01. Falk Lieder and Thomas L. Griffiths

02. Abstract word count: **60**; Main text word count: **981**; References word count: **369**; Entire text (total + addresses etc.) word count: **1459**

03. Sampling as a resource-rational constraint

04. Adam N. Sanborn; Jianqiao Zhu; Jake Spicer; Nick Chater

05. University of Warwick; University of Warwick; University of Warwick; Warwick Business School

06. University of Warwick, CV4 7AL, UK; Warwick Business School, University of Warwick, CV4 7AL, UK

07. Sanborn: +44 24 761 51354

08. a.n.sanborn@warwick.ac.uk; J.Zhu@warwick.ac.uk; j.spicer@warwick.ac.uk; Nick.Chater@wbs.ac.uk

09. https://warwick.ac.uk/fac/sci/psych/people/asanborn/; https://warwick.ac.uk/fac/sci/psych/people/zjianqiao/; n/a; https://www.wbs.ac.uk/about/person/nick-chater/

10. Resource rationality is useful for choosing between models with the same cognitive constraints but cannot settle fundamental disagreements about what those constraints are. We argue that *sampling* is an especially compelling constraint, as optimizing accumulation of evidence or hypotheses minimizes the cost of time, and there are well-established models for doing so which have had tremendous success explaining human behavior.

11. In the target article, the case for resource rational analyses is made in general terms: it is a widely-applicable method for identifying how to best use cognitive resources given a set of cognitive constraints, and the long list of successes of this approach shows how resource rational analyses explain a wide range of behavior. We are sympathetic to the overall thrust of the article, and particularly the argument that resource rational analyses are useful for choosing between models with common cognitive constraints. Resource rationality provides a principled method for identifying how cognitive resources are used to solve tasks while assisting in identifying the important cognitive constraints.

But a key challenge for resource rational analyses, which was highlighted in the target article, is identifying what the key cognitive constraints are. The long list of success in the target article is a heterogeneous one – it is comprised of many different approaches that are responding to different cognitive constraints, including neural constraints, representational constraints, time constraints, and attentional constraints, amongst others.

Researchers have tended to focus on a single constraint, rather than looking at them jointly. And indeed, different constraints do not necessarily all sit comfortably with one another, nor are they jointly necessary to explain behavioral biases. For concreteness, we focus on one of the topics discussed in the target article: biases in human probability judgments (Tversky & Kahneman, 1974).

Several explanations have been advanced for these biases which appeal to resource rationality as a justification. One of the most influential explanations is that these biases are the result of estimating the probability of complex events (i.e., conjunctions and disjunctions of events) by averaging individual event probabilities together, rather than combining them correctly (Fantino, Kulik, Stolarz-Fantino, & Wright, 1997). A resource rational justification for averaging is that it is more accurate in the presence of internal or external noise than the correct combination rule (Juslin, Nilsson, & Winman, 2009).

Models based on quantum probability have also been used to explain these behavioral biases, and make predictions that are similar to those of averaging. However, the underlying mechanisms of these models are very different from averaging, and also have a different resource rational justification: instead of appealing to robustness to noise, they are justified as conserving representational resources (Busemeyer, Pothos, Franco, & Trueblood, 2011).

The third approach is covered in the target article: that people do not have access to their subjective probabilities, but are able to generate samples of events from either memory or an internal model. After an infinite number of samples, people could in principle recover their subjective probabilities exactly; but sampling is slow and effortful. With small samples, biases are introduced according to where sampling begins and by how small samples are converted into estimates. The resource rational justification here is that generating samples takes time and effort – people make judgments and decisions with a small number of samples to optimally allocate time between different opportunities and challenges (Dasgupta, Schulz, & Gershman, 2017; Sanborn & Chater, 2016; Zhu, Sanborn, & Chater, 2018a).

These three explanations appeal to very different, and likely mutually exclusive, cognitive constraints. As a result, resource rationality cannot be used to directly adjudicate between them. The best way to do so remains designing clever experiments for which the candidate models make different predictions. However, we argue that because resource rationality is part of the argument for each of these explanations, it is still useful to evaluate how compelling the cognitive constraints are and how well resource rationality has been applied.

We believe the cognitive constraint of sampling (in a broad sense, e.g., generating evidence or hypotheses in proportion to underlying probabilities) is especially compelling, as it is well-established both theoretically and empirically. In many contexts, the sampling process is slow and serial (Maylor, Chater, & Jones, 2001), and therefore it is clearly important to optimize this time cost. Resource rationality is a starting point for many models using sampling: how to optimally stop sampling is well established, both for accumulating until a target confidence is reached and for stopping as soon as the expected cost exceeds the expected gain (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Vul, Goodman, Griffiths, & Tenenbaum, 2014; Wald, 1950). Models based on sequential sampling and optimal stopping have been extremely successful in both perceptual decision-making, and in wider forms of decision making (Ratcliff & McKoon, 2008; Shadlen & Shohamy, 2016). Indeed, sampling limitations underlie other examples discussed in the target article: why people probability match, and why very good and very bad events are over-weighted.

Other constraints, such as representational or process noise constraints, are less well-attested and their consequences less clear cut. For example, applying representational constraints require first establishing what the representations are, and the nature of cognitive representations is often controversial (Spicer & Sanborn, 2019). Internal noise is commonly used as a constraint - and indeed individual neurons are noisy - but in aggregate this noise may be less important than it seems (Beck, Ma, Pitkow, Latham, & Pouget, 2012), and its consequences again depend on the form of the representation. While some aspects of the sampling process do also depend on the representation (Dasgupta, et al., 2017; Zhu, Sanborn, & Chater, 2018b), the fundamental goal of minimizing the number of samples remains.

Finally, beyond its usefulness as a cognitive constraint, sampling also satisfies other desiderata of the resource rationality approach. As resource rational analyses start from formulating a computational solution to a problem, sampling from the posterior is a useful algorithmic constraint to consider, because samplers are general algorithms for performing inference. Sampling models also have a clear connection to AI and statistics, where these methods are widely used in Bayesian inference, and as a result can ease transfer of knowledge between these fields and the cognitive and brain sciences. For these reasons and those above, sampling is a very compelling cognitive constraint for resource rationality to target.

12. Alphabetical reference list

Beck, J. M., Ma, W. J., Pitkow, X., Latham, P. E., & Pouget, A. (2012). Not noisy, just wrong: the role of suboptimal inference in behavioral variability. *Neuron*, *74*(1), 30-39.

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*(4), 700-765.

Busemeyer, J. R., Pothos, E. M., Franco, R., & Trueblood, J. S. (2011). A quantum theoretical explanation for probability judgment errors. *Psychological Review*, *118*(2), 193-218.

Dasgupta, I., Schulz, E., & Gershman, S. J. (2017). Where do hypotheses come from? *Cognitive Psychology, 96,* 1-25.

Fantino, E., Kulik, J., Stolarz-Fantino, S., & Wright, W. (1997). The conjunction fallacy: A test of averaging hypotheses. *Psychonomic Bulletin & Review*, *4*(1), 96-101.

Juslin, P., Nilsson, H., & Winman, A. (2009). Probability theory, not the very guide of life. *Psychological Review*, *116*(4), 856-874.

Maylor, E. A., Chater, N., & Jones, G. V. (2001). Searching for two things at once: Evidence of exclusivity in semantic and autobiographical memory retrieval. *Memory & Cognition*, *29*(8), 1185-1195.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873-922.

Sanborn, A. N., & Chater, N. (2016). Bayesian brains without probabilities. *Trends in Cognitive Sciences*, *20*(12), 883-893.

Shadlen, M. N., & Shohamy, D. (2016). Decision making and sequential sampling from memory. *Neuron*, *90*(5), 927-939.

Spicer, J. & Sanborn, A.N. (2019). What does the mind learn? A comparison of human and machine learning representations. *Current Opinion in Neurobiology, 55*, 97-102.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124-1131.

Vul, E., Goodman, N., Griffiths, T. L., & Tenenbaum, J. B. (2014). One and done? Optimal decisions from very few samples. *Cognitive Science, 38*(4), 599-637.

Wald, A. (1950). *Statistical decision functions.* John Wiley & Sons.

Zhu, J.-Q., Sanborn, A. N., & Chater, N. (2018a). The Bayesian sampler: generic Bayesian inference causes incoherence in human probability judgments. https://doi.org/10.31234/osf.io/af9vy

Zhu, J.-Q., Sanborn, A. N., Chater, N. (2018b). Mental sampling in multimodal representations. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, & R. Garnett (Eds.) *Advances in Neural Information Processing Systems 31*, 5752-5763.