*What's the main question being asked in this study?*

When facing difficult decisions with many options and limited time, people will generate a small subset of those options – a choice set – to actually evaluate, and they will be more likely to include options that were good in the past. Moreover, when asked to quickly judge which options were "possible" to choose, people will selectively judge low-value options – the ones that tended to be excluded from their choice sets – to have been impossible to choose. This pattern will disappear be exhibited when participants are given time to reflect before responding. We test this in an experiment with three stages. In Stage 1, people are trained to associate a series of fourteen English words with different values. In Stage 2, people judge whether each word was possible to choose in Stage 1, either under time pressure or time delay. In Stage 3, people use these words as potential answers to a series of difficult decisions.

*Describe the key dependent variables specifying how they will be measured.*

The first dependent variable is possibility judgment in Stage 2. Participants will judge whether each of the fourteen words from Stage 1, along with seven words that were absent in Stage 1, were possible to choose in Stage 1. The words will be presented one at a time, in a random order. Participants in the time pressure condition will have to respond within one second; participants in the time delay condition will not be able to respond until one second has passed.

The second dependent variable is word choice in Stage 3. There will be eight decisions total. For each one, we will describe the decision that needs to be made (e.g. "Give us a word from Stage 1 with the most number of horizontal lines. You'll get 10 points for each horizontal line in the word you describe"), give an example, a comprehension check, and then a textbox to submit a response. To parse participants' answers, we will compare their response to the list of Stage 1 words using the Optimal String Alignment method in the "amatch" function of R package "stringdist" (with a maximum distance of 2). Any response that fails to match a Stage 1 word will be coded as NA. All decisions will have time limits, which were calibrated a priori to the difficulty of the question.

*How many and which conditions will participants be assigned to?*

For each participant, half of the Stage 1 words will be randomly assigned to be high point value, and half to be low point value. (We treat Stage 1 value as a dichotomous variable.) Half of participants will be assigned to the time pressure condition in Stage 2, and half to time delay.

*Specify exactly which analyses you will conduct to examine the main question.*

First, for each participant, we will compute the percentage of high-value and low-value words judged impossible. We predict an interaction between word value and time condition: People under time pressure will be more likely to judge low-value words impossible. We will test this with a within-between ANOVA.

Second, we will fit twelve computational models to trial-level choices. Six models will sample a subset of options (using a softmax function over Stage 1 value, negative Stage 1 value, Stage 3 value, a linear combination of Stage 1 and Stage 3 value, a linear combination of negative Stage 1 value and Stage 3 value, or uniformly at random), and then choose among those options with another softmax over Stage 3 value (with an independent temperature parameter). The other six models will sample a single option (with the same variants). We will find the maximum a posteriori parameters for each model (with a Gamma(4.5, 1) prior and [0, 10] range for the inverse temperatures, and a uniform prior and {2, 3, 4} range for the subset size). We will then compute the Laplace approximation to the model-evidence for each subject for each model, and compare them using Bayesian model selection. (If the Hessian is degenerate for any subject, we will replace the Laplace approximation for that subject with the BIC.) We predict that the choice set model which uses both Stage 1 and Stage 3 value to construct choice sets will have the

highest protected exceedance probability, and will be used by the highest expected posterior percentage of the population.

Third, we will test for a correlation between individual differences in the first and second analyses. For each participant, we will compute the percentage of low-value words judged impossible minus the percentage of high-value words judged impossible. Then, across participants, we will correlate this measure with the MAP estimate of the Stage 1 mixture weight from the computational model above. We predict, specifically for participants in the time pressure condition, a positive correlation between the two.

To test for a basic influence of Stage 1 and Stage 3 values on Stage 3 choice, we compute, for each participant, (a) the percentage of choices with high Stage 1 values and (b) the average Stage 3 rank of their choices. We predict that both will be significantly above chance, measured with one-tailed one-sample t-tests. At the end of Stage 3, we give people a free recall test for the words in Stage 1. We predict that people will remember words with a high Stage 1 value earlier then words with the low Stage 1 value. We will test this with a linear mixed effects model, regressing free recall order on Stage 1 value (with all random effects for subject and word).

*How many observations will be collected?*

N = 300

*Anything else you would like to preregister?*

We will exclude participants for whom any of the following is true: They don't complete the study, they successfully rewrite less than 75% of the words or values during Stage 1 training, they show a Pearson correlation between Stage 1 value and Stage 1 choices of less than .75, they fail to give a Stage 1 word for more than 2 of the 8 Stage 3 trials, they repeat an answer in Stage 3 more than twice (people were not allowed to repeat words on consecutive trials), they pass less than 50% of the Stage 2 comprehension checks, they recall less than 5 words in the free recall question, or they write things down physically during the experiment (as measured by a probe at the end). We will also exclude any Stage 3 trials in which the participant did not give a response that matched a Stage 1 word.