Dissociation of category-learning mechanisms via brain potentials

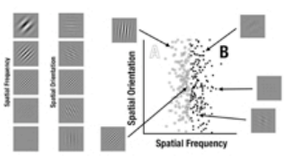
Robert G. Morrison — *Loyola University Chicago*

Paul J. Reber, Ken A. Paller — *Northwestern University*

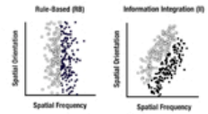
Introduction

- Behavioral (see Ashby & Maddox, 2005), neuropsychological (see Kéri, 2003), and neuroimaging (e.g., Nomura et al., 2007) evidence suggests that categories can be learned via explicit and/or implicit mechanisms.
- Ashby and Maddox (2005) described a feedback category-learning paradigm with different category distributions to selectively encourage one of the two types of learning mechanisms: Explicit or Rule Based (RB) versus Implicit or Information Integration (II).
- These two types of strategies have been dissociated behaviorally using working memory dual-tasks (e.g., Zeithamova & Maddox, 2006), feedback delay (e.g., Maddox, Ashby, & Bohil, 2003), and procedural interference (e.g., Ashby, Ell and Waldron, 2003).
- Nomura and colleagues (2007) used fMRI to demonstrate that RB category learning is dependent on areas in Prefrontal Cortex and Medial Temporal Lobe, whereas II category learning is dependent on areas in the Basal Ganglia and Occipital Cortex.
- Building on the success of previous neuroimaging efforts to dissociate RB and II category learning (e.g., Nomura et al., 2007), we utilized EEG/ERP methods to obtain additional evidence about this neural dissociation with greater temporal precision.

Features Determine Categories

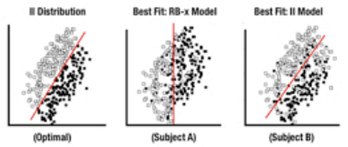


Different Strategies



Category distributions were created by sampling from two bivariate normal distributions. The variance of the distributions and the separation between them (d-prime) were adjusted based on pilot testing to match difficulty of categorization in both conditions.

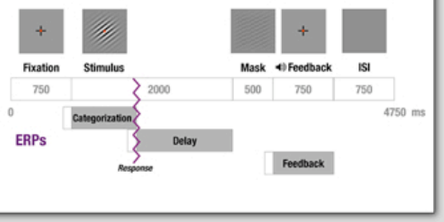
Models of Participant Responses



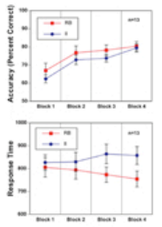
Methods

- In separate sessions, each participant learned an RB category defined by the spatial frequency of Gabor patches and an II category defined by a diagonal threshold based on both spatial frequency and spatial orientation (session order counterbalanced, ≥1 week apart). Participants performed 320 trials during each session divided into four equal blocks, and were debriefed after the second testing session.
- Participants received no instructions about the nature of the categories, but rather discovered the categories via feedback given 2.5 s after stimulus onset.
- Mathematical models derived from Decision Bound Theory (DBT; see Maddox, Ashby, & Bohil, 2003) were fit to each participant's responses to get a more detailed picture of how they were categorizing the stimuli.
- At the beginning of the first session, participants passively viewed 80 RB and 80 II stimuli from the actual stimulus distributions to ensure that there were no systematic differences in ERPs for the two types of distributions. None were found.
- EEG data were collected for 64 channels (bandpass 0.05-200 Hz, sampling rate 1000 Hz), including 4 channels used for rejecting trials with EOG artifacts and ocular artifact correction.
- ERPs elicited by Gabor patches throughout learning were averaged separately for each response type (correct and incorrect) and as a function of whether they were near or far to the category boundary. Initial statistical comparisons focused on amplitudes that were visibly different for RB and II waveforms. Formal comparisons of ERP amplitude were made using repeated-measures ANOVA. Waveforms were smoothed with a 30-Hz low-pass-zero-phase-shift Butterworth filter for presentation purposes only. Time-Frequency analyses of these data are on going and not reported at this time.

Trial Design



Accuracy and Response Time



Participants

Of the 28 participants run in the two-part experiment, 15 II distribution response profiles were best fit by an II DBT model with 2 eliminated because of an inadequate number of incorrect trials. The following data is based on the remaining 13 participants.

Accuracy

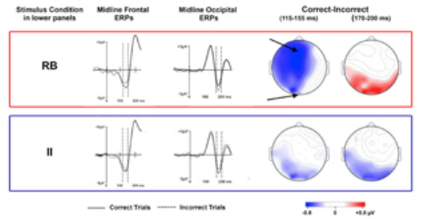
Accuracy in RB and II conditions did not reliably differ [F (1,12) = 1.5, *ns*]. There was a main effect of block [F (3,36) = 29, p < .001] and no interaction [F (3,36) < 1, *ns*].

Response Time

Response time in RB and II conditions did reliably differ [F (1,12) = 5.1, p = .043]. There was no main effect of block [F (3,36) < 1, *ns*] and a reliable interaction [F (3,36) = 3.8, p = .018].

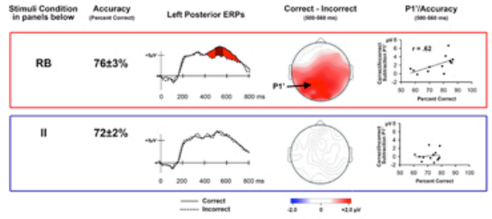
Results

Dissociations in Early ERPs



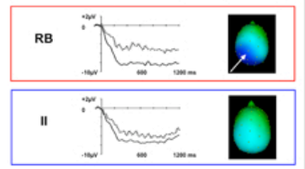
An early (115-155 ms) negative frontocentral ERP was predictive of correct categorization in the RB [F (1,12) = 8.1, p = .015], but not the II [F (1,12) = .07, *ns*] condition. A slightly later (170-200 ms) negative occipitoparietal ERP is modulated by categorization condition, with correct RB trials less negative than incorrect RB trials [F (1,12) = 8.2, p = .014], but with correct II trials more negative than incorrect trials [F (1,12) = 7.9, p = .016].

Dissociations in Late ERPs



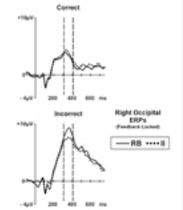
A late (400-700 ms) positive parietal ERP was predictive of correct categorization in the RB condition [F (1,12) = 9.0, p = .011], but not in the II condition [F (1,12) = .077, *ns*]. The magnitude of the correct/incorrect waveform subtraction (as measured in a 60-ms window at its maximum) was reliably correlated with RB performance [r(11) = .62, p = .02], but not with II performance [r(11) = .11, *ns*].

ERP Differences during Delay



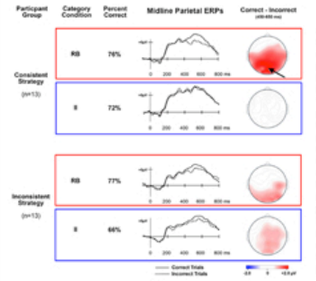
An ERP (0-1200ms response locked) largest in magnitude in left parietal regions showed differential amplitudes for correct and incorrect trials in RB [F (1,12) = 15, p = .002] and II [F (1,12) = 4, p = .06] conditions. There was also a significant interaction [F (1,12) = 7, p = .02] due to larger correct/incorrect differences in the RB condition than in the II condition.

Feedback ERP



Both correct and incorrect trials showed P300s with broad topographies centered around 350 ms. Whereas ERPs for RB and II conditions did not differ in the 310 to 410 ms range for correct trials [F (1,12) = .01, ns], they did differ for incorrect trials [F (1,12) = 5.7, p = .03].

Effect of Strategy on Late ERPs



As mentioned previously, the behavior of approximately half of the participants in the II condition was best fit by RB DBT models. ERPs for these "Inconsistent-Strategy" participants suggested that they engaged similar processing in the II and RB conditions. Correct-incorrect ERPs differed across RB and II conditions for Consistent-Strategy participants [F (1,12) = 10, p = .007], but not for Inconsistent-Strategy participants [F (1,12) = .05, *ns*], yielding a reliable three-way interaction [F (1,24) = 4.3, p = .049].

Discussion

- Our results implicate distinct category-learning mechanisms for rule-based versus information-integration conditions, consistent with previous behavioral and fMRI studies.
- Differential ERPs were observed in RB and II category learning but not during passive viewing of the same stimuli.
  1) An enhanced frontocentral ERP at 115-155 ms for correct trials in the RB condition may reflect early top-down allocation of attentional resources in the RB strategy.
  2) An occipitoparietal ERP at 170-200 ms showed the opposite pattern for RB and II conditions with respect to correct and incorrect trials.
  3) A left parietal positivity at 400-700 ms was enhanced for correct trials in the RB condition. This effect was present in subjects using rule-based behavior (identified by DBT model fits) in both the RB and II conditions, and it may reflect memory retrieval, as in fMRI studies and a previous ERP study of category learning (Curran, Tanaka, & Weiskopf, 2002) or alternatively greater decision confidence in the more explicit RB condition.
  4) A correct/incorrect subtraction of response locked CNV ERPs differed across category conditions with a larger subtraction for RB than II trials. This ERP may reflect differences in working memory demands during delay, consistent with working memory dual-task results reported by Zeithamova and Maddox (2006).
  5) Participants showed a larger incorrect P300 ERP recorded during feedback in the RB than in the II condition. The same ERP from correct trials was not differential across category conditions. This suggests participants have greater explicit awareness of the meaning of feedback during RB learning than during II learning. This result provides additional support for the neural dissociation between RB and II category learning, consistent with participants' self reports.
- ERP findings appear to reflect multiple neurocognitive processes engaged during category learning. Whereas processes critical for both RB and II learning may be operational, certain processes are used selectively.

Acknowledgements

The presenters wish to thank Emi Nomura for programming the DBT Models, Joel Voss and John Rudoy for technical assistance, and Courtney Clark and Ilya Bendich for assistance in data collection. We would also like to thank the Northwestern University Mechanisms of Aging and Dementia Training Grant from the NIA (2T32AG020506, RGM), and the National Science Foundation (0518800 and 0818912, KAP) for their generous support.