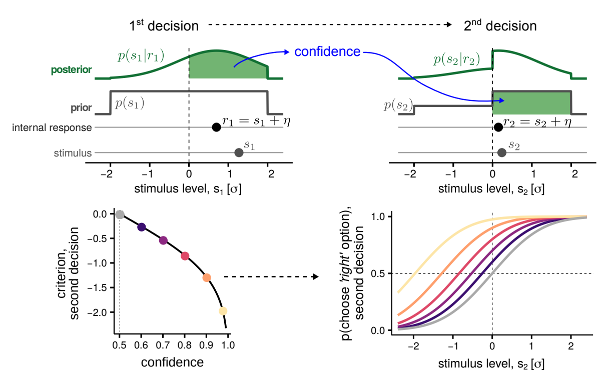

Figure 2: Idealised Bayesian model for the dual-decision task. The confidence in the first decision is used as a prior in the second decision (upper panels), and this process amount to shifting the decision criterion according to uncertainty in individual trials (lower panels).

We assume that participants base their decisions on r_i, a noisy internal estimate of the true stimulus value s_i, representing the difference between the two stimuli (such as the log GDP ratio of the two countries in the GDP task). This estimate is defined as $r_i = s_i + \eta$, where the subscript $i = 1, 2$ denotes the first and second decision, respectively, and η is Gaussian noise with variance σ^2, reflecting sensory or knowledge-related uncertainty. The task reduces to determining whether r_i is greater or less than zero.

The standard Bayesian model (Fig. 2) assumes that this type-1 noise σ^2 is the only source of error and that decisions follow a likelihood ratio rule. As shown by Lisi et al. (2021), this model corresponds to comparing r_2 to a decision criterion θ_2, which is a function of both the internal noise σ^2 and the confidence in the first decision, c_1 (defined as the Bayesian posterior probability that the first choice was correct given the evidence). The decision criterion is computed as:

$$\theta_2 = \sqrt{2}\sigma \, \text{erf}^{-1}(1 - 2c_1)$$

We consider two possible ways in which participants may deviate from the ideal Bayesian model. First, they may use a biased estimate of their own uncertainty, such that $\hat{\sigma} \neq \sigma$, to compute confidence. This *biased-Bayesian* model introduces one additional parameter:

$$\text{confidence bias} = \frac{\hat{\sigma}}{\sigma}$$

Second, we consider the possibility that participants' estimates of their own uncertainty fluctuate randomly from trial to trial. Following Boundy-Singer et al. (2023), we assume that these fluctuations follow a log-normal distribution:

$$\hat{\sigma} \sim \text{LogNormal}\left(\log\left(\frac{\sigma^2}{\sqrt{\sigma_m^2 + \sigma^2}}\right), \log\left(\frac{\sigma_m^2}{\sigma^2} + 1\right)\right)$$

Here, σ_m is the *confidence noise* parameter, which quantifies the amount of additional variability introduced in metacognitive computations.

Results

Estimated confidence parameters varied systematically across tasks (Fig. 3). Perceptual tasks were associated with stronger under-confidence (higher confidence bias; $t = 2.75$, $p = .013$) and lower confidence noise ($t = 3.93$, $p < .001$), based on linear mixed-effects models fit to log-transformed parameters with random intercepts for participant and study.

Model comparison results (Fig. 4) showed that, overall, the confidence bias model provided the best fit. This preference was more pronounced for perceptual tasks, where its advantage over the confidence noise model was approximately twice as large as in knowledge-based tasks.

Discussion

Previous work using this paradigm (Lisi, 2021; Constant et al., 2023) found consistent under-confidence in perceptual decisions. We extend this to knowledge-based tasks, where confidence bias is reduced or absent, aligning with prior findings of

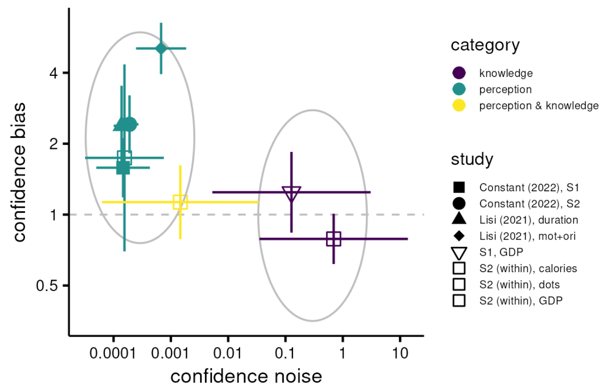

Figure 3: Confidence bias vs. confidence noise across tasks. Values of confidence bias > 1 indicate under-confidence. Points show participant medians; error bars are bootstrapped standard errors of the median.

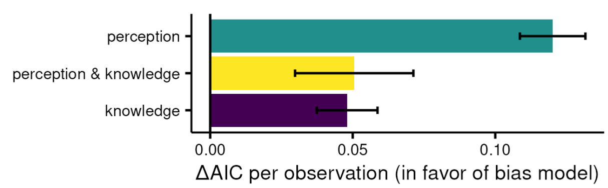

Figure 4: Model comparison. Positive values indicate better per-observation AIC for the confidence bias model relative to the confidence noise model.

high subjective confidence in semantic judgments (Fischhoff et al., 1977), but confidence noise is greater.

This pattern suggests domain-specific confidence strategies: under-confidence in perception may reflect an adaptive mechanism to down-weight noisy estimates, whereas knowledge-based decisions show less bias but greater susceptibility to confidence noise. These findings challenge models assuming a unified metacognitive process and highlight the need for domain-sensitive accounts of confidence computation.

Acknowledgments

This work was funded by Research Initiative Funds (RIF) from Royal Holloway, University of London.

References

Boundy-Singer, Z. M., Ziemba, C. M., & Goris, R. L. T. (2023). Confidence reflects a noisy decision reliability estimate. *Nature Human Behaviour, 7*(1), 142–154. https://doi.org/10.1038/s41562-022-01464-x

Charbonnier, L., van Meer, F., van der Laan, L. N., Viergever, M. A., & Smeets, P. A. (2016). Standardized food images: A photographing protocol and image database. *Appetite, 96*, 166–173. https://doi.org/10.1016/j.appet.2015.08.041

Constant, M., Pereira, M., Faivre, N., & Filevich, E. (2023). Prior information differentially affects discrimination decisions and subjective confidence reports. *Nature Communications, 14*(1), 5473. https://doi.org/10.1038/s41467-023-41112-0

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance, 3*(4), 552–564. https://doi.org/10.1037/0096-1523.3.4.552

Lisi, M., Mongillo, G., Milne, G., Dekker, T., & Gorea, A. (2021). Discrete confidence levels revealed by sequential decisions. *Nature Human Behaviour, 5*(2), 273–280. https://doi.org/10.1038/s41562-020-00953-1

Lund, A. E., Correa, C., Fardo, F., Fleming, S. M., & Allen, M. (2023, April 21). Domain generality in metacognitive ability: A confirmatory study across visual perception, memory, and general knowledge. *PsyArXiv*. https://doi.org/10.31234/osf.io/p97s5

Smeets, P. A. (2022, February 9). *Full4Health Image Collection*. https://doi.org/10.17605/OSF.IO/CX7TP

Shekhar, M., & Rahnev, D. (2021). Sources of metacognitive inefficiency. *Trends in Cognitive Sciences, 25*(1), 12–23. https://doi.org/10.1016/j.tics.2020.10.007

World Bank. (2024). *World Development Indicators: GDP per capita (current US$)*. Retrieved from https://data.worldbank.org/indicator/NY.GDP.PCAP.CD