Title: The temporal dynamics of value-based decisions in humans

Abbreviated Title: Dynamics of value-based decisions

Article Reviewed: Gluth, S., Rieskamp, J., & Buchel, C. (2012). Deciding When to Decide: Time-Variant Sequential Sampling Models Explain the Emergence of Value-Based Decisions in the Human Brain. *The Journal of Neuroscience*, *32*(31), 10686–10698. doi:10.1523/JNEUROSCI.0727-12.2012

Authors: Aaron M. Bornstein, New York University, 4 Washington Pl. Suite 888, New York, NY 10003; Sara M. Constantino, New York University, 4 Washington Pl. Suite 888 New York, NY 10003

Corresponding Authors: Aaron M. Bornstein, New York University, 4 Washington Pl. Suite 888 New York, NY 10003 email: <aaronb@nyu.edu>, phone: 617.216.6546; Sara M. Constantino, New York University, 4 Washington Pl. Suite 888 New York, NY 10003 email: <smc517@nyu.edu>

Number of Figures: 0

Number of Pages: 4 (including references)

Number of Words: 1415

Keywords: bayesian learning; value-based decisions; sequential sampling; drift-diffusion; neuroeconomics; striatum

Acknowledgements: A.M.B. was supported by National Institute of Mental Health Predoctoral Research Fellowship 1F31MH095501. S.M.C. was supported by a McCracken Graduate Fellowship from New York University.

Humans and animals must often choose between unfamiliar options. To make these decisions, they have to *estimate* the value of each option, using any available information. Acquiring this information might itself have a cost -- at a minimum, waiting to acquire more information can delay the reward -- against which the decision-maker has to weigh the value of the outcome. In a recent issue of the *Journal of Neuroscience*, Gluth et al. (2012) employ the tools of Bayesian estimation as a model of how we trade off the value of more information, which could result in a more accurate estimation, against the value of deciding right now. In other words, they attempt to model the process of “how we decide when to decide.” In doing so, they bring closer into contact two distinct streams of decision-making research, and establish a common framework for future exploration.

The value-based decision-making literature has largely ignored the dynamic nature of within-trial value formation, focusing instead on characterizing patterns of static preferences from choices or the learning of values from feedback in previous trials. Controlled observations of the temporal evolution of value in single trials require a different experimental paradigm than those traditionally employed in studies of value-based decision making. In particular, control over the sequence of available information and measures such as reaction time or gaze can shed light on the unfolding process.

There is a rich tradition in the perceptual decision-making literature of doing just this, using sequential presentations of noisy evidence (most famously, ‘random dot motion’ or RDM) and reaction time measures to examine the emergence of visuomotor commands leading to a decision. Gluth and colleagues introduce a clever adaptation of these sorts of tasks to economic decisions. Their task allows the researchers precise control of the cost, quality and temporal sequence of evidence leading to a value-based decision. They found behavior in this task to be consistent with value constructed by a noisy, sequential integration of evidence, suggesting this may be a general mechanism to construct decision variables under uncertainty.

In each trial subjects had to decide whether to buy or reject a stock that could turn out to be either good (positive reward) or bad (negative reward). They were sequentially presented with costly rating information from six fictitious rating companies and could choose at any time to interrupt the sequence (and, thus, avoid further cost) by buying or rejecting the stock. Ratings were independent, probabilistic and varying in degree or quality of evidence. In general, waiting longer incurs a higher information cost but leads to a more accurate estimate of stock value, so subjects had to decide both when and what to decide.

In the task, objective information changes randomly over the course of a single trial leading to different sequences of estimated value on each trial. The sequential nature of the information is analogous to the accumulation of evidence in the RDM task, although the noisiness is now embedded in the value representation itself rather than resulting from noisy sensory inputs. This parallel extends to the best fitting model. The authors find that the linear decrease in required evidence and in reaction time is well described by a sequential sampling model

(SSM) with a stochastically evolving decision variable that moves towards a time varying decision boundary with a speed that depends on the strength of observed ratings. Asymmetric reject/buy bounds predict the tendency of subjects to reject an offer and to require less evidence for a reject decision.

Critically, the authors observe that information about the estimated value and accumulated evidence is transmitted continuously to cortical motor regions. Interestingly, trial-by-trial differences in caudate, pre-SMA, and insula activity at the start of accumulation -- that is, *before* the presentation of the first sample of evidence -- predict choice times above and beyond the variance explained by model parameters. This observation suggests that the value samples within a trial may not be the only information affecting choices.

The suggestion that noisy value estimates are integrated over time is of a piece with several recent theoretical proposals and empirical observations. Indeed, it has been proposed (Rangel et al., 2008) that integrating over value estimates derived from noisy samples is a general mechanism for situations where cached action values are not available.

How might investigators build on this result? Extending the Bayesian metaphor employed by the authors, decision-consequential activity in integration regions at the start of a trial could be thought of as a *prior* about the choice outcome. Modeling this activity as a per-trial prior opens the question of what information is used to compute the prior. Is the prior a near-optimal integration of expectations inferred from past experience? Is it a prior over the reliability of the ratings agencies? Is it bias towards a motor action given some other kind of contextual cue? Further exploration of this activity might yield new observations about the mechanisms of action selection.

Broadly, the authors have demonstrated that value information is integrated as it is received, at least when presented serially and through a single sensory input stream. But an arguably broader class of decisions involve information that is *not* presented sequentially, but rather available in parallel from a variety of input modalities, including internal states and expectations. It seems that such information *must* be of consequence, even in well-controlled settings such as the current experiment, as in the case of the per-trial baseline fluctuations described above.

Using the tools provided by Gluth and colleagues, we can ask several novel questions of the value estimation mechanism in other, more naturalistic settings. For instance: From what sources are evidence samples drawn when information is not explicitly provided and what affects the integration process? Previous work has suggested that evidence accumulation accelerates when an option is visually attended (Krajbich et al., 2010) because its internal value representation is updated. Separately, an extensive literature points to episodic memory as a source of information for the estimation process. Indeed, memory judgements were the first application of SSM's (Ratcliff, 1978). Ido Erev and colleagues have argued that value estimations in decisions from experience show distinct signatures of a "contingent sampler" mechanism (Erev et al., 2008), where samples of previous episodes are drawn from memories cued by the current state, and used to update the value estimate. These memory-driven expectations might be constructed 'online', at the time of decision (Johnson & Redish, 2007), or during offline periods, such as during sleep or rest (Johnson & Redish, 2005).

Gluth et al. firmly establish sequential sampling with decreasing bounds as a plausible mechanism for incorporating noisy information into value-based choice in a nearly Bayes-optimal manner. In doing so, they provide a key step towards understanding the trade-offs that actors must make when deciding using estimated values based on uncertain information. Critically, they show that value estimates are continually updated to match available information and that these estimates are transmitted to motor regions. In addition to the value estimates and accumulated evidence, behavior appears to also be influenced by pre-trial activity in value and motor areas. A fruitful line of further exploration could extend the framework offered by Gluth et al. to characterize multiple influences on choice behavior in a common Bayesian vocabulary, identifying them as sources of information that affect how, when, and even what we decide.

References

Erev, I. D. O., Ert, E., & Yechiam, E. (2008). Loss Aversion, Diminishing Sensitivity, and the Effect of Experience on Repeated Decisions. *Journal of Behavioral Decision Making*, *21(*May), 575–597. doi:10.1002/bdm

Gluth, S., Rieskamp, J., & Buchel, C. (2012). Deciding When to Decide: Time-Variant Sequential Sampling Models Explain the Emergence of Value-Based Decisions in the Human Brain. *The Journal of Neuroscience*, *32*(31), 10686–10698. doi:10.1523/JNEUROSCI.0727-12.2012

Johnson, A., & Redish, A. D. (2005). Hippocampal replay contributes to within session learning in a temporal difference reinforcement learning model. *Neural Networks*, *18*(9), 1163–71. doi:10.1016/j.neunet.2005.08.009

Johnson, A., & Redish, A. D. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *Journal of Neuroscience*, *27*(45), 12176–89. doi:10.1523/JNEUROSCI.3761-07.2007

Krajbich, I., Armel. C., Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*, 1292-98.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*(7), 545–56. doi:10.1038/nrn2357

Ratcliff, R. (1978). A Theory of Memory Retrieval. *Psychological Review*, *85*(2), 59–108.