SVDW Recognition: Preregistration

Kaitlyn Fallow (kmfallow@uvic.ca)

2019-03-27

Contents

Background

This brief document describes the full follow-up version of the pilot study described here: https://osf.io/5yf6e/

After pilot testing 10 participants in one-on-one sessions, an error was fixed in the stimulus numbering that may have led to 1-2 pictures being presented in both word and picture formats for a couple of participants. Other than this, there were no noticeable issues in the data – overall performance was high but not worryingly so, with no instances of hit or correct rejection rates at ceiling and most falling below 90%. Because there were only a few weeks of data collection left in the semester at the point of this initial analysis, we decided to move ahead to the "full" version and attempt to get as many participants as possible rather than proceeding with the originally planned pilot sample of 20-25. Methodological details remain the same, except participants are now mostly being tested in groups (group sizes vary up to a maximum of 10/session) and participants are not verbally interviewed for feedback at the end.

Hypotheses

The primary hypotheses are that, unlike the words and pictures we have used previously, the SVDW stimuli will *not* produce a significant materials-based difference in recognition memory response bias, nor will the bias for pictures be significantly conservative.

We also anticipate the usual picture superiority effect obtained with these stimuli, but this is not of central interest. In the pilot study this trend was apparent in d' and both hit and false alarm rates but was not statistically evaluated.

Planned analyses

Our primary dependent measures of interest will be hit and false alarm rates (determined by collapsing our 1-6 response scale into a binary one, with 1-3 = "new" and 4-6 = "old"), and corresponding signal detection theory (SDT)-based measures of response bias (c) and sensitivity (d'). Ceiling and floor rates will be replaced according to Macmillan & Kaplan (1985) to enable calculation of c and d'. Given long-known issues with the application of c and d' to recognition memory data, namely the tendency for data to be inconsistent with the assumption of equal variance of the old and new item distributions (conditions under which c and d' become confounded), we will also conduct receiver operating characteristic (ROC) analyses with the goal of estimating the extent of this violation in our data and calculating corrected measures of

sensitivity (da and de) and bias (ca and ce). The plan is to apply these corrections at the participant level based on participant-level ROCs, but if results suggest these curves are ill-fitting (e.g., if a lot of people have primarily used the extreme ends of the confidence scale) we will instead apply corrections based on aggregate by-condition ROCs. These alternate measures will be subject to the same analyses as c and d', and all will be reported; for brevity, I will refer only to c and d'.

Because our primary hypotheses regarding response bias (c) are null hypotheses, these data will be subject to equivalence testing with bounds of -0.2 > d < 0.2. We may also supplement these analyses with Bayes factors.

Conventional paired t-tests (alpha = .05) on d', hit rates, and false alarm rates will evaluate the picture superiority effect.

In addition to ROCs, we will likely conduct a number of other exploratory analyses, such as evaluating our dependent measures at the test quartile level and for individual items, in the latter case using generalized linear mixed models.

Exclusion criteria

Participants with d' < 0.2 (in either condition) will be excluded from further analyses; performance using this experimental setup is usually quite high, so this cutoff is intended to exclude participants who are responding essentially at random (or who adopt a strategy of allocating attention to only pictures or words).

I have no *a priori* plan to exclude outliers. If data checks (e.g. frequency histograms) suggest there are some that may be unduly influential, these will be dealt with on a case-by-case basis, most likely by running the analyses of primary interest with and without them and reporting both sets of results.

References

Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin*, *98*(1), 185–199. https://doi.org/10.1037/0033-2909.98.1.185