Using EEG pattern classification to track competition in negative priming

Ehren L Newman & Ken Norman

Princeton University, Princeton NJ, USA

Princeton Computational Memory Lab

Abstract

Competition in cognitive processing has lasting consequences for the subsequent accessibility of competing representations. Negative priming (NP) demonstrates that, when representations compete, the representations that lose the competition are subsequently harder to access.

To better understand the competitive dynamics that generate these effects, we developed a method of tracking the activation of the competing representations at the sub-trial time scale. Our methods rely on a pattern classification analysis of EEG data.

We found that when a subject views an image, we were significantly above chance at classifying which one (of four) image types the subject was viewing based upon the EEG signal. We also show that when images from different categories (e.g. a face over a house) were superimposed the classifiers were significantly above chance at predicting the class of both images. Using this, we tracked the activation of each stimulus in a negative priming task.

We then tested for connections between how much the to-be-ignored stimulus is processed (as detected by the classifiers) and how fluidly it is processed in the future. We present preliminary evidence relating classifier activity to subsequent reaction times.

Introduction

Making a choice has consequences
- Chosen item is subsequently stronger
- Non-chosen item is subsequently weaker
- Examples
  - Retrieval Induced Forgetting (Anderson & Neely, 1996)
    - Memories compete to be retrieved
    - Non-retrieved memories are less likely to be retrieved later
  - Negative Priming (NP) (Tipper, 1985)
    - Visual stimuli compete for attention
    - Non-attended stimuli are slower to be attended to later

In retrieval induced forgetting:
- Non-chosen item must activate to show subsequent weakening

Perhaps the same is true for negative priming
- If we can detect perceptual processing, we could test for this

Negative priming basics
- Basic design -
  - Two images are simultaneously presented to the subject on each trial
  - Subject is cued to respond to one image and ignore the other
  - Subject is then asked to respond to the ignored image on next trial
- Basic result -
  - Subject is slower to name previously ignored image than novel image

Our Goals

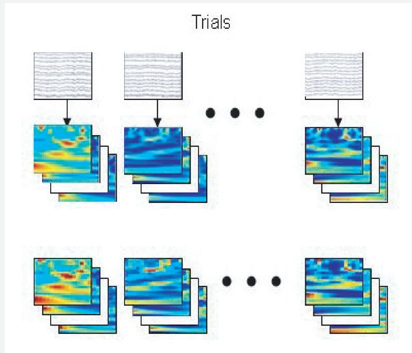
Detect how much the subject processed the distractor during each trial

Predict reaction times for each NP trial from amount distractor was processed

Decoding EEG via Distributed Pattern Analysis

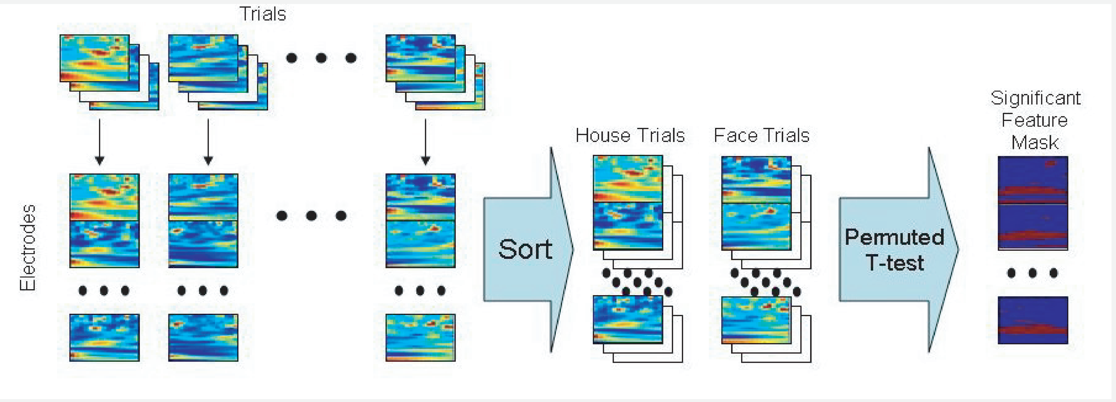
Data preparation -

- Collect data with 79 electrode cap (1000Hz sampling rate)
- Remove trials with excessive noise or blinks
- Perform frequency decomposition
  - Wavelet decompsition (6 cycle Morlet wavelet)
  - 49 frequency bands between 2 & 128Hz
  - Extract power of each frequency band
- Collapse data to form 20ms time bins (averaging)

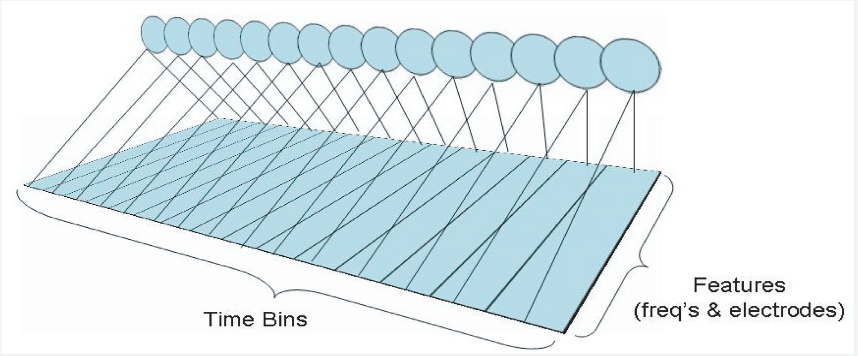


Classification preparation -

- Perform feature selection across *time bins / frequencies / electrodes*
  - Compute non-parametric *p*-value for each combination
  - Include features with $p<0.05$ as an input feature



- Build a ridge regression classifier for each time bin
  - Input patterns -
    - Significantly discriminating frequency / electrode combinations
  - Output patterns -
    - Binary regressors



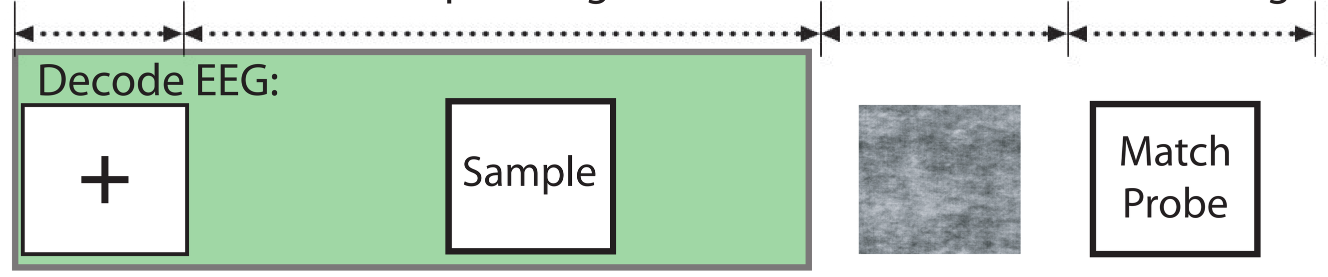
Classification procedure -

- Training the classifiers (for each time bin)
  - Use ridge regression learning algorithm
  - Use N-1 approach validation approach
    - Train on 9/10th of the trials
    - Test on remaining 1/10th
    - Repeat 10x

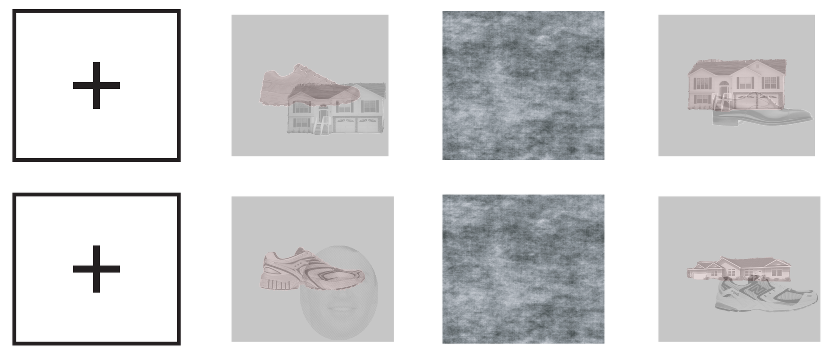
Experiment Design

Task design -
Simuli consist of shoes, faces, chairs, & houses

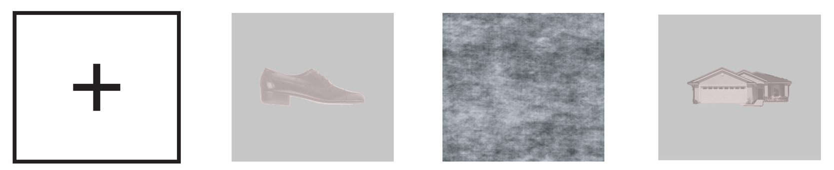
Delayed match to sample task (over 2 sessions)



1st session - superimposed stimuli
- Added superimposed image, ask subjects to ignore it
- Sample image tinted red to guide subjects
- Second image is always from a different category



2nd session - pure stimuli

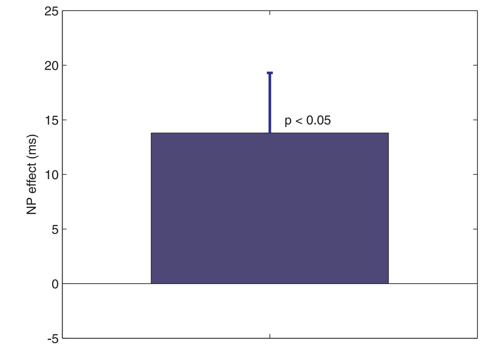


Basic logic -
- Train classifiers on the category of the target image
- Use trained classifiers on superimposed stimuli trials
  - Check that classifier can detect both images
  - Compare classifier ouput on slow vs fast trials

Behavioral Results

Subjects were slower to name ignored stimuli
M = 13.8ms, SEM = 5.5ms

t(17) = 2.57, p = 0.02

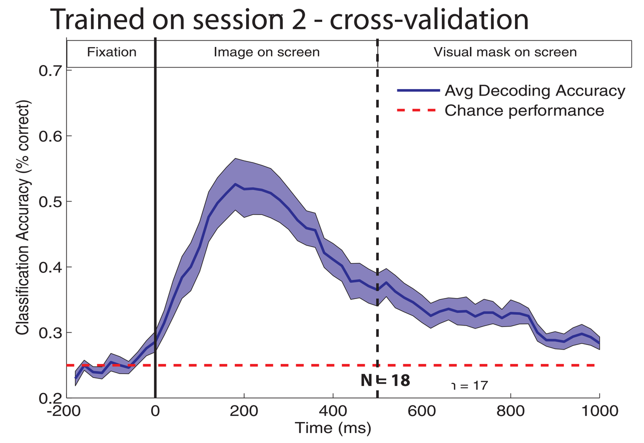


Classification Results

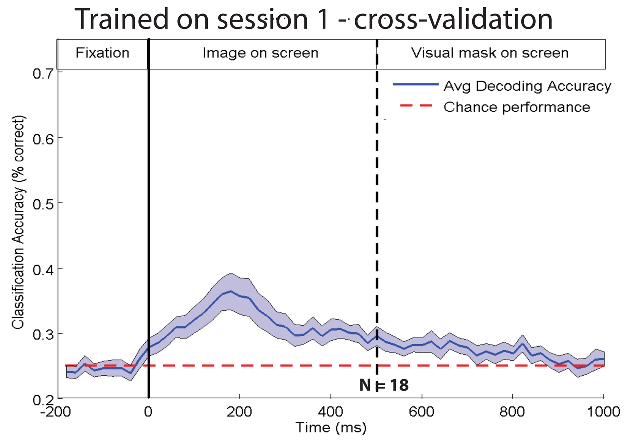
We compared the performance of classifiers that were trained on -
the pure stimuli of the 2nd session | the superimposed stimuli of the 1st session

Classifier training performance

Trained on session 2 - cross-validation

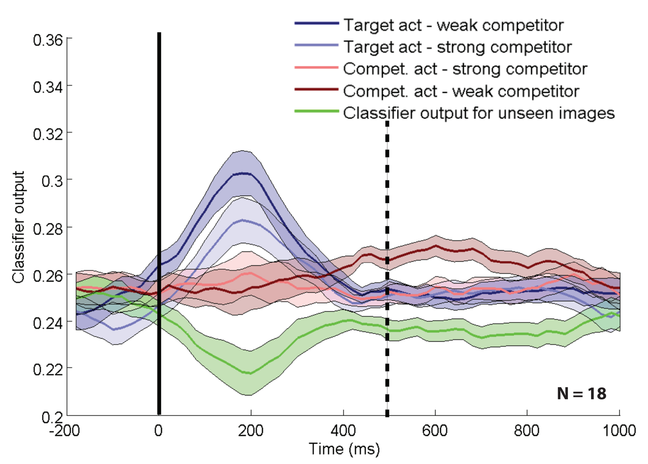


Trained on session 1 - cross-validation

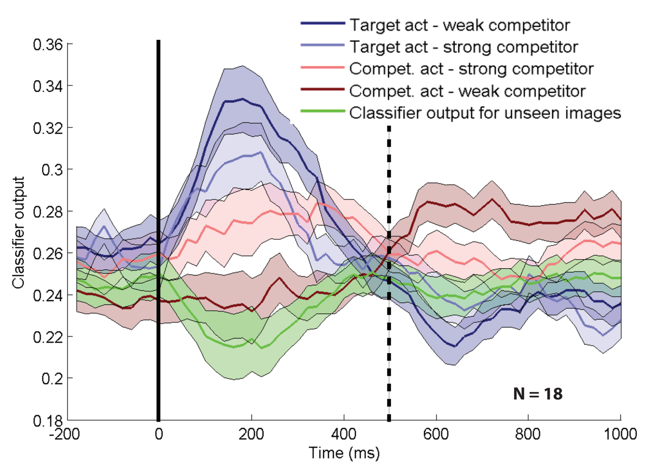


Classifier generalization performance

Trained on session 2 - generalized to session 1



Trained on session 1 - generalized to session 1



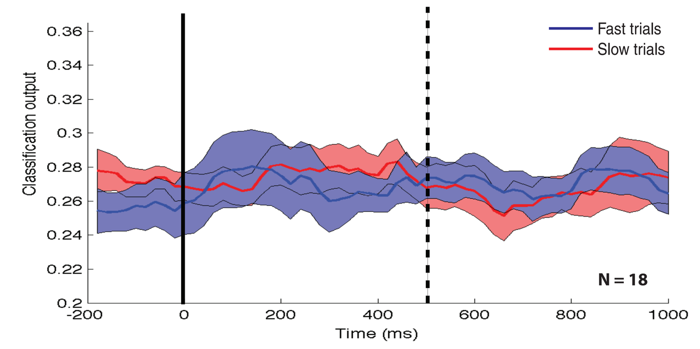
Classifier trained on session 2:
Better cross-validation
Worse distractor detection

Classifier trained on session 1:
Worse cross-validation
Better distractor detection

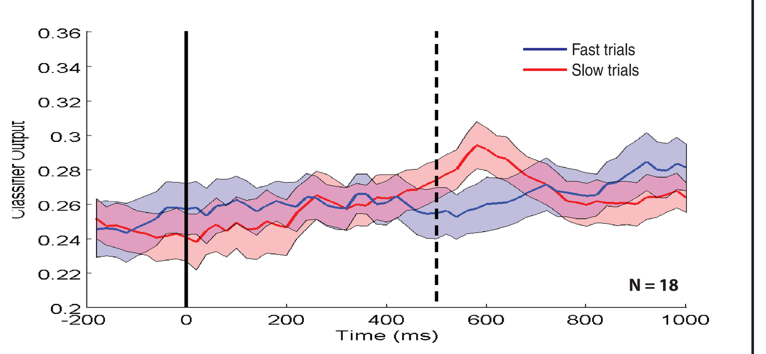
Preliminary results

Use median split on NP trials to identify fast and slow trials
Compute average activation of target and competitor

Trained on session 2 - distractor act. for fast vs slow trials



Trained on session 1 - distractor act. for fast vs slow trials



No significant differences for fast vs. slow

Trend toward significant differences for fast vs. slow

Results & Discussion

It is possible to decode which image category the subject is viewing.
The trained decoders can detect the category of multiple presented images
The strength of the classifier output varies with manipulation of the stimuli.
Preliminary evidence suggests that these decoders will be useful to examine difficult to observe dynamics such as distractor activation in a NP study.

Funding for this research was provided by
NIMH grant R01 MH069456 awarded to KAN & NIMH grant F31 MH077469 awarded to ELN