Improving metacognitive ability with practice

Zoe M. Boundy-Singer (zoebsinger@utexas.edu)
Center for Perceptual Systems, 180 E Dean Keeton St
Austin, Texas 78712, USA

Akash Raj G P (akashrajgp@utexas.edu)
Center for Perceptual Systems, 180 E Dean Keeton St
Austin, Texas 78712, USA

Corey M. Ziemba (ziemba@utexas.edu)
Center for Perceptual Systems, 180 E Dean Keeton St
Austin, Texas 78712, USA

Robbe L.T. Goris (robbe.goris@utexas.edu)
Center for Perceptual Systems, 180 E Dean Keeton St
Austin, Texas 78712, USA

Abstract

Observers are aware of the fallibility of perception. When we feel confident in a perceptual interpretation, it is more likely to be correct. However, such metacognitive judgements are imperfect. Performance in difficult perceptual and cognitive tasks can improve with practice. Might metacognitive abilities improve with practice as well? We hypothesized that metacognitive learning may occur when (1) subjects perform a difficult metacognitive task and (2) are provided with useful feedback signals about their metacognitive judgements. To test this, we conducted a series of perceptual confidence experiments and manipulated both metacognitive task difficulty and feedback availability. We fit each subject's data with a novel dynamic model of perceptual confidence and studied the temporal evolution of metacognitive ability over the course of the task. We found that metacognitive learning is a robust, prominent, and general phenomenon, especially in difficult metacognitive tasks that include trial-by-trial metacognitive feedback.

Keywords: metacognition; confidence; process model

Introduction

Humans and other animals can meaningfully introspect about the quality of their actions and decisions. In general, we feel more confident in easy than in difficult decisions. However, confidence reports do not perfectly track decision accuracy, but also reflect response biases and difficulty misjudgements. This last factor limits the quality of confidence reports. The brain learns from experience, as is evident across a wide range of perceptual, cognitive, and motor tasks. We wondered whether this is also true of metacognitive judgements in perception. Does the quality of perceptual confidence reports improve as subjects acquire more experience with a perceptual decision-making task? We hypothesized that metacognitive learning may occur in tasks where confidence judgments are difficult, such that there is room for improvement, and in which subjects are provided with useful feedback about their confidence assessments, such that there is a supervisory signal.

Results

A statistic that can measure metacognitive learning

To test these hypotheses, we conducted several experiments in which human subjects jointly reported a binary decision about a sensory stimulus (belongs to 'Category A' vs 'Category B') and their confidence in this decision ('high' vs 'low'; Fig. 1a). The difficulty of the perceptual decision was varied by manipulating stimulus strength (e.g., stimulus orientation in an orientation discrimination task) and stimulus reliability (e.g., stimulus contrast in the orientation discrimination task). To provide metacognitive feedback, we rewarded correct high confidence responses more generously than correct low confidence responses. However, incorrect high confidence responses incurred a penalty, making the high confidence option risky (Fig. 1a).

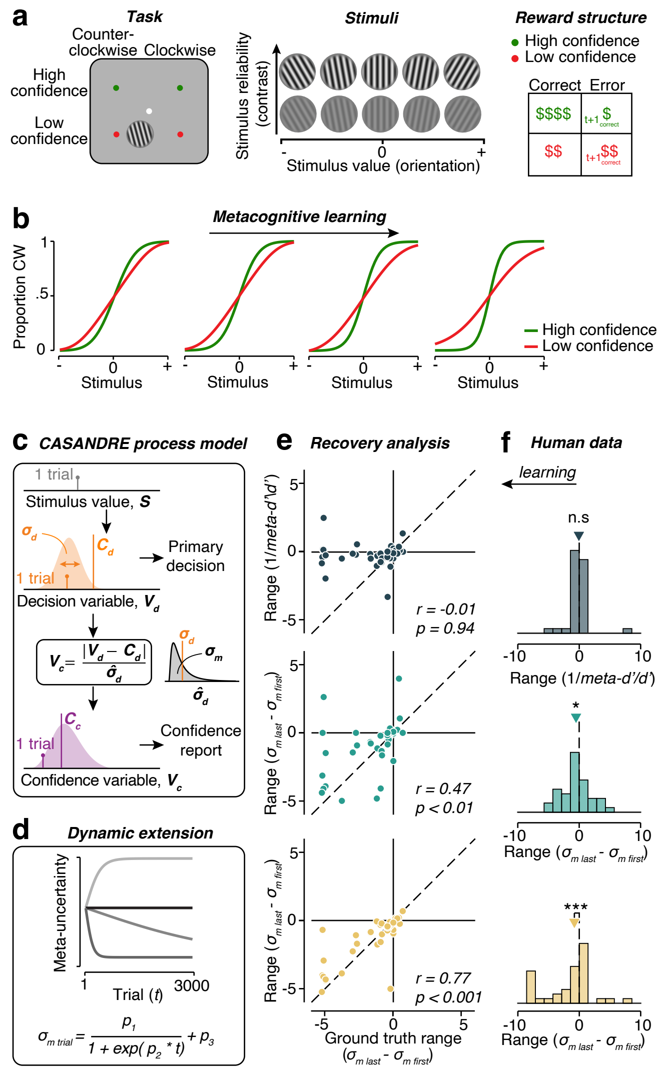

Figure 1: **a** (Left) Perceptual confidence task. (Middle) Task difficulty depends on stimulus value (x-axis) and reliability (y-axis). (Right) Reward structure. The better a subject is able to judge the reliability of a perceptual decision, the more reward they can earn. **b** Signature of metacognitive learning. **c** Schematic of CASANDRE, a two stage hierarchical model of confidence. On a single trial, confidence is the result of normalizing perceived stimulus strength ($|V_d - C_d|$) by an estimate of perceptual uncertainty ($\hat{\sigma}_d$). Meta-uncertainty, σ_m, determines metacognitive ability. **d** In dynamic-CASANDRE, meta-uncertainty is allowed to vary over time. **e** Each dot represents a simulated observer. In all panels, negative values indicate metacognitive learning. In the top panel, the y-axis is the difference in 1/(meta-d'/d') between the first and last quartile of trials. In the middle and bottom panel, it is the difference in static- and dynamic-CASANDRE's meta-uncertainty parameter, respectively. **f** Distribution of these three statistics for human subjects (N = 38, experiments 1,2,3,5,6). Triangles indicate population median.

Conceptually, metacognitive learning will manifest as a change in the slope of the psychometric functions conditioned on the subject's confidence report. Initially, when metacognitive ability is poor, there will not be a clear distinction between the reliability of "confident" and "not confident" perceptual choices (Fig. 1b, left; green vs red). But as introspective abilities improve over time, this distinction will grow (Fig. 1b, right). Recognizing this pattern unfold over the course of a limited number of trials (a few hundred at most) presents a statistical challenge. We first developed a method capable of exactly this. To this end, we used a recently proposed process model of confidence, CASANDRE (Boundy-Singer, Ziemba, & Goris, 2023), to generate synthetic choice-confidence data and varied the amount of metacognitive learning. One widely used descriptive metric of metacognitive ability is meta-d'/d' (Maniscalco & Lau, 2012). We computed this statistic on the first and last quartile of the synthetic data and plotted the difference against the ground truth. The correlation was small (r = -0.01, P = 0.94; Fig. 1e, top), demonstrating the unsuitability of this approach. We then estimated metacognitive ability using CASANDRE's meta-uncertainty parameter. This yielded a modest correlation (r = 0.47, P $<$ 0.01; Fig 1e, middle), showcasing the potential of a model-based analysis. Finally, we developed a dynamic extension of CASANDRE, holding on to the generative process, but allowing the parameters to drift over time (Fig. 1d). This modification yielded a substantially higher correlation (r = 0.77, P $<$ 0.001, Fig. 1e, bottom), and thus offers the first plausible statistical tool to study metacognitive learning.

Analysis of human data in six experiments

We conducted a series of perceptual confidence experiments in a controlled laboratory setting. All participants (N = 46) were naive to the purpose of our study, none had previously participated in psychophysical experiments, and each subject participated in only a single experiment. Consider the temporal evolution of meta-uncertainty across all experiments that involved metacognitive feedback (Fig. 1f, middle and bottom). The median difference in meta-uncertainty between the first and last trial was -0.80, indicating that metacognitive ability typically improved over the course of the experiment (P $<$ 0.001, Wilcoxon signed rank test). Importantly, this is not the conclusion we would have reached had we relied on a non-process model based statistic. The change in meta-d'/d' was not significant (median = -0.05, P = 0.22; Fig. 1f, top).

Which factors drive metacognitive learning? We hypothesized that metacognitive learning may be more prominent when metacognitive decisions are difficult and are followed by useful feedback. To explore the role of task difficulty, we conducted two orientation discrimination experiments that differed in the number of levels of stimulus reliability. It has previously been shown that this is a key factor that determines metacognitive task difficulty (Boundy-Singer et al., 2023). Expt. 1 involved six levels of stimulus contrast, expt. 2 only two. As predicted, metacognitive learning was substantial in the first experiment (median = -4.90, P $<$ 0.01), but weaker in the second (median = -0.59, P = 0.13), though note that this difference did not reach statistical significance (P = 0.12). Nevertheless, the trend confirms our hypothesis. To explore the role of feedback, we conducted two further experiments with six levels of contrast. Critically, the earned reward was not communicated on every trial but only after a block of 50 trials (expt. 3) or not at all (expt. 4). As expected, metacognitive learning was weak in both experiments that lacked immediate reward updates (expt. 3: median = -1.19, P = 0.03; expt. 4: median = -0.11, P = 0.23). These values differ significantly from expt. 1 (1 vs 3: P = 0.05; 1 vs 4: P = 0.05). Thus, in the absence of immediate feedback about the quality of a confidence assessment, metacognitive learning may not occur. Finally, we hypothesized that metacognitive learning is not specific to orientation discrimination but generalizes to other perceptual decision-making tasks. We conducted two more experiments, in which subjects either judged the orientation of noisy oriented stimuli (expt. 5) or the category of noisy visual textures (expt. 6). As predicted, we found a similar amount of metacognitive learning in both tasks (expt. 5: median -1.19, P = 0.02; expt. 6: median -1.20, P $<$ 0.01; P = 0.36).

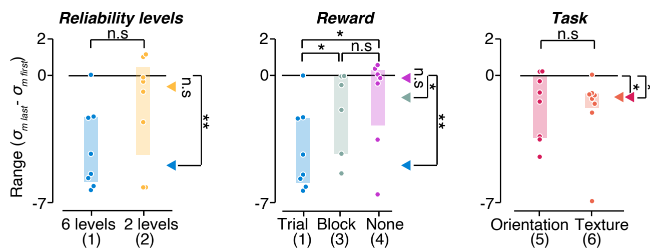

Figure 2: The y-axis is meta-uncertainty range value from a fit of dynamic-CASANDRE minus the same statistic computed on a null-model.

Conclusions

There is a growing interest in understanding the factors that govern metacognition (Rahnev et al., 2022). Previous attempts to study metacognitive learning yielded inconsistent results (Carpenter et al., 2019; Rouy et al., 2022). Our study is the first to show that metacognitive ability is not a static property of human subjects. As subjects accrue more experience in difficult tasks and receive informative feedback signals, the quality of their confidence assessments progressively improves. In this sense, metacognition resembles learning in perception (Goldstone, 1998), cognition (Jaeggi, Buschkuehl, Jonides, & Shah, 2011), and action (Newell, 1991). Our discovery was enabled by the development of a dynamic process model of confidence and a principled recovery analysis. It opens the door to a range of natural follow-up questions. Is the learning specific to the experienced stimulus conditions, or does it generalize to other stimuli, modalities, and tasks? Does metacognitive learning vary across domains, individuals, clinical states, and development? Answering these questions will be made easier based on the principled and proven analysis presented here.

Acknowledgments

This work was supported by the US National Institutes of Health (grant nos. T32 EY021462 to Z.B.S, T32 EY021462 and K99 EY032102 to C.M.Z., and EY032999 to R.L.T.G.).

References

Boundy-Singer, Z. M., Ziemba, C. M., & Goris, R. L. T. (2023, January). Confidence reflects a noisy decision reliability estimate. *Nature Human Behaviour*, *7*(1), 142–154.

Carpenter, J., Sherman, M. T., Kievit, R. A., Seth, A. K., Lau, H., & Fleming, S. M. (2019, January). Domain-general enhancements of metacognitive ability through adaptive training. *Journal of Experimental Psychology. General*, *148*(1), 51–64.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, *49*, 585–612.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011, June). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences*, *108*(25), 10081–10086.

Maniscalco, B., & Lau, H. (2012, March). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, *21*(1), 422–430.

Newell, M. (1991). Motor skill acquisition.

Rahnev, D., Balsdon, T., Charles, L., Gardelle, V. d., Denison, R., Desender, K., . . . Zylberberg, A. (2022, July). Consensus Goals in the Field of Visual Metacognition:. *Perspectives on Psychological Science*.

Rouy, M., de Gardelle, V., Reyes, G., Sackur, J., Vergnaud, J. C., Filevich, E., & Faivre, N. (2022). Metacognitive improvement: Disentangling adaptive training from experimental confounds. *Journal of Experimental Psychology: General*, *151*(9), 2083–2091.