**Introduction**

Humans are bombarded with immensely complicated decisions. Imagine a psychology student – Sally – deciding where to eat lunch. Sally has certain preferences (e.g. she likes Mexican food, dislikes walking long distances) and constraints (e.g. she’s deathly allergic to walnuts) that she should factor into their decision. Ideally, she would carefully evaluate all her options, and choose the one with highest expected value. For example, she might appraise each option based on how close it is to her office, how Mexican its cuisine is, the likelihood that it uses walnuts, etc., and then choose the option with the highest aggregated value. This process – computing the expected values of options at decision time by planning over a causal model of the environment – has been intensively studied, and we have some idea how it could be accomplished for a small set of options (Dolan & Dayan, 2013; Doll, Simon, & Daw, 2012).

But in any real-world decision, there are an overwhelming number of potential options (cite). There are hundreds of restaurants in Harvard Square, and thousands in the greater Boston area. And the problem is even worse than this, because Sally has more options than just restaurants: She could also grow the crops herself, or catch a wild animal to eat, or steal food from the communal refrigerator, etc. She couldn’t possibly plan over all her options – she would die of starvation before she finished.

Yet people like Sally are able to make these decisions with speed and ease. How? Intuitively, people don’t consider all possible options – they construct a small set of options to evaluate, and ignore all the rest. For instance, Sally might only consider Chipotle and Taco Bell, and choose one of those. The process by which people narrow down the enormous set of potential options to a small set of relevant choices is known as *choice set construction* (cite).

The aim of this paper is to characterize how choice sets are constructed. Not all options are equally likely to make it into someone’s choice set; people clearly favor some options (e.g. Darwin’s) over others (e.g. catching a wild animal). What determines which options make the cut? We focus on one potentially important factor: how good an option has been in the past (i.e. the option’s past value). Options that have been good in the past tend to be good in the future. Moreover, prior research has demonstrated that people spontaneously compute and maintain a representation of how good an option has been, on average, in the past (cite); the past values of options are pre-computed, or “cached”, before decision time. (In the reinforcement learning framework, this process is often called “model-free learning” (cite).) Hence, the mechanism that constructs choice sets might be designed to propose options with high past values – narrowing down options to consider without incurring a high computational cost.

We explore this idea in two ways. First, we construct a computational model of value-guided choice set construction, and simulate its performance (and the performance of two alternative models) in various environments. When options that have been good in the past tend to be good in the future, the choice set model achieves good accuracy at low computational cost.

Second, we present an experimental paradigm designed to elucidate the role of value-guided choice set construction in decision-making. We fit the model to people’s choices in the experiment, and find that the best-fitting model constructs choice sets guided by the prior value of options. These results suggest that people spontaneously construct choice sets when faced with difficult decisions, and are often more likely to include options with high prior values in those choice sets.

Third, we connect our model to recent work on modal cognition. **Jonathan write this!**

**Computational model**

A schematic of the choice set model is depicted in Figure 1. There is a large pool of *N* potential options, each marked with a pre-computed past value. We assume that agents have learned these values from past experience, and do not explicitly model the learning process. The agent samples a small number of *K* options, without replacement, from this pool.

The sampling process is non-uniform, and is more likely to sample options with high past values. Let $Q^p\_i$ be the cached the past value of option *i*. Thenthe probability of sampling an option *i* is:

where is an inverse temperature parameter controlling the degree to which sampling is biased towards high-value options. (This formula employs a *softmax function* over the past values; we will use this terminology to describe it later.)\footnote{Non-uniform sampling without replacement is tricky, because maintaining stable sampling probabilities as the pool shrinks seems to require a costly renormalization after every sample. Fortunately, there is a simple, highly parallelizable algorithm that can sample without replacement in one pass over the options, without having to constantly renormalize. The agent simulates an exponentially distributed random number (with rate parameter 1) for each option, divides it by the option’s probability , and chooses the *K* options with the lowest resulting numbers. This algorithm achieves the desired weighted sample. See Efraimidis & Spirakis, 2006 and Müller, 2016 for proof and elaboration.}

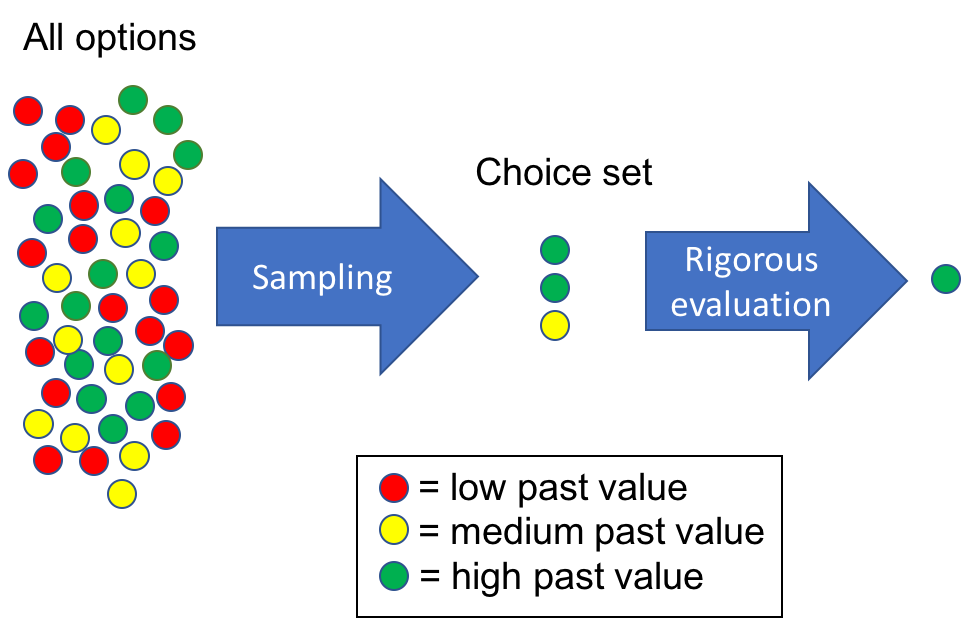


Figure 1: A schematic depiction of the choice set construction model.

Once the choice set is sampled, the agent uses a laborious planning process to compute the current value $Q^c\_i$ of each option in the choice set, and chooses probabilistically among them with another softmax function:

We do not explicitly model the planning process. Instead, we treat it as a black box that the agent can use to compute the current values of options (at high computational cost).

*Alternative models*

We compare the choice set model to two alternatives, which anchor the two ends of a spectrum of computational complexity (Figure 2). The “no planning” model does not perform any forward planning or evaluation of options in the current context, and simply samples an option with probability proportional to its past value. Because past values were cached before decision time, this process is computationally cheap, but can be highly inaccurate if circumstances change.

In contrast, the “optimal planning” model plans over all options in the current context, and chooses the best. By planning over a causal model of the environment, this approach achieves high accuracy – but at a high computational cost.

The choice set model falls somewhere between these two extremes; it evaluates some options via planning, but much less than the optimal planning model. We show that, for a plausible range of environments, the choice set model provides major gains in accuracy over the no planning model, at a fraction of the cost of optimal planning.

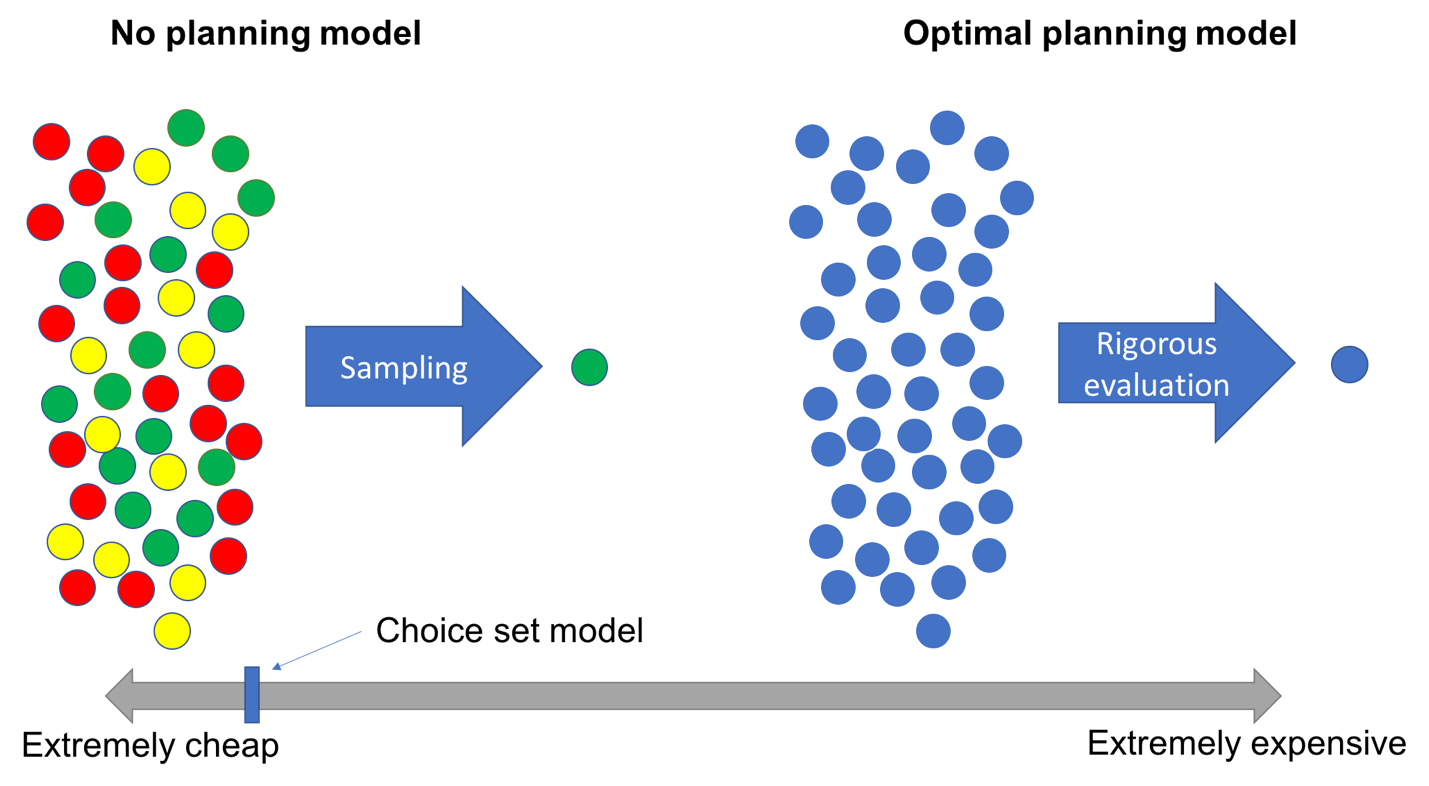


Figure 2: A schematic depiction of two alternative models. The no planning model is extremely computationally cheap; the optimal planning model is extremely expensive; and our choice set model falls somewhere in between.

*Simulation setup*

To show this, for each of the three models, we simulated 10,000 agents using that model to make decisions in five different environments. Each agent made a single decision in each environment, which consisted of choosing between *N =* 1000 options based on their past and/or present value. The past and current value of each option were simulated anew for each agent-environment pair. The no planning model sampled according to the options’ past values; the optimal planning model deterministically chose the option with the highest current value; and the choice set model used the past values to construct a choice set of size *K =* 10, from which it chose an option with probability proportional to the current values. Again, we assumed that actual agents would acquire these values through prior learning episodes (for the past values) or online planning (for the current values), but we did not explicitly model the learning/planning processes.

The five environments differed solely in the simulated correlation between past and current values. The values were drawn from lognormal distributions with correlation either 0, .25, .5, .75, or 1. (The lognormal distribution embodied the assumption that most of the options available to us are bad, while only a few are good.)

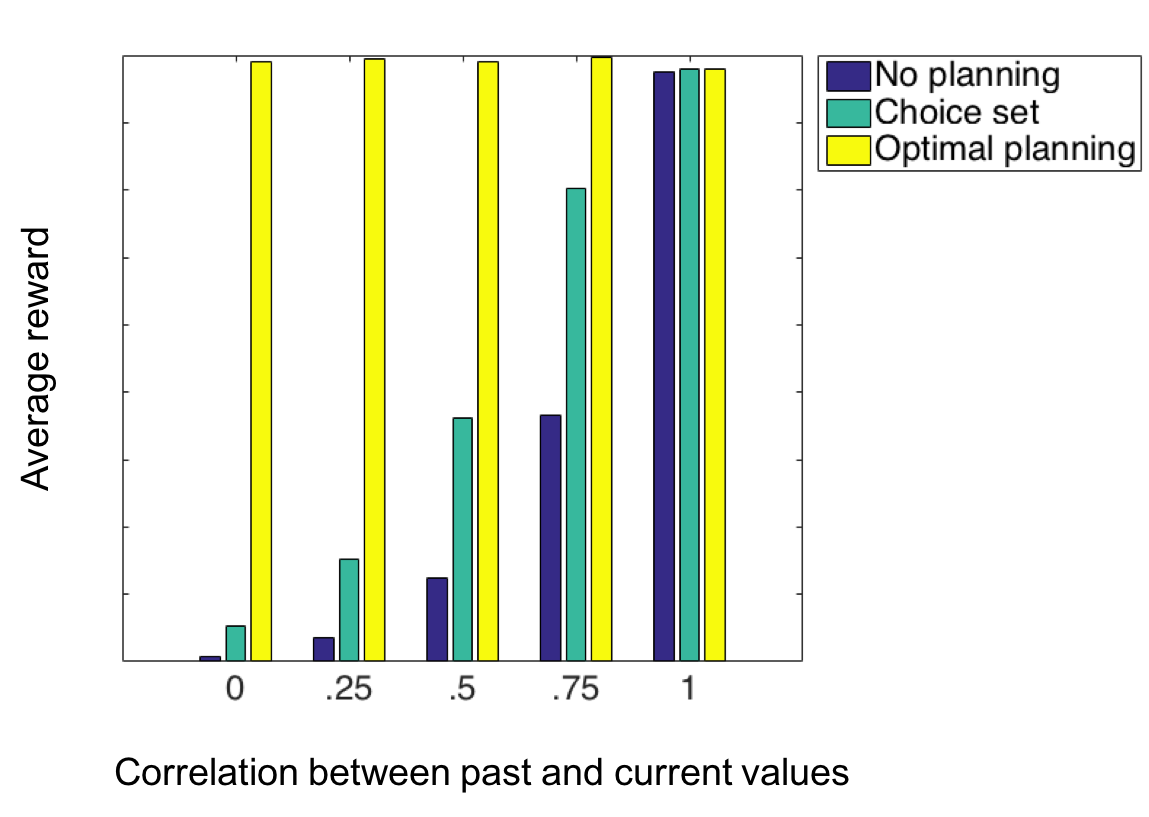


Figure 3: Simulated earnings of the three models across five environments.

*Simulation results*

The simulation results are shown in Figure 3. As expected, the optimal planning model achieves the highest accuracy across all environments (but at an extreme computational cost). Moreover, when the correlation between past and current values is low – i.e. when the past value of options is not indicative of their current value – then both the no planning and choice set models perform poorly.

Crucially, however, there is a sweet spot where choice set models are adaptive. When past values are highly but not perfectly predictive (e.g. *r* = .75), the choice set model performs almost twice as well as the no planning model, with accuracy approaching the optimal planning model, at a low computational cost. This result suggests that, as long as decision environments do not change too quickly, constructing value-guided choice sets is an efficient way to make effective decisions.

**Behavioral experiment**

Next, we tested whether people construct value-guided choice sets when faced with difficult decisions. To test this, we employed an experiment with two stages. The idea was to expose people to a large set of different-value options in Stage 1, and then ask them to make decisions using those options in Stage 2. The resulting decision patterns could be tested for signatures of value-based choice set construction.

*Design*

In Stage 1 of the experiment, participants were exposed to a set of fourteen common English nouns (e.g. “basket”, “community”, “machine”). Each word was associated with some amount of bonus points. For instance, “basket” might have been worth 10 points, and “community” might have been worth 0. Half of the words were randomly chosen to have a low point value (either 0, 1, or 2 points), and half to have a high point value (either 8, 9, or 10 points). (Points were translated into bonus money at the end of the experiment.)

In order to learn these word-value associations, in Stage 1 participants played a game where they repeatedly chose between a word and a fixed number of points (Fig. 4). For instance, on one trial, a person might have had to choose between “basket” and 5 points. If they chose the word, they earned however many points it’s worth. If they chose the fixed number of points, they received that many points. Thus, participants were incentivized to learn the word-value associations and use that knowledge to win more bonus points throughout the game.

Participants completed 8 trials per word, for a total of 112 trials. Importantly, no matter what they chose, we showed the word’s point value on each trial. This procedure guaranteed that people were exposed to each word an identical number of times. To further ensure that people learned the word-value associations, we asked participants to retype the word and its value after each trial.

Then, in Stage 2, participants faced a series of decisions like: “Give us a word from Stage 1 with the most number of vertical lines in its letters. You’ll win 10 points for each vertical line in the letter of your word.” In these questions, the potential options were the words from Stage 1, and each option’s current value (e.g. the number of vertical lines in the word) was difficult to evaluate. There were 8 decisions in total. For each decision, participants were given an example and a comprehension check. All decisions had a time limit, which was calibrated *a priori* to the difficulty of each decision (e.g. the vertical lines decision had a time limit of thirty seconds.)

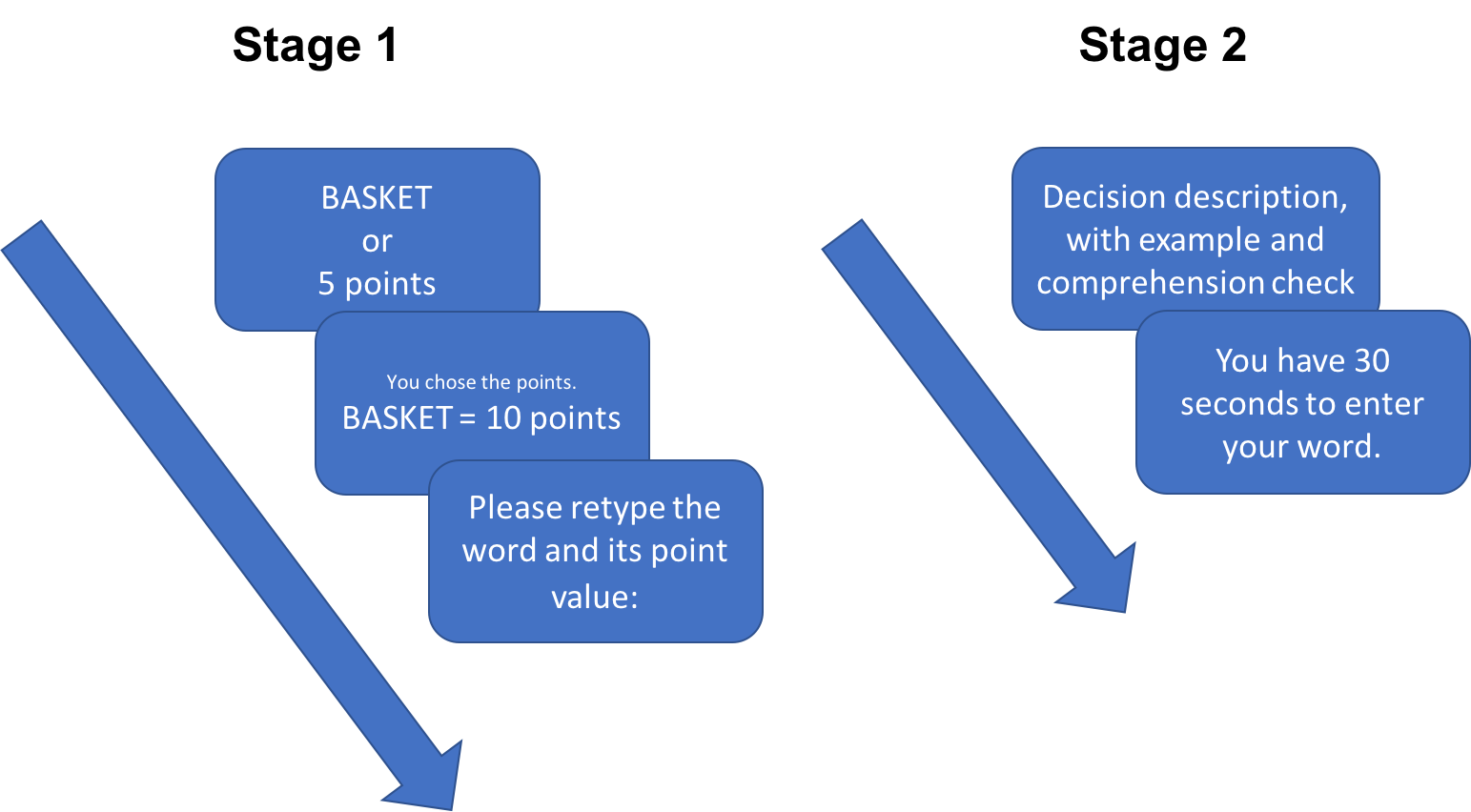


Figure 4: Design of the two-stage experiment.

Each option’s current value in the Stage 2 decisions (e.g. the number vertical lines in the word) were uncorrelated with its past value in Stage 1. Participants were explicitly instructed of this fact. Nonetheless, we hypothesized that, to make these decisions, people would construct a small set of words to evaluate, and that words with high Stage 1 point values would be more likely to enter this choice set.

*Results*

X participants completed the experiment. We excluded anyone who… (The sample size and exclusion criteria were decided in advance.) After exclusion, X participants remained. All participants give informed consent, and the study was approved by Harvard’s Committee on the Use of Human Subjects.

We tested for the presence of value-guided choice sets in two ways. First, if people are constructing value-guided choice sets, then their word choices in Stage 2 should show an influence of both past value (i.e. the point values in Stage 1) and current value (i.e. the values of the words in the current decision).

To test for an influence of current value, for each Stage 2 decision, we ranked all the words according to their current values (from 1, the worst word, to 14, the best word). Then, we computed the average rank of each participant’s Stage 2 choices (Fig. X). People’s choices were ranked significantly above chance, demonstrating that they were influenced by the current values of the words (stats here).

To test for an influence of past value, we employed a similar procedure. We computed the percentage of each participant’s word choices which had high point values in Stage 1 (figure X). People chose words with high Stage 1 values significantly more than chance, suggesting that they were also influenced by the past values of the word (stats here).

Of course, the fact that people are influenced by both the past and current values of the words does not prove that they are constructing value-guided choice sets. There are other ways that the past and current word values could combine to influence choice. People could be alternating between the two approaches, computing the current values in some trials and relying on past values in others; or, people could be basing their choices on a linear mixture of past and current values.

To rule out these alternatives, we fit our choice set model and non-choice-set alternatives to people’s choices, and performed formal model comparison. For each type of model (choice set, no choice set), we fit several variants, shown in Table 1. We computed the maximum *a posteriori* estimates for all parameters, using a Gamma prior for the inverse temperatures and a uniform prior for the mixture weights and choice set size. The possible choice set sizes were restricted to {2, 3, 4}. We then performed Bayesian model selection by entering the model evidences (computed with the Laplace approximation) into the spm\_BMS routine in SPM8.

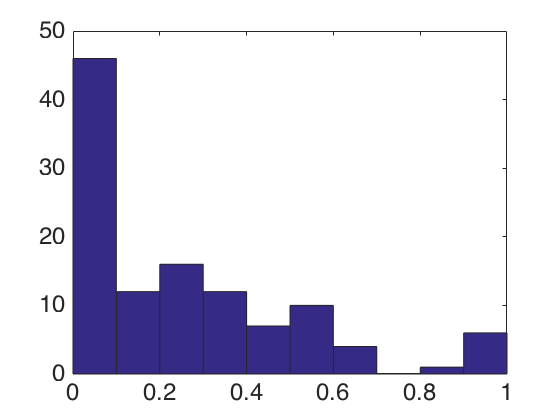
The overwhelmingly preferred model, with an exceedance probability of over .999, was the choice set model which used both Stage 1 (past) and Stage 2 (current) value to construct choice sets. This result suggests two things. First, as hypothesized, people were using past value – the point values from Stage 1 – to guide their choice set construction.

Second, before going through the laborious computation of current values (e.g. counting the number of vertical lines), people likely had access to some cue that correlated with the current value of the word (e.g. the word’s length). By using this cue as an additional influence on choice set construction, people’s choice sets also appeared to be influenced by Stage 2 value. \footnote{The existence of these cues does not stop a person from planning once the choice set is constructed; the cues would be very rough estimates of current value, and further evaluation of the options in the choice set would still be beneficial.}

According to the preferred model, when sampling a choice set, people employed a weighted mixture of Stage 1 and Stage 2 values: math here, where $w$ captures the influence of Stage 1 value. As a final test, we extracted each participant’s best-fit $w$ and examined the distribution (Figure X). About half of participants showed little influence of Stage 1 value ($w < 0.1$), but the other half showed a range of influence. The mean $w$ was , indicating that, on average, Stage 1 value made a meaningful contribution to choice set construction.

|  |  |  |  |
| --- | --- | --- | --- |
| **Family** | **Model** | **Description** | **Parameters** |
| Choice set | Stage 1 value | Used Stage 1 value to sample a choice set. Chose among the options in choice set in proportional to their Stage 2 value. | inverse temperature for choice set sampling  inverse temperature for subsequent choice  choice set size |
| Choice set | Stage 2 value | Same as above, but used Stage 2 value to construct choice set. | Same as above |
| **Choice set** | **Both** | **Same as above, but used a mixture of Stage 1 and Stage 2 value to construct choice set.** | **Same as above, with:**  **mixture weights for choice set softmax** |
| Choice set | Random | Same as above, but sampled choice set uniformly. | inverse temperature for choice softmax |
| Non-choice-set | Stage 1 value | Chose among options in proportion to their Stage 1 value (i.e. no planning model). | Same as above |
| Non-choice-set | Stage 2 value | Chose among options in proportion to their Stage 2 value (i.e. optimal planning model). | Same as above |
| Non-choice-set | Both | Chose among options in proportion to a mixture of Stage 1and Stage 2 value. | Same as above, with:  mixture weights for choice softmax |
| Non-choice-set | Random | Chose among options uniformly at random. | None |

Table 1: Models fit to participant choices. The bolded model was preferred.



Weight on Stage 1 value

Frequency

Fig. 6: Distribution of weights of the influence of Stage 1 value on choice set sampling, across subjects.

**Conclusion**

In sum, decisions with large option sets can be tackled by constructing choice sets, and using the prior value of options to bias the choice set sampling procedure. I showed that, under certain environmental conditions, this model can achieve major accuracy gains at low computational cost. Moreover, in a novel behavioral experiment, I showed that a large percentage of people appear to spontaneously employ this decision strategy. This finding is a step towards understanding how people make quick, effective decisions in environments of real-world complexity.

**References**

Dolan, R. J., & Dayan, P. (2013). Goals and Habits in the Brain. *Neuron*, *80*(2), 312–325. https://doi.org/10.1016/j.neuron.2013.09.007

Doll, B. B., Simon, D. A., & Daw, N. D. (2012). The ubiquity of model-based reinforcement learning. *Current Opinion in Neurobiology*, *22*(6), 1075–1081. https://doi.org/10.1016/j.conb.2012.08.003

Efraimidis, P. S., & Spirakis, P. G. (2006). Weighted random sampling with a reservoir. *Information Processing Letters*, *97*(5), 181–185. https://doi.org/10.1016/j.ipl.2005.11.003

Müller, K. (2016). Accelerating weighted random sampling without replacement. *Arbeitsberichte Verkehrs- Und Raumplanung*, *1141*. Retrieved from http://e-citations.ethbib.ethz.ch/view/pub:176429