The Journal of Open Source Software

Psifr: Analysis and visualization of free recall data

Neal W Morton[1]

1 Research Fellow, Center for Learning and Memory, The University of Texas at Austin

DOI: 10.21105/joss.02669

Software

- Review
- Repository
- Archive

Editor: Christopher R. Madan

Reviewers:

- @samhforbes
- @paxtonfitzpatrick

Submitted: 28 August 2020
Published: 15 October 2020

Summary

Research on human memory has been strongly influenced by data from free recall experiments, wherein participants study a list of items (such as words) and then freely recall them in any order they wish (Murdock, 1962). Free recall provides an extremely rich dataset that not only reflects which items were recalled but also the order in which they were recalled. However, analysis of free recall data is difficult, as many different influences on recall must be taken into account. For example, one influential analysis, conditional response probability as a function of lag (lag-CRP), has been used to measure the tendency of participants to successively recall items that were originally presented near to each other in time (Kahana, 1996). This analysis requires taking into account the items that are still available for recall at each transition between recalled items. This analysis may need to be made conditional on other factors, such as the category of the items being recalled (Morton & Polyn, 2017; Polyn, Erlikhman, & Kahana, 2011), thus complicating the analysis further.

`Psifr` (pronounced like "cypher") was developed to consolidate a number of free recall analysis methods (often implemented in MATLAB) within a flexible Python package. The `Psifr` package includes core utilities that simplify and standardize a number of different analyses of recall sequences, including analyses focused on serial position (Murdock, 1962), temporal order (Kahana, 1996; Polyn et al., 2011), stimulus category (Morton & Polyn, 2016; Polyn, Norman, & Kahana, 2009), and the semantic meaning of presented items (Howard & Kahana, 2002). The core utilities are also designed to facilitate implementation of extensions to tailor analyses for specific experiments.

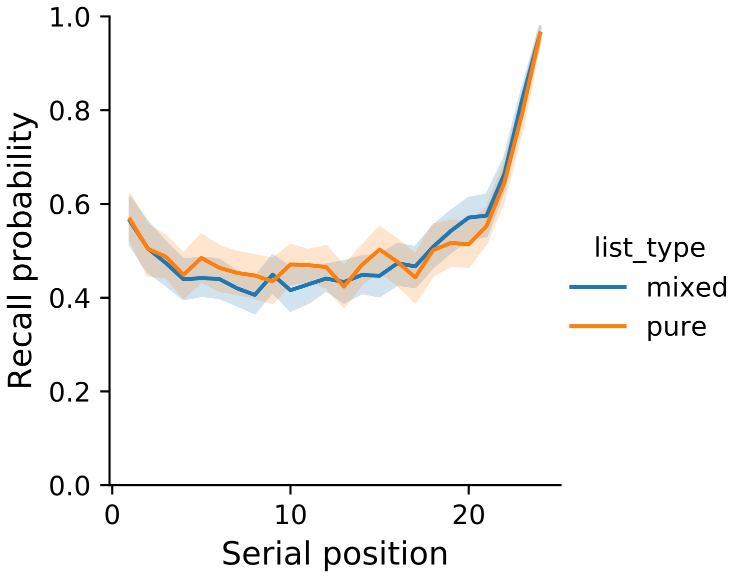

Figure 1: Example of a serial position curve showing the probability of recalling an item based on its position in the list. Plots may be flexibly divided by condition using grouping semantics supported by Seaborn. In this case, different list types (mixed-category or pure-category) are plotted as separate curves.

Statement of Need

Existing packages for analysis of free recall data include `EMBAM` and `Quail`. EMBAM, which is a fork of the Behavioral Toolbox, is implemented in MATLAB, making it difficult to use with the extensive data science ecosystem in Python. The pybeh package is a Python port of `EMBAM` written using `numpy`. As it is a fairly direct port of EMBAM, `pybeh` does not make use of some of the advantages of Python, such as the advanced data science packages of `Pandas` and `Seaborn` (Reback et al., 2020; Waskom et al., 2020). `Quail`, a Python package, provides some similar functionality to `Psifr`, including analysis of recall order (Heusser, Fitzpatrick, Field, Ziman, & Manning, 2017). However, while `Quail` uses a custom data structure to store free recall sequences, `Psifr` uses `Pandas DataFrame` objects. This design makes it possible for the user to make full use of the split-apply-combine operations of `Pandas` to quickly run complex analyses.

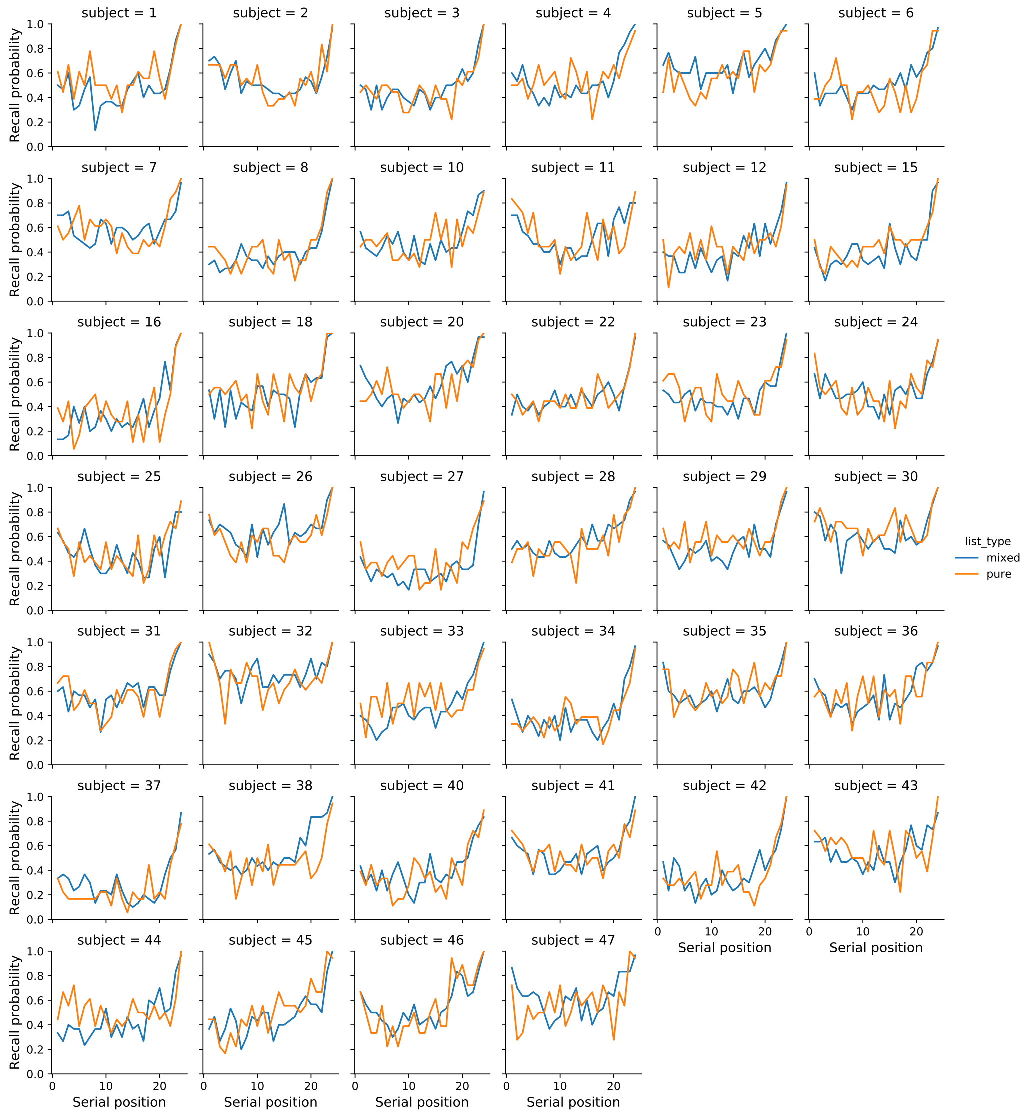

Figure 2: Serial position curve split by list type, with a separate panel for each participant in an experiment.

Similarly, `Psifr` makes available the full power of the `Seaborn` visualization package to provide expressive visualization capabilities. The plotting functions in `Psifr` allow the user to easily view analysis results in different ways; for example, an analysis of recall by serial position can be visualized either as a single plot with error bars (Figure 1) or as a grid of individual plots for each participant in the experiment (Figure 2). `Psifr` also supports creation of raster plots (Figure 3), a method for visualizing whole free recall datasets to facilitate quick discovery of patterns in the order of recalls (Romani, Katkov, & Tsodyks, 2016).

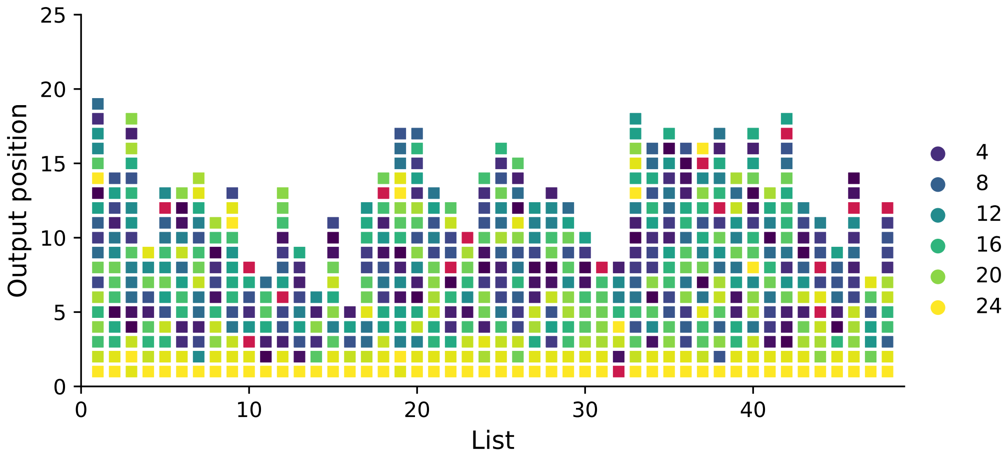

Figure 3: Raster plot displaying the order of every recall for one participant. Each marker indicates one recall, and the color of the marker reflects the serial position of the recalled item. Red markers indicate intrusions of items not on the studied list.

`Psifr` was designed to be used by memory researchers and students. It is currently being used in two ongoing projects that require advanced analysis and visualization. The interface is designed to simplify common tasks while also allowing for substantial customization to facilitate analysis of specific episodic memory experiments. Advanced visualization further helps to support better understanding of complex datasets. The source code for `Psifr` has been archived to Zenodo with the linked DOI: 10.5281/zenodo.4086188.

Acknowledgements

Thanks to Sean Polyn for helpful discussions. `Psifr` is inspired by functions initially developed for `EMBAM`, which was developed by Sean Polyn, Neal Morton, Richard Lawrence, Joshua McCluey, Karl Healey, Lynn Lohnas, and Jin Jeon. `Psifr` also takes inspiration from the `Quail` package developed by Andrew Heusser, Paxton Fitzpatrick, Campbell Field, Kirsten Ziman, and Jeremy Manning.

References

Heusser, A. C., Fitzpatrick, P. C., Field, C. E., Ziman, K., & Manning, J. R. (2017). Quail: A python toolbox for analyzing and plotting free recall data. *The Journal of Open Source Software*, *2*(18). doi:10.21105/joss.00424

Howard, M. W., & Kahana, M. J. (2002). When Does Semantic Similarity Help Episodic Retrieval? *Journal of Memory and Language*, *46*(1), 85–98. doi:10.1006/jmla.2001.2798

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, *24*(1), 103 109. doi:10.3758/bf03197276

Morton, N. W., & Polyn, S. M. (2016). A predictive framework for evaluating models of semantic organization in free recall. *Journal of Memory and Language*, *86*, 119 140. doi:10.1016/j.jml.2015.10.002

Morton, N. W., & Polyn, S. M. (2017). Beta-band activity represents the recent past during episodic encoding. *NeuroImage*, *147*, 692 702. doi:10.1016/j.neuroimage.2016.12.049

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*(5), 482 488. doi:10.1037/h0045106

Polyn, S. M., Erlikhman, G., & Kahana, M. J. (2011). Semantic cuing and the scale insensitivity of recency and contiguity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 766. doi:10.1037/a0022475

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review*, *116*(1), 129 156. doi:10.1037/a0014420

Reback, J., McKinney, W., jbrockmendel, Bossche, J. V. den, Augspurger, T., Cloud, P., gfyoung, et al. (2020). *Pandas-dev/pandas: Pandas 1.1.1*. Zenodo. doi:10.5281/zenodo.3993412

Romani, S., Katkov, M., & Tsodyks, M. (2016). Practice makes perfect in memory recall. *Learning & Memory*, *23*(4), 169–173. doi:10.1101/lm.041178.115

Waskom, M., Botvinnik, O., Ostblom, J., Gelbart, M., Lukauskas, S., Hobson, P., Gemperline, D. C., et al. (2020). *Mwaskom/seaborn: V0.10.1 (april 2020)*. Zenodo. doi:10.5281/zenodo.3767070