*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 consideration set – to actually evaluate, and they will be more likely to include options that were good in the past. We test this in an experiment with two stages. In Stage 1, people are trained to associate a series of well-known twelve English words – the names of the 12 months – with different values. In Stage 2, people use these words as potential answers to a difficult decision. Finally, people will be asked which words they considered (which words “came to mind”) during the decision.

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

There are two dependent variables: which word people choose in Stage 2, and which words they consider during the decision. First, we will describe the decision that needs to be made ("Give us a word from Stage 1 whose third letter is late in the alphabet. You'll win points based on the position of the word in the alphabet (i.e. A = 1, Z = 26)."), give an example, two comprehension checks, and then a textbox to submit a response. (Participants will also get a “scratchpad” textbox to help them think.) 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). If their response fails to match a Stage 1 word, we will try the same parsing procedure on the scratchpad. If both responses fail to match a Stage 1 word, the trial will be coded as NA. The decision will have a 25-second time limit.

Second, we will ask people which words they considered while making their decision. We will present the words one at a time with “Yes” and “No” radio boxes below them, and ask people to select “Yes” if the word came to mind at all during the decision.

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

There are no between-subject conditions. The values of the month names in Stage 1 will be 1, 2, 3, …, 12, and will be randomly assigned for each participant.

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

(A) We predict that words with higher Stage 1 values will be more likely to come to mind during the decision. To test this, we will estimate a logistic mixed effects regression model, regressing people’s answer to whether each word came to mind on the Stage 1 value of the word. We will estimate both random intercepts and slopes, and calculate a one-tailed p value for the Stage 1 value coefficient.

(B) We predict that, of the words which came to mind during the decision, words with higher Stage 2 values will be more likely selected. To test this, we will fit a multinomial logistic regression model to people’s choices (using the R package “mlogit”), regressing Stage 2 choice on Stage 2 value. (We will omit word-specific intercepts. The regression analysis will only include, for each participant, the words that the participant reported coming to mind during the decision.) We will calculate a one-tailed p value for the Stage 2 value coefficient.

*How many observations will be collected?*

N = 500 (before exclusion). Our simulations indicated that this sample size would give us over 80% power.

*Anything else you would like to preregister?*

We will conduct three additional analyses. First, we will test whether Stage 2 value influences which words people consider. To test this, we will estimate the same logistic mixed effects model as in (A), except with Stage 2 value instead of Stage 1 value as the regressor. Second, we will test whether Stage 1 value influences selection out of the consideration set. To test this, we will fit the same multinomial logistic regression model as in (B), except with Stage 1 value instead of Stage 2 value as the regressor. Third, we will

We will exclude participants for whom any of the following is true: They don't complete the study, they choose the better alternative in Stage 1 training on less than 70% of trials, they fail to give a word within the time limit in Stage 2, they fail the Stage 2 comprehension checks, or they write things down physically during the experiment (as measured by a probe at the end).