Neuropsychologia 93 (2016) 128–141

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

A single trial analysis of EEG in recognition memory: Tracking the neural correlates of memory strength

Roger Ratcliff*, Per B. Sederberg, Troy A. Smith, Russ Childers

The Ohio State University, United States

ARTICLE INFO

Keywords:
Diffusion model
Recognition memory
EEG
Single-trial regressor
Reaction time

ABSTRACT

Recent work in perceptual decision-making has shown that although two distinct neural components differentiate experimental conditions (e.g., did you see a face or a car), only one tracked the evidence guiding the decision process. In the memory literature, there is a distinction between a fronto-central evoked potential measured with EEG beginning at 350 ms that seems to track familiarity and a late parietal evoked potential that peaks at 600 ms that tracks recollection. Here, we applied single-trial regressor analysis (similar to multivariate pattern analysis, MVPA) and diffusion decision modeling to EEG and behavioral data from two recognition memory experiments to test whether these two components contribute to the recognition decision process. The regressor analysis only involved whether an item was studied or not and did not involve any use of the behavioral data. Only late EEG activity distinguishes studied from not studied items that peaks at about 600 ms following each test item onset predicted the diffusion model drift rate derived from the behavioral choice and reaction times (but only for studied items). When drift rate was made a linear function of the trial-level regressor values, the estimate for studied items was different than zero. This showed that the later EEG activity indexed the trial-to-trial variability in drift rate for studied items. Our results provide strong evidence that only a single EEG component reflects evidence being used in the recognition decision process.

1. Introduction

In studies examining the neural processes used in decision making, it is often assumed that if a neural response varies as a function of some perceptual or memory variable, it tracks the evidence being used to make a decision. However, this view does not discriminate between early representations of stimuli and evidence extracted from them that is used in the decision process.

In previous work, single-trial analysis of EEG in concert with diffusion decision modeling of choice and reaction time (RT) has been applied to examine the neural components of perceptual decisions. In a face/car discrimination task, Philiastides et al. (2006) used a single-trial analysis, also known as multivariate pattern analysis (e.g., Blankertz et al., 2011; Pereira et al., 2009) that weighted signals from an array of electrodes to produce a single component value that represented how car-like or face-like was the EEG signal. Two components were obtained that tracked stimulus quality, one at around 180 ms and one at around 380 ms (by component, it is meant a significant single trial regressor signal as in Philiastides et al., 2006). Ratcliff et al. (2009) sorted the behavioral data for each condition into halves based solely on the EEG component value to which they fit the diffusion model. Drift rate (quality of evidence used in the decision process) differed for the two halves, but only for the late component. This suggested that the later component tracks information being used in the decision, but the earlier component represents the quality of the stimulus encoding from which decision-relevant information is extracted.

In the behavioral memory literature, there is considerable debate about whether one or two processes are involved in making recognition decisions. The single-process view, reflected in most computational modeling approaches, proposes that recognition decisions are made based on either a single source of information or on information from multiple sources that is combined into a single continuous measure of memory strength (Cohen et al., 2008; Dennis and Humphreys, 2001; Dunn, 2004; Gillund and Shiffrin, 1984; Hintzman, 1984; Shiffrin and Steyvers, 1997; Starns and Ratcliff, 2008, Starns et al., 2012; Wixted, 2007). By contrast, the dual-process view proposes that there are two distinct decision processes that are used in recognition– a sense of familiarity that is a continuous variable and an all-or-none recollection component based on details about the encoding event (Rotello et al., 2004; Yonelinas, 1994, 1997; see also Buchler et al., 2008).

Key evidence for the dual process view comes from studies

* Correspondence to: Department of Psychology, The Ohio State University, Columbus, OH 43210, United States.
E-mail address: ratcliff.22@osu.edu (R. Ratcliff).

http://dx.doi.org/10.1016/j.neuropsychologia.2016.09.026
Received 27 June 2016; Received in revised form 26 September 2016; Accepted 29 September 2016
Available online 29 September 2016
0028-3932/

examining event-related potential (ERP) components when subjects are making recognition decisions (Rugg, 1995; Rugg and Curran, 2007). Specifically, it has been argued that a frontal component occurring in the range of 300–500 ms after onset of a test item can be identified with familiarity while a parietal component occurring in the range of 400–600 ms after onset can be identified with recollection (see Eichenbaum et al., 2007; Rugg and Yonelinas, 2003; but see Paller et al., 2007).

In this article, we apply this single trial regressor and diffusion modeling approach to data from a recognition memory task. This allows us to determine whether these proposed neural correlates of the memory processes guide the recognition decision.

2. Methods

We report the results of two recognition memory experiments, the second experiment serving as an independent replication of the first experiment, with a simplified design. The two experiments both used a standard item recognition procedure in which a list of words is presented followed by a test list. For every test word, subjects have to decide whether it was studied or not. In many prior experiments, response time and accuracy data are well-explained by the diffusion decision model (Ratcliff et al., 2004, 2010, 2011). In Experiment 1 we manipulated number of repetitions of study words and word frequency of studied and new words and in Experiment 2 we manipulated number of repetitions of study words. These manipulations were done to provide constraints on modeling and to demonstrate that the experiments provided standard results from this task.

2.1. Subjects

All subjects were recruited from the university community at The Ohio State University and spoke and read English fluently. Subjects provided written consent in accordance with requirements of the local institutional review board and were paid $20 for their time. Experiment 1 comprised twenty-five right-handed subjects (14 female) between 18 and 38 years of age (M =21.6). Data from subjects for whom there were excessive motion artifacts or recording noise (n =5) or who failed to follow instructions (n =2) are not included in the analyses. Experiment 2 comprised fourteen right-handed subjects, however three were discarded due to experimental error, excessive motion artifacts, and/or recording noise, leaving 11 subjects (3 female) between 18 and 24 years of age (M =20.6).

2.2. Stimuli

Stimuli were drawn from the SUBTLEXUS database (Brysbaert and New, 2009). A total of 549 high-frequency words (between 40 and 400 occurrences per million, M=122.31, SD=84.52) and 553 low-frequency words (between 1 and 7 occurrences per million, M=1.44, SD =0.39) formed the stimulus pool. All words in the pool had between 4 and 7 letters, and word length was equated across the frequencies (M=5.03 and 4.99, respectively).

2.3. Design

The first experiment used a 2 (Word Frequency: high vs. low) × 3 (Item Strength: strong vs. weak vs. new) factorial within-subject and within-list design. Each subject studied and was tested on 13 lists of words. Each study list was constructed by randomly selecting 9 words of each frequency from the pool to be studied one time (the “weak” condition) and 9 words of each frequency to be studied three times (the “strong” condition). Thus, there were 36 unique words and 72 item presentations in each study list, with the order of the presentations pseudo-randomized such that there were no immediate repetitions of the strong items. A matching test list consisting of all 36 studied words along with a set of 36 matching lures was constructed for each study list. The order of words on the test lists was randomized in 2 blocks with the first 18 unique studied words tested in the first block and the second 18 tested in the second block. This was done to ensure that the end of list items were never tested immediately, thereby reducing the probability of recency effects. After discounting the first list, which was used for practice, this design yielded a total of 108 target trials and 108 lure trials for each experimental condition.

Experiment 2 had only the item strength manipulation, giving rise to a 3 factorial (strong vs. weak vs. new), within-subject and within-list design. Each study list was constructed from a pool excluding low-frequency words, but without the word length restriction applied in Experiment 1. Subjects studied and were tested on 11 lists, each comprising 15 strong (presented three times), 15 weak (presented one time), and two once-presented buffer words at the end of the list, giving rise to 62 item presentations. Each study list had a matching test list, where each of the 30 unique target words was presented in random order along with 30 lure words. The two buffer target items, along with two buffer lure items were tested at the start of the list to account for recency effects. After discounting the first list, which was used for practice, the design of Experiment 2 yielded 150 trials in each of the strong, weak, and lure conditions.

2.4. Equipment

For Experiment 1, a desktop computer with a 17″ LCD display was used to present the stimuli and collect subject responses. A custom program written using the Python experiment programming library (PyEPL; Geller et al., 2007) was used to generate the study lists for each subject, control the timing of the tasks, and record subject responses. For Experiment 2, a desktop computer with a 24″ LCD display, running at 120 Hz was used to present the stimuli and collect subject responses. A custom Python program written using the State Machine Interface Library for Experiments (SMILE; https://github.com/compmem/smile) generated the study lists, controlled the timing of the tasks, and recorded subject responses.

For both experiments, scalp EEG data were sampled and recorded reference-free at 1000 Hz using a DC-powered actiCHamp amplifier/analog-to-digital converter connected to a desktop PC equipped with PyCorder software (BrainProducts GmbH, Munich, Germany). Prior to receiving instructions, each subject was fitted with an elastic cap containing 64 active electrodes arrayed in an extended 10-20 layout. Electrode impedances were reduced to less than 25 K ohms in accordance with operating instructions for the actiCAP system (BrainProducts GmbH, Munich, Germany).

2.5. Procedure

After informed consent was obtained, subjects were fitted with the EEG cap, seated in front of a computer, and given instructions by the experimenter. Subjects were told that they would be studying lists of words and would be given a recognition memory test after each study list. They were informed that some of the study words would be repeated, but were not told that this was an experimental manipulation. Subjects were also instructed to try to avoid thinking back to previous words on the list during the study session. Subjects were then given the first study and test list as a practice list. The experimenter answered any questions and then started the experiment proper.

The precise timing of study and test item presentations was similar, but not identical for the two experiments. In Experiment 1, prior to each study list, an orientation cross was displayed in the center of the screen for 2600–3000 ms. Then the list of words were displayed, one at a time, in the center of the screen for 1600 ms followed by a blank screen for 300–700 ms. Following the study list, an orientation cross was displayed for 1000–1200 ms to signify the test was about to begin. On each test trial, the probe word was displayed at the center of the