Frugal preference formation

Nisheeth Srivastava (nsrivastava@ucsd.edu)
Department of Psychology, University of California San Diego
San Diego, CA 92093 USA

Paul R Schrater (schrater@umn.edu)
Department of Psychology, University of Minnesota
Minneapolis, MN 55455 USA

Abstract

Most theories explaining how animals form preferences for their actions agree upon a basic outline: animals discover what is preferable through interactions with the world, store this information in memory, and recall it to help them decide what to do in a new situation. However, no single theory currently explains both how preferences are learned, and how they are recalled in a way that is compatible with empirical data. We advance precisely such a proposal in the form of a stochastic choice model where the decision agent learns what to do based on scale-free comparisons between options it observes in the world and at each decision instance recalls a subset of these comparison experiences in a manner that minimizes the metabolic costs of memory recall. In simulation, this model makes qualitatively accurate predictions connecting agent choices with various dynamic choice correlates documented in the literature on choice process models.

Keywords: Decision-making; cognitive science; Bayesian modeling; computer simulation; learning; memory; mathematical modeling; artificial intelligence

Introduction

While economists are primarily concerned with representing decision-makers' static preferences, psychologists are more interested in investigating the *processes* by which animals make the choices they do. Naturally, there are also differences in the analysis tools that the two disciplines bring to bear in addressing decision-making under uncertainty. Economic theories of choice tend to axiomatically impose conditions on when subjects' preferences can be said to be originating from some latent function measuring the desirability of various options. In contrast, cognitive theories of the choice process emphasize algorithmically modelling the deliberative process by which agents accumulate evidence and make choices. Theories of the choice process, therefore, differ substantially from static theories of decision-making under uncertainty both in means and methods.

Choice process models have a long history. Procedural theories like elimination-by-aspects, suppresion-of-aspects, lexicographic heuristics etc. can be considered the earliest choice process theories. Computational theories of the decision process originate in the duelling visions of two separate research programs: Busemeyer's decision field theory (DFT) (Busemeyer & Townsend, 1993), and McClelland's leaky competing accumulator models (McClelland, 2001). The basic insight shared by both frameworks, and indeed most succeeding models, is that evidence (as a function of payoffs) for decisions accumulates independently for each option, and that once the quantity of evidence reaches a certain threshold, a decision is made. The biggest difference between the two approaches is that whereas accumulator models are deterministic in selecting the first option to attain a preset evidence threshold, random walk theories like DFT require the difference between multiple options' evidence to reach a preset threshold in order to output a stochastic choice.

The latest generation of process theories borrows the basic evidence accumulation structure of its forebears, but differs in being more specific in defining the nature of evidence being compiled. For example, Chater and colleagues (Stewart, Chater, & Brown, 2006) have proposed a system of inferring choices that dispenses with the need to map payoffs to intrinsic value scales. In their proposal, the payoff's subjective value emerges from ordinal binary comparisons to a sample of payoffs drawn from memory. The means by which certain payoffs are preferentially recalled, though, are left unspecified, so this model cannot predict the behavior of choice process correlates like reaction time and neural activation. A more recent proposal, due to Dickhaut and colleagues (Dickhaut, Rustichini, & Smith, 2009), uses a signal detection analogy to describe the choice process, with the agent's goal being to estimate utility from noisy observations of the world state with minimum cognitive effort. Crucially, existing process models leave the question of the origin of preferences unanswered.

In this paper, we develop a theory of the choice process that specifies both the manner in which preferences are learned from the world, and the manner in which they are recalled for future decisions. The basis for this development is a recent theoretical advance (Srivastava & Schrater, 2012), where we demonstrated that optimal Bayesian inference about which options in the world an agent has to select between, as well as which options in these option subsets are the best, yields preferences that are economically rationalizable under some general constraints on the agent's observation history. In our present work, we build upon this theory of preference formation by adding embodiment constraints via a requirement of effort-sensitive computation. By doing so, we obtain a computationally tractable choice model that can jointly make predictions both for the static revealed preferences of subjects and dynamic process correlates like error rate, reaction time, and neural activation.

Modeling the choice process

Our theory can be distilled into three specific claims:

1. Animals infer what to do from choices they have made in the past concerning different stimulus configurations.

2. Animals perform this inference frugally, using a limited subset of past experiences.

3. Animals are smart about which experiences to use - they select them in order of expected informativeness.

In the three subsections below, we operationalize each of these claims computationally.

Bayes-optimal preference formation

Real subjects, unlike traditional economic agents, need not consider the set of all possible options in the world at every decision instance. Instead, options tend to have typical co-occurrence relations - such that *which* options co-occur becomes itself a signal that allows the agent to establish the *context* in which an option's desirability is to be evaluated. Therefore, in (Srivastava & Schrater, 2012), we outlined a theory of preference formation wherein the basic atom of value includes both information about which option(s) are desirable and the set of options it is desirable amongst, using modal frames as a mathematical formalism for expressing an instantaneous judgment of desirability for human subjects. Compiling desirability across multiple such objects to retrieve an analogue of traditional value computation required the development of a new framework for inferring preferences which we detail in (Srivastava & Schrater, 2012) and briefly review here.

Consider a standard choice formulation where an agent is presented with the set of options $\mathcal{X}$. Traditional treatments of preference learning assume that there is some hidden state function $U : \mathcal{X} \to \mathbb{R}_+$ such that $x \succ y$ iff $U(x) > U(y)$. Preference learning, in such settings, is reduced to a task of statistically estimating a monotone distortion of U. In (Srivastava & Schrater, 2012), we showed that assuming the existence of such a U is incompatible with a number of behavioral results on choice behavior. Our alternative strategy was to demonstrate that the set of options $o \in \mathcal{P}(\mathcal{X})$[1] actually observed at any decision instance can be used instead to directly infer future preference using preferences revealed at previous decision instances without having to resort to intermediate utility computations.

Formally, we introduce preference information into our system via a desirability function d that simply *points* to the best option in a given context, i.e. $d^{(o)} = B$, where B is an accessibility relation (o,o,m) corresponding to the Kripke frame $\langle o,B \rangle$, designed to point to the best option in the observed set by defining $m_i = 1$ iff $o_i \succ o_{i'} \forall o_{i'} \in o \setminus \{o_i\}$ and zero otherwise. The desirability indicated by $d^{(o)}$ can be remapped on to the larger set of options by defining a relative desirability across all possibilities $r(x,o) = m, x \in o$ and zero otherwise.

We further assume that agents infer the situation of the world that they are required to respond to using the option sets they encounter, both as a way to orient themselves with respect to the environment and be able to respond flexibly to novel situations. While the set of contexts inferred $\mathcal{C}$ by an agent from its observation history can be latent in general, a crucial novelty of our framework was to restrict the nature of these contexts as bijective maps of observed option sets, i.e. $\mathcal{C} \subseteq \mathcal{P}(\mathcal{X})$. In our framework, the computation corresponding to utility is desirability $p(r|x,o)$, which is obtained by marginalizing over $\mathcal{C}$,

$$D(x) = p(r|x,o) = \frac{\sum_c^{\mathcal{C}} p(r|x,c)p(x|c)p(c)}{\sum_c^{\mathcal{C}} p(x|c)p(c|o)}, \qquad (1)$$

where it is understood that the *context* probability $p(c|o) = p(c|\{o_1, o_2, \cdots, o_{t-1}\})$ is a distribution on the set of all possible contexts incrementally inferred from the agent's observation history. From the definition of desirability, we can also obtain a simple definition of the *desirability* probability $p(r|x,c)$ as $p(r_i|x_i,c) = 1$ iff $r_i x_i = 1$ and zero otherwise.

To instantiate equation (1) concretely, we must also instantiate the *observation* probability $p(o|c)$. Multiple definitions that obtain the highest possible value for $c = o$ and penalize mismatches in set membership are plausible. This likelihood function is used to update the agent's posterior belief about the contexts it considers viable at decision instance t, given its observation history as,

$$p(c^{(t)}|o^{(1:t)}) = \frac{p(o^{(t)}|c)p(c|o^{(1:t-1)})}{\sum_c^{\mathcal{C}} p(o^{(t)}|c)p(c|o^{(1:t-1)})}, \qquad (2)$$

which, in conjunction with the desirability based state preference update, and a simple decision rule (e.g. MAP, softmax) yields a complete decision theory.

Effort-sensitive preference formation

Whereas the computational aspect of our theory is fully specified in terms of a Bayesian agent seeking to accumulate evidence for future choices based on past choice experiences, at the mechanistic level, there are further considerations that are expected to constrain it. The principal constraint we consider is a requirement of biological organisms to reduce the metabolic costs of thinking.

Cognitive dynamics form a large fraction of the body's basal metabolic requirements. It is reasonable, therefore, to assume that achieving efficiency in cognitive processing would promote natural selection. Thus, it is not unnatural to assume, as we do in this paper, that animals' mechanisms of cognitive dynamics have evolved to be sensitive to the metabolic costs of choice selection, and that animals allocate resources for cognitive processing rationally.

Given the computational goal of a statistically optimal preference-learning agent outlined in the preceding section,

[1] $\mathcal{P}(\cdot)$ references the power set operation.

we develop a model of effort-sensitive preference formation by assuming that while the brain does perform the Bayesian updates we have described above, it does so using a subset of previous observations, not the entire history to impute desirability. Thus, we model the choice process as animals trying to solve a tradeoff between the amount of choice-relevant information they must recall to choose wisely in any particular decision instance and the amount of metabolic effort they must incur in doing so.

We model the process of memory recall as the activation of a subset $\mathcal{M}$ of all decision-relevant memory particles. Using this notation, a general belief formation model could be expressed as,

$$p(x) = \sum_{m \in \mathcal{M}} p(x|m)p(m), \tag{3}$$

where $x \in \mathcal{X}$ are the choices available to the agent, and $m \in \mathcal{M}$ are memory *particles* corresponding to past choice selections. Here, the probability distribution $p(m)$ - which we call the *memory prior* - encodes the likelihood of recalling the memory of experience m, while the distribution $p(x|m)$ encodes beliefs about outcomes learned during the experience corresponding to the memory particle m.

Instead of the idealized context inference in Equation 2, the memory-constrained preference learning agent will employ an update,

$$\begin{aligned} p(c|o,m) &= \frac{p(o|c,m)p(c|\mathcal{M}_{old})}{\sum_c^{\mathcal{C}} p(o|c,m)p(c|\mathcal{M}_{old})}, \\ \Rightarrow p_k(c|o^{(t)}) &= \sum_m^{\mathcal{M}_k} \frac{p(o|c,m)p(c|m^{(1:k-1)})p(m)}{\sum_c^{\mathcal{C}} p(o|c,m)p(c|m^{(1:k-1)})}, \\ &= \sum_m^{\mathcal{M}_k} \frac{p(o|c)p(c|m^{(1:k-1)})p(m)}{\sum_c^{\mathcal{C}} p(o|c)p(c|m^{(1:k-1)})}, \end{aligned} \tag{4}$$

where $p(o|c,m) = p(o|c)$ follows from the fact that the *observation* likelihood of a particular option set o conditioned on having seen the same option set before in context c will be independent of which memory particle was responsible for recalling context c. Note that the index k in Equation 4 indicates the temporal order in which evidence from various memory particles is accumulated during preference formation during a particular decision. Also note that, while the set $\mathcal{C}$ still retains its original definition as the set of inferred contexts, this set will now be determined by the set of memory particles presently activated, not by directly indexing past choice sets as was possible earlier.

The relative desirability computation in Equation 1 also changes to reflect the dependence of the computation on the result of a race between multiple memory particles to influence the agent's choice. In particular, the k^{th} arrival will determine that the relative desirability of option x is,

$$D_k(x) = \frac{\sum_c^{\mathcal{C}} p(r|x,c)p(x|c)p_k(c|o)}{\sum_c^{\mathcal{C}} p(x|c)p_k(c|o)}, \tag{5}$$

where k is a positive integer-valued parameter controlling the amount of deliberation an agent is willing to undergo before making an explicit choice.

An animal will express desirability D_{k^*} as an observable decision when it believes it has constructed a sufficiently useful preference. Modeling the transition from fluid internal preferences to static revealed preference is equivalent to estimating k^* in our framework. As we describe below, and illustrate in Figure 1, we believe that animals are sensitive to the amount of information that particular memory particles can bring to bear on a decision problem, and that they likely select the particles they use to form preferences in order of informativeness. Thus, a reasonable form for k^* inference would be a stopping rule

$$k^* = \underset{k}{\text{argmax}}\, KL(p_k(r|x) \| p_{k-\delta}(r|x)) < \varepsilon, \tag{6}$$

reflecting diminishing marginal informativeness of incremental evidence accumulation.

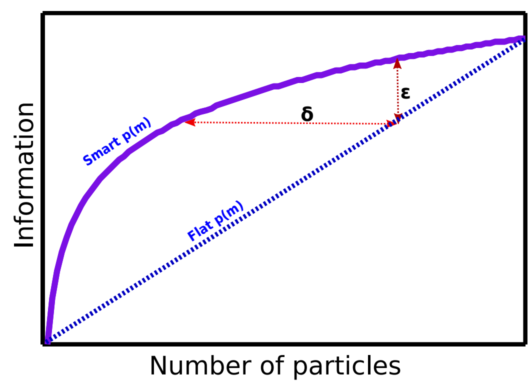

Figure 1: Smartly selecting memory particles for preference formation leads to frugal but accurate computations.

The amount of cognitive effort an animal is willing to invest in a decision will depend both on the quality of its constructed preference and the existence of other concurrent goals that it is currently seeking to fulfill. This intuition is naturally integrated in our model by permitting the *effort* parameters $\{\delta, \varepsilon\}$ to be calculated by a higher level hierarchical controller. Once the high-level controller has assigned the effort parameters, the agent constructs the best possible preference within the effort constraint and outputs it as a decision.

For our present purposes, we assume that we already possess accurate point estimates for these parameters. Realistically, estimates of the effort parameters may also be uncertain. Weak posterior distributions on the effort parameters will yield behavior in our model resembling the very human ability of contradictory decisions being made on very short timescales[2]. However, exact modelling of the higher-level

[2]Waiter: "What dressing would you like with your salad?". Diner: "Umm, French. No wait, ranch. Uhh..., actually let's just go with French." Waiter: "Ok, I'll be right back." Diner (muttering):"I should have asked for vinaigrette."

controller lies outside the scope of this article.

Sparse but smart preference formation

The final piece of our modelling is defining the criterion that agents use to activate memory particles, practically reflected in the choice of $p(m)$ in Equation 4. A flat $p(m)$ would correspond to an agent that is indifferent to the information content of various memory particles. Such an agent though, would be information-theoretically inefficient, as outlined in Figure 1. A more useful strategy would be to selectively use maximally informative particles.

We use a particular information-theoretic construction to operationalize this assumption. Deviations in the amount of information communicated by a particular memory particle is hypothesized to correspond to its decision salience in neural computation. We implement this strategy in our model in the following way. We begin with a standard specification of prediction error using an information divergence,

$$R(p(r|x), p(r|x,m)) = \sum_{x \in X} p(r|x) \log \frac{p(r|x)}{p(r|x,m)}, \qquad (7)$$

where, $p(r|x)$ is the currently computed desirability, and $p(r|x,m)$ is the desirability content of memory particle m, measured as the relative desirability we would impute in Equation 5 using $\mathcal{M} = \{m\}$ in Equation 4. We instantiate the memory prior as a softmax function of particle salience,

$$p(m) = \frac{\exp(\lambda A(m))}{\sum_{m \in \mathcal{M}} \exp(\lambda A(m))}, \qquad (8)$$

where λ is a parameter controlling the scale of salience and the particle salience $A(m)$ itself is instantiated as a convex function of the prediction error, e.g.

$$A(m) = \cos(\alpha \pi R(p(r|x), p(r|x,m))), \qquad (9)$$

where α is a normalizing constant[3]. The intuition behind this relationship between salience and prediction error derives from the need for cognitively salient memories to either reinforce existing policies, or support switching away from them. This dual requirement privileges both low error conditions (generated by reinforcing particles) and high error ones (generated by highly contrasting ones). In contrast, in a domain where on-task behavior is relatively automatic (eye movements), only high contrast samples need be considered salient, which is indeed what is seen empirically (Itti & Koch, 2001).

Algorithmically, the salience computation arises dynamically as the contexts are being compiled to generate relative desirability at the present decision instance. When the set of compiled contexts is empty, there are no salience weights on any of the memory particles. Once contexts begin being compiled, we use our definitions of prediction error and salience anchoring on the incomplete context frame set to sequentially update weights for memory particles.

[3] In practice, α is an uninteresting parameter. We simply normalize with the largest value we know the KL between two vectors of size $|\mathcal{X}|$ will take.

Results

Due to its commitment to a particular embodied form of preference dynamics, our model requires relatively few parameters to relate dynamic choice correlates to static choice selections. The present specification uses four parameters: the mismatch penalty weight β used in computing observation likelihood $p(o|c)$, the salience scaling parameter λ, and the effort parameters $\{\delta, \epsilon\}$, controlling the amount of evidence an agent will recall before revealing its preference. Unlike existing process models, the size of our parameter set does not increase with the number of choice options. In the experiments conducted below, λ and β were found to not influence the results significantly, and remained fixed at the value of 3 in all cases.

The ideal version of preference inference has already been shown to conform with normative rational expectations of choice behavior in (Srivastava & Schrater, 2012). In this paper, we focus on demonstrating that our model of effort-constrained preference inference generates the right profiles of dynamic choice correlates that other process models have sought to model. Specifically, we demonstrate the ability of our model to qualitatively replicate the absolute identification results of (Lacouture & Marley, 2004), which (Brown & Heathcote, 2008) have set up as a benchmark for testing accumulator models.

The basic setup of the experiment involves showing subjects 40 copies each of n stimuli, and asking them to assign number labels $1 \cdots n$ to them. The underlying assumption in previous models has been that subjects possess some internal scale to which they map stimuli lengths, with some amount of noise inherent to this process. We assume, on the other hand, that subjects can learn relative magnitude information by comparing between the stimuli they are seeing. For analytical tractability, we restricted ourselves to considering only pairwise comparisons, i.e. we assumed that agents are comparing the presently seen stimulus to the stimulus seen immediately prior. On each trial, agents updated their estimates for $p(r|x_t, o = \{x_t, x_{t-1}\})$ following our model. Note that the only comparison permitted in our model was judging which one was longer. No absolute magnitude information was stored.

We found that simulated agents operating under these constraints were able to learn the relative ordering of stimuli lengths using a relatively small number of comparisons (40 presentations of each stimulus, as used in (Lacouture & Marley, 2004)). While our results are presented using a smaller stimulus set than the original experiment, the qualitative trends match exactly (see Figure 2). The fraction of correct responses appears as a convex function of stimulus order, and the response time appears as a concave function of stimulus order, precisely matching the profiles observed by (Lacouture & Marley, 2004) (also see Figure 10 in (Brown & Heathcote, 2008)). Additionally, manipulating the effort parameters permits us to make predictions about the behavior of the response time and error curves respectively in speed-emphasis and accuracy-emphasis. Our results are qualitatively in agree-

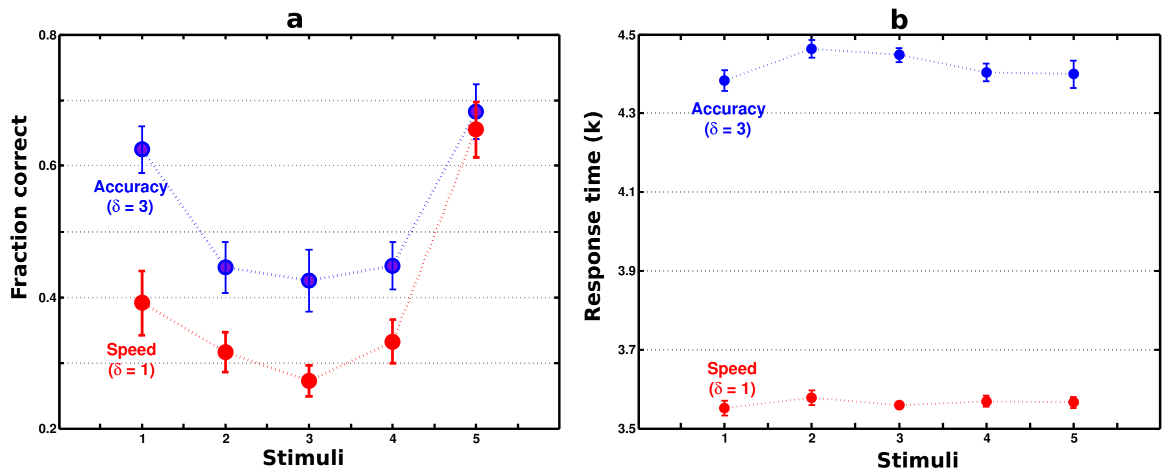

Figure 2: Our model replicates patterns of error rate and response time as a function of stimulus magnitude order previously observed in human subjects in an absolute identification task. Task parameters were chosen to replicate Experiment 1 in (Lacouture & Marley, 2004) - 40 trials per stimulus per session; results averaged across 12 simulation sessions. **(a)** error rate is lowest for extreme stimuli; intermediate magnitude stimuli are predicted incorrectly significantly more often. Error bars are 2 SD across. **(b)** response time is also lowest for extreme-valued stimuli, and increases for intermediate stimuli. Error bars are 2 SE across (repeated measurements). In both cases, simulations emphasizing speed over accuracy change the profile of the curves in the same manner observed in human subjects by (Ratcliff & Rouder, 1998).

ment with data for such a manipulation collected by (Ratcliff & Rouder, 1998) (see Figure 7 in (Brown & Heathcote, 2008) for a visual comparison). The major quantitative difference is that whereas human subjects do not lose significant accuracy in the speed condition, our artificial subjects, trained on a limited set of trials, lose considerably more.

The model's capacity to identify absolute stimuli arises from its ability to accumulate evidence from pairwise comparisons and recall a small number of these comparisons to dynamically estimate the order of the present stimulus. This linkage of preference with history renders transparent the relationship between the psychophysical bowtie effect seen in identification experiments (Lacouture & Marley, 2004) and the economic distance effect (Dickhaut, Smith, Xin, & Rustichini, 2013) observed in multiple behavioral economics experiments where it is found that extreme choice valence (distance in utility) appears to be correlated with lower error rate, response times and interestingly, levels of neuronal activation as measured by fMRI (Dickhaut et al., 2013).

The same simulation also replicates a complementary decision-theoretic observation - high response times are usually associated with high error rates, and in turn with largely indifferent choice (Dickhaut et al., 2013). The working of our model illustrates that differences in the amount of cognitive effort required to disambiguate options that are close in value arise from the fact that such options have both been previously chosen while members of different option sets. Assimilating conflicting evidence requires more samples for the desirability distribution to stabilize, resulting in greater effort, correlated with higher RT and brain activation.

Having modeled a difficult benchmark for existing process models, we also demonstrate an effect that accumulator models and decision field theory find difficult to account for, but originates endogenously in ours - the increase in choice response time as a function of choice set size. Since other computational models of the choice process model evidence accumulation for each option individually, they end up predicting that response time should be independent of choice set size. Instead, as Hick's law formalizes, RT is empirically seen to increase logarithmically with set size (Hick, 1952). Accumulator models have sought to remove this discrepancy by assuming that the response threshold increases logarithmically with the option set size, but rationales for such assumptions (which verge upon assuming the consequent) are unclear.

Response times increase naturally in our model, though the rate of increase appears to be closer to linear than sub-linear (see Figure 3). That RT should increase with choice set size is a natural consequence of our model; larger option sets need more comparison information to draw useful inferences from. The fact that our model fails to show a sub-linear Hicksian trend suggests that there are aspects of hierarchical clustering of option sets that our current framework does not accommodate.

Discussion

Earlier efforts at designing choice models have almost uniformly reified the notion of value as an easily accessible external signal. If value is easily accessible in the form of payoffs, then we need not worry about learning it. However, evidence is compiling that value learning is not simple, nor unneces-

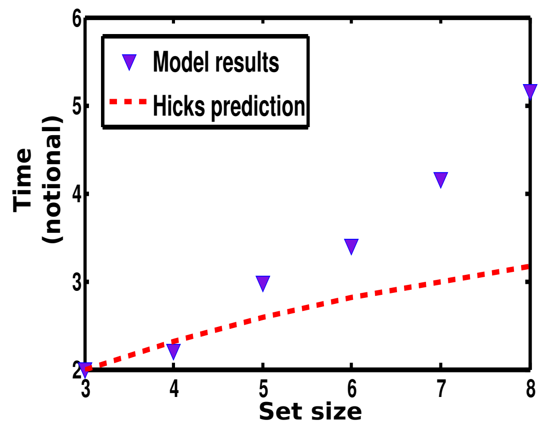

Figure 3: While our model differs from existing process models in predicting increasing response time as a function of choice set size, it does not converge to Hick's law's prediction of sub-linear increases.

sary, and in fact, it is not clear if we should be thinking in terms of option-specific value signals at all (Vlaev, Chater, Stewart, & Brown, 2011). Recognizing this ambiguity, our own earlier efforts have focused on developing a rational theory of preference learning that does not require intermediate value computations. In this paper we have connected these ideas with limited, frugal computation to generate an integrated theory of metabolically frugal preference learning. The resulting theory can both predict which option an observer will select, and also how much time (relative to other options) they will take to do so, depending on subjects' history of experience. Crucially, ours is a rationalizable theory of the choice process - it can both explain what choices an animal will make, and *why*, by outlining the mechanism by which previous choice inform future ones.

Sensitivity to initial conditions, while possibly a limitation, is a very interesting feature of our model. Its path-dependent construction of preferences constrains us to predict that the options that animals predict will depend significantly on the order of evidence. Since we model the activation propensity of memory particles as a function of the preference being formed, which particles arrive first strongly influence which ones will be recalled later. In a binary choice, for instance, a rare particle (pointing to an infrequently chosen option) arriving first will essentially make it impossible for further rare particles to participate. Conversely, if the initial particle set is typical (pointing to the frequently chosen option), rare particles will have a greater impetus to participate in preference construction. It is not unreasonable to expect animals to behave this way; and predictions from this aspect of our model are testable. For example, in economic choices between a safe and a risky option, our model predicts that subject populations will segment into two cohorts: one that select the risky option with probabilities matching empirical frequencies of success, and one that overweights the probability of success in their choice decisions, a prediction that is supported empirically (Bruhin, Fehr-Duda, & Epper, 2010).

In summary, we have modelled the choice process of animals as statistically efficient memory retrieval, and shown that such a model can explain important regularities in the relationship between memory samples, choice predictions, and choice prediction response times. Our results improve upon previous explanations for a number of response time effects, because they connect prior choice history to these variables. Our model is not just rationalizable, it is also substantially normative; optimizing a dual objective of adapting to environmental frequencies while reducing the internal processing costs of such adaptation. Outlining a tractable form for this optimization is the key contribution of this work.

References

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive psychology*, *57*(3), 153–178.

Bruhin, A., Fehr-Duda, H., & Epper, T. (2010). Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica*, *78*(4), 1375–1412.

Busemeyer, J., & Townsend, J. (1993). Decision field theory: A dynamic cognition approach to decision making. *Psychological Review*, *100*, 432-459.

Dickhaut, J., Rustichini, A., & Smith, V. (2009). A neuroeconomic theory of the decision process. *PNAS*, *106*(52), 22145-22150.

Dickhaut, J., Smith, V., Xin, B., & Rustichini, A. (2013). Human economic choice as costly information processing. *Journal of Economic Behavior & Organization*, *94*, 206–221.

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, *4*(1), 11–26.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature reviews neuroscience*, *2*(3), 194–203.

Lacouture, Y., & Marley, A. (2004). Choice and response time processes in the identification and categorization of unidimensional stimuli. *Perception & psychophysics*, *66*(7), 1206–1226.

McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological review*, *108*(3), 550–592.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*(5), 347–356.

Srivastava, N., & Schrater, P. (2012). Rational inference of relative preferences. In *Proceedings of advances in neural information processing systems 25.*

Stewart, N., Chater, N., & Brown, G. D. (2006). Decision by sampling. *Cognitive Psychology*, *53*(1), 1 - 26.

Vlaev, I., Chater, N., Stewart, N., & Brown, G. (2011). Does the brain calculate value? *Trends in Cognitive Sciences*, *15*(11), 546 - 554.