Erroneous classification of false memories as veridical: Possible limits of ERP based memory assessments

Ralf Mertens, John J.B. Allen

The University of Arizona

Introduction

ERPs hold promise for assessing memory in clinical and forensic settings, providing high sensitivity and specificity to identify familiar material in individual subjects. Traditional methods of analysis (e.g. ANOVA, however, are not adequate for providing outcomes for *individual* subjects.

Bayesian classification provides a methodology to classify stimuli as familiar or unfamiliar to individual subjects, based solely on features of the ERP waveform (Allen, Iacono, & Danielson, 1992). It is untested, however, whether such ERP assessment procedures produce false positive errors in response to events individuals believe to have experienced when in fact they did not (i.e., false memories).

A paradigm developed by Deese (1959) and extended by Roediger & McDermott (1995) is of particular interest in this line of research as it is recognized for eliciting high rates of false recognition under controlled experimental conditions.

Two previous studies (Mertens & Allen, submitted) found that P3 amplitude in response to false recognition is highly similar to P3 amplitude in response to truly recognized materials. It is of interest for the present investigation to determine whether the ERP assessment would effectively differentiate falsely recollected materials from learned and unlearned materials under these experimental conditions.

Methods

Participants

Data from two previous studies that only differed in the way decisions about stimuli were made (i.e., yes/no vs. confidence intervals) were combined (n=20, n=23) for the present study. All participants were native English speakers, reported normal or corrected-to-normal vision and were free of psychiatric and neurological disorders.

Procedure

Participants listened to 31 lists of words; items within each list were highly associated to a never presented common word (lure). Immediately following each list, participants engaged in a recognition task by viewing a second randomized list of words composed of ten unlearned words, one previously heard word (learned), and the lure. Words were presented for 300ms with an ISI of 2700 ms. Unlearned words were matched for word frequency to lure and learned words (Polich & Donchin, 1988)

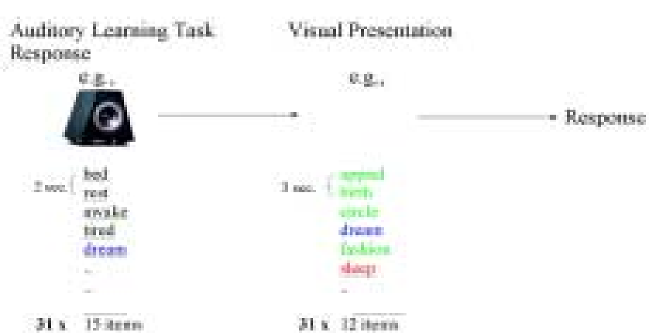


Figure 1. List learning procedure

<u>Stimulus characteristics</u>

<u>Category</u>	<u>Frequency of occurrence</u>	<u>Description</u>
Learned	1/12	Item heard
Lure	1/12	Item never presented
Distractors	1/12	Task irrelevant item

Table 1. Frequency of occurrence and characteristics of word stimuli

Methods

EEG

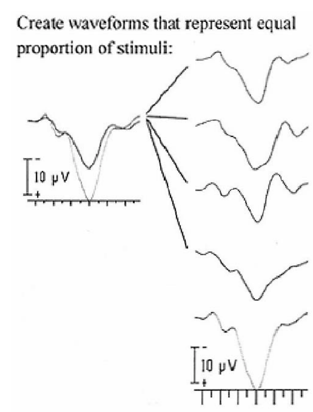
EEG was recorded by using a standard Lycra cap (Electrocap) with 28 tin electrodes, references online at Cz and offline to linked mastoids.

- EEG was digitized continuously at 200 Hz, then epoched from -250 to +1750 msec, with a 150ms baseline
- EEG was analyzed off-line. Ocular artifacts were removed and signals were filtered with a digital band-pass filter (12.5 hz, 96 db low pass)

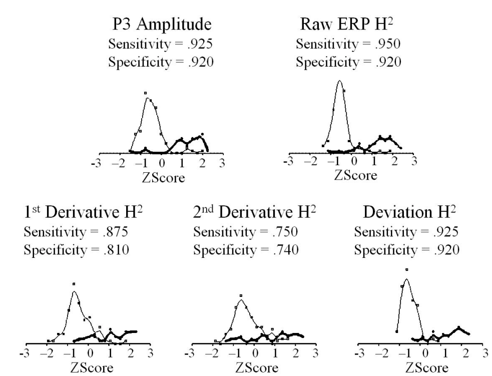
Bayesian analysis

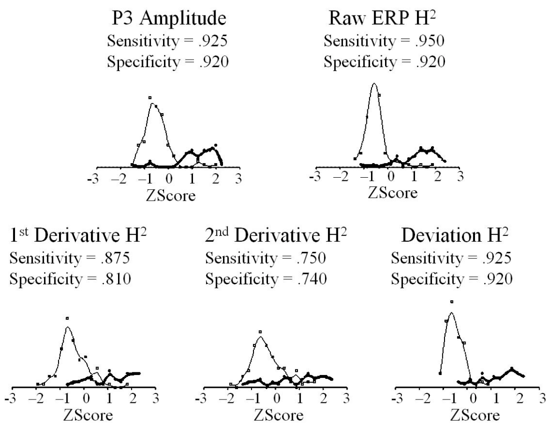
To derive statistically supported decisions about learned vs. unlearned materials for each individual subject, Bayesian analysis was employed in the following fashion:

- Computation of grand average waveforms for each stimuli type (1 learned, 1 lure, 10 unlearned).
- Extraction of 5 measures based on these waveforms with the latter four through factor analysis based on the entire epoch (-150 to +1050)
  - P3 amplitude
  - four communality measures
  - raw waveforms
    - 1st derivative
    - 2nd derivative
    - deviation waveform (grand average across all conditions per participant subtracted from each of the 12 item-type specific waveforms).



- utilization of previously validated cutpoints on an intra-individual basis that determine when a specific indicator is "large" enough to differentiate learned from unlearned





Bayesian analysis (cont.)

- Computation of Bayesian Posterior probabilities using above mentioned indicators beyond cutpoints considered "large".

Then determine the Bayesian Posterior probability that a list was learned given the presence of the indicator(s) (e.g., large P3):

$$\frac{Ll}{Ll + Uu}$$

Where:

L = Proportion of Learned items
l = Sensitivity of indicator
U = Proportion of Unlearned Items
u = (1- Specificity) of indicator

Results

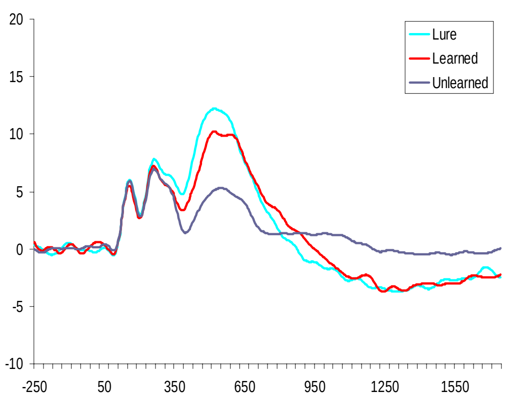


*Figure 2. Grand averaged ERP waveform depicting falsely recognized **Lures**, correctly recognized **Learned** and correctly rejected **Distractors** items at Pz. Grand average waveforms are plotted with positive amplitude up.*

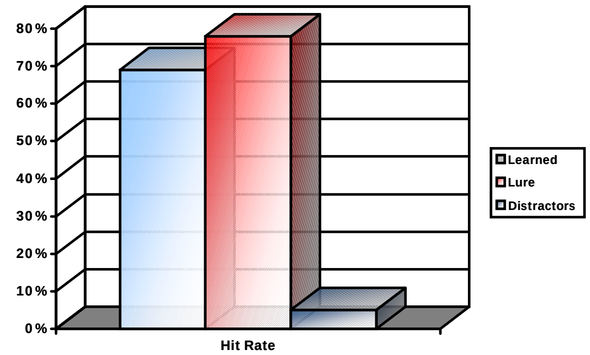


Figure 3. Reported rates of recognition

	Test Verdict	
Ground Truth	Recognition	No Recognition
Learned	56%	44%
Lure	72%	28%
Distractors	5%	95%

Table 1. Rates of Bayesian classification for Learned, Lure and Distractor items

Discussion

- ERP grand averages reveal highly similar ERPs for veridcally recognized items and falsely recognized items.
- At the level of individual classifications, falsely recognized material is classified as familiar at very high rates. Unlearned items, however still identified as unfamiliar at rates seen in other studies.
- These findings suggest that newer "alternative" deception detection methods may have limitations, apparent in situations during which participants subjectively come to falsely believe that they experienced a particular stimulus or event.

  - This might occur in forensic context after participants have been subjected to suggestive interviewing techniques or repeated exposure to critical items on the test.
  - This could be relevant in memory assessments in clinical populations vulnerable to suggestibility
  - These findings suggest that ERP memory assessment procedures will be of limited (if any) use in assessing the validity of disputed recollections (e.g., a litigant claiming memories of abuse, with the defendant denying them).

  - Further research is essential to explore the situations under which ERP-based memory assessments appear valid, and those where limitations may be apparent:

  - Effects of substances or stress during memory encoding
  - Effects of time between initial encoding and subsequent recall, especially if others variables interact with memory recall.
  - Effects of countermeasures designed to interfere with the procedure.