

Extending the SIMPLE model to forced-choice tasks

Sean Conway

5/10/2022

Introduction

Historically speaking, computational models of memory have typically fallen into one of two camps: long-term or short-term memory. Brown, Neath, & Chater (2007) attempted to bridge this divide with the Scale-Invariant Memory, Perception, and Learning (SIMPLE) model, which posits a single system that can account for a number of empirical memory effects.

SIMPLE proposes that the timescale of memory is best understood on a logarithmic scale. Transforming the scale to time causes greater “compression” for events further in the past compared to those more recently experienced. Figure 1 shows an example.

In this example, each point on the graph is an experienced event (e.g., words seen in a memory experiment). The x axis represents the number of seconds since the words were studied, while the y axis represents the logarithmically transformed number of seconds since study. Note that both pairs of dots have the same distance between them on the x axis. Then, note that on the y axis, the blue dots are further spread out than are the red dots. Brown et al. (2007) referred to this as “Weberian compression” after the famous psychophysicist Ernst Weber. Psychophysics research had long established Weberian compression in perception research, but Brown et al. (2007)’s extended the phenomenon to the timescale of memory. Crucially, their SIMPLE model proposes that more recent events are more psychologically discriminable from one another compared to more distant events. In terms of the Figure 1 example, people should be better at discriminating the red dots compared to the blue dots.

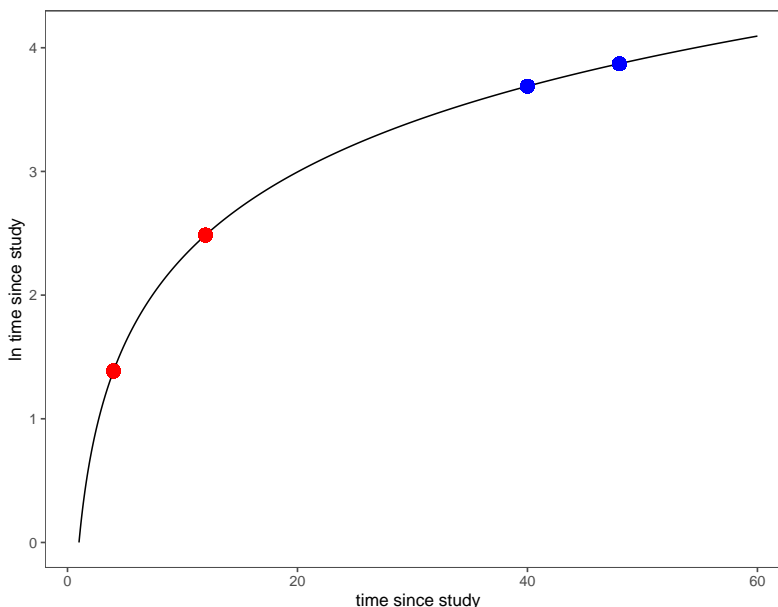


Figure 1: Weberian compression example.

The SIMPLE model borrows from work in categorization, identification, and choice modeling (Debreu, 1960; Nosofsky, 1986). The model proposes that stored memories differ from one another on at least

one dimension: a temporal dimension.¹ A memory’s value on this dimension is simply how much time has passed since the event was experienced.

SIMPLE proposes that one memory’s discriminability (the extent to which the observe can tell it apart from other memories) is based on its similarity to all other memories. The authors computed similarity as an exponentially decreasing function of (log-transformed) temporal distance. This measure is a well-known method of computing similarity (Shepard, 1987).

Given a set of n items in memory, each item’s discriminability can be computed as:

$$D_i = \frac{1}{\sum_{j=1}^n \eta_{i,j}} \quad (1)$$

where $\eta_{i,j}$ is the similarity between the canonical stimulus i and a given stimulus j and is computed as:

$$\eta_{i,j} = e^{-c*|Mi-Mj|^\alpha} \quad (2)$$

where c (often thought of as sensitivity to psychological distance) is a free parameter of the model and α is a constant that indicates which distance metric to be used. Brown et al. (2007) used $\alpha = 1$, known as a city-block distance metric. All simulations reported in this paper use city-block distance.

The model can make predictions for various types of memory tasks, including serial ordering and free recall. However, the equations for these tasks will not be shown, as the focus of this report is to demonstrate an application of this model to a novel paradigm: 2 & 3-alternative forced choice (2 & 3 AFC). Before proceeding to the work at hand, however, we should note two important effects the model can account for: primacy and recency in free recall.

In a free recall task, participants attempt to list studied items in their canonical positions, with no cue (that is, relying only on their memory). Murdock Jr (1962) first demonstrated primacy and recency effects in free recall (see Figure 2). Primacy occurs when the item in the first position is correctly recalled more than its neighbors, while recency occurs when the most recently studied item is recalled correctly more often than any item in the study set. These effects are well-documented in memory research and are crucial for understanding the predictions of the model proposed by this report.

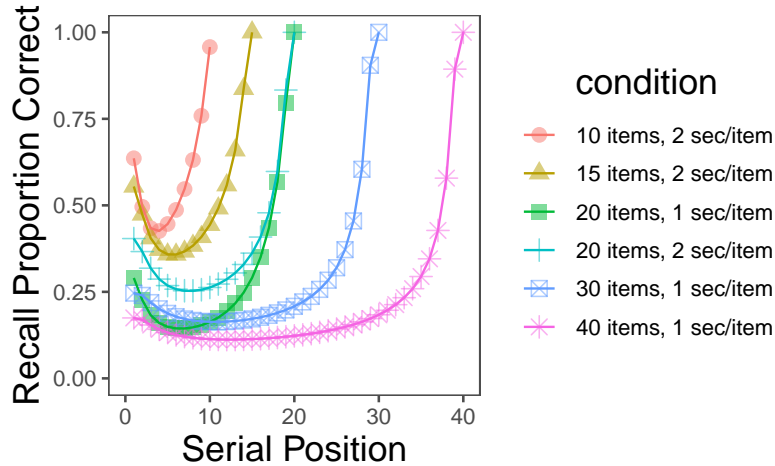


Figure 2: Murdock 1962 serial position effects/SIMPLE predictions.

This project’s purpose is to extend the SIMPLE model to 2 & 3 AFC tasks and explore its predictions. However, before exploring model details, we must understand the hypothetical choice task.

¹While Brown et al. (2007) included simulations with multiple dimensions, this report will focus exclusively on the temporal dimension.

Choice Task

The proposed experiment will consist of a number of trials. Each trial will begin with a study phase, wherein participants study a list of stimuli (e.g., words), presented sequentially. After a short time passes (i.e., 20-30 seconds), the participant will see two (or three) stimuli from the list. They will then identify which word was shown in position **X**.

For example, imagine the participant saw the list [hope, bear, coin, cars, town, palm, tarp, wane, lake, brow]. At test, the participant saw the choice set [bear, coin] along with the prompt “Please identify the word that was shown in position 2”. Ideally, the participant will choose “bear” more often than “coin”, though with some non-zero probability they will choose incorrectly (as is common in memory experiments).

This hypothetical task will continue for a number of trials (enough to achieve stable choice probabilities at the individual participant level). The term “test position” will be used to refer to the item position participants respond to. The term “canonical item” will be used to the true studied item at a given position. For example, in the above scenario, the test position is 2, and the canonical item is “bear”.

Proposed Model

To generate predictions, we must compute the similarity of each item in the choice set to the item in the studied position (using Equation 2). To compute these similarity values for the above prompt (bear vs. coin), we assume that “bear” was studied in position 2, 33 seconds before test, and “coin” was studied in position 3, 31 seconds before test.

We then compute $\eta_{bear,bear}$ as:²

$$\eta_{bear,bear} = e^{-c*|log(33)-log(33)|} = 1 \quad (3)$$

We next compute $\eta_{bear,coin}$ as:

$$\eta_{bear,coin} = e^{-c*|log(33)-log(31)|} = .47 \quad (4)$$

To generate choice predictions, we use Luce’s Choice Rule (LCR) (Debreu, 1960). LCR predicts that an item’s choice probability is proportional to its value (or “utility”) compared to the utility of all items in the choice set.

For example, according to LCR, the probability that a decision-maker selects item A out of a choice set $[A,B]$ is:

$$P(A|A, B) = \frac{U_A}{U_A + U_B} \quad (5)$$

where U_i indicates the utility of option i .

Returning to the choice task at hand, the probability of selecting option i in a choice set of size n , given test position k , can be computed (using LCR) as:

$$P(i|i...n) = \frac{\eta_{i,k}}{\sum_{j=1}^n \eta_{j,k}} \quad (6)$$

that is, the numerator is similarity of option i to the canonical item k , while the denominator is the summed similarity of *all* options to the canonical item k .

Finally, we can compute the probability of responding “bear” in the above example as:

$$P(bear|bear, coin) = \frac{1}{1 + .47} = .68 \quad (7)$$

The probability of responding “coin” follows logically:

$$P(coin|bear, coin) = \frac{.47}{1 + .47} = .32 \quad (8)$$

Note that the choice probabilities for a given choice set will always sum to 1, given that choices are forced.

²For this and all subsequent model equations, c is set to 12.1 per Brown et al. (2007)

The Stimuli

We use the Murdock Jr (1962) stimulus values, as described by Brown et al. (2007). These stimuli consisted of 10 studied words. We assume that there is once choice prompt per study set. We also assume that this prompt comes 20 seconds after onset of the last stimulus, as also assumed by Brown et al. (2007). Finally, in all choice sets, the correct item will **always** be in the choice set. We show the raw and log transformed stimulus values in Figure 3.

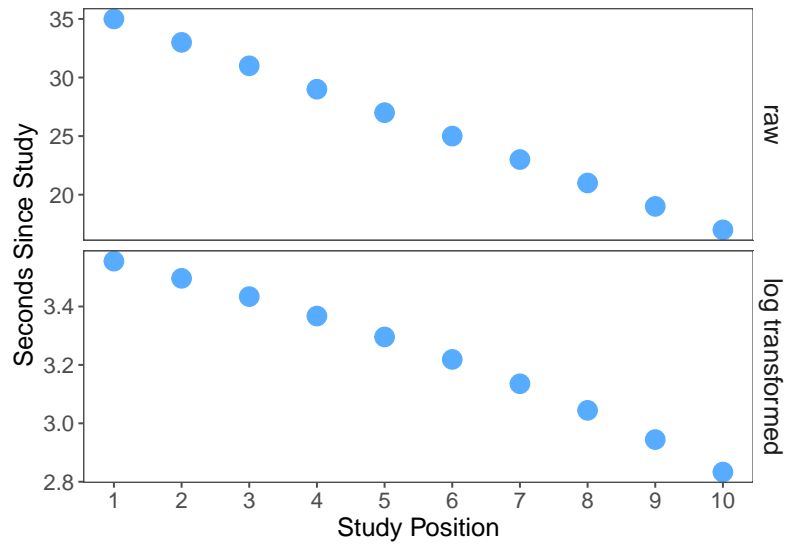


Figure 3: Stimulus values: raw and log transformed timescales.

Model Predictions

2-Backwards Prompt

For the first set of model predictions, referred to as the “2-Backwards Prompt”, the choice set includes each item in its canonical position and the item studied in the position before (except for position 1, which did not have a prior item). Predictions are shown in Figure 4. Test position is indicated on the x-axis, choice probability is indicated on the y-axis.

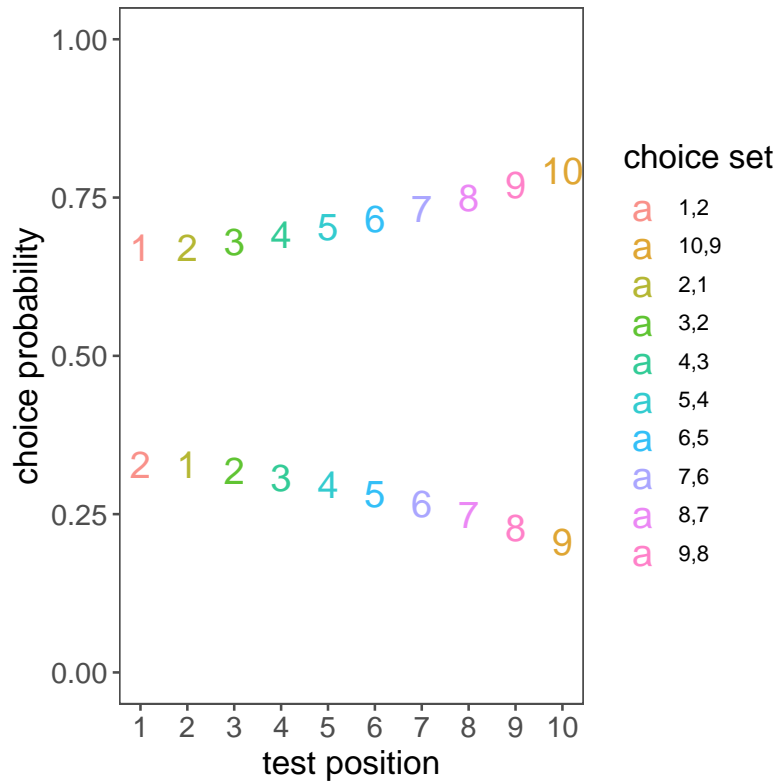


Figure 4: Model predictions - 2-Backwards Prompt

The model predicts a recency effect (higher canonical choice proportion for item 10, the last item seen) but not a primacy effect (no advantage for item 1, the first item seen). This is due in part to the fact that item 1 has no prior neighbor. Item 10 is highly discriminable from its neighbor, item 9, while item 1 is less discriminable from its neighbor, item 2 (which follows rather than precedes it).

2-Forwards Prompt

Next, the choice set includes each item in its canonical position and the item studied after (except for position 10, which did not have a neighbor after it). Predictions are shown in Figure 5.

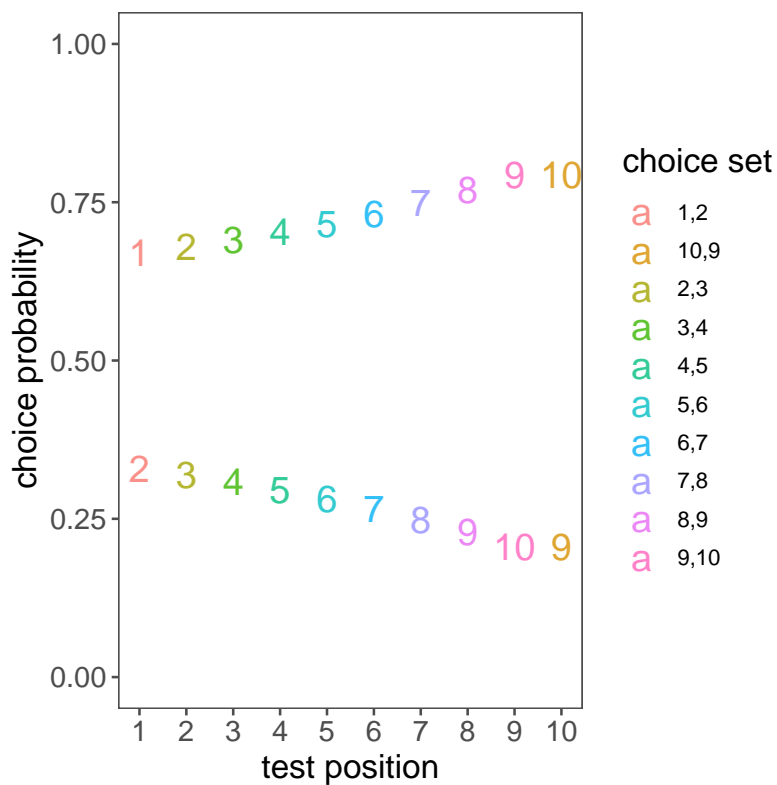


Figure 5: Model predictions - 2-Forwards Prompt

The model no longer predicts a recency effect in this scenario. While items later in the study set have higher choice proportions than those earlier in the study set, the advantage for item 10 is now eliminated. This absence of an advantage comes from the identical choice set for positions 9 and 10. This removes any recency advantage garnered from item 10.

3-Backwards Prompt

Next, the choice set includes each item in its canonical position and the two preceding items (except for items 1 & 2). Predictions are shown in Figure 6.

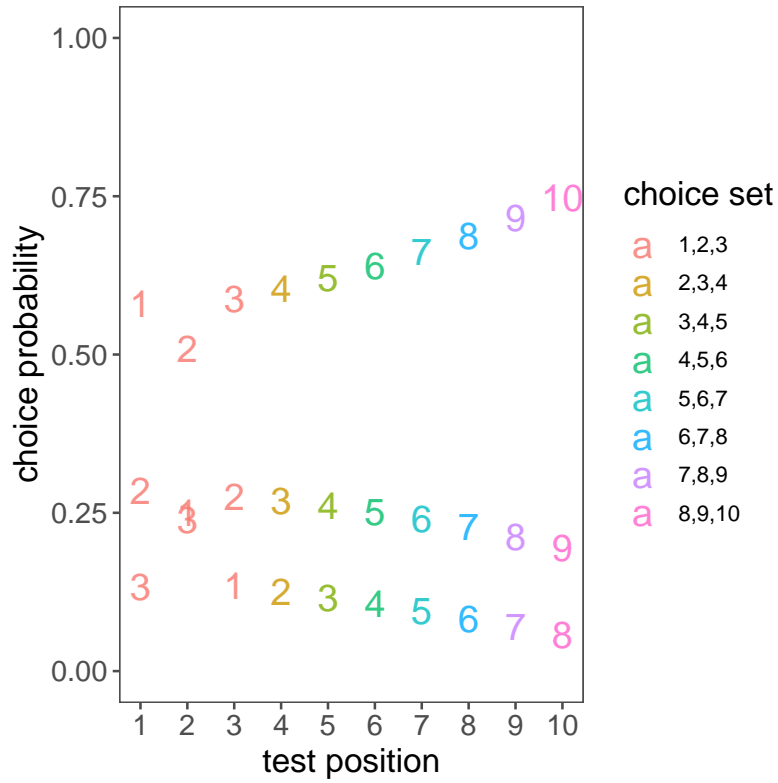


Figure 6: Model predictions - 3-Backwards Prompt

Here, the recency effect returns, as item 10 is more discriminable from the 2 preceding items than any other canonical item in the study list (due to Weberian compression). A small primacy effect appears, wherein item 1 has a small advantage over item 2.

3-Forwards Prompt

Next, the choice set includes each item in its canonical position and the two following items (except for positions 9 & 10). Predictions are shown in Figure 7.

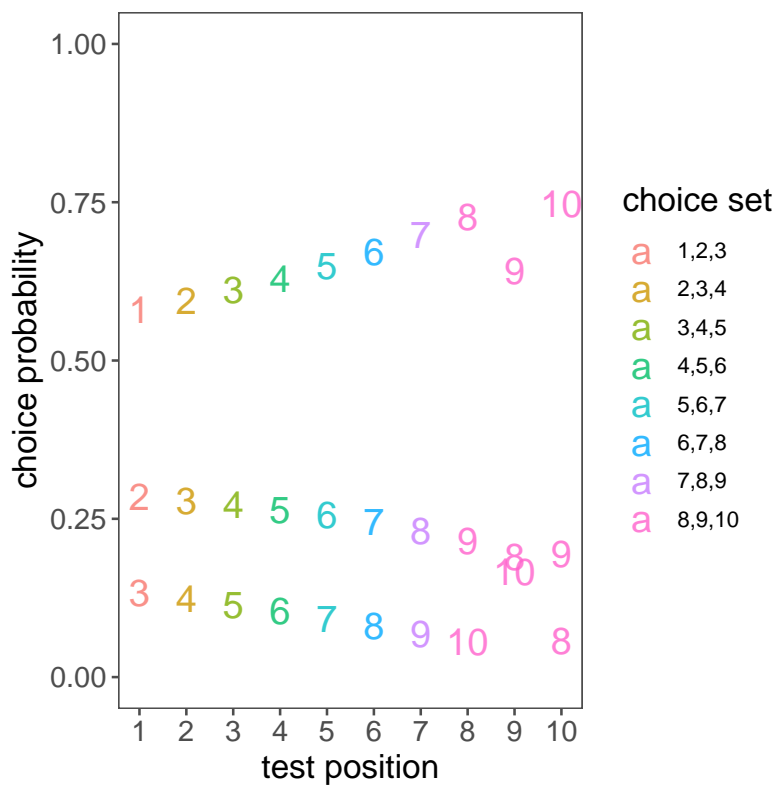


Figure 7: Model predictions - 3-Forwards Prompt

No primacy effect appears, as Weberian compression “hurts” the first studied item more than any other item. A small recency effect appears as well. Interestingly, a slight dip in choice proportion is seen at item 9. This occurs because item 9 is tested against the items before *and* after it. These items are more confusable to item 9 than are the other items in the other choice sets and therefore cause a drop in choice probability for item 9.

3-Middle Prompt

Next, the choice set includes each item in its canonical position along with its neighbors to either side (except for positions 1 & 10). Predictions are shown in Figure 8.

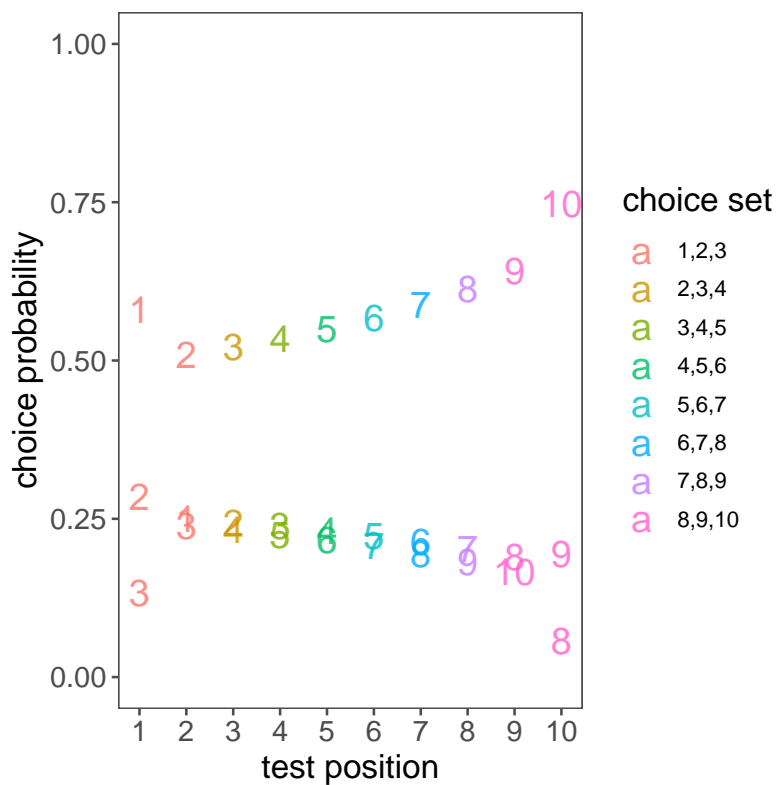


Figure 8: Model predictions - 3-Middle Prompt

Here, we see both primacy and recency effects, though these effects only appear because items 1 & 10 each lack a neighbor to one side. This makes the third item in the choice set (i.e., item 3 for position 1 & item 8 for position 10) less confusable than is the norm in these trials. This leads to a jump in probability for both items 1 and 10.

Primacy vs. Recency

Lastly, we explore the primacy-recency issue in greater depth. We tested the first and last item in the study set (e.g., items 1 & 10) against all other items in the choice set. Predictions are shown in Figure 9. Here, x-axis refers not to test position but rather to comparison option. For example, if the x axis value is 4, this indicates that both items 1 and 10 were tested in their canonical positions against item 4.

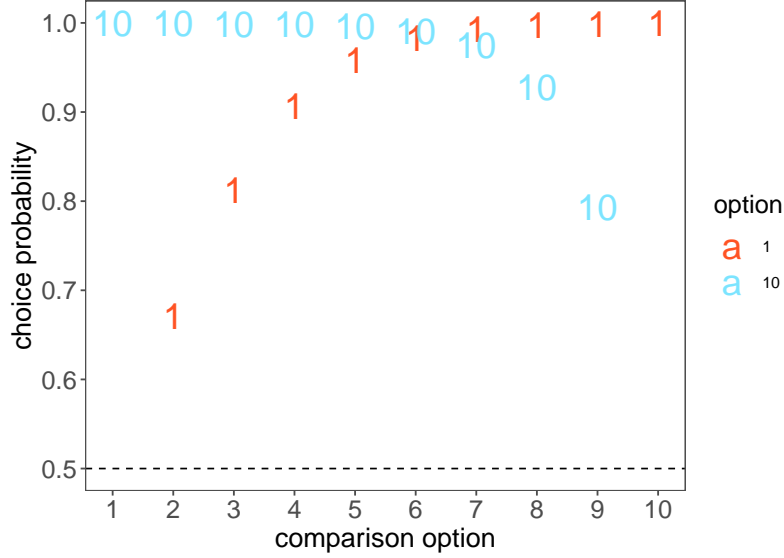


Figure 9: Model predictions - Primacy/Recency

Here, the recency advantage on display. When compared to their nearest neighbors (i.e., item 2 for item 1 & item 9 for item 10), item 10 has a distinct advantage over item 1. This occurs due to Weberian compression. When we plot item-item similarity (rather than choice probability; see Figure 10), we see that item 1 is more similar to item 2 than is item 10 to item 9. This is because items that follow a stimulus are less discriminable than those that precede it, as well as because items further in the past are less discriminable than those more recently seen.

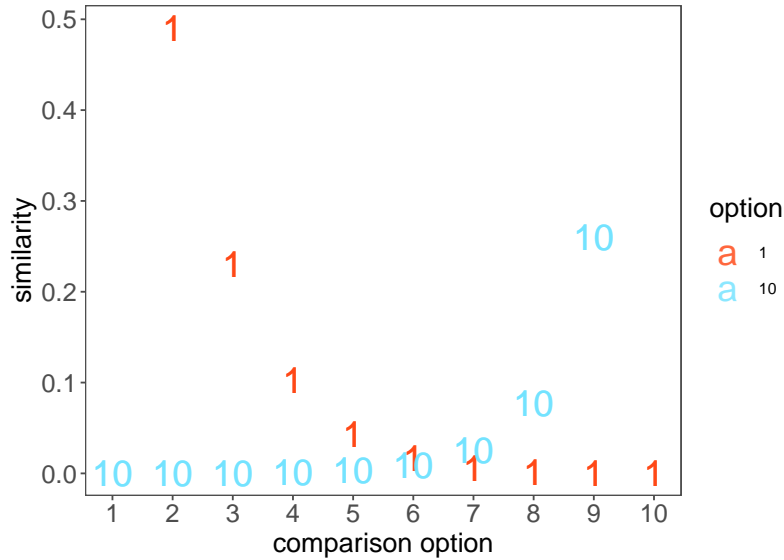


Figure 10: Similarity values - Primacy/Recency

Conclusion

We extended SIMPLE to a novel task in which participants make forced choices regarding which of a set of items was studied in a particular position. Model predictions yielded meaningful and testable results. Notably, primacy and recency effects (a well-observed phenomenon in psychology) can disappear and reappear depending on the choice set.

Future modeling work should extend this work to different types of choice sets (e.g., when the correct option is not in the set) and to different types of choice problems (e.g., when options are presented but a choice is not forced). Finally, the benefit of this work is to generate testable predictions, so empirical tests should take place.

References

- Brown, G. D., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, 114(3), 539.
- Debreu, G. (1960). Review of RD Luce, individual choice behavior: A theoretical analysis. *American Economic Review*, 50(1), 186–188.
- Murdock Jr, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, 64(5), 482.
- Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, 115(1), 39.
- Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, 237(4820), 1317–1323.