Efficient Neural Compression of Sensory Information during Human Category-Based Decisions

Julie Drevet (julie.drevet@inserm.fr)
Laboratoire de Neurosciences Cognitives et Computationnelles,
Institut National de la Santé et de la Recherche Médicale,
Département d'Études Cognitives, École Normale Supérieure, Université PSL,
Paris 75005, France

Jan Drugowitsch (Jan_Drugowitsch@hms.harvard.edu)
Department of Neurobiology,
Harvard Medical School,
Boston, MA, USA

Valentin Wyart (valentin.wyart@ens.psl.eu)
Laboratoire de Neurosciences Cognitives et Computationnelles,
Institut National de la Santé et de la Recherche Médicale,
Département d'Études Cognitives, École Normale Supérieure, Université PSL,
Paris 75005, France

Abstract:

Human visual categorization relies on an inference process that extracts the statistics of ambiguous sensory observations through imprecise computations. But theories diverge regarding whether this imprecise inference process integrates sensory information in its native stimulus space or in a compressed space centered on decision-relevant categories. Here we designed a visual categorization task in which we manipulated the space in which human observers can perform inferences. We found that humans perform more accurate inferences when integrating sensory information in category space. Concurrent magnetoencephalographic recordings showed accuracy-predictive signatures of compressed neural representations of sensory information in conditions where humans can perform inferences in category space. Together, these findings indicate that humans mitigate the costs of imprecise inferences by focusing limited computational resources on decision-relevant information.

Keywords: human; inference; visual categorization; magnetoencephalography; computational model

Introduction

Making efficient categorical decisions based on uncertain sensory information requires accumulating information over time before committing to a decision, a process described in normative terms as statistical inference (Bogacz et al., 2006). It has recently been shown that human statistical inference is inherently variable (Drugowitsch et al., 2016). To better understand how human decision-makers handle this computation noise, we examined how sensory information is represented during the inference process, reasoning that computation noise might constrain the representation format. Surprisingly, this format has long remained under-specified: some theories propose an implementation by sensory cortices in the 'native' sensory format that represents stimulus features (Denève Latham, & Pouget, 2001; Jazayeri & Movshon, 2006), whereas other theoretical accounts propose an 'output' format that encodes the action plan resulting from the decision (Shadlen et al., 2008). Here, we propose that humans preferentially represent sensory information in an abstract 'category' space defined by current decision alternatives, allowing to accumulate only the evidence that is relevant for the upcoming decision.

Methods and Results

Task. We designed a probabilistic cue combination task causally manipulating the ability of human observers to represent sensory information in its native stimulus space or in a compressed category space. Participants were asked to categorize sequences of oriented bars on a gray background (Figure 1a). The orientation sequences were drawn from one of two overlapping probability distributions centered on orthogonal orientations (blue or orange, Figure 1b). Participants had to accumulate the evidence provided by each stimulus to infer its generative category. We manipulated how this evidence could be represented in two fully matched conditions. POST: participants had to track the sequence mean orientation and compare it to the two choice alternatives according to a category-mapping cue presented after the sequence (accumulation in stimulus space); PRE: the category-mapping cue was presented before the sequence, participants had therefore the possibility to relate the orientation of each stimulus to the two choice alternatives (accumulation in category space), or they could disregard the category-mapping cue until the end of the sequence, and track the sequence mean orientation (as in POST: accumulation in stimulus space). While normative models predict the exact same behavior in both conditions, these two modes of evidence accumulation take place in different representation spaces, requiring distinct computations, that might be differentially affected by inference noise.

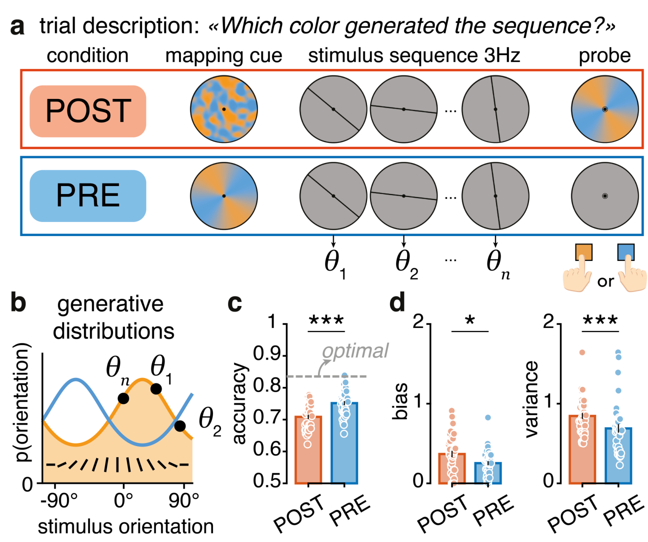

Figure 1: *Task design and human performance. (a) Trial description. (b) Category distributions. (c) Decision accuracy. (d) Bias-variance decomposition.*

Behavioral results. Participants ($N = 37$) achieved better accuracy when the categories were presented before accumulating the evidence (Figure 1c; POST: 0.709; PRE: 0.752; $p < 0.001$), suggesting a distinct accumulation strategy in PRE by making use of the category-mapping cue. They were suboptimal in both conditions (optimal: 0.835) and we partitioned choice errors through a ‘bias-variance trade-off’ decomposition method (Drugowitsch et al., 2016; Wyart, 2018). Participants were more biased and more variable in POST (Figure 1d; all $p < 0.001$), indicating that knowledge of categories before accumulation reduced both systematic and random errors during inference.

Computational modeling. We derived a computational model (Figure 2a) implementing both deviations from Bayes-optimality observed in behavior – systematic errors governed by a prior imbalance parameter γ and random errors governed by an inference noise parameter σ_{inf}. Selection noise σ_{sel} was not considered as source of suboptimal choices after having been ruled out by a factorized Bayesian model selection (Figure 2b). Model and parameter recovery attested the validity of our model. The fitted parameters (Figure 2c) showed that participants had a stronger underweighting of previously accumulated evidence ($p < 0.001$) and a larger spread of inference noise ($p < 0.05$) in POST than in PRE, confirming that inference is more suboptimal when performed in sensory space.

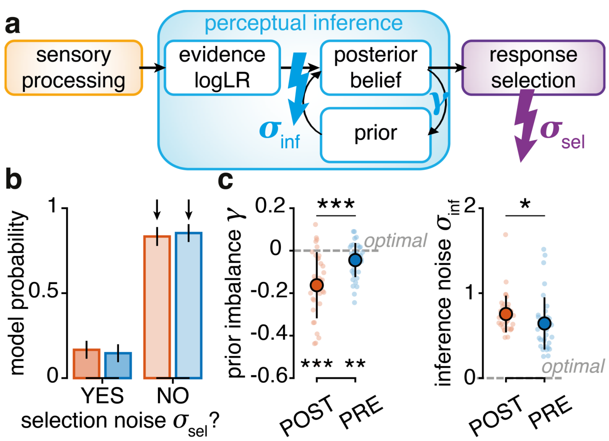

***Figure 2:** Suboptimal Bayesian inference model. (a) Suboptimal inference model. (b) Factorized Bayesian exclusion of selection noise. (c) Best-fitting parameters*

Neural correlates of compression in category space. We used a 306-channel whole-head Elekta Neuromag TRIUX system to record participants’ MEG activity while performing the task. We aligned the MEG signals to the stimulus onset and applied a regularized (ridge regression) approach (Weiss et al., 2021) to ‘decode’ the dynamic MEG patterns associated with the accumulation of sensory evidence in category space. First, we could decode the category information (log-likelihood) provided by individual stimuli in PRE. Importantly, the precision of this neural representation was predictive of participants’ accuracy (Figure 3a). To establish that it is used for evidence accumulation, we decoded the consistency between this category information and the previously accumulated (ongoing) belief. We found that participants selectively ‘compress’ the information conveyed by the stimulus, by amplifying relevant and suppressing irrelevant information, before comparing it to their ongoing belief. The more this compression was reflected in neural activity, the more accurate participants’ choices were (Figure 3b left). Finally, interindividual differences in this neural compression correlated with interindividual differences in inference noise measured from behavior: the stronger the compression in category space, the more precise the resulting inferences (Figure 3b right; $r^2 = 0.476$, $p < 0.001$). Together, our results suggest that compressing information during category-based inferences constitutes an efficient strategy to compensate for the costs of imprecise inferences, by focusing limited resources on decision-relevant information.

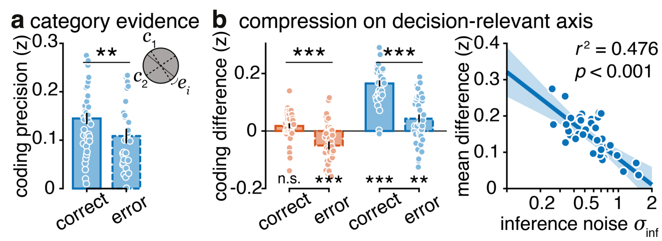

***Figure 3:** Neural signatures of efficient compression. (a) Category evidence. (b) Compression on decision-relevant axis (coding difference relevant vs irrelevant)*

Discussion

We provide converging behavioral and neural evidence that the compression of evidence in category space is an efficient (and preferred) human strategy for evidence accumulation of even low-level sensory information (orientation). But does this compression reflect the by-product of the adaptation to other (real-life) conditions that are not captured by our laboratory task, or a genuinely efficient strategy in the 2AFC categorization task tested here? An idea could be to train recurrent neural networks (RNNs) on the PRE and POST conditions, to see whether they develop the same compression of evidence when trained in the PRE condition. Recent studies have started to use RNNs in this way to better understand the nature of animal and human generalization abilities (Findling & Wyart, 2020) and decision idiosyncrasies (Molano-Mazón et al., 2023).

Acknowledgments

This work was supported by the European Research Council (starting grant No. ERC-StG-759341 to V.W.), the National Institute of Mental Health and the Agence Nationale de la Recherche (US–France collaborative research grant No. 1R01MH115554-01 / ANR-17-NEUC-0001-02 to J.Drugowitsch and V.W), and an institutional grant awarded to the Department of Cognitive Sciences of the Ecole Normale Supérieure (No. ANR-17-EURE-0017).

References

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*, 700–765.

Denève,S., Latham, P.E., & Pouget, A. (2001). Efficient computation and cue integration with noisy population codes. *Nature Neuroscience*, *4*, 826–831.

Drugowitsch, J., Wyart, V., Devauchelle, A.-D., & Koechlin, E. (2016). Computational precision of mental inference as critical source of human choice suboptimality. *Neuron*, *92*, 1398–1411.

Findling, C., & Wyart, V. (2020). Computation noise promotes cognitive resilience to adverse conditions during decision-making. *bioRxiv*.

Jazayeri, M., & Movshon, J. A. (2006). Optimal representation of sensory information by neural populations. *Nature Neuroscience*, *9*, 690–696.

Molano-Mazón, M., Shao, Y., Duque, D., Yang, G. R., Ostojic, S., & de la Rocha, J. (2023). Recurrent networks endowed with structural priors explain suboptimal animal behavior. *Current Biology*, *33*, 622–638.e7.

Shadlen, M. N., Kiani, R., Hanks, T. D., & Churchland, A. K. (2008). Neurobiology of decision making: An intentional framework. In C. Engel & W. Singer (Eds.), *Better than conscious?* Cambridge: MIT Press.

Weiss, A., Chambon, V., Lee, J. K., Drugowitsch, J., & Wyart, V. (2021). Interacting with volatile environments stabilizes hidden-state inference and its brain signatures. *Nature Communications*, *12*, 2228.

Wyart, V. (2018). Leveraging decision consistency to decompose suboptimality in terms of its ultimate predictability. *The Behavioral and Brain Sciences*, *41*, e248.