Testing predictions of a model for flexible goal-directed decision-making

Romy Froemer (r.froemer@bham.ac.uk)
School of Psychology, Centre for Human Brain Health, University of Birmingam, UK

Chih-Chung Ting (chihchung.ting@uni-hamburg.de)
Department of Psychology, University of Hamburg, Hamburg, GER

Sebastian Gluth (sebastian.gluth@uni-hamburg.de)
Department of Psychology, University of Hamburg, Hamburg, GER

Amitai Shenhav (amitai_shenhav@brown.edu)
Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence, RI, USA, 02912

Abstract

Every day, humans flexibly make a broad range of decisions, including choosing the item they like most or least, or assigning a value to their option set as a whole. We recently showed that a single sequential sampling model could flexibly accommodate these and other types of decisions. We developed a theoretical framework that formalizes the necessary representations that align sequential sampling and evidence accumulation with one's current choice goals. We implemented this framework within an extended leaky competing accumulator model and showed that model simulations can parsimoniously explain behavior across a range of different choice goals, while also generating predictions for previously untested choice goals. Here we test behavioral predictions of our model and show that human behavior matches the predicted patterns.

Keywords: Value-based decision making; Leaky Competing Accumulator; Mutual Inhibition

Introduction and Framework

Humans can flexibly adjust *how* they make decisions to achieve even the most arbitrary goals. In isolation, sequential sampling models have been able to capture a wide variety of decisions (Busemeyer, Gluth, Rieskamp, & Turner, 2019), including choosing the best among sets of options (Thomas, Molter, & Krajbich, 2021), and assigning a rating to options (Smith, 2016). Here, we test predictions of a model that links research into the computational mechanisms of decision-making with research into cognitive control to explain how humans can flexibly make *any* of those decisions within the same cognitive architecture (Frömer, Gluth, & Shenhav, 2022).

Our framework integrates a sequential sampling process into a cognitive control architecture that sets the parameters of this process according to the decision-maker's current choice goal and the characteristics of the current decision (Fig. 1). Parameters specify a) the relevant feature dimension upon which to decide (e.g., size versus value), and how this property translates into evidence for the b) current goal (e.g., finding largest vs smallest) as well as c) the response structure (e.g., discrete choices vs. ordinal ratings).

Model implementation and Simulations

We implemented this framework by extending a biologically plausible sequential sampling model, the leaky competing accumulator model (LCA) (Usher & McClelland, 2001). In the LCA, evidence at each time step t is accumulated as

$$A_t = A_{t-1} + EI - kA_{t-1} - wWA_{t-1} + sN \tag{1}$$

where A is the vector of response activations, I is an input vector containing the evidence assigned to each response via excitation matrix E, k is a leak parameter that scales how much evidence is "forgotten", w is a scalar on mutual inhibition W, and s is a scalar on normally distributed noise (N) for each option.

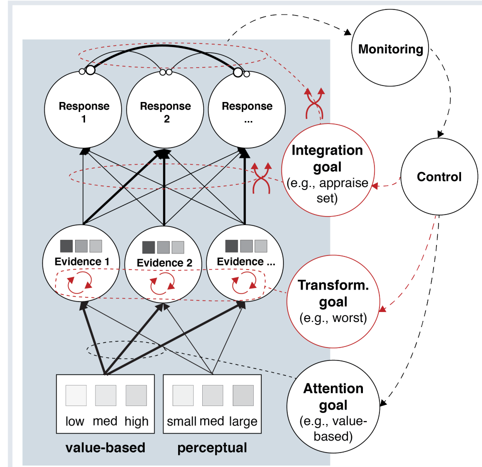

Figure 1: **Flexible decision-making architecture.** Control mechanisms set the parameters on the embedded sequential sampling process. These determine (a) which type of information is selected (attention goal), (b) how this information is transformed into evidence (transformation goal), and (c) how the information is integrated to select a response (integration goal).

Integrating evidence in accord with relevant outputs

In a typical value-based decision-making task, A and I are vectors with each entry corresponding to one option. The excitation matrix E ensures that the inputs in I are added to each accumulator independently. Since all options are equally mutually exclusive, the inhibition matrix W makes them inhibit each other with a constant weight (Fig. 2 top). When appraising one or multiple options (e.g., rating their value), responses do not map onto concrete options (e.g., "choose the bottle"), but onto discrete levels of the relevant dimension (e.g., "choose the highest value level"). These levels (e.g., ratings) are not independent; rather, neighboring levels are more similar than levels that are farther apart. To account for this structure (Fig. 2 bottom) inputs I are determined by mapping samples for each option onto the response space (e.g., 5 ordinal ratings) and integrating across options. Evidence is added to all accumulators proportional to their distance from the input (e.g., rating 2 evidence also activates rating 1 and 3 via E), and mutual inhibition increases with response distance (via W).

Our simulations of choices and appraisals capture canonical behavioral findings: Choices are faster and more consistent as the value difference between options increases, and response times further decrease as the overall value of options increases. Appraisal ratings increase with the overall value of the set (Frömer, Dean Wolf, & Shenhav, 2019; Shenhav & Karmarkar, 2019), and response times are faster when giving an appraisal rating closer to the extremes rather than the center of the scale (Lebreton, Abitbol, Daunizeau, & Pes-

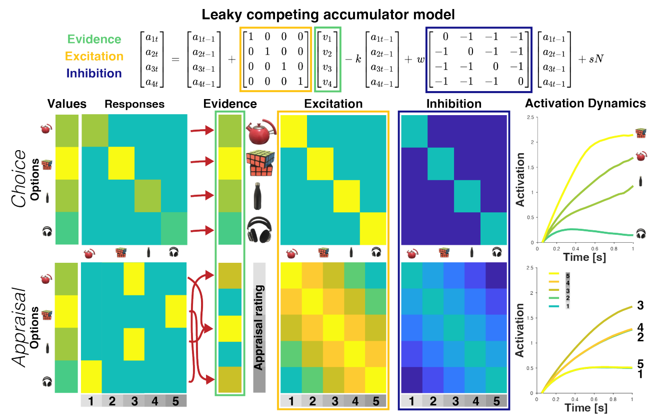

Figure 2: **Reconfiguration of the accumulator structure flexibly affords choice and appraisal of the same options.**

siglione, 2015; Shenhav & Karmarkar, 2019).

Transforming sampled values into suitable inputs

We frequently need to select options other than the best, i.e. the smallest item or one that has exactly the right size (or value). When participants choose the worst instead of the best item, the typical response speeding influence of overall value reverses (Frömer et al., 2019). A hidden layer in which value information (v_i) is transformed into goal-dependent evidence (i_i) can parsimoniously reproduce behavior in best/worst choice, *and* combined with the above changes to integration, generate novel behavioral predictions for liking vs disliking appraisals (Fig. 3 left).

Our simulations reproduce our previous findings for best/worst choice and show that our same architecture can generate goal-congruent appraisal ratings (Fig. 3). Since choice RTs speed up with increasing *magnitude* of the inputs (Frömer et al., 2019), their relationship with overall value reverses as the goal changes from choosing the best to choosing the worst. However, appraisals are more sensitive to the *consistency* of the input (Lebreton et al., 2015). Since inverting the appraisal goal does not change this consistency, our model predicts that this goal manipulation should not affect appraisal RTs.

Empirical findings match model predictions

To test these predictions, we had 44 participants (37 female, M_{age} = 24, SD_{age} = 4) choose or appraise option sets under different transformation goals. Participants were familiarized with the items, then rated them in isolation. Based on these ratings, choice sets were generated to vary in overall value and value difference. The same options were shown twice, once in a choice condition, once in an appraisal condition. Participants used e, f, j, and i keys on a standard keyboard to either choose the option on the screen (positions were matched to key locations) or rate the options on the relevant scale using the keys from left to right. The 4 blocks (choose best, choose worst, appraise liking, appraise disliking) were counterbalanced across participants and participants practiced the respective button mapping before each block. Participants performed up to 60 trials per block, depending on how many choices sets could be generated based on their ratings ($min_{totaltrials}$ = 120, $median_{totaltrials}$ = 208, $SD_{totaltrials}$=24).

Consistent with model predictions, we found dissociable behavioral patterns across the four conditions. The probability of choosing the most goal-congruent options increased with the difference between the most congruent option's value and the average remaining option values (value difference, b = 0.33, $p < .001$), and participants were similarly accurate across 'choose best' and 'choose worst', and across the range of overall value ($ps > .1$). When participants appraised liking, ratings increased with increasing overall value, (b = 0.26, $p < .001$) and consistent with our model predictions, this effect reversed when participants appraised disliking (p = -0.26, $p < .001$). Importantly, we found the expected dissociable effects of transformation goals on RT. As in our previous studies, participants responses speeded up with increasing overall value when choosing the best option, and slowed down when choosing the worst options, with a significant interaction of goal and overall value, b = -0.08, $p < .001$. In contrast, and as predicted by our model, appraisal RT showed a quadratic effect of overall value on RT (b = -0.01, $p < .001$), that did not significantly vary by condition (b = -0.00, p = .118). We also found that participants were significantly slower when either choosing (b = -0.05, p = .003) or appraising (b = -0.06, p = .010) under a negative frame (worst, disliking). This finding, which is not currently accounted for by our model, could reflect additional inputs on action selection, for instance from a Pavlovian system that facilitates approach towards rewards.

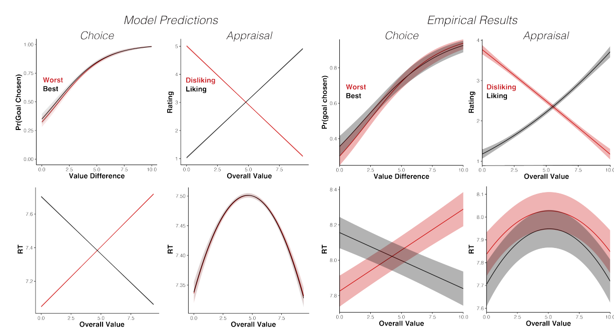

Figure 3: **Transformation goals shape behavior.** Our model predicts that transformation goals have dissociable effects on choice RTs compared to appraisal RTs. Our empirical results show the predicted pattern.

Conclusion

Our model integrates insights from cognitive control with a biologically inspired computational model of decision-making to accommodate a range of different decisions. It offers insights into how humans flexibly align how they decide to reach their current goals and generates testable predictions for novel goals.

Acknowledgments

This work was supported by NIH grant 1R01MH124849-01, NSF CRCNS Award 2309022, and by ERC Starting Grant grant 948545.

References

Busemeyer, J. R., Gluth, S., Rieskamp, J., & Turner, B. M. (2019). Cognitive and neural bases of multi-attribute, multi-alternative, value-based decisions [Journal Article]. *Trends Cogn Sci*, *23*(3), 251-263. doi: 10.1016/j.tics.2018.12.003

Frömer, R., Dean Wolf, C. K., & Shenhav, A. (2019). Goal congruency dominates reward value in accounting for behavioral and neural correlates of value-based decision-making [Journal Article]. *Nature Communications*, *10*(1), 4926. doi: 10.1038/s41467-019-12931-x

Frömer, R., Gluth, S., & Shenhav, A. (2022). Hidden knobs: Representations for flexible goal-directed decision-making [Conference Abstract]. *Multidisciplinary Conference on Reinforcement Learning and Decision Making*.

Lebreton, M., Abitbol, R., Daunizeau, J., & Pessiglione, M. (2015). Automatic integration of confidence in the brain valuation signal [Journal Article]. *Nat Neurosci*, *18*(8), 1159-67. doi: 10.1038/nn.4064

Shenhav, A., & Karmarkar, U. R. (2019). Dissociable components of the reward circuit are involved in appraisal versus choice [Journal Article]. *Sci Rep*, *9*(1), 1958. doi: 10.1038/s41598-019-38927-7

Smith, P. L. (2016). Diffusion theory of decision making in continuous report [Journal Article]. *Psychol Rev*, *123*(4), 425-51. doi: 10.1037/rev0000023

Thomas, A. W., Molter, F., & Krajbich, I. (2021). Uncovering the computational mechanisms underlying many-alternative choice [Journal Article]. *Elife*, *10*, e57012. doi: 10.7554/eLife.57012

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model [Journal Article]. *Psychological Review*, *108*(3), 550-592. doi: 10.1037/0033-295X.108.3.550