How you see is what you get: A deep learning look at how cognitive state is variably informed by key aspects of raw eye movement data

Zachary J. Cole, Karl M. Kuntzelman, Michael D. Dodd, & Matthew R. Johnson

BACKGROUND

The relationship between mental state and eye movement patterns was first demonstrated by Yarbus [1] (see Figure 1). But more recent classification models have faced mixed results [2].

Purpose

Systematically characterize the contributions of eye movement variables (x-coordinates, y-coordinates, pupil size) and data types (image, timeline) to cognitive task predictions.

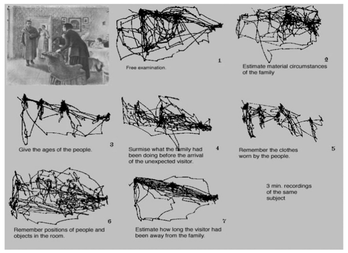


Figure 1. Figure from MacInnes et al. [3], originally adapted from Yarbus [1].

DATASET

Experiment 1

- Participants (N = 124) searched, memorized, or rated scene images (Figure 2).
- Eye movements during the first 6s of each trial were tracked with an SR Research EyeLink 2 eye tracker (1000Hz).
- After removing bad trials, N = 12634 trials were analyzed.

Experiment 9

- A separate set of participants (N = 77) completed the same tasks as the participants in Experiment 1.
- After removing bad trials, N = 8301 trials were analyzed.

Deep Learning Classification

- Implemented a convolutional neural network classifier using DeLINEATE, a deep learning toolbox [4].
- Data were split: Training: 70%; Test: 15%; Validation: 15%.
- Eye tracking coordinates were converted to Plot Images (Figure 3).
- Timeline Eye Tracking data were classified using X and Y coordinates, and Pupil Size data. Additionally, these data were systematically classified with No X, No Y, and No Pupil Size information.

Example Scenes

Figure 2. Scenes did not show any people or faces.

Example Images

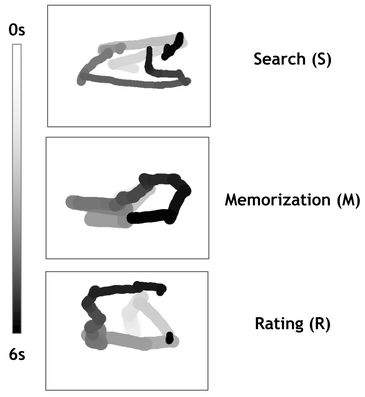
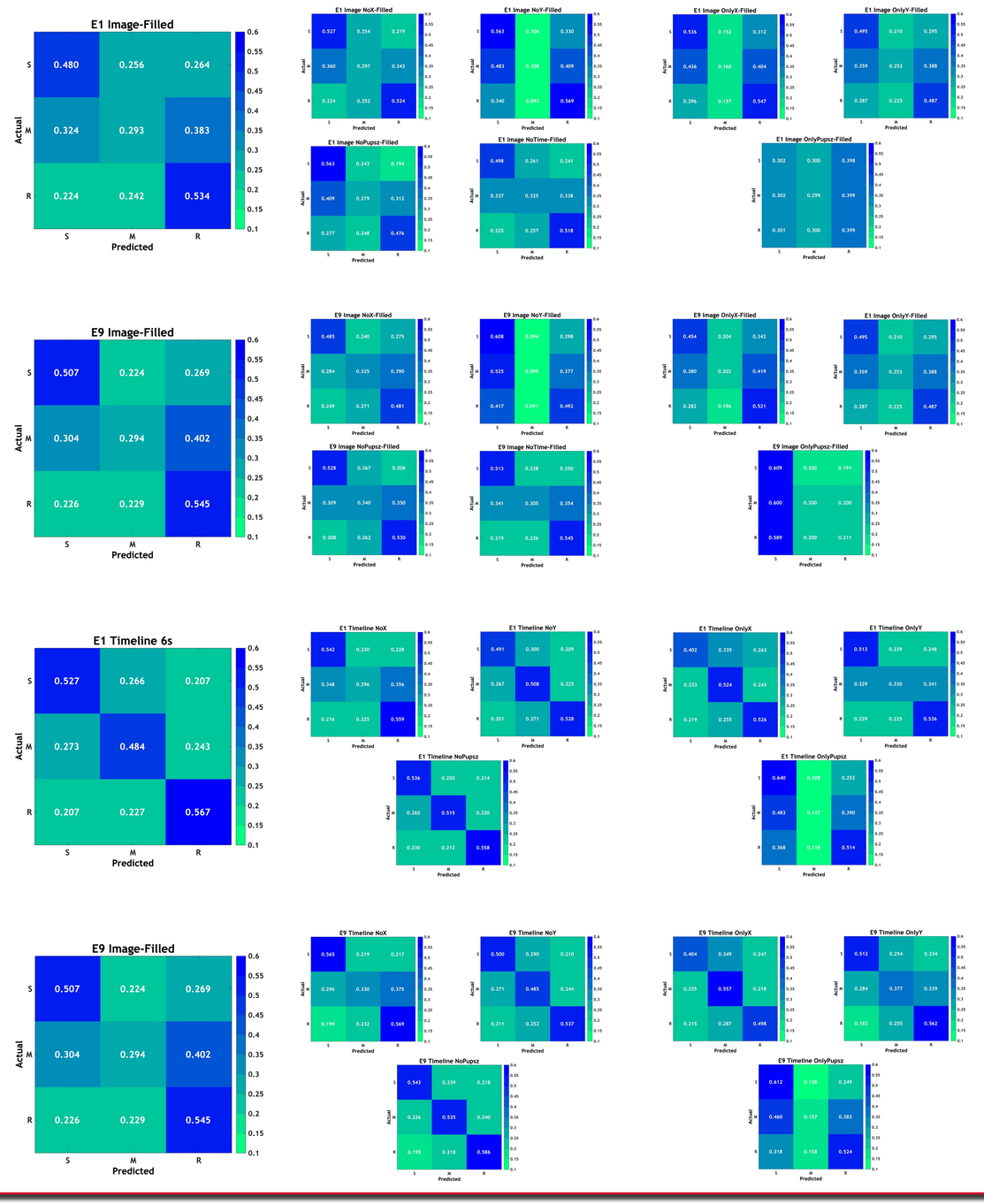


Figure 3. Plot diameter indicates pupil size measurements.

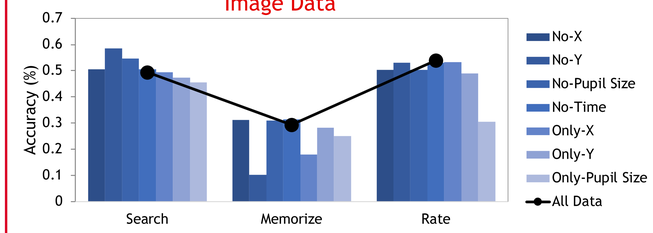
CLASSIFICATION

- Confusion matrices (below) represent the probability of the actual trial type being classified (predicted) as a Search, Memorization, or Rating trial.

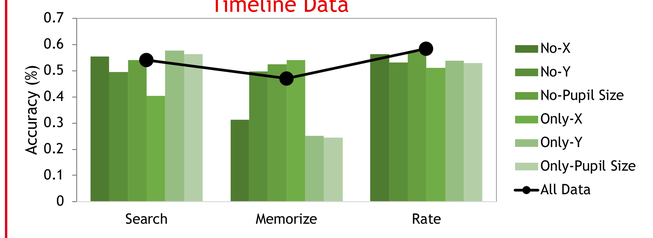


ACCURACY

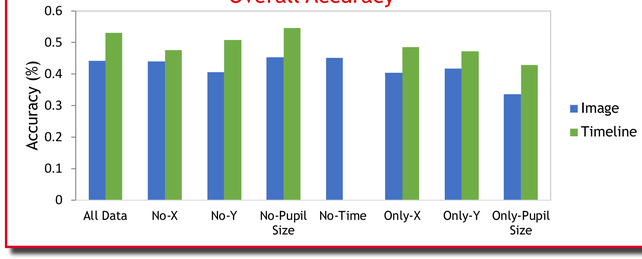
Image Data



Timeline Data



Overall Accuracy



CONCLUSIONS

- Timeline data more accurate than image data — potentially due to information loss from the overlap of data in the images.
- Memorization most often confused with other tasks.
- Pupil size and time are least informative variables. Y-coordinates are most informative.
- The shape of scene images and the natural distribution of objects within these images could be a factor in the apparent importance of the horizontal eye movements.
- These findings suggest that deep learning models can extract a surprising amount of useful information from nearly-raw eye tracking data with minimal human guidance.

REFERENCES & ACKNOWLEDGEMENTS

[1]Yarbus AL. Eye Movements and Vision, New York: Plenum. (Originally published in Russian 1965); 1967.

[2]Greene MR, Liu T, Wolfe JM. Reconsidering Yarbus: A failure to predict observers' task from eye movement patterns. *Vision Research*. 2012; 62: 1-8.

[3]MacInnes JW, Hunt AR, Clarke ADF, Dodd MD. A generative model of cognitive state from task and eye movement. *Cognitive Computation*. 2018; 10: 703-717.

[4]DeLINEATE: A deep learning toolbox for neuroimaging data analysis. Kuntzelman K, Williams JM, Samal A, Rao PK, Johnson MR. 2019. *Cognitive Neuroscience Society 26th Annual Meeting*, Poster B105. http://delineate.it

**Supported by NSF/EPSCoR grant #1632849 and NIH grant GM130461 to MRJ and colleagues