

Metacognition in multialternative choices is based on more information than type-1 decisions.

Nicolás A. Comay^{a+}; Guillermo Solovey^{b*}; Pablo Barttfeld^{a*}

a- Cognitive Science Group. Instituto de Investigaciones Psicológicas (IIPsi, CONICET-UNC), Facultad de Psicología, Universidad Nacional de Córdoba, Córdoba, Argentina. nicocomay@gmail.com

b- Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, UBA-CONICET, Buenos Aires, Argentina

+ *corresponding author*

* *last authors*

Abstract

Humans often face decisions between multiple alternatives. In these contexts, some evidence suggests that only the alternative with highest evidence is represented by the decision system. However, other findings indicate that information from unchosen alternatives remains accessible for decision computations. Furthermore, the amount of information that reaches metacognitive levels remains unexplored. By fitting competing computational models to data from two pre-registered experiments (N=32), we found that noisy evidence from unchosen alternatives reaches the decision system. We extended this by fitting the models to two previously published datasets involving different stimuli (N=17) and found concordance evidence. Moreover, allowing a separate channel of evidence not corrupted by noise for metacognitive computations improves model fitting, which suggests that the metacognitive system has access to more evidence than the decision system. These results act as a bridge between previous conflicting findings in multialternative decisions and extend them by also modelling the metacognitive stage.

Keywords: *perceptual decision making; metacognition; confidence; multialternative decisions; computational modelling.*

Statement of relevance

Understanding how humans decide and evaluate their decisions is a key goal of cognitive psychology. To enhance the comprehension of these processes, there is a growing interest in multialternative decisions, which are thought to reflect a more realistic, real-life scenario. In this realm, there is conflicting evidence regarding whether humans can encompass information from all available alternatives or merely represent the option with highest evidence. Moreover, the amount of information that reaches metacognitive levels remains untested. We therefore compared competing models of these processes in four experiments involving different stimuli and amount of alternatives at play. We found that the metacognitive system can access more information than the decision system, as the evidence that reaches the latter—but not the former—is corrupted by noise. Our results serve to reconcile previous conflicting findings in multialternative decision-making and further expand on them by addressing this issue at the metacognitive stage.

Introduction

Perceptual decision making, i.e. decisions human observers make about sensory information (Hanks & Summerfield, 2017; Heekeren et al., 2008), have been extensively studied through 2-alternative forced choice (2-AFC) tasks. Empirically driven computational models suggest that in these tasks, the observer takes into account the evidence for each of the two competing alternatives and the decision aligns with the alternative with the highest evidence (Shadlen & Kiani, 2013). However, many real-world decisions are not binary. There is a growing interest in understanding the nuances of multi-alternative decisions (Busemeyer et al., 2019; Churchland et al., 2008; Rahnev et al., 2022; Turner et al., 2018), that are thought to reflect more naturalistic scenarios than 2-AFC paradigms (Yeon & Rahnev, 2020).

A key question in multi-alternative decision making is whether humans hold individual representations of all available alternatives—an assumption of virtually all computational models of decision-making—or, alternatively, they create an abbreviated representation, that only encodes the most subjectively salient or valuable stimuli. While it is reasonable that in simple 2-AFC tasks the two competing alternatives can be represented by the decision system, a recent study using multialternative tasks found evidence for a “Summary” model, where only the information from the chosen alternative is included in the decision representation (Yeon & Rahnev, 2020). This result suggests that the decision system does not replicate the sensory information exactly, implying a suboptimal use of information in perceptual decision making. However, another study involving multialternative decisions found evidence that the decision system has access to the evidence from unselected alternatives (McLean et al., 2020)—a result inline with previous multi-alternative studies that modelled the decisions assuming a competition within the decision system among the evidence for each alternative (Busemeyer et al., 2019; Churchland et al., 2008; Dumbalska et al., 2020; Niwa & Ditterich, 2008a; Turner et al., 2018).

In addition, the impact of these two scenarios on metacognition—the ability to monitor our own cognitive processes (Fleming, 2024)—remains unexplored. Traditionally, metacognition is evaluated by asking participants how confident they are in their decision accuracy (Fleming & Lau, 2014). In scenarios where only the chosen alternative's evidence is encoded, this evidence alone might underpin confidence judgments. However, recent studies in multi-alternative decision-making suggest that evidence from unchosen alternatives can also influence this metacognitive feeling of confidence (Comay et al., 2023; Li & Ma, 2020). These results, however, are not entirely inconsistent, as previous metacognition models suggest a separate route of evidence for metacognitive computations (known as “dual-channel” models of metacognition; Mamassian, 2016; Mamassian & De Gardelle, 2022; Maniscalco & Lau, 2016) which could lead to dissociations between the information used by the decision system and by the metacognitive system (Fleming & Dolan, 2012).

To clarify the nature of the multi-alternative decision and metacognitive representations, we carried out two pre-registered experiments that varied the number of alternatives. We put to test three computational models (Figure 1b) with different degrees of information loss. We tested these models against data from our experiments and two previously published datasets involving different stimuli (from McLean et al., 2020 and Yeon & Rahnev, 2020). All experiments employed a “second-guess” paradigm (McLean et al., 2020 and Experiment 3 in Yeon & Rahnev, 2020; Figure 1). In our task, participants had to identify the largest geometrical figure from a set and were given a second opportunity to make a choice if their initial selection was incorrect. If no evidence from initially unselected options is available at the decision stage (i.e., summary encoding), then chance-level performance should be found in these second judgments. If, however, evidence from initially unselected options can access the decision stage, performance should exceed chance levels. In addition, participants were requested to report their confidence in this second judgement, which allowed us to test whether a dissociation between decision and metacognitive levels regarding the amount and type of evidence used for computations was present (Fleming & Dolan, 2012).

To summarise, our results indicate that evidence from unchosen alternatives reaches both the decision-making and metacognitive levels. However, the information accessed at each level is not identical: interestingly, more information reaches the metacognitive system than the decision-making

level. Indeed, we found that behaviour is best captured by a model where the evidence that accesses the decision system for performing second-guesses is corrupted by noise —implying a loss of the initial information which was consistent in all datasets. However, a dual-channel version of this model where the metacognitive channel is not corrupted by this noise fits the data substantially better, meaning that more information reaches the metacognitive stage.

Methods

Experiments were programmed in JavaScript using the library jQuery and ran on a JATOS (Lange et al., 2015) server. The experimental protocol was approved by the ethical committee of the Psychological Research Institute (National University of Córdoba & National Scientific and Technical Research Council – Córdoba, Argentina). Experiments were pre-registered <https://osf.io/9w6ju>.

Participants

18 participants took part in Experiment 1 (13 females; $M_{age} = 24.5$; $SD_{age} = 3.37$) and in Experiment 2 (14 females; $M_{age} = 23.78$; $SD_{age} = 2.94$). Sample size was calculated using the GPower software to reach >80% power to detect a significant difference in the t-tests performed (see the pre-registration at <https://osf.io/9w6ju> for details). Participants read and accepted an informed consent sheet prior to the experiment. All participants reported no psychiatric or neurological history and no chronic consumption of psychoactive substances.

Procedure and experiment design

Data collection was conducted locally in experimental rooms dedicated to this purpose. Experiment 1 (Figure 1) involved a 4-alternative perceptual decision making task. Participants sat 81 centimetres away from the screen, and completed two experimental sessions on different days, each including 10 practice trials and 350 experimental trials. First, participants were presented with a fixation dot displayed on the centre of the screen for 800ms, followed by a stimuli array consisting of geometrical shapes (squares and circles) in a 2x2 grid for 500 ms (stimuli were separated both vertically and horizontally by 6.31° of visual angle). The task was to identify the largest shape among randomly presented squares and circles. Only a single figure was the largest, and the others were equally sized. The largest figure had an equal probability of appearing in any position. The largest figure had a mean size of 1.33° of visual angle, with a standard deviation of 0.33° of visual angle. To pick a figure, participants clicked on the position where the figure they believed was the largest one was present (positions were signalled with small dots after stimuli disappearance). If the decision was correct, the next trial began automatically. Otherwise, they had a second opportunity to choose one of the remaining figures. Chance level for this second decision is therefore defined as a proportion of correct trials of $\frac{1}{3}$, since it is made on 3 alternatives. After each second judgement, participants reported their confidence on being correct on a 4-point scale by clicking on any of four buttons representing the scale. The spanish phrases “nada seguro” and “completamente seguro” (translating to “not sure at all” and “completely sure”, respectively) were displayed below the 1 and 4 buttons, respectively. A 1up/1down staircase was implemented on the first 60 trials of the task to obtain 50% accuracy on the first choice, to ensure a sufficient number of second decisions. The variable controlled by the staircase was the size of the incorrect alternatives, defined as a proportion of the area of the correct alternative. After each incorrect choice the size of the incorrect alternatives decreased by a proportion of the area of the correct alternative equals to .005, whereas after each correct choice this proportion increased by .005. After the first 60 trials, the average size difference from trials 50 to 60 was computed and this was the size difference between correct and incorrect figures used in the rest of the trials. In each session, two rest pauses were included at trials 120 and 240. In those breaks, a message informing participants that they can take a couple of minutes to rest was presented on the screen. Participants clicked anywhere on the screen to continue with the experiment after the rest period.

Experiment 2 (Figure 1) was identical to Experiment 1 but with twelve instead of four alternatives. The procedure was exactly the same as in Experiment 1: participants sat 81cm from the screen and

completed two sessions of 360 trials each in two different days. Figures were displayed on a 3x4 grid, and were separated (vertically and horizontally) by the same degrees of visual angle as in Experiment 1. As a result, the farthest from the centre that an alternative could be was 9.46° horizontally and 6.31° vertically. Chance level for the second guess performance was equal to $\frac{1}{11}$, since 11 alternatives are left for the second judgement.

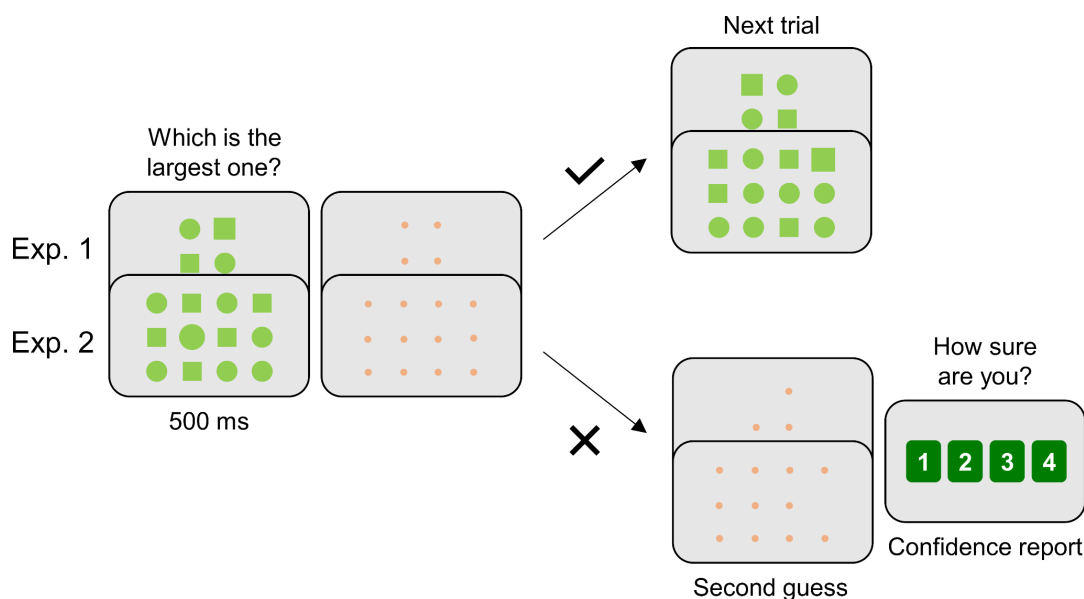


Figure 1 – Size discrimination task. We employed a “second guess” experimental paradigm: experiments consisted of a size discrimination task where participants had to identify the largest geometrical shape within a set of 4 alternatives (Experiment 1) or 12 alternatives (Experiment 2). If the decision was correct, the next trial began automatically. Otherwise, participants had a new chance to choose which figure was the largest (critically, stimuli did not appear again on screen). After this second guess, participants reported their confidence on being correct on a 4-point y scale.

Data analysis

We excluded trials with response times (RT) larger than 8 s or shorter than 150 ms on any of the responses (first and second decision, and confidence report). We discarded the first 60 trials (which included the practice trials and the trials with the staircase). No subject was excluded. Our predefined alpha level was .05.

For each subject, we computed the proportion of correct responses on the first decision (first decision performance), the proportion of correct responses on the second decision (second decision performance) and metacognitive sensitivity on the second decision. We operationalized metacognitive sensitivity as the area under a Receiver Operating Characteristic curve (AUROC-2; Fleming & Lau, 2014).

We used one-tailed t-tests to compare the second decision performance and the metacognitive sensitivity level to chance levels ($\frac{1}{3}$ for second guesses in Experiment 1, $\frac{1}{11}$ for second guesses in Experiment 2, and $\frac{1}{2}$ for metacognition). We also explored, using linear regression models, if second decision performance was predicted by first decision performance and if metacognitive sensitivity was predicted by second decision performance.

All of these analyses and criteria were pre-registered and can be found at <https://osf.io/9w6ju>.

Computational modelling

We fitted the data from our experiments, McLean et al. (2020) data and Yeon & Rahnev (2020) Experiment 3 data (note that we do not include model fitting results from Yeon & Rahnev Experiments 1, 2 and 4 since its design does not allow to recovery the Population+noise model, see the Supplementary material at <https://osf.io/d5qyp/>). In Experiment 3 from Yeon & Rahnev study participants (N=10) had to decide which among six symbols were presented more times in a 7x7 array of symbols presented for 500 ms. In 40% of the wrong responses, a second opportunity to decide was given to the participants. In McLean et al. (2020) tasks, participants (N=7) faced a random dot motion (RDM) task between three different directions and had to choose which direction the majority of the dots moved. Participants were told to report an additional second guess after each decision. Participants could watch the stimuli as long as they wanted up to 5 s. Further details of the experimental designs can be found in the original publications.

We compared three different computational models to evaluate different possible information processing scenarios regarding the evidence available in multi-alternative perceptual decision making. Each model points to a different degree of loss of information. All models start similarly: random samples (one for each alternative) are generated from a Gaussian distribution with mean μ_i and standard deviation σ . While μ_i are fixed, σ is fitted to each participant by maximising its log-likelihood given the data. Note that in our experimental design, μ_i is fixed at 1 for the largest alternative and to the proportion of the area of the largest figure obtained for each subject for the rest of the alternatives. In Yeon & Rahnev (2020) design, μ_i is fixed at 1 for the largest alternative and to .9 for the rest of the alternatives. In McLean et al. (2020) design, μ_i is fixed at the coherence value used in the random dot motion task.

The first decision corresponds to the alternative associated with the highest sample obtained. If the decision is correct, the trial ends. If the decision is incorrect, three possibilities arise, each corresponding to a different model. We next describe each model in detail.

Summary model. The Summary model proposes that the observer arrives at a decision taking into account only the information of the alternative with the highest obtained sample, following the proposal of Yeon and Rahnev (2020). Hence, when prompted for a second decision, the observer has no information about non-chosen alternatives, leading to both a random decision and a random confidence level. For fitting Experiment 3 of Yeon & Rahnev (2020) we also tested the **Summary+strategic choice** model (proposed by Yeon & Rahnev, 2020), where participants remember one random non-chosen symbol from the stimuli array and choose its category in the second guess stage instead of randomly picking any of the initially non-selected alternatives.

Population model. The Population model proposes that at the decision instance the activity of all the alternatives is represented and sustained until the end of the decision process (McLean et al., 2020; Yeon & Rahnev, 2020). Therefore, when prompted for a second decision, the observer will choose the alternative whose associated evidence is maximum out of all previously non selected options (i.e., the second highest sample overall). Confidence on this second decision under this model can have several mappings: it can reflect the level of activation of the alternative chosen (*“Max model”*; Zylberberg et al., 2012), the difference between the two highest activations (*“Balance of evidence model”*; Li & Ma, 2020; Mamassian, 2016), the sum of the differences between the highest activation and the rest (*“Contrast model”*; Comay et al., 2023), and the difference between the activation of chosen alternative and the mean of the rest activations (*“Average-residual model”*; Comay et al., 2023).

Population + noise model. The Population + noise model proposes that, as in the Population model, at the decision instance the activity of all the alternatives is represented. However, random Gaussian noise with mean 0 and standard deviation σ_2 corrupts the samples at the second decision stage. σ_2 is a free

parameter fitted to each subject by maximising its log-likelihood. Confidence has the same possible mappings as in the Population model.

In order to compute metacognitive sensitivity in the size discrimination tasks, in all models a confidence criterion β parameter is fitted to each subject by maximising its log-likelihood. This parameter categorises models' predicted confidence into high and low, and this transformed confidence is used to compute the AUROC-2 predicted by the model.

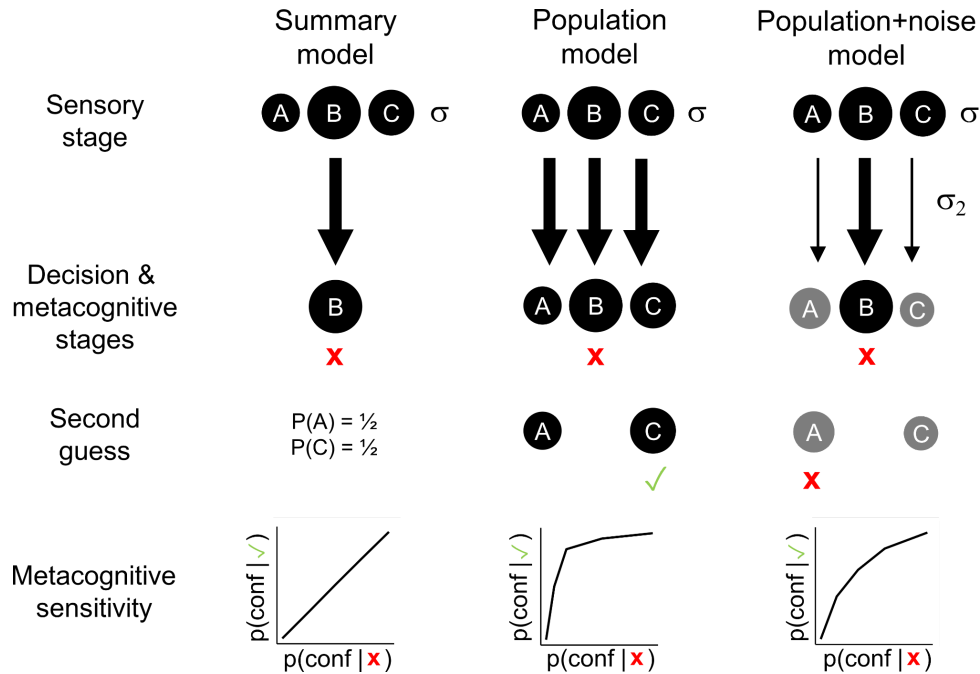


Figure 2 – Computational models and their predictions. We compared three models with different degrees of information loss. Assuming a choice between 3 alternatives (“A”, “B” and “C”) the Summary model proposes that only the evidence of the alternative with highest activation (“B” in the example, as illustrated by its larger size) reaches the decision and metacognitive stages. The Population model proposes that information that reaches the decision and metacognitive systems is an exact copy of the sensory information. The Population+noise model lays on between the two, proposing that noisy evidence (with noise controlled by an extra parameter: σ_2) from initially unchosen options reaches decision and metacognitive systems. According to the Summary model, no evidence from initially unchosen options is present at the decision level, therefore if an initial decision is incorrect and a second judgement is solicited, then decision and metacognitive performance in this second-guess should be at chance level, as no information from initially unchosen options is available. The Population and Population+noise models, however, predict above chance levels in both second decisions and metacognitive performance, as information by unchosen options can reach the decision and metacognitive systems. Moreover, higher second guess performance and metacognitive sensitivity is predicted by the Population model when compared to the Population+noise model, as this model is the model with highest information available.

Model fitting procedure

We fitted the free parameters (σ , β and σ_2 in the case of the Population + noise model) by maximising their log-likelihood. To compute the log-likelihood we simulated 2000 trials and computed the model's probability of being correct on the first and second decisions and the predicted metacognitive sensitivity. We repeated this process 10 times, and calculated the mean of the 10 probabilities of being correct on the first and second decisions to approximate the true probabilities predicted by the model. We computed the probability of the data given the parameters values using the binomial data model and the previously calculated probabilities using the *dbinom* function in R. We followed a similar procedure to fit Yeon & Rahnev (2020) and McLean et al. (2020) data, but simply simulating 30000 trials to approximate the probabilities of being correct predicted by the models. To model metacognitive sensitivity (only in our experiments) we followed a similar approach but using a normal data model: using the *dnorm* function in

R we computed the probability of the metacognitive sensitivity data given the parameters. To calculate the mean and standard deviation of the normal distribution we computed the mean metacognitive sensitivity of the 10 simulations and the standard deviation of those simulated metacognitive sensitivities, respectively.

We started the fitting procedure with a coarse grid search to find sensible initial values for the parameters and then runned 3 gradient descent routines using the *optim* function in R. In order to have different initial values, for each routine we slightly corrupted the initial values of the parameters with Gaussian noise with mean zero and 0.05 standard deviation.

For the Population and the Population + noise model we fitted four different variations of the model, each differing in the confidence mappings. The best fitting variation (i.e., the one with maximum log-likelihood) was selected for comparison between models.

For the dual-channel version of the Population + noise model we simply allowed confidence judgments to be made with the alternatives evidence before being corrupted by σ_2 .

Parameter recovery

We ran a parameter recovery analysis and found that our method was able to recover the true parameters values. We obtained significant Pearson correlations between the true and the recovered parameters of 0.80 ($p < .001$, $CI = [0.71; 0.86]$, $df = 98$) for the σ parameter, to 0.36 ($p < .001$, $CI = [0.17; 0.52]$, $df = 98$) for the σ_2 parameter and to 0.32 ($p = .001$, $CI = [0.14; 0.49]$, $df = 98$) for the β parameter.

Model comparison

We compared the performance of the models by using the Bayesian Information Criterion (BIC). The formula used for the BIC was:

$$BIC_i = k \log(n) - 2 \log(L_i)$$

Where k is the number of free parameters (2 for the Summary and Population model, 3 for the Population + noise model), n is the number of data points (600), L is the likelihood of the parameters given the data and i indexes the subjects. We summed the BIC across subjects.

Results

Behavioural results

In Experiment 1 participants had a mean performance for the first decision of 0.68 ($SD = 0.06$; Figure 3a). Importantly and as predicted in the pre-registration, participants showed above chance second decision performance ($t_{17} = 10.03$, $p < .001$, $d = 2.36$, $M = 0.54$, $CI = [0.51; \text{Inf}]$; Figure 3a) and also above chance metacognitive sensitivity ($t_{17} = 8.92$, $p < .001$, $d = 2.10$, $M = 0.61$, $CI = [0.59; \text{Inf}]$; Figure 5a). We also explored how these variables relate to each other. We found that first decision performance significantly predicted performance in the second decision ($F_{1,16} = 14.5$, $\beta_{\text{first decision}} = 0.95$, $SE = 0.25$, $p = .002$, $R^2 = 0.44$; Figure 3a) and second decision performance significantly predicted metacognitive sensitivity ($F_{1,16} = 9.48$, $\beta_{\text{second guess}} = 0.36$, $SE = 0.12$, $p = .007$, $R^2 = 0.33$; Figure 5a).

In spite of the much larger number of alternatives, similar results were obtained in Experiment 2. Performance on the first decision had a mean of 0.60 ($SD = 0.06$; Figure 3b). Participants again showed above chance second guess performance ($t_{17} = 9.40$, $p < .001$, $d = 2.22$, $M = 0.20$, $CI = [0.18; \text{Inf}]$; Figure 3b) and also above chance metacognitive sensitivity ($t_{17} = 6.89$, $p < .001$, $d = 1.62$, $M = 0.59$, $CI = [0.57; \text{Inf}]$; Figure 5a). Note that, as predicted due to a larger loss of information associated with a larger number of alternatives, effect sizes were smaller on this experiment. Exploring the relationship between the variables, we found that —contrary to Experiment 1 results— performance on the first decision did not predict second decision performance ($F_{1,16} = 2.62$, $\beta_{\text{first decision}} = 0.28$, $SE = 0.17$, $p = 0.12$, $R^2 = 0.09$;

Figure 3b) and performance on the second decision did not predict metacognitive sensitivity ($F_{1,16} = 0.01$, $\beta_{\text{second guess}} = -0.03$, $SE = 0.28$, $p = .92$, $R^2 = -0.06$; Figure 5a).

Computational modelling

Predicting second decision performance

Our model fitting results suggest that while the Summary model underestimated second decision performance, the Population model overestimated it. On the other hand, the Population+noise model could better capture participants' behaviour in the two experiments (Figure 3a and Figure 3b, bottom rows). By comparing the models using the BIC to correct for the number of parameters, we found that the Population model was the best model in Experiment 1 and the Population+noise model was the best model in Experiment 2 (Table 1; although note that the Population model losses against a dual-channel version of the Population+noise model). Importantly, and in line with our prediction that more information will be lost when more alternatives are involved, the Population model was the worst fitting model in Experiment 2.

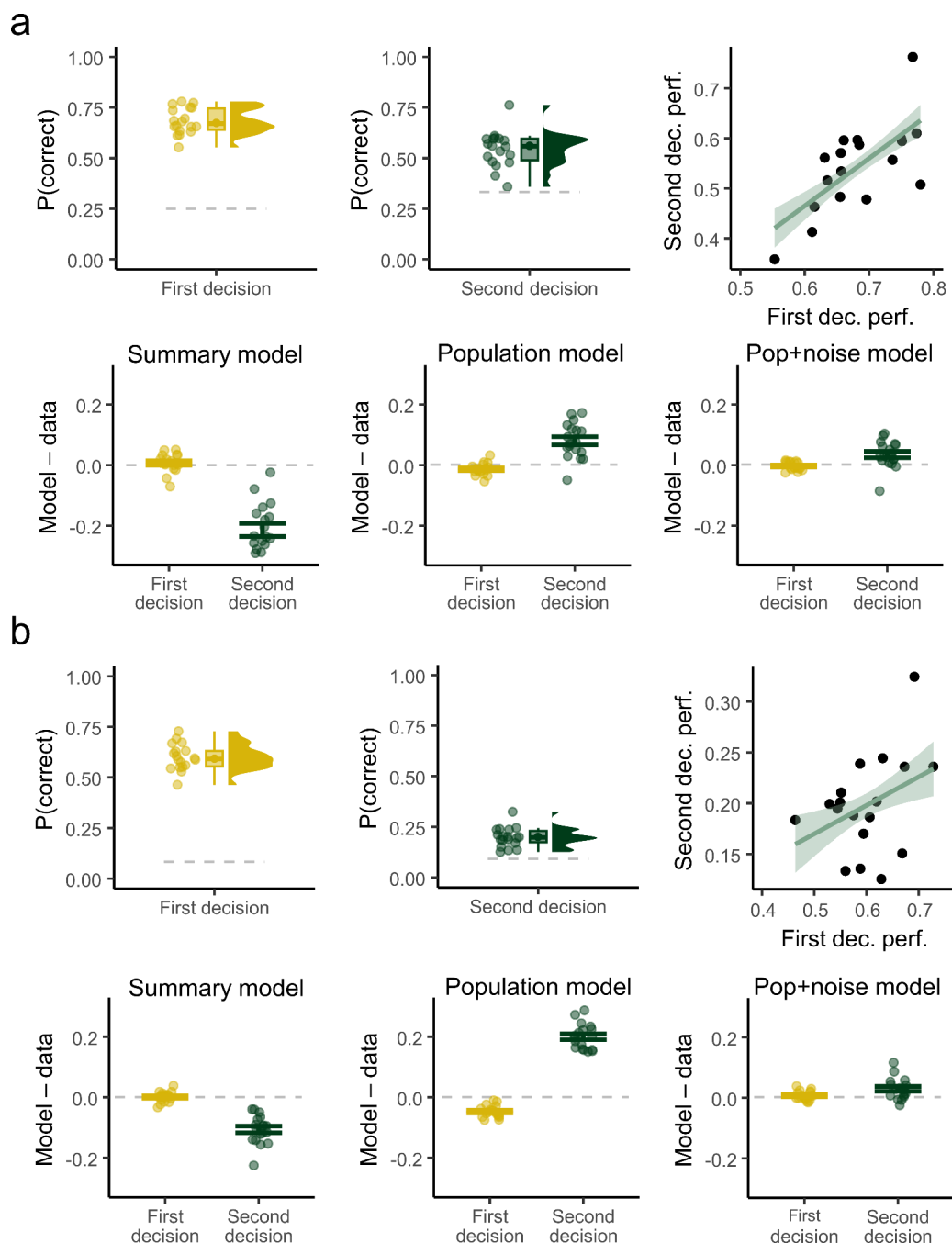


Figure 3 – Size discrimination task. Behavioural and model fitting results. a) Experiment 1 (four

alternatives) results. Participants were above chance in both first and second decisions (chance level is illustrated with the grey dotted line). Moreover, an increase in first decision performance significantly predicted an increase in second decision performance (first row). Regarding models' predictions (second row), while accurately predicting the performance on the first judgement, the Summary model underestimated second decision performance. In contrast, both the Population model and the Population+noise model fitted the data better, with the Population model overestimating more second decision performance. b) Experiment 2 (twelve alternatives) results. Similar to the results on Experiment 1, participants were above chance in both first and second decisions. However, an increase in first decision performance did not predict an increase in second decision performance (first row). With respect to models' predictions, the Summary model again underestimated participants' performance on the second decisions. Illustrating a greater loss of information due to the increased number of alternatives, the Population model was the worst fitting model in this second experiment, underestimating first decision performance and largely overestimating second decision performance. The Population + noise was the best-fitting model, accurately predicting first and second decision performance. In this and all figures dots represent individuals and error bars represent SEM.

We then extended our results by fitting the models to two previously published datasets that also employed a second-guess paradigm (McLean et al., 2020; Yeon & Rahnev, 2020). Regarding McLean et al. (2020) data, we again found that the Population+noise model was the best fitting model (Table 1 and Figure 4a). In Yeon & Rahnev (2020) data, we found that another model with intermediate loss of information, the Summary+strategic choice model, was the best fitting model, closely followed by the Population+noise model (Table 1 and Figure 4b). The pattern of the results was similar to the previous experiments: while the Summary model underestimated second guesses accuracy, the Population model overestimated it. On the other hand, the Population+noise could capture the behaviour of the participants in both first and second decisions. Combining all the datasets together the Population+noise model is selected as the best fitting model (Table 1).

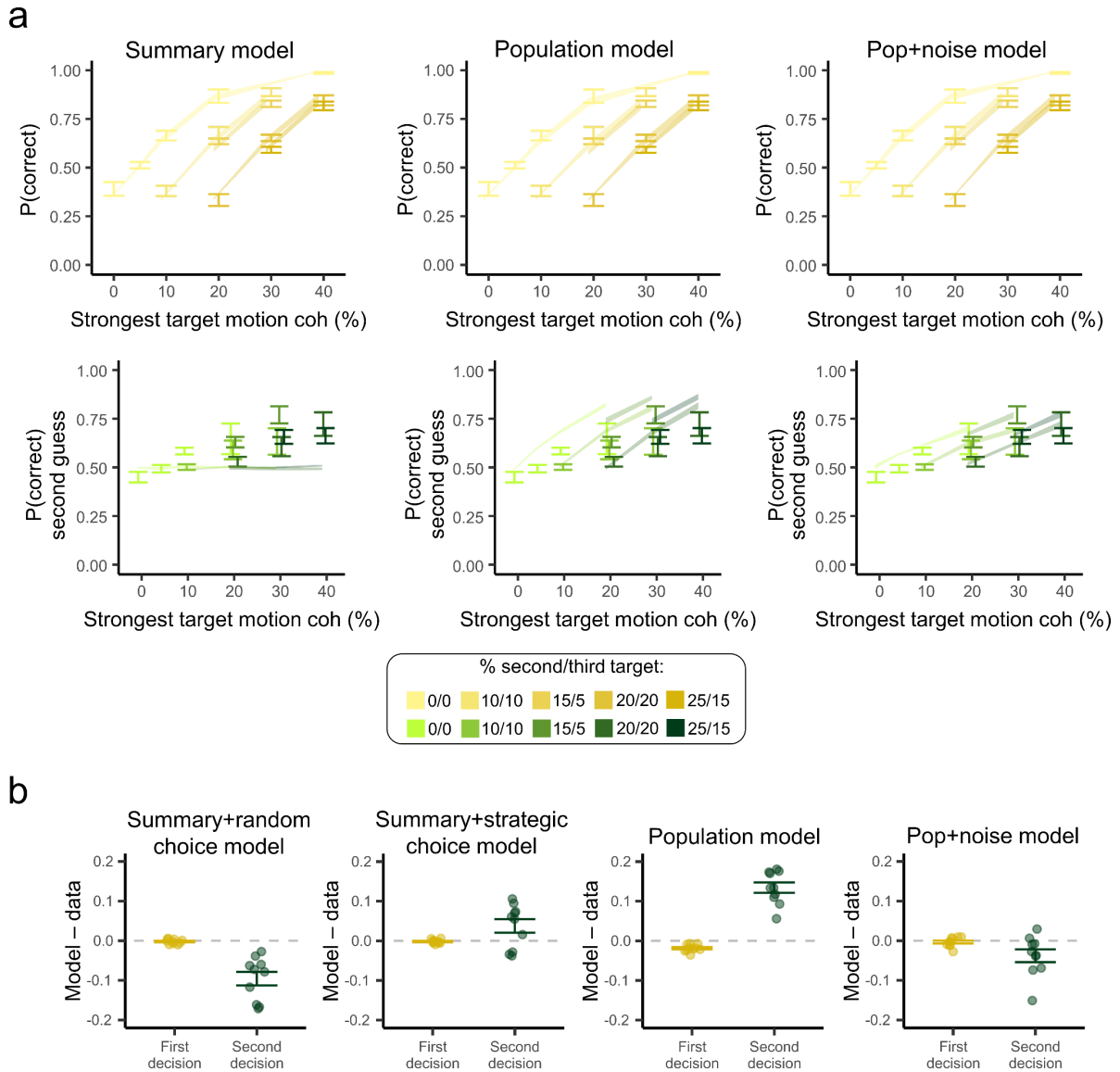


Figure 4 – Model fitting results to (a) McLean et al. (2020) data and (b) Yeon & Rahnev (2020) Experiment 3 data. The three models accurately predicted first decision behaviour. However, the Summary model greatly underestimates second decision performance and the Population model consistently overestimates it. The Population+noise model accurately predicts both first and second decisions. Interestingly, the Summary+strategic choice model was the best-fitting model in the data from Yeon & Rahnev (2020), suggesting that the loss of information is stronger in that experiment. In panel a, shaded regions represent the SEM of models' predictions.

Comparing single-channel and dual-channel models for metacognition

The modelling results suggest that unchosen options evidence accesses the decision system but in a suboptimal manner, as this evidence is corrupted by noise. To evaluate whether the metacognitive system suffers the same loss of information we compared the predictions of a single-channel and a dual-channel version of the Population+noise model. Specifically, in this implementation of the dual-channel model it is assumed that confidence evidence—from which metacognition is computed—has a separate channel of evidence that it is not corrupted by σ_2 . The modelling results suggest that this model architecture fits the data from both experiments significantly better than the single-channel version (Table 1 and Figure 5b).

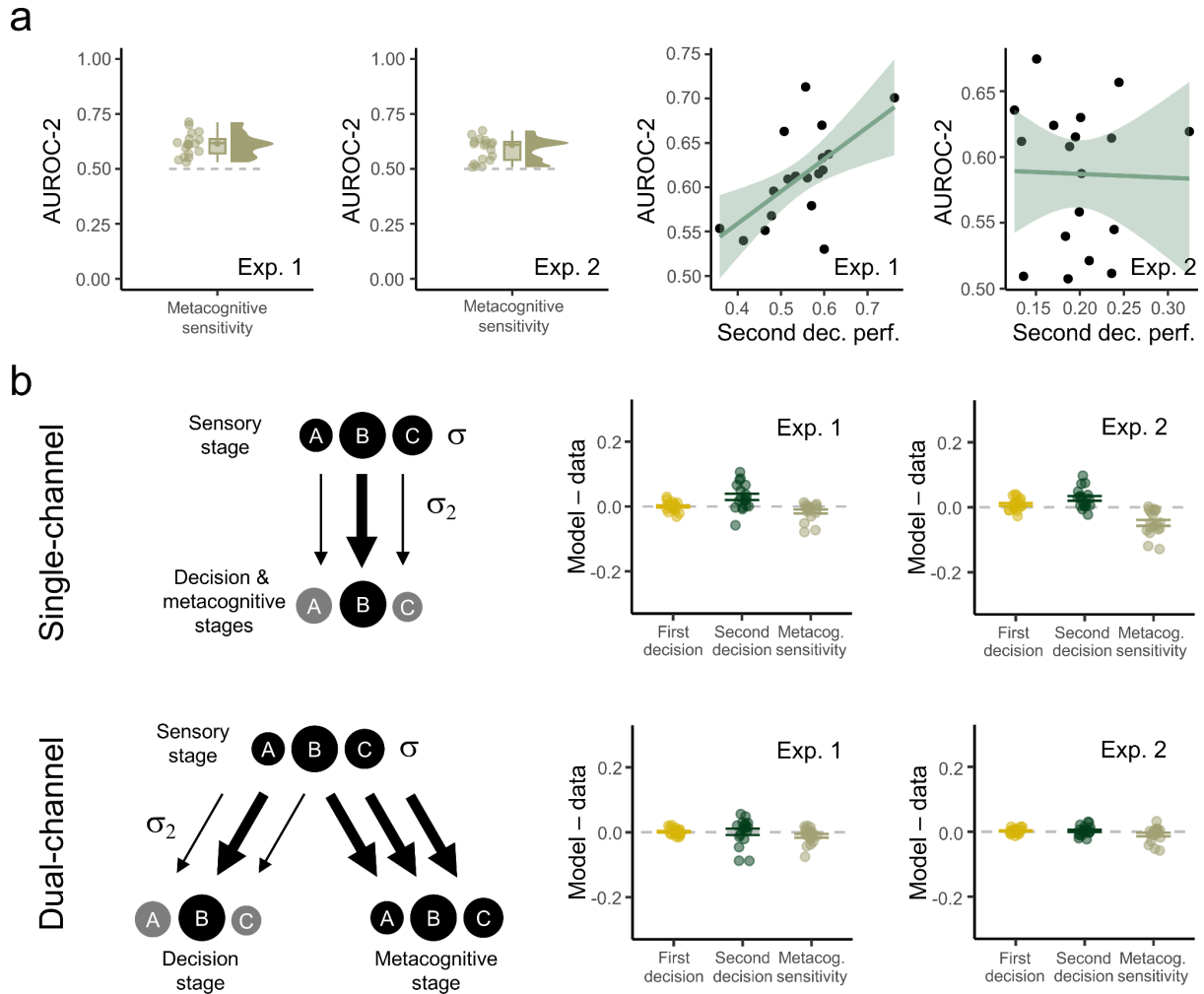


Figure 5 – Behavioural and model fitting results in metacognitive sensitivity. a) Behavioural results. As predicted, metacognitive sensitivity was above chance in both experiments, suggesting that information from initially unchosen options can reach the metacognitive level. Participants with higher second decision performance also had higher metacognition in Experiment 1 but, interestingly, this relationship was not present in Experiment 2. b) A dual-channel (second row) architecture that allows for a separate route of evidence where unchosen options' information is not corrupted by σ_2 fits the data substantially better than a single-channel model (first row). These results are evidence that the metacognitive system can access more information than the decision system. Interestingly, allowing for a separate route of evidence for metacognitive judgments also made the model's predictions regarding second guess performance more accurate, as the model did not need to overestimate accuracy to achieve higher metacognitive levels.

Table 1 – Bayesian Information Criterion model comparison.

	Summary model	Population model	Population+noise model
Experiment 1	1274.89	455.63	494.01 (454.20)
Experiment 2	911.03	1141.99	546.12 (445.09)
McLean et al. (2020)	1333.09	1304.02	1166.84
Yeon & Rahnev (2020) Experiment 3	528.66 (336.14)	652.57	396.02
All data (summed BIC)	4020.67 (3828.15)	3554.21	2602.99 (2462.15)

Notes: the winning model has its BIC value in bold. Values in parentheses under the Summary model reflect the fit of the Summary+strategic choice model. Values in parentheses under the Population+noise model reflect the fit of the dual-channel version.

Discussion

In the present work we have studied the multialternative representations that sustain decision-making and metacognition. The underlying motivations were that contradictory results were found regarding the amount of information available at the decision stage (McLean et al., 2020; Yeon & Rahnev, 2020) and, moreover, there was an open question about whether this—or different—information reaches the metacognitive level.

Our results suggest that human decision-makers can recover information from unchosen alternatives in order to give meaningful second guesses and confidence ratings, even in complex multialternative contexts. Nevertheless, the idea of a decision system carrying on an exact copy of the sensory information is not supported by the data, as the Population model was outperformed by the Population + noise model. This latter model includes extra noise that corrupts the decision representation at the second decision stage, meaning that some—but not all—information gets lost at this level, thus conciliating previously contradictory results. In addition, metacognitive computations seem to access more information than the decision system, as a dual-channel version of the mentioned model fitted the data substantially better. In this implementation, the channel of evidence for confidence judgments did not include extra noise, therefore not underestimating the metacognitive ability of the participants.

This suboptimal use of information by the decision system is found not only in our experiments but also in two previously published datasets involving different stimuli. Importantly, even in not-so-complex decisions between 3 (McLean et al., 2020) or 4 alternatives (our Experiment 1) some information is lost. The findings in McLean et al. (2020) are particularly notable since this experiment not only has only one extra alternative compared to classic 2AFC tasks, but participants could see the stimuli for up to 5 seconds, allowing ample time to retain the sensory evidence. Therefore, this suboptimal use of information in multialternative decisions seems to be a fundamental limitation of the decision system (Yeon & Rahnev, 2020). As Yeon and Rahnev (2020) suggest, an exact copy of the sensory information may be present in more automated processes—such as multisensory integration—or in simple 2AFC tasks, but for multialternative explicit decisions, the decision making system systematically loses available information.

How much information is lost can depend on a multitude of factors, not just the amount of alternatives. In the visual domain, Rosenholtz (2020) indicates that two main factors influence the loss of information in visual perception: the limits that peripheral vision gives for performing certain tasks and the limits of the decision mechanisms that cannot perform arbitrarily complex tasks. For instance, the Exp. 3 from Yeon & Rahnev (2020) study involves more complex stimuli (symbols) whose information can be already lost in peripheral encoding as the distinction of features that requires binding usually needs selective attention (Rosenholtz, 2016, 2020). Moreover, this task has considerable decision complexity as it involves both recognizing each individual stimulus and estimating their frequency while taking into account the frequency of the other stimulus categories. In comparison, one can argue that our size discrimination task is simpler both perceptually (alternatives are distinguishable by their size) and conceptually (only one variable—i.e., the size of alternatives—is relevant for performing successfully). In this sense, further research could take advantage of the extra-noise parameter of the Population + noise model which can be useful to quantify which task factor (e.g. stimulus complexity, number of alternatives, tasks demands, etc.) is inducing more or less information loss in the decision system. This extra-noise parameter can also be useful to quantify individual differences in the capacity limits of the decision system.

One might argue that this second guess design is not a fair test for the Summary model, as participants can in principle quickly make two decisions and, in case of being asked for a second-guess, report their second choice. While this cognitive strategy is in principle possible, model recovery analysis suggests that this design was the best to capture the hybrid possibility proposed by the Population+noise model (see Supplementary material). Despite this cognitive strategy possibility, our results still point out that this second judgement is performed suboptimally—in agreement with the Summary model.

Traditionally, perceptual decision making and metacognition have been studied using 2AFC tasks. The computational models developed under this approach assume that the information from the competing alternatives is represented by the observer, and the decision is made by judging the relative evidence for each alternative (Shadlen & Kiani, 2013). How do our results impact this notion? Our results suggest that this assumption extends to multialternative decisions, as even in a 12-alternative decision making task human observers can, although suboptimally, recover information from unchosen options. The fact that—although noisy—information from unchosen alternatives accesses the decision stage is also inline with previous multialternative work (Churchland et al., 2008; Li & Ma, 2020; McLean et al., 2020; Niwa & Ditterich, 2008) and with several contextual effects that arise in multi-alternative decision making tasks both in decisions (Busemeyer et al., 2019; Trueblood et al., 2013) and in confidence judgments (Comay et al., 2023). As most computational accounts of these empirical results rely on the assumption that information of all alternatives is available and combined in a specific way (Busemeyer et al., 2019; Comay et al., 2023; Dumbalska et al., 2020; Li & Ma, 2020; Niwa & Ditterich, 2008; Turner et al., 2018), the evidence reported here is then critical to sustain those explanations.

A related current discussion is whether the same circuits that encode information for choice also, and at the same time, encode information that is read-out to convey confidence judgments. The evidence in favour of circuit sharing mostly comes from experiments in macaques that show that the same neurons that are thought to encode choice formation also, and at the same time, affect confidence (Kiani & Shadlen, 2009; Kiani et al., 2014). However, if confidence is merely encoded in the same circuitry as choice it is unclear how there are several experiments that show clear dissociations between decision accuracy and confidence (Graziano & Sigman, 2009; Graziano et al., 2015; Maniscalco & Lau, 2016; Zylberberg et al., 2014). In this sense, some computational models explicitly suggest a separate line of evidence for metacognitive judgments such as confidence ratings (Mamassian, 2016; Mamassian & De Gardelle, 2022) and others propose hierarchical structures where “type 2” judgments evaluate the quality of the “type 1” (e.g. decision) information (Maniscalco & Lau, 2016). Our results are inline with these views as the single-channel version of the Population+noise model underestimated the metacognitive sensitivity found in our data, specially in Experiment 2. As this model implementation only has access to the information of the type 1 judgement for computing metacognitive sensitivity, an upper bound for the predicted metacognitive ability is then established. Consequently, the higher metacognitive levels found compared to the model's predictions suggest that confidence judgments had extra or different information than the one that supports type 1 judgments, resulting in a boosted metacognitive sensitivity. This is sustained by the fact that a dual-channel version of the Population+noise model fits the data significantly better and does not underestimate metacognition. In other words, our results suggest that metacognitive computations have access to extra information than the decision computations, and as this information is not corrupted by noise, it can be considered that the metacognitive system is less suboptimal than the decision system in multialternative choices.

In line with this idea, on Experiment 2 there was not an association between performance in the second decision and metacognition, suggesting that participants that lost more sensory information to inform their decisions nevertheless had information to inform their metacognitive judgments. One limitation of this finding is that the AUROC-2 measure can be affected by task performance (Fleming & Lau, 2014; Maniscalco & Lau, 2012). Unfortunately, no alternative method has been developed yet to address metacognition independently of performance in multialternative decision tasks, and even methods that seem promising in controlling for performance confounds in 2-AFC tasks have been shown to fail with respect to that aim (Rahnev, 2023). Further research is needed to precisely evaluate metacognition independently of performance in multialternative decisions.

In conclusion, here we found consistent evidence for a suboptimal use of unchosen options' information by the decision system in multi-alternative decisions. Moreover, the metacognitive system appears to access more information than the decision system, as a model without noise corrupting the evidence that reaches this stage fitted the data significantly better. Our results serve to reconcile previous contradictory

findings in multi-alternative decisions by proposing a model with an intermediate degree of loss of information and further expand on them by incorporating a model for the metacognitive stage.

Acknowledgements

The authors want to thank Dr. Jochen Ditterich for kindly sharing the dataset from McLean et al. (2020) study. N. A. Comay has a PhD scholarship from the National Scientific and Technical Research Council of Argentina. The work was sustained with two grants from the National Scientific and Technical Research Council of Argentina (Grant #2018-03614 and Grant #2021-0083) and a grant from the University of Buenos Aires (UBACyT 20020220400271BA).

Open data and materials

Data & scripts that reproduce the reported results are available at: <https://osf.io/d5qyp/>.

References

- Busemeyer, J. R., Gluth, S., Rieskamp, J., & Turner, B. M. (2019). Cognitive and Neural Bases of Multi-Attribute, Multi-Alternative, Value-based Decisions. *Trends in Cognitive Sciences*, 23(3), 251–263. <https://doi.org/10.1016/j.tics.2018.12.003>
- Churchland, A. K., Kiani, R., & Shadlen, M. N. (2008). Decision-making with multiple alternatives. *Nature Neuroscience*, 11(6), 693–702. <https://doi.org/10.1038/nn.2123>
- Comay, N. A., Della Bella, G., Lamberti, P., Sigman, M., Solovey, G., & Barttfeld, P. (2023). The presence of irrelevant alternatives paradoxically increases confidence in perceptual decisions. *Cognition*, 234, 105377. <https://doi.org/10.1016/j.cognition.2023.105377>
- Dumbalska, T., Li, V., Tsetsos, K., & Summerfield, C. (2020). A map of decoy influence in human multialternative choice. *Proceedings of the National Academy of Sciences*, 117(40), 25169–25178. <https://doi.org/10.1073/pnas.2005058117>
- Fleming, S. M. (2024). Metacognition and Confidence: A Review and Synthesis. *Annual Review of Psychology*, 75(1), annurev-psych-022423-032425. <https://doi.org/10.1146/annurev-psych-022423-032425>
- Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1594), 1338–1349. <https://doi.org/10.1098/rstb.2011.0417>
- Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, 8. <https://doi.org/10.3389/fnhum.2014.00443>
- Graziano, M., & Sigman, M. (2009). The spatial and temporal construction of confidence in the visual scene. *PLoS One*, 4(3), e4909.

- Graziano, M., Parra, L. C., & Sigman, M. (2015). Neural correlates of perceived confidence in a partial report paradigm. *Journal of Cognitive Neuroscience*, 27(6), 1090-1103.
- Hanks, T. D., & Summerfield, C. (2017). Perceptual Decision Making in Rodents, Monkeys, and Humans. *Neuron*, 93(1), 15–31. <https://doi.org/10.1016/j.neuron.2016.12.003>
- Heekeren, H. R., Marrett, S., & Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nature Reviews Neuroscience*, 9(6), 467–479.
<https://doi.org/10.1038/nrn2374>
- Kiani, R. & Shadlen, M. N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. *Science*, 324(5928), 759-764.
- Kiani, R., Corthell, L., & Shadlen, M. N. (2014). Choice certainty is informed by both evidence and decision time. *Neuron*, 84(6), 1329-1342.
- Lange, K., Kühn, S., & Filevich, E. (2015). "Just Another Tool for Online Studies" (JATOS): An Easy Solution for Setup and Management of Web Servers Supporting Online Studies. *PLOS ONE*, 10(6), e0130834. <https://doi.org/10.1371/journal.pone.0130834>
- Li, H.-H., & Ma, W. J. (2020). Confidence reports in decision-making with multiple alternatives violate the Bayesian confidence hypothesis. *Nature Communications*, 11(1), 2004.
<https://doi.org/10.1038/s41467-020-15581-6>
- Mamassian, P. (2016). Visual Confidence. *Annual Review of Vision Science*, 2(1), 459–481.
<https://doi.org/10.1146/annurev-vision-111815-114630>
- Mamassian, P., & De Gardelle, V. (2022). Modeling perceptual confidence and the confidence forced-choice paradigm. *Psychological Review*, 129(5), 976–998.
<https://doi.org/10.1037/rev0000312>
- Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, 21(1), 422–430.
<https://doi.org/10.1016/j.concog.2011.09.021>
- Maniscalco, B., & Lau, H. (2016). The signal processing architecture underlying subjective reports of sensory awareness. *Neuroscience of Consciousness*, 2016(1). <https://doi.org/10.1093/nc/niw002>
- Maniscalco, B., Peters, M. A. K., & Lau, H. (2016). Heuristic use of perceptual evidence leads to dissociation between performance and metacognitive sensitivity. *Attention, Perception, & Psychophysics*, 78(3), 923–937. <https://doi.org/10.3758/s13414-016-1059-x>

- McLean, C. S., Ouyang, B., & Ditterich, J. (2020). Second Guessing in Perceptual Decision-Making. *The Journal of Neuroscience*, 40(26), 5078–5089. <https://doi.org/10.1523/JNEUROSCI.2787-19.2020>
- Niwa, M., & Ditterich, J. (2008b). Perceptual Decisions between Multiple Directions of Visual Motion. *The Journal of Neuroscience*, 28(17), 4435–4445. <https://doi.org/10.1523/JNEUROSCI.5564-07.2008>
- Rahnev, D. (2023). *Measuring metacognition: A comprehensive assessment of current methods*. PsyArXiv. <https://doi.org/10.31234/osf.io/waz9h>
- Rahnev, D., Balsdon, T., Charles, L., De Gardelle, V., Denison, R., Desender, K., Faivre, N., Filevich, E., Fleming, S. M., Jehee, J., Lau, H., Lee, A. L. F., Locke, S. M., Mamassian, P., Odegaard, B., Peters, M., Reyes, G., Rouault, M., Sackur, J., ... Zylberberg, A. (2022). Consensus Goals in the Field of Visual Metacognition. *Perspectives on Psychological Science*, 17(6), 1746–1765. <https://doi.org/10.1177/17456916221075615>
- Rosenholtz, R. (2016). Capabilities and Limitations of Peripheral Vision. *Annual Review of Vision Science*, 2(1), 437–457. <https://doi.org/10.1146/annurev-vision-082114-035733>
- Rosenholtz, R. (2020). Demystifying visual awareness: Peripheral encoding plus limited decision complexity resolve the paradox of rich visual experience and curious perceptual failures. *Attention, Perception, & Psychophysics*, 82(3), 901–925. <https://doi.org/10.3758/s13414-019-01968-1>
- Shadlen, M. N., & Kiani, R. (2013). Decision Making as a Window on Cognition. *Neuron*, 80(3), 791–806. <https://doi.org/10.1016/j.neuron.2013.10.047>
- Trueblood, J. S., Brown, S. D., Heathcote, A., & Busemeyer, J. R. (2013). Not Just for Consumers: Context Effects Are Fundamental to Decision Making. *Psychological Science*, 24(6), 901–908. <https://doi.org/10.1177/0956797612464241>
- Turner, B. M., Schley, D. R., Muller, C., & Tsetsos, K. (2018). Competing theories of multialternative, multiattribute preferential choice. *Psychological Review*, 125(3), 329–362. <https://doi.org/10.1037/rev0000089>
- Yeon, J., & Rahnev, D. (2020). The suboptimality of perceptual decision making with multiple alternatives. *Nature Communications*, 11(1), 3857. <https://doi.org/10.1038/s41467-020-17661-z>
- Zylberberg, A., Barttfeld, P., & Sigman, M. (2012). The construction of confidence in a perceptual decision. *Frontiers in Integrative Neuroscience*, 6. <https://doi.org/10.3389/fnint.2012.00079>
- Zylberberg, A., Roelfsema, P. R., & Sigman, M. (2014). Variance misperception explains illusions of

confidence in simple perceptual decisions. *Consciousness and cognition*, 27, 246-253.