Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biological Psychology 71 (2006) 278–288
BIOLOGICAL PSYCHOLOGY
www.elsevier.com/locate/biopsycho

Electrophysiological correlates of categorization: P300 amplitude as index of target similarity

A. Azizian [a,*], A.L. Freitas [b], T.D. Watson [b], N.K. Squires [b]

[a] *The Scripps Research Institute, Cognitive Electrophysiology Laboratory, Department of Neuropharmacology, 10550 North Torrey Pines Road, TPC-10, La Jolla, CA 92037, USA*
[b] *Stony Brook University, Department of Psychology, Stony Brook, NY, USA*

Received 18 January 2005; accepted 30 May 2005
Available online 25 July 2005

Abstract

Two experiments examined event-related potentials (ERPs) and behavioral correlates of categorizing stimuli varying in perceptual similarity to targets. Participants performed a target-detection task in which non-target stimuli varied in target similarity but occurred with equivalent probability. The stimuli were variations of a schematic human face comprised of eight distinct features: two eyes, two eyebrows, one nose, one mouth, and two ears. Non-target stimuli that were perceptually similar to targets produced larger P300-like neurophysiological responses than did other non-target stimuli. These effects emerged whether participants' target was relatively complex (eight features) or quite simple (zero features). Accordingly, the presence of many constituent elements of a test stimulus does not appear necessary to trigger increases in categorical processing of non-targets that are similar to a target. The data further suggest that the P300 amplitude may be used as a good index of perceptual similarity between target and non-target stimuli.

Keywords: P300 amplitude; Categorization; Similarity; Event-related potentials; Stimulus evaluation; Schematic faces

Thinking about or acting upon environmental stimuli requires at least a cursory comprehension of what those stimuli are. Psychological and neural mechanisms of categorization thus underlie a broad range of human behavior, and extensive efforts have aimed to elucidate them (e.g., Anderson, 1991; Ashby, 1992; Delorme and Thorpe, 2001; Medin and Schaffer, 1978; Nosofsky, 1986).

In this vein, an enduring conceptual question has been whether categorization entails some sort of tabulation of the discrete, separable elements comprising a stimulus representation (Fodor, 1970; Mervis and Rosch, 1981; Tversky, 1977). On the one hand, reaching a decision criterion might require repeatedly sampling elements of a test stimulus, while continually evaluating those elements' similarities to exemplars of candidate categories (Lamberts, 2000). On the other hand, a test item's degree of similarity to exemplars of different categories should impact the likelihood of retrieving those exemplars from memory. Records of such retrievals themselves, rather than successive sampling of constituent elements of the test item, might then form the basis of categorization, with decisions in favor of the category of the most-quickly retrieved exemplars (Nosofsky and Palmeri, 1997). Current cognitive theories thus agree that extensive categorical processing is needed when a test item is highly similar to more than one category, but they differ in the degree to which such computations are presumed to entail processing constituent features of stimulus representations.

Having extended from explaining simple target and non-target classification to explaining categorical processing of complex stimuli, current neural and electrophysiological work is well positioned to examine how constituent properties of a stimulus impact categorization of it. In agreement with the classic ERP literature (for review see Sutton, 1979) recent studies have demonstrated that neural processes associated with rule-based categorization are affected by the intrinsic properties of stimuli (e.g., Fize et al.,

* Corresponding author. Tel.: +1 858 784 8423; fax: +1 858 784 9293.
E-mail address: azizian@scripps.edu (A. Azizian).

0301-0511/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.biopsycho.2005.05.002

2000; Fujihara et al., 1998; Koenig et al., 2005; Thorpe et al., 1996). Importantly, however, the bulk of this work has focused on differences in brain responses to stimuli of different categories (such as animal-present scenes versus animal-absent scenes; Thorpe et al., 1996) rather than on differences in brain responses to heterogeneous stimuli within the same category. One view, an important instance of within-category heterogeneity is when non-target stimuli vary in similarity to a target. Such cases can be especially informative in that they allow simultaneous investigation of the effects of complexity and target similarity of non-targets on categorical processing. If processing features of stimulus representations constitutes the basis of categorical processing, then larger effects of target similarity on brain responses to non-targets should emerge for complex non-targets than for simple non-targets, because complex stimuli have a greater number of constituent elements than do simple stimuli.

Pursuing this issue, we capitalized on extensive research on the P300 component, an endogenous late-positive component of the event-related potential (ERP; Sutton et al., 1965). Characterized in terms of amplitude, latency, and scalp topography, this component has proved an important predictor of numerous cognitive and behavioral phenomena (e.g., Squires et al., 1975; Picton, 1992). A central functional correlate of P300 is classification of improbable task-relevant stimuli into generally defined ''target'' categories (Duncan-Johnson and Donchin, 1982). When improbable target stimuli are detected, target stimuli produce P300s that are topographically maximal in the central-parietal electrode sites (e.g., Duncan-Johnson and Donchin, 1977). Most importantly to the current investigation, the P300 is larger in amplitude for low-probable stimuli that are covertly or overtly categorized as targets. In traditional two-stimulus oddball paradigms, target and non-target stimuli differ on the basis of a single feature presented by standard stimuli (e.g., pitch or color). In conceptual versions of the oddball paradigm, stimuli are classified according to their intrinsic characteristics (e.g., Kutas et al., 1977; Kotchoubey and Lang, 2001). In such task designs, participants classify not standard-invariant, but within-class variable stimuli. This effect was first demonstrated by Kutas et al. (1977) in which participants were instructed to count variable female names intermixed with male names. Kotchoubey and Lang (2001) investigated the P300 in semantic classification tasks in which standards and targets differed in their semantic qualities (e.g., body parts, plants). In both studies, with the exception of longer latency, typical P300 components were recorded for stimuli belonging to the target category. Together, those findings strongly support interpreting the P300 as reflecting neurophysiological mechanisms of categorical processing.

We designed the present experiments to address more directly the role of P300-like neurophysiological responses during categorization of simple and complex stimuli. Given extensive previous evidence that stimuli categorized as targets produce larger P300s than stimuli not categorized as targets, we hypothesized that non-targets highest in similarity to targets would produce the largest P300-like responses, due to increased categorical processing of them. Most importantly, we tested whether this hypothesized effect (larger P300s to target-like non-targets) would obtain whether the most highly similar non-targets were composed of relatively many or relatively few features.

1. Current studies

Towards this end, we constructed a stimulus set that varied in the number of constituent elements its members contained, and we manipulated whether participants' target category was the most complex or most simple member of the set. Accordingly, when making judgments about whether or not a test item matched the most complex member of the set, the most similar non-target items would contain relatively many constituent elements. In contrast, when making judgments about whether or not a test item matched the least complex member of the set, the most similar non-target items would contain relatively few constituent elements. If repeated sampling of the constituent elements of test items is how people resolve difficult classifications, then larger processing differences between similar and dissimilar items should emerge in the former case than in the latter case.

1.1. Materials

The stimuli were variations of a schematic human face that was comprised of eight distinct features: two eyes, two eyebrows, one nose, one mouth, and two ears (see Fig. 1). There were nine categories of stimuli, each defined by the number of features present. In the first study, the target stimulus included all eight features; in the second study, the target stimulus was the oval shape with no features. With the exception of the eight-feature face and the zero-feature oval, there were eight distinct stimuli per category, each containing a randomly selected set of features. All categories were presented in random order with equal probability ($p = .11$). There were 32 stimulus presentations per category for a total of 288 stimuli per session, broken down into two experimental blocks. The stimulus duration was 500 ms, and the inter-stimulus interval was 1000 ms.

1.2. Procedure

In the ERP studies, participants were instructed to keep a silent mental count of the number of times that the target stimulus was presented. At the end of each block, participants were asked to report their count. The experiments were conducted in a sound-attenuating chamber with the lights off, with participants were seated comfortably in a reclining chair, approximately 61 cm from the screen. The stimuli were presented on a flat-panel LCD computer

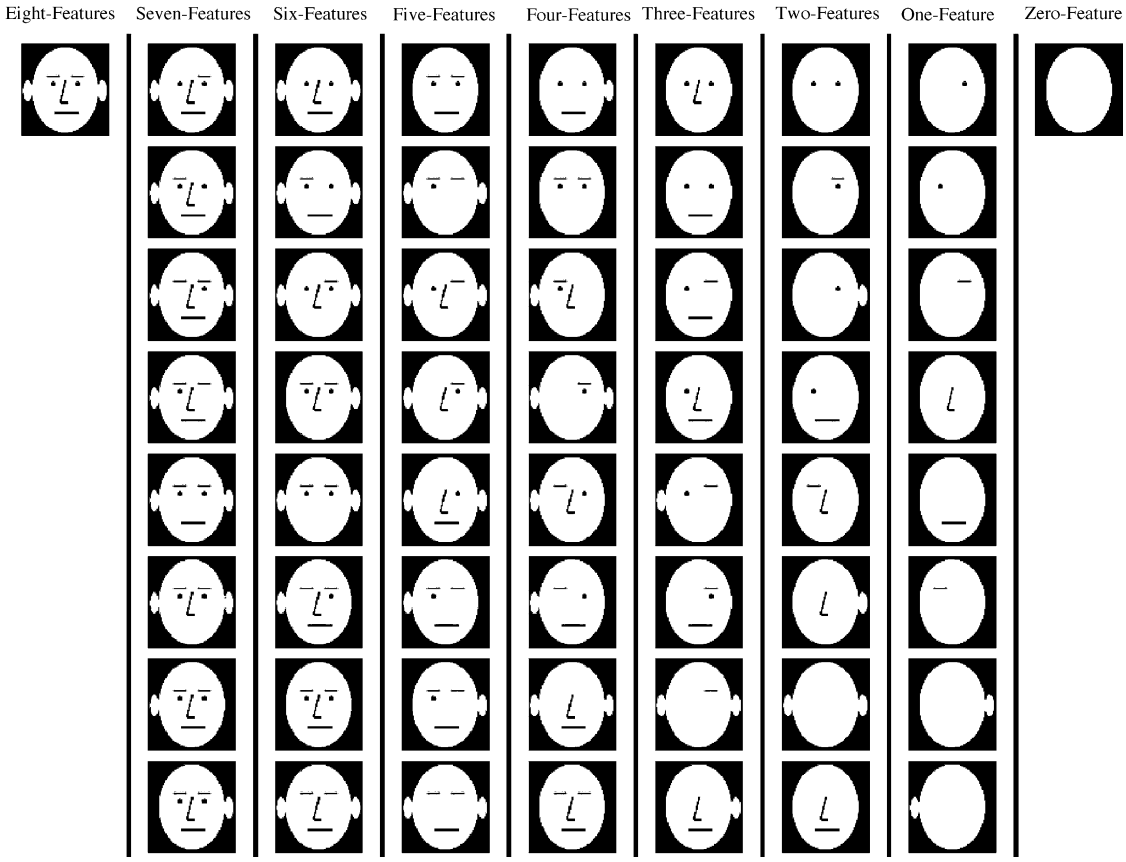

Fig. 1. Schematic face stimuli used in Experiment 1 (in which eight-feature "face" stimulus was designated as a target) and Experiment 2 (in which zero-feature "oval" stimulus was designated as a target).

monitor. The stimuli were 5.08 cm high × 5.08 cm wide. Participants were instructed to remain as still as possible, and to minimize eye blinks throughout the experiment.

The behavioral study used materials and procedures identical to those described above, except that: (a) participants were instructed to quickly and accurately press one of two computer keys to discriminate between target and non-target stimuli; (b) the experiments were conducted in small single-participant rooms of standard construction; (c) the stimuli were displayed on CRT monitors.

1.3. Electrophysiological recording

The EEG was recorded continuously using a 64-channel electrode cap (Neuroscan Inc., Sterling, USA). All recordings were performed using a fronto-central electrode as ground, and electronically linked mastoid electrodes as reference. The horizontal EOG was monitored from electrodes at the outer canthi of the eyes, and the vertical EOG was monitored from electrodes above and below the orbital region of the left eye. Impedances for all electrodes were kept below 10 kΩ. The EEG and EOG signals were digitized at 1000 Hz, and were amplified with a gain of 500. The filter bandpass was .01–30 Hz. To eliminate EOG artifact, trials with EEG voltages exceeding 50 μV were rejected from the average. Artifact rejection and averaging were done offline. Approximately 20% of the trials were excluded due to artifacts. ERP epochs began 100 ms prior to stimulus onset and continued for 900 ms thereafter.

1.4. ERP analysis

Individual ERP averages were created for each stimulus category. The P300 amplitude was measured as the maximum peak occurring in 300–700 ms latency window. Peak amplitude was measured relative to the pre-stimulus baseline. Because the primary purpose of the present study was to assess the P300 amplitude as an index of non-target similarity, other ERP components and characteristics will not be discussed. Sixty-four channel scalp topographical maps at the time point of maximum amplitude were taken to present scalp distribution. To reduce the number of statistical comparisons made in this study, and to allow for analysis of differences in amplitude in the anterior–posterior dimensions, the data from only 25 electrodes were analyzed statistically. The electrodes selected correspond to maximal voltage displayed in the topographical maps and are shown in Fig. 2. In the advent of a bad electrode (three instances), the group-mean amplitude was substituted. The Greenhouse–Geisser (Greenhouse and Geisser, 1959) correction was used for all comparisons with more than two levels, and an alpha level of .05 was used for all analyses in this study.

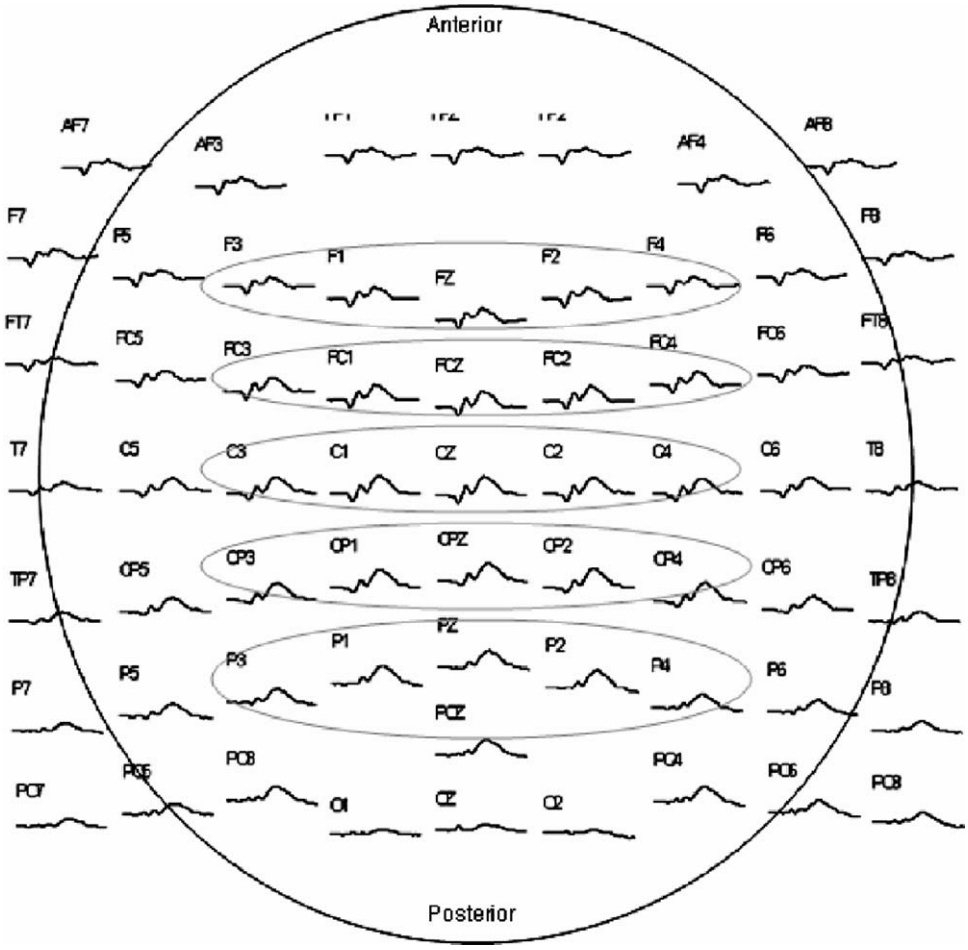

Fig. 2. Layout of the electrode array utilized for statistical analysis.

2. Behavioral pilot experiment

Responding to the stimulus set described above (see Fig. 1), half of participants distinguished the eight-feature ''face'' target from all other stimuli, and half distinguished the zero-feature ''oval'' target from all other stimuli. Across both experimental conditions, we expected that increases in target similarity would predict increased response times to non-target stimuli.

2.1. Participants

Participants were 26 Stony Brook University students, aged 17–37 (M = 19.8 years, S.D. = 3.8), of whom 17 were women. They gave written informed consent and received credit towards fulfillment of undergraduate psychology course requirements. All participants reported normal or corrected-to-normal vision.

2.2. Results

Participants' responses were correct on 98.67% of trials. Response times on incorrect trials were not analyzed, nor were response times less than 200 ms or greater than 1000 ms, resulting in valid response time measures from 95.13% of trials. Response times were analyzed in a 2 (target complexity: zero-feature ''oval'' target versus eight-feature ''face'' target) × 9 (target similarity, with nine levels corresponding to similarity to target) ANOVA, with repeated measures on the last factor. Consistent with the expectation of slower response times to target-similar stimuli (see Table 1), there was a main effect of target similarity, $F(8, 192) = 107.44$, $p < .0001$, $\eta_p^2 = .82$.

2.2.1. Planned comparisons

We next conducted planned comparisons focusing specifically on response times to the eight different types of non-targets. To do so, we computed the seven subtraction scores between response times to non-target types successively similar to targets. In other words, for participants with the eight-feature target, we subtracted the time taken to respond to the six-feature non-targets from time taken to respond to seven-feature non-targets, the time taken to respond to the five-feature non-targets from time taken to respond to six-feature non-targets, the time taken to respond to the four-feature non-targets from time taken to respond to five-feature non-targets, etc. On the other hand, for participants with the zero-feature target, we subtracted

Table 1
Average response times (and standard deviations) on categorization task, from behavioral pilot study

Stimulus	Target assigned	
	Complex	Simple
Target	690.82 (76.00)	548.67 (75.23)
Sim8	625.8 (80.49)	532.8 (63.89)
Sim7	557.98 (96.97)	508.04 (64.23)
Sim6	498.76 (79.63)	479.76 (71.56)
Sim5	487.13 (80.35)	456.69 (63.95)
Sim4	481.47 (78.57)	447.65 (68.37)
Sim3	475.16 (83.00)	447.56 (76.60)
Sim2	456.21 (86.25)	444.06 (76.41)
Sim1	454.25 (70.83)	437.97 (58.35)

Note: Target denotes schematic face stimulus in complex target condition and oval stimulus in simple target condition. Sim1 through Sim8 denote varying degrees of target similarity of non-targets, with higher suffixes signaling higher target similarity, $N = 26$.

the time taken to respond to the two-feature non-targets from time taken to respond to non-targets stimulus types, the time taken to respond to the three-feature non-targets from time taken to respond to two-feature non-targets, the time taken to respond to the four-feature non-targets from time taken to respond to three-feature non-targets, etc. Thus, higher values on these difference scores reflect longer time taken to respond to target-like non-targets. We averaged the seven difference scores, to form a single index of the effect of increasing target similarity of non-targets on response time. Analyzing this index while collapsing across both experimental conditions, an increase by one feature in target similarity caused an average increase in response time of 21.70 ms (S.D. = 10.47) to non-target stimuli; this effect differed significantly from zero, $t(25) = 10.57$, $p < .0001$, $\eta_p^2 = .52$.

We next recomputed the above contrasts separately for participants assigned to the two experimental conditions. An increase by one feature in target similarity caused an average increase of 29.57 ms (S.D. = 9.12) to respond to non-target stimuli ($t(13) = 11.68$, $p < .0001$, $\eta_p^2 = .72$) among participants assigned the eight-feature ''face'' target and an average increase of 13.84 ms (S.D. = 3.31) to respond to non-target stimuli ($t(13) = 15.07$, $p < .0001$, $\eta_p^2 = .81$) among participants assigned to the zero-feature target. As noted, there was greater variability in this effect among participants assigned the eight-feature target (S.D. = 9.12) than among participants assigned the zero-feature target (S.D. = 3.31), resulting in a significant departure from equality of variances across experimental conditions, $F(12, 12) = 7.60$, $p < .01$. Accordingly, prior to testing whether target complexity significantly moderated the effect of target similarity on response times to non-targets, we translated the raw effects into the metric of Cohen's d (see, e.g., Greenwald et al., 1996). That is, we divided each participant's overall contrast effect by the standard deviation of that effect among participants in the same experimental condition. As indicated by an independent-samples t-test ($t(24) = 2.40$, $p < .05$), the effect size was somewhat larger among participants with the zero-feature oval target ($M = 4.18$, S.D. = 1.00) than among participants with the eight-feature face target ($M = 3.24$, S.D. = 1.00). This difference in standardized effect sizes echoes the above-reported higher t and η_p^2 values among participants assigned the zero-feature oval target than among participants assigned the eight-feature face target.

2.3. Discussion

Results from a behavioral study strongly confirmed that increasing target similarity increases response times to non-targets. Most importantly, this effect emerged both among participants assigned a relatively complex target and among participants assigned a relatively simple target. In terms of raw milliseconds, the effect appeared larger among participants assigned a relatively complex (eight-feature) ''face'' target than among participants assigned a (relatively simple) zero-feature ''oval'' target. However, target similarity explained a slightly larger portion of the variance in response times to non-targets among participants assigned the simple target than among participants assigned the complex target. These findings underscore the need for investigating the effects of graded target similarity of non-targets on not only behavioral but also electrophysiological correlates of categorization. A major advantage of this approach is decomposing response times into subcomponents that relate to different mental operations.

3. ERP experiments

Electrophysiological data were recorded while participants performed an oddball task entailing keeping a mental count of target appearances. In both experiments, a standard target stimulus was embedded among the heterogeneous non-target stimuli pictured in Fig. 1. In Experiment 1, participants' target was the eight-feature ''face'' stimulus; in Experiment 2, participants' target was the zero-feature ''oval'' stimulus. Across both experiments, we tested whether non-target stimuli similar to targets would elicit higher amplitude electrophysiological responses than would non-target stimuli that were dissimilar to targets.

3.1. Participants

Participants were Stony Brook University students, all of whom reported normal or corrected-to-normal vision. They gave written informed consent and received credit towards fulfillment of undergraduate psychology course requirements. Nineteen participants (eight females; three left-handed; $M = 22.3$ years) were recruited for Experiment 1. Thirteen participants (four females; all right-handed; $M = 20.83$ years) were recruited for Experiment 2.

3.2. Results

3.2.1. Counting performance

Data of participants reporting a count greater than 10% different from the correct count were excluded from excluded from analysis. The data of two participants from Experiment 1 and one participant from Experiment 2 were excluded for this reason.

3.2.2. ERP waveforms

Figs. 3 and 4 show grand-average ERP waveforms for each stimulus category in Experiments 1 and 2, respectively. Across both experiments, the target stimulus elicited the largest P300 component at all electrode sites, with maximum amplitude at central electrode sites. Non-target stimuli most highly similar to targets elicited intermediate P300s that were topographically target-like, but smaller than those elicited by target stimuli. Figs. 5 and 6 display 64-channel topographical voltage-maps across the scalp for target and non-target stimuli, for Experiments 1 and 2, respectively. Each map is shown at the time point at which positive peak is maximum, between the 300 and 700 ms window. Across both experiments, visual inspection of the topographical maps for target and similar-to-target categories shows a maximum amplitude at central-parietal areas.

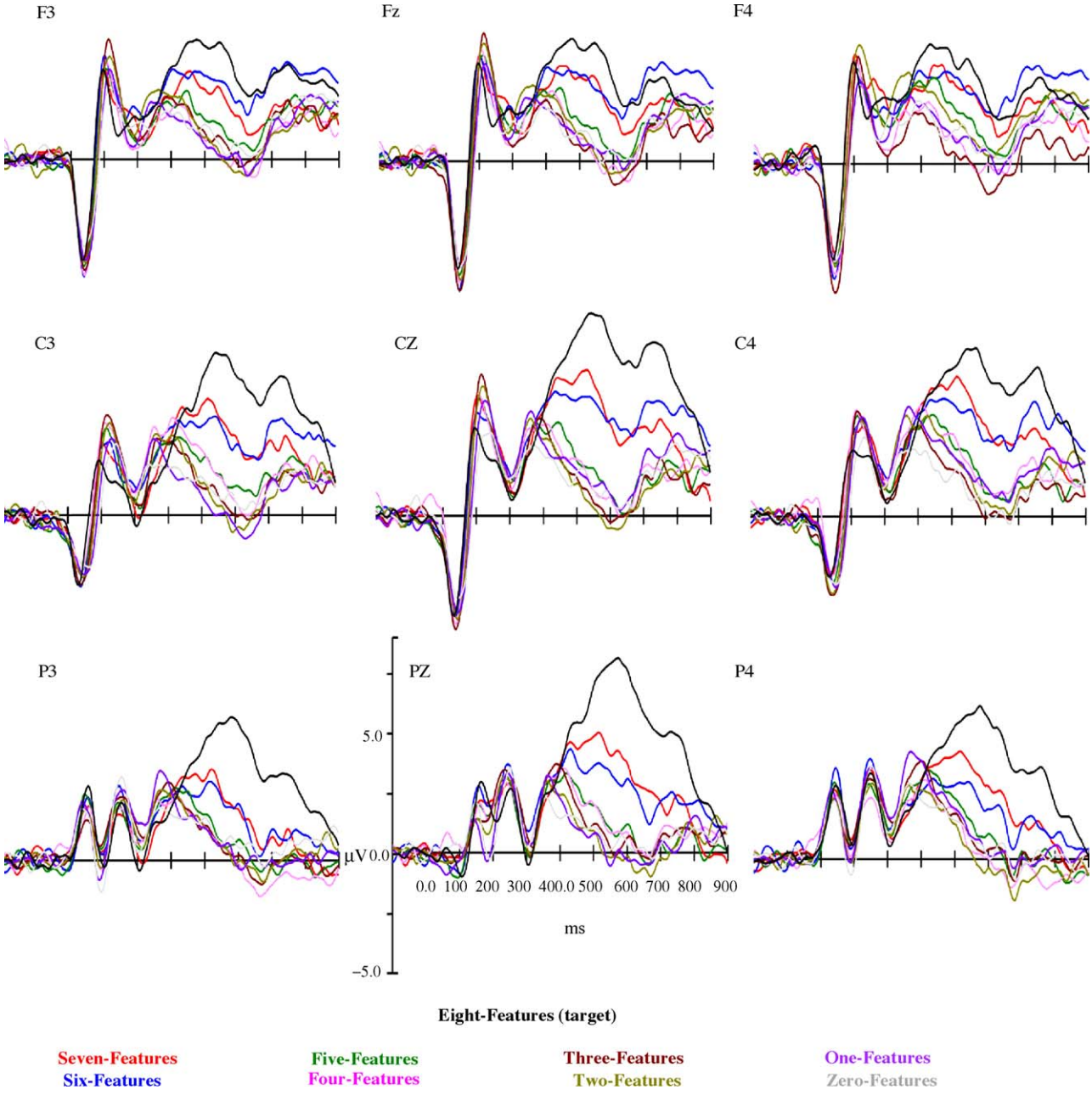

Fig. 3. Grand-average ERPs for each stimulus category superimposed at left, midline and right electrode sites, Experiment 1, $N = 17$.

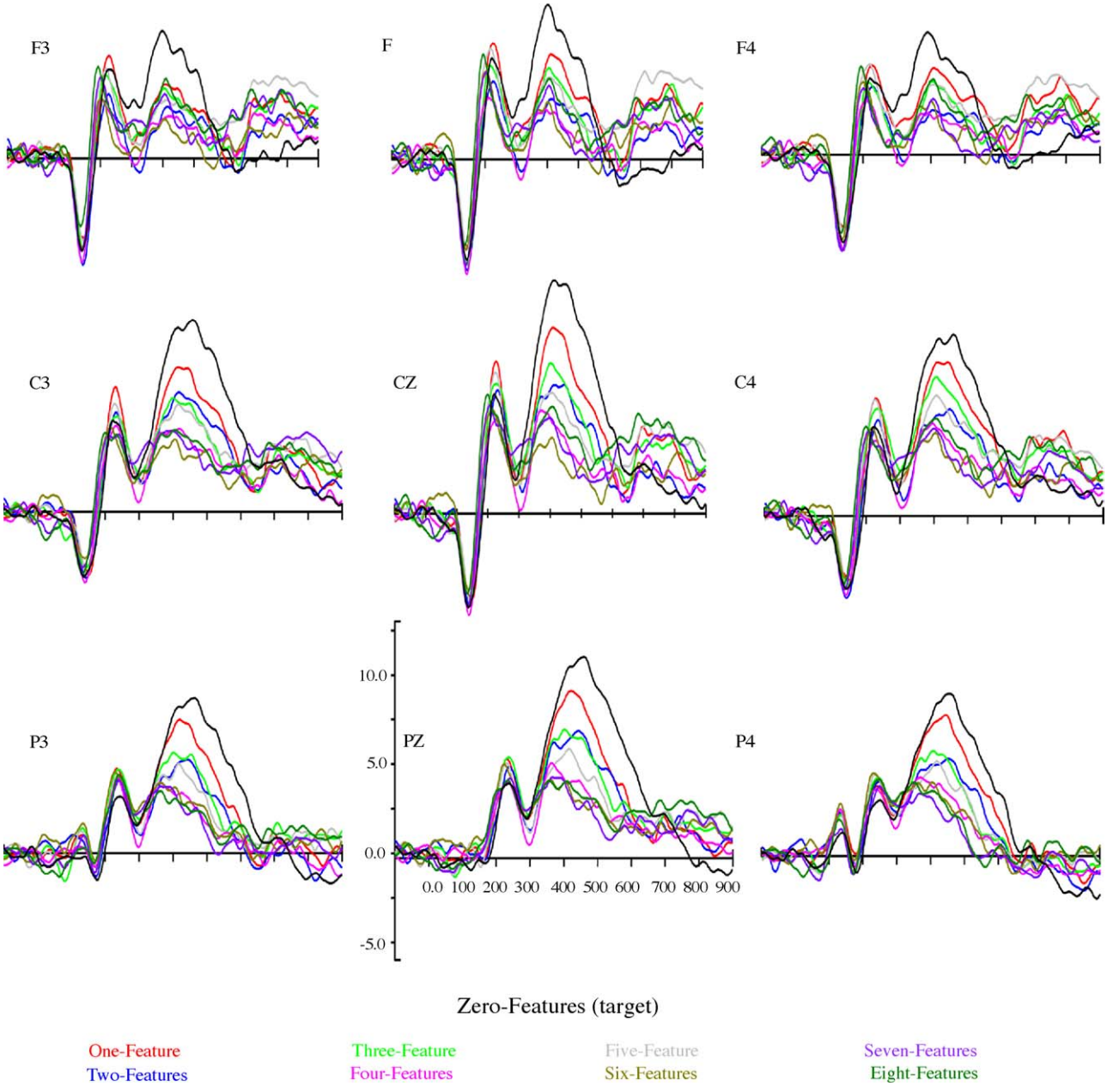

Fig. 4. Grand-average ERPs for each stimulus category superimposed at left, midline and right electrode sites, Experiment 2, $N = 12$.

3.2.3. Omnibus analyses of ERP amplitudes

Mean ERP amplitudes across the 300–700 ms window from both experiments were analyzed in a 2 (target complexity) × 9 (target similarity) × 5 (row) ANOVA, with repeated measures on the last two factors. Most importantly, there was a significant effect of target similarity, $F(8, 216) = 19.56$, $p < .0001$, $\eta_p^2 = .42$, and there was no evidence of moderation by target complexity, in that the target complexity × target similarity interaction was not significant, $F(8, 216) = 1.50$, $p > .19$, $\eta_p^2 = .05$. Of lesser significance to current aims, effects of row ($F(4, 108) = 6.81$, $p < .0001$, $\eta_p^2 = .20$) and of the row × similar-similarity interaction also emerged ($F(32, 864) = 3.04$, $p < .005$, $\eta_p^2 = .10$), reflecting higher-amplitude responses (and a larger effect of target similarity) at the central electrode sites. There also was a significant row × target complexity interaction ($F(4, 108) = 3.01$, $p = .05$, $\eta_p^2 = .10$), reflecting a somewhat more frontal distribution in Experiment 2 (with the zero-feature "oval" target) than in Experiment 1 (with the eight-feature "face" target). Again, however, there was no evidence of moderation by target complexity, in that the three-way target similarity × row × target complexity interaction was not significant ($F(32, 864) = 1.04$, n.s., $\eta_p^2 = .04$).

3.2.4. Focused comparisons of ERP amplitudes

In the most relevant tests of our hypotheses, we next conducted planned comparisons focusing specifically on

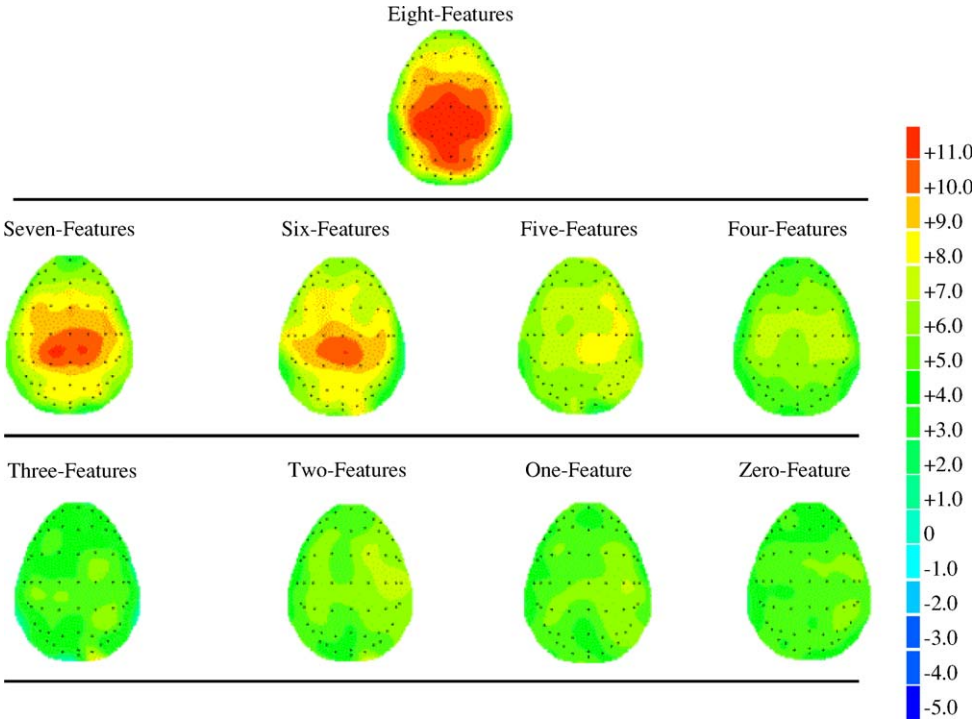

Fig. 5. Sixty-four channel ERP topographical maps shown at the time point at which the positive peak between 300 and 700 latency is maximum at Cz electrode site, Experiment 1.

mean amplitude of ERPs to the eight different types of non-targets. Just as in the behavioral study (see above for details), we did so by computing the seven subtraction scores between mean amplitude of ERPs to non-target types successively similar to targets. We then averaged the seven difference scores, to form a single index of the effect of increasing target similarity of non-targets on ERP amplitude. As before, with this approach, differences between

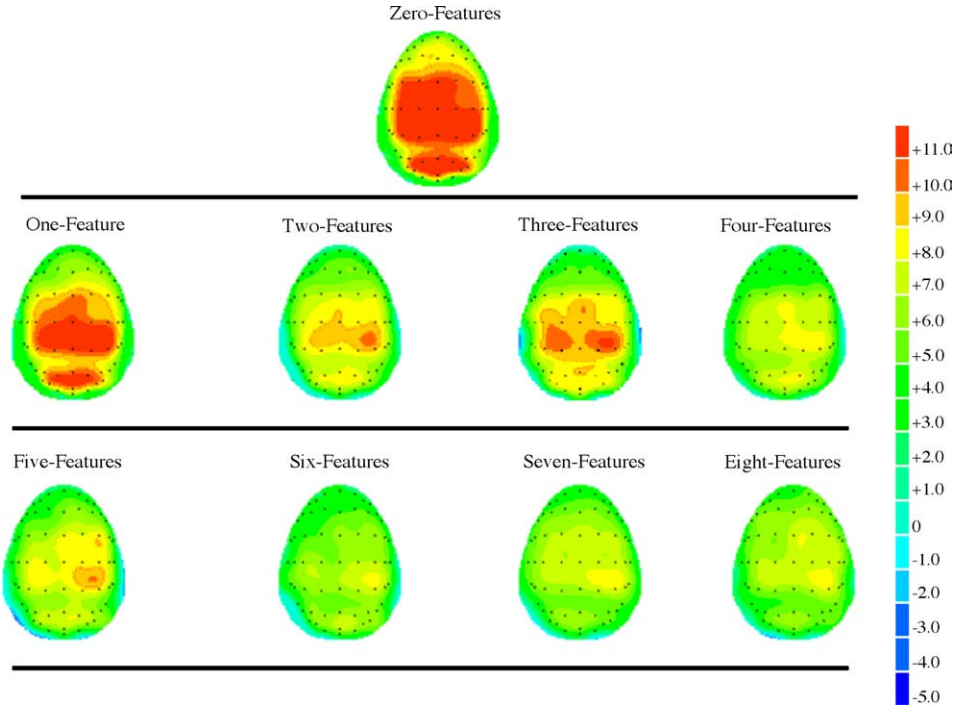

Fig. 6. Sixty-four channel ERP topographical maps shown at the time point at which the positive peak between 300 and 700 latency is maximum at Cz electrode site, Experiment 2.

every successive level of target similarity all are considered simultaneously in a single analysis. Collapsing across Experiments 1 and 2, an increase by one feature in target similarity caused an average increase in ERP amplitude of .41 μV (S.D. = .42 μV) to non-target stimuli; this effect differed significantly from zero, $t(28) = 5.29$, $p < .0001$, $\eta_p^2 = .195$.

We next recomputed the above contrast separately for participants in Experiment 1 (with the relatively complex "face" target) and Experiment 2 (with the relatively simple "oval" target). An increase by one feature in target similarity caused an average increase in ERP amplitude of .33 μV (S.D. = .39 μV) to non-target stimuli ($t(16) = t3.43$, $p < .01$, $\eta_p^2 = .15$) among participants assigned the eight-feature "face" target and an average increase of .52 μV (S.D. = .44 μV) to non-target stimuli ($t(11) = 4.15$, $p < .01$, $\eta_p^2 = .26$) among participants assigned to the zero-feature target. The sizes of these effects did not differ significantly as a function of Target Complexity ($t(27) = 1.24$, n.s.), nor was there a significant departure from equality of variances, $F(11, 16) = 1.21$, n.s. These findings suggest a robust impact of target similarity on amplitude of responses to non-targets, irrespective of target complexity.

3.2.5. Analyses of ERP latencies

Although not central to our hypotheses, latency to peak ERP amplitude also warrants investigation, given much previous evidence that increasing stimulus complexity and compatibility increases ERP latency (e.g., McCarthy and Donchin, 1981). The time point at which P300 amplitude reached its maximal peak across the 300–700 ms window was defined as P300 latency. Peak latencies from both experiments were analyzed in a 2 (target complexity) × 9 (target similarity) × 5 (row) ANOVA, with repeated measures on the last two factors. There was a significant effect of Target Similarity, $F(8, 216) = 12.34$, $p < .0001$, $\eta_p^2 = .31$, but it was moderated by a target complexity × target similarity interaction, $F(8, 216) = 5.51$, $p < .005$, $\eta_p^2 = .17$. Clarifying the nature of this interaction, separate analyses of the two experiments revealed a significant effect of similarity on ERP latency among participants (in Experiment 1) assigned a relatively complex target ($F(8, 128) = 14.94$, $p < .0001$, $\eta_p^2 = .48$) but not among participants (in Experiment 2) assigned a relatively simple target ($F(8, 88) = 1.72$, n.s.). In Experiment 1, peak latencies to the (complex) target, at 554.68 ms (S.D. = 62.36), were later by 101.48 ms (S.D. = 53.38) than were peak latencies to the combined average of all other stimuli; this effect differed significantly from zero, $t(16) = 7.84$, $p < .0001$, $\eta_p^2 = .47$. In contrast, in Experiment 2, peak latencies to the (simple) target, at 468.50 ms (S.D. = 103.06), were slower by 2.14 ms (S.D. = 47.83) than were peak latencies to the combined average of all other stimuli, which did not differ significantly from zero, $t(11) = .15$, n.s.

Accordingly, as in much previous research (e.g., McCarthy and Donchin, 1981), peak latency was longer for the complex target (in Experiment 1) than for the simple target (in Experiment 2), $t(27) = 5.34$, $p < .0001$. In contrast, when the same (zero-feature versus eight-feature) stimuli served as non-targets for participants of Experiment 1 ($M = 495.44$, S.D. = 125.11) and Experiment 2 ($M = 468.50$, S.D. = 103.06), respectively, there was no significant difference in peak latency between them, $t(27) = .63$, n.s.

4. General discussion

In the two experiments described in this article, participants performed a target-detection task in which non-target stimuli varied in target similarity but occurred with equal probability. The most important finding to emerge was that non-target stimuli that were perceptually similar to targets produced larger P300-like neurophysiological responses than did other non-target stimuli. These effects were remarkably similar whether participants' target was a (relatively complex) face (Experiment 1) or an (extremely simple) empty oval (Experiment 2). Non-targets' degree of similarity to targets, then, rather than their perceptual properties, appears to underlie these effects. Moreover, because target-similar non-targets in Experiment 2 contained only an oval plus a single feature, the presence of many constituent features does not appear necessary to trigger increases in categorical processing of non-targets that are similar to a target.

4.1. Implications for P300 research and categorization

Following work by Courchesne and colleagues, infrequent non-target processing has been examined almost exclusively in the context of the three-stimulus oddball paradigm (Courchesne et al., 1975). This design has provided much of the neurophysiological data examining stimulus deviation and novelty (e.g., Polich and Comerchero, 2003). A methodological limitation of extant applications of that experimental paradigm, however, is that deviation is limited between three sets of stimuli: targets ($p = .10$), non-targets ($p = .10$), and standards ($p = .80$). A problem in this design is the dissociation between non-target probability and intrinsic characteristics. When low probable non-target stimuli are interspersed between targets and standards, non-targets produce P300-like components regardless of context and perceptual characteristics. The non-target positive components are driven by low probability and their scalp distribution is influenced on the basis of stimulus novelty and context. In the present experiment, non-target categories contained equal number of stimuli occurring with the same probability. This approach allowed a systematical comparison between non-targets that deviated systematically on the basis of their target-like characteristics.

While target stimuli were influenced by active detection and low probability, non-target categories were independent of these factors and occurred under identical task conditions.

4.2. P300 amplitude and non-target similarity

Past studies have shown that perceptual and attentional processes affect the P300 amplitude (for a detailed discussion, see Donchin and Coles, 1988). One important question in regards to the present non-target P300s is the eliciting function. Given that P300 served as an index of target evaluation, why were intermediate P300s produced to correctly classified non-targets? We propose that intermediate non-target P300s reflected a cognitive process that was related to the ease of target and non-target discrimination. When targets and non-targets were dissimilar, classification was simple and no P300 was elicited. Conversely, stimuli with target-like properties captured attention and required additional mental effort for discrimi- nation. Consequently, these mental processes manifested in topographically target-like, but smaller in amplitude P300s. This view is also in accordance with past studies indicating that P300 amplitude is attenuated or diminished when the eliciting stimulus is non-informative or predictable (e.g., Donchin and Israel, 1980). In this perspective, it can be stated that equally probable stimuli that were easily evaluated as non-target required less mental work for discrimination and produced no P300-like components.

4.3. P300 latency and non-target similarity

There is strong evidence that P300 latency reflects speed of information processing independent of response selection (Kutas et al., 1977). This observation is known in the literature as the stimulus evaluation hypothesis. More specifically, P300 latencies are often prolonged when the discrimination between stimuli is made difficult (e.g., Ritter et al., 1979). Our findings show that P300 latency was affected on the basis of target complexity. The mean P300 latency for complex targets was 545 and 446 ms for the simple targets. The same pattern was observed in the pilot RT experiment. As stimulus evaluation is necessary for categorization, P300 latency was affected according to the complexity of stimulus discrimination. In the complex-target experiment, stimuli had to be scanned or computed for the presence or absence of features. It is likely that this process prolonged evaluation manifesting long P300 latencies. In the simple-target experiment, target evaluation required detecting an oval shape with no features. In this task, irregardless of number, the presence of features facilitated discrimination between target and non-target stimuli.

4.4. Similarity-based categorization

Similarity is a central factor in classification and most psychological models of categorization are based on similarity. Tversky (1977) introduced the first comprehensive model of similarity and described similarity as a feature-matching process. According to this model, objects are represented by distinctive features and similarity increases with the addition or deletion of distinct features. This concept has been fundamental in developing generalization in learning models, memory templates, and recognition. Traditionally, the most common approaches to study of the categorization have been stimulus rating, sorting and naming of objects, and time required to respond whether stimuli are similar or different (e.g., Podgorny and Garner, 1979). These approaches have been employed to explain the cognitive bases of similarity such as common and distinctive features, conceptual knowledge, effects of context, and other phenomena. The utility of the ERP methodology in combination with the oddball paradigm is advantageous in several aspects. First, categorization is measured indirectly and similarity judgments are not a requisite of the task. That is, participants are not asked to make spontaneous decisions about whether ‘chess’ is a ‘game’ or a ‘sport’. Thus, the brain’s metric for similarity can be assessed without biasing the subjects’ judgment. The present findings validate that brain responses to non-target stimuli are not arbitrary and can be utilized as indices of similarity-based categorization. These findings were consistent across stimulus materials that varied in complexity and simplicity.

Acknowledgements

The authors would like to thank Dr. Andreas Keil and two anonymous reviewers for insightful comments and suggestions on a previous version of this manuscript.

References

Anderson, J.R., 1991. The adaptive nature of human categorization. Psychological Review 98, 409–429.

Ashby, F.G., 1992. Multidimensional models of categorization. In: Ashby, F.G. (Ed.), Multidimensional Models of Perception and Cognition. Erlbaum, Hillsdale, NJ, pp. 449–483.

Courchesne, E., Hillyard, S.A., Galambos, R., 1975. Stimulus novelty, task relevance and the visual evoked potential in man. Electroencephalography and Clinical Neurophysiology 39, 131–143.

Delorme, A., Thorpe, S.J., 2001. Face identification using one spike per neuron: resistance to image degradations. Neural Networks 14, 795–803.

Donchin, E., Coles, M.G.H., 1988. Is the P300 component a manifestation of context updating? Behavioral and Brain Sciences 11, 357–374.

Donchin, E., Israel, J.B., 1980. Event-related brain potentials and psychological theory. In: Kornhuber, H.H., Deecke, L. (Eds.), Motivation, Motor and Sensory Processes of the Brain: Electrical Potentials, Behavior Clinical Use. Elsevier/North Holland, Amsterdam, pp. 697–715.

Duncan-Johnson, C.C., Donchin, E., 1977. On quantifying surprise: the variation of event-related potentials with subjective probability. Psychophysiology 14, 456–467.

Duncan-Johnson, C.C., Donchin, E., 1982. The P300 component of event-related brain potential as an index of information processing. Biological Psychology 14, 1–52.

Fize, D., Boulanouar, K., Chatel, Y., Ranjeva, J.P., Fabre-Thorpe, M., Thorpe, S., 2000. Brain areas involved in rapid categorization of natural images: an event-related fMRI study. Neuroimage 11, 634–643.

Fodor, J.A., 1970. Three reasons for not deriving "kill" from "cause to die". Linguistic Inquiry 1.4, 429–438.

Fujihara, N., Nageishi, Y., Koyama, S., Nakajima, Y., 1998. Electrophysiological evidence for the typicality effect of human cognitive categorization. International Journal of Psychophysiology 29, 65–75.

Greenhouse, W.W., Geisser, S., 1959. On methods in the analysis of profile data. Psychometrika 24, 95–112.

Greenwald, A.G., Gonzalez, R., Harris, R.J., Guthrie, D., 1996. Effect sizes and p values: what should be reported and what should be replicated? Psychophysiology 33, 175–183.

Koenig, P., Smith, E.E., Glosser, G., DeVita, C., Moore, P., McMillan, C., Gee, J., Grossman, M., 2005. The neural basis for novel semantic categorization. Neuroimage 15, 369–383.

Kotchoubey, B., Lang, S., 2001. Event-related potentials in an auditory semantic oddball task in human. Neuroscience Letters 310, 93–96.

Kutas, M., McCarthy, G., Donchin, E., 1977. Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. Science 197, 792–795.

Lamberts, K., 2000. Information–accumulation theory of speeded categorization. Psychological Review 107, 227–260.

McCarthy, G., Donchin, E., 1981. A metric for thought: a comparison of P300 latency and reaction time. Science 211, 77–79.

Medin, D.L., Schaffer, M.M., 1978. Context theory of classification learning. Psychological Review 85, 207–238.

Mervis, C.B., Rosch, E., 1981. Categorization of natural objects. Annual Review of Psychology 32, 89–115.

Nosofsky, R.M., 1986. Attention, similarity, and the identification-categorization relationship. Journal of Experimental Psychology: General 115, 39–57.

Nosofsky, R.M., Palmeri, T.J., 1997. An exemplar-based random walk model of speeded classification. Psychological Review 104, 266–300.

Picton, T.W., 1992. The P300 wave of the human event-related potential. Journal of Clinical Neurophysiology 9, 456–479.

Podgorny, P., Garner, W.R., 1979. Reaction time as a measure of inter- and intraobject visual similarity: letters of the alphabet. Perception and Psychophysics 26, 37–52.

Polich, J., Comerchero, 2003. P3a from visual stimuli: typicality, task, and topography. Brain Topography 15, 141–152.

Ritter, W., Simson, R., Vaughan, H.G., Friedman, D.A., 1979. A brain event related to the making of a sensory discrimination. Science 203, 1358–1361.

Squires, N.K., Squires, K.C., Hillyard, S.A., 1975. Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. Electroencephalography and Clinical Neurophysiology 38, 387–401.

Sutton, S., 1979. P300-13 years later. In: Begleiter, H. (Ed.), Evoked Brain Potentials and Behavior. Plenum Press, New York, pp. 107–126.

Sutton, S., Braren, M., Zubin, J., John, E.R., 1965. Evoked-potential correlates of stimulus uncertainty. Science 150, 1187–1188.

Thorpe, S., Fize, D., Marlot, C., 1996. Speed of processing in the human visual system. Nature 381, 520–522.

Tversky, A., 1977. Features of similarity. Psychological Review 84, 327–350.