

**Novelty and conflict in the categorization of complex stimuli**

Jonathan R. Folstein (1,2), Cyma Van Petten (1,3), and Scott A. Rose (4)

Running head: Novelty and conflict in the categorization of complex stimuli

1. Department of Psychology, The University of Arizona, Tucson, AZ
2. Department of Psychology, Vanderbilt University, Nashville, TN
3. Department of Psychology, Binghamton University, Binghamton, NY
4. Department of Psychology, The Queens College, Flushing, NY

Correspondence:  
Jonathan Folstein  
Psychology Department  
Vanderbilt University  
301 Wilson Hall  
111 21st Avenue South  
Nashville, TN 37203

### **ABSTRACT**

We manipulated categorical typicality and the presence of conflicting information as participants categorized multi-featured artificial animals. In Exp. 1, rule-irrelevant features were correlated with particular categories during training. In the test phase, participants applied a one-dimensional rule to stimuli with rule-irrelevant features that were category-congruent, category-incongruent, or novel. Category-incongruent and novel features delayed RT and P3 latency, but had no effect on the N2. Exp. 2 used a two-dimensional rule to create conflict between rule-relevant features. Conflict resulted in prolonged RTs and larger amplitudes of a prefrontal positive component, but had no impact on the N2. Stimuli with novel features did elicit a larger N2 than those with frequent features. These results suggest limitations on the generality of the N2's sensitivity to conflicting information, while confirming its sensitivity to attended visual novelty.

The study of perceptual categorization has long been concerned with the effects of typicality, selective attention to an object's features, and the need to suppress the influence of features suggesting the incorrect category (Allen & Brooks, 1991; Nosofsky, 1986; Regehr & Brooks, 1993; Rosch & Mervis, 1975; Smith, Patalano, & Jonides, 1998). The current study focuses on the influences of typicality and conflicting information during visual categorization, and explores the degree to which ERP measures will reflect and perhaps help tease apart these two factors.

For natural categories, typical exemplars (e.g., SPARROW as a type of bird) elicit faster verification responses than atypical exemplars (e.g., CHICKEN). Typicality can be thought of as the extent to which an item overlaps with the central tendency of a category, or the amount of positive evidence supporting the assignment of an item to a category. Category verification times are also influenced by the presence of conflicting information, or the amount of evidence suggesting assignment to the wrong category. For natural categories, this is most easily seen in prolonged RTs to reject items that share features with a target category, but are not in fact members (as in rejecting BAT as a type of bird; Heinze, Münte, & Kutas, 1998; McCloskey & Glucksberg, 1979). Despite these examples, it is fairly difficult to tease apart the influences of typicality and conflicting information, particularly for trials that receive the same response. For instance, as compared to sparrows, chickens are poor birds because they lack some common bird attributes like the ability to fly (placing them far from the central tendency of the "bird" category), but their use as a food animal may also create links with alternate categories whose members are primarily mammals (e.g., cows and pigs) such that access to the "used for food" property of chickens might pose a conflict with categorizing them as birds.

When well-known exemplars are used as stimuli, it is also likely that participants draw on a variety of different sorts of information to make category judgments, some more perceptual in nature, some more closely related to function (e.g., use for food), and some that are difficult to classify (e.g., typical geographical location). Categorization researchers have frequently turned to the use of artificial stimuli, for which categories are learned *de novo* in the laboratory, judgments can be more closely tied to perceptual features of the stimuli, and which may allow a cleaner distinction between typicality and conflicting information.

Allen and Brooks (1991) used artificial stimuli to demonstrate that atypical and/or conflicting exemplars resulted in more errors and slowed category decisions, even when participants were instructed in a rule that could yield perfect classification (see also Patalano, Smith, Jonides, & Koeppel, 2001 for a partial replication). During initial training, participants learned to categorize cartoon animals as "diggers" or "builders" based on a two-out-of-three rule – that an animal belonged to a given category if it possessed at least two of three diagnostic features (e.g., long/short neck, angular/rounded body, spots/no-spots). The animals possessed two other features that were nondiagnostic because they occurred equally often in diggers and builders (e.g., number and length of legs). Although the nondiagnostic features were irrelevant to the assigned rule, their existence allowed manipulations of similarity between members of opposite categories, when the similarity was irrelevant to the instructed rule. During the test phase, participants viewed the trained exemplars, and new instances of both categories. Critically, some of the new exemplars belonged to one category according to the rule, but were very similar to a trained item in the opposite category because they were identical on four of five features — *negative matches*. *Positive matches* were also identical to a trained exemplar in four of five features, but belonged to the same category according to the rule. Negative matches elicited more categorization errors, and longer RTs even when they were correctly classified.

Allen and Brooks (1991) conclusively demonstrated that rule-irrelevant similarity between exemplars can interfere with rule application in categorization. However, the difference between positive- and negative-match stimuli could be attributed to differential typicality, or to decision conflict arising from misleading features. A more detailed analysis of their design indicates that positive-match items differed from trained exemplars of their category in only a single rule-relevant feature, whereas negative-match items differed from trained exemplars of their own category in two features, albeit nondiagnostic features that should have been irrelevant to rule application. If we use capital letters to denote rule-relevant features of Categories A and B, and numerals to denote nondiagnostic features, then a positive match might be an AAA22, differing in one rule-relevant feature from the closest trained exemplar of the same category, ABA22. In addition to its similarity to a trained “A”, the positive match also falls under the rule-based definition of having at least two A features. In contrast, a negative match could have a feature list of AAB12, when the closest trained member of its own “A” category was AAB21, and the closest trained member of the opposite Category B was ABB12. The negative match case is thus two features away from the closest trained member of its own category, in addition to its misleading similarity with the opposite category. This association between categorical atypicality and presence of misleading information cannot be avoided when all features are equally familiar from initial categorization training.

In the current experiments, we introduce new features during the test phases in order to disentangle the influences of typicality and conflicting information. Figure 1 illustrates the design of Experiment 1. Participants were instructed to categorize artificial animals according to one dimension, the type of body markings (spots or stripes). During training with feedback on this categorization rule, the two varieties of body markings were perfectly correlated with global body shape – humanoid versus horse-like creatures. During the test phase without feedback, these now *Typical* exemplars were intermixed with two other conditions. *Conflicting* exemplars were composed of familiar features (spots, stripes, humanoid and horse shapes), but in new combinations. If participants had learned to rely on feature combinations during the training phase – despite the fact that body shape is irrelevant to the rule – the *Conflicting* condition should lead to reduced accuracy and prolonged RTs, much as Allen and Brooks (1991) negative match condition. *Novel* exemplars contained the same rule-relevant features (body markings) as all the others, but a new body shape was introduced during the test phase. Both the *Conflicting* and *Novel* exemplars are thus atypical of their categories, but only the *Conflicting* items contain information that had previously been associated with a different categorization decision that is now incorrect for the test phase. The *Novel* exemplars are dissimilar from other members of their category, but contain no information suggesting a different category. We expected both of these conditions to elicit slower behavioral responses than *Typical* exemplars. Event-related potentials (ERPs) were recorded to shed additional light on whether perceptual typicality and decision conflict elicit similar or different brain activity during categorization. Below, we briefly review functional aspects of the P3 and N2 components that make them potentially useful tools for this endeavor.

---

Insert Figure 1 about here

---

Since they were first observed by Sutton and colleagues (Sutton, Braren, Zubin, & John, 1965), variation in the amplitude and/or latency of the P3 and N2 components of the event-related potential have been observed in a wide variety of tasks. When participants must classify

visual stimuli into a small number of categories (typically only two), stimuli that exemplify a less-probable category elicit larger P3s than those in a more-probable category (Breton, Ritter, Simson, & Vaughn, 1988; Courchesne, Hillyard, & Galambos, 1975; Simson, Vaughn, & Ritter, 1977; Squires, Donchin, Squires, & Grossberg, 1977). In the current experiments, response probability is equated across conditions of interest, so that variations in P3 amplitude are of somewhat less interest than variations in latency. Multiple studies offer strong support for the conclusion that P3 latency tracks the time required for evaluation of a stimulus with regard to the instructed classification principle (see Verleger, 1997 for a review). Presenting multiple, physically distinct examples of two categories as compared to only a single repeated exemplar of each category prolongs P3 latency (Kutas, McCarthy, & Donchin, 1977), as does embedding a stimulus in perceptual noise as compared to clear presentation (McCarthy & Donchin, 1981). We predicted that the Novel condition of Exp. 1 would elicit later P3 latencies than the Typical, because the perceptual variability of the irrelevant features would slow stimulus evaluation. P3 predictions for the Conflicting condition are less obvious, because the component features of these items are familiar, but it remains possible that the conflicting information would elicit prolonged stimulus evaluation.

The label “N2” encompasses what are clearly multiple negative-going voltage deflections peaking 200 to 400 ms after stimulus presentation (Folstein & Van Petten, in press; Pritchard, Shappell, & Brandt, 1991). We focus here on N2s with frontocentral scalp distributions, which are modulated by two rather different sorts of experimental manipulations that are both relevant for the current experiments. On the one hand, large anterior N2s are elicited by visual stimuli that are very distinct from other stimuli in an experiment (*novel*, typically unique on each trial) regardless of their target status (Courchesne et al., 1975; Czigler & Balázs, 2005; Daffner et al., 2000; Suwazono et al., 2000; Thomas & Nelson, 1996). On the other hand, anterior N2s are also larger in conditions that require inhibition of a prepared response (*no-go* trials), and/or contain elements suggesting two conflicting responses, as compared to conditions without response inhibition or response conflict (Nieuwenhuis, Yeung, van den Wildenberg, & Ridderinkhof, 2003; Pfefferbaum, Ford, Weller, & Kopell, 1985). The most commonly used paradigm for manipulating response conflict has been the Eriksen flanker task, in which participants make a choice response to a central letter that is “flanked” on both sides by letters associated with the same response or the opposite response (Gehring, Gratton, Coles, & Donchin, 1992; Heil, Osman, Wiegmann, Rolke, & Hennighausen, 2000; Kopp, Rist, & Mattler, 1996).

Manipulations of perceptual novelty and “cognitive control” (response inhibition, detection of conflict, monitoring or regulation of strategy) have largely taken place in different studies with relatively little crosstalk (but see Holroyd, 2004; Nieuwenhuis et al., 2003). In a more extensive review, we have suggested that the influences of stimulus novelty and response conflict on frontocentral negative peaks are sufficiently dissociable as to suggest that the anterior N2 consists of two different components (Folstein & Van Petten, in press). However, both should be useful in the study of perceptual categorization. At minimum, typical exemplars that are both familiar from the training phase and contain information consistent with one response should elicit smaller anterior N2s than stimuli that are perceptually novel and/or elicit a response conflict due to the presence of features associated with different responses. If both categorical atypicality and conflict can be shown to influence N2 amplitudes in the same experiment, it will be of some interest to compare the topographies and latencies of the two effects.

To date, studies showing novelty N2 effects have employed simple target detection tasks, while studies of cognitive control have used paradigms with very simple categorization rules and familiar stimuli such as the Eriksen flanker task. In the General Discussion, we take up a handful of recent studies from other laboratories that used more complex rules and/or stimuli. Our own initial study explored the ERP correlates of rule- and similarity-based categorization of complex multi-featured cartoon animals (Folstein & Van Petten, 2004). We observed a frontal negative peak from 200 to 600 ms that was larger to exemplars containing more features consistent with a different category than for those with fewer such features. Due to its frontal distribution and insensitivity to feature frequency, this effect was interpreted as sensitive to the initial detection of conflicting information, linking it with the family of cognitive control N2 components. It is also possible, however, that the effect reflected the atypicality of some exemplars as members of their category, linking it instead with the novelty N2.

To distinguish between these two possibilities, and to further explore ERP correlates of rule-based perceptual categorization, we report two experiments comparing the influence of conflicting category information with that of novel stimulus features. In Experiment 1, participants categorized cartoon animals using a single stimulus dimension. Like the studies of Brooks and colleagues, categories were also associated with rule-irrelevant stimulus features during a training phase. At test, the rule-relevant features were presented either with the same rule-irrelevant features as training (Typical), with the rule-irrelevant features associated with the opposite category during training (Conflicting), or with novel rule-irrelevant features not seen during training (Novel). In the second experiment, the assigned rule was two-dimensional, and the conflicting feature was relevant to the categorization rule.

## EXPERIMENT 1

### Methods

#### Participants

Eighteen healthy young adults (9 men, 9 women, mean age of 25 years) were paid for their participation after providing informed consent. All were free of neurological or psychiatric disorders by self-report, and also reported no medications likely to affect the central nervous system. They had a mean of 17 years of formal education. Sixteen participants were right-handed, two were left-handed. An additional five individuals participated, but generated unusable data: three due to an excessive (greater than 50%) number of trials contaminated by eyeblink or movement artifact, one who was found to have participated in a previous experiment with similar stimuli (reported in Folstein, 2007), and one whose accuracy in all conditions was close to zero.

#### Materials

Stimuli were 192 cartoon creatures that spanned 3.7 x 3.0 degrees of visual angle on average. Examples are shown in Figure 1. These were described to participants as creatures from a different planet, or “aliens”, so that we also use this term below. The central body of each alien contained either spots or stripes (equal numbers of each), and this dimension of body-marking determined category membership. Participants learned that *Mogs* contained spots, and *Nibs* stripes, or vice-versa (counterbalanced across participants). In addition to a body-marking, each alien had a global shape so that individuals shared a family resemblance of appearing humanoid, horse-like, or fish-like, although this family resemblance – or “species” -- was irrelevant to the categorization rule. Prior to the test phase, participants viewed stimuli in which

body-marking and global shape were correlated (e.g., Mogs are not only spotted, but always humanoid). At test, these were repeated as Typical exemplars. Combinations of body-marking and global shape that were the opposite of those viewed in training appeared in the Conflicting condition. The Novel condition was defined by combinations of spots/stripes and a global shape not experienced prior to the test phase. Across participants, “species” were rotated across conditions, so that horse-like stimuli appeared equally often as Typical Mogs (and Conflicting Nibs), Typical Nibs (and Conflicting Mogs), and Novels equally often.

Each alien was comprised of a color (eight distinct colors) plus four rule-irrelevant dimensions that collectively made up the global body shape: the exact shape of the central body, legs (or tail fin for alien fish), head (or eye-stalks for alien fish), and one additional appendage (arms for humanoids, tails for horses, and dorsal fin for fish). For both the Typical and Conflicting exemplars, there were two alternate features for each dimension: two different horse heads, two different horse tails, etc. These were assembled such that each Typical combination of body marking and global shape was exemplified by four unique combinations of body parts, each occurring in eight colors, for a total of 32 individuals (i.e., Typical Mogs, with another 32 Typical Nibs). As Figure 1 implies, every Conflicting stimulus corresponded to a Typical stimulus that was identical in every way except for body markings (64 individual Conflicting stimuli).

Novel stimuli were created in a similar manner, except that there were eight versions of each individual rule-irrelevant feature (i.e., shape of the dorsal fin for the fish, or shape of the arms for the humanoids). This allowed for the creation of 64 unique shapes for the Novel items (although all possessing the same global body shape, or “species”), which were evenly divided across the 8 colors. The large variety of specific features used in the rule-irrelevant dimensions also meant that the Novel features had a low probability of occurrence across the experiment, although each Novel stimulus included a high-frequency, rule-relevant body marking. Average frequencies of the rule-relevant (spots, stripes) and rule-irrelevant features are shown in Table 1 for each condition.

---

Insert Table 1 about here

---

## **Procedure**

After electrode application, participants sat in a reclining chair ~100 cm from a 20” (50.8 cm) video monitor under the control of a PC computer. Key-presses with the left and right index fingers were used to signal decisions; mapping between response hand and decision was counterbalanced across participants. The experimental session was divided into three phases: pre-exposure to Typical stimuli while performing a feature verification task, training on the categorization rule, and finally a test phase which generated accuracy, reaction time, and ERP data.

The purpose of the initial feature verification task was to induce participants to attend to the rule-irrelevant features of the stimuli, and to associate them with the rule-relevant body-markings of their categories. Participants viewed a Typical alien (duration 2 sec) followed by a 3 sec retention interval, then one of the four rule-irrelevant features presented in isolation (duration 1 sec). Participants then indicated whether or not the part was included in the sample alien and received accuracy feedback. Participants performed feature verification on each Typical alien twice for a total of 128 trials. The four stimulus dimensions were presented with approximately equal frequency (25 or 26 times) on verification trials.

After the feature verification task, participants were informed that the aliens could be divided into two categories, called Mogs and Nibs, based on their body markings. They then performed a categorization task on the 64 Typical aliens (32 Mogs and 32 Nibs). Stimulus duration was 600 ms, and each alien was followed by the correct category name three seconds later. Participants were instructed to take their time, but to respond before the feedback appeared on the screen. Participants always learned the difference between Mogs and Nibs within the first few trials.

During the test phase, participants categorized a total of 256 aliens without feedback: 128 Typical (2 presentations of each of the 64), 64 Novel, and 64 Conflicting. Stimulus duration was 600 ms and inter-stimulus interval was 4 sec. The larger number of Typical aliens was to insure that the associations between irrelevant features and their categories formed in training were maintained to some extent during test.

### **Electrophysiological Methods**

The electroencephalogram (EEG) was recorded from tin electrodes in an elastic cap (Electrocap International, Eaton, OH). Seven electrodes spanned the midline of the scalp from prefrontal to occipital (Fpz, Fz, Fcz, Cz, Cpz, Pz, Oz). Seven dorsal lateral pairs were used (Fp1/2, F3/4, Fc3/4, C3'/4' (placed 4 cm lateral to Cz), Cp3/4, P3/4, and O1/2), and four ventral lateral pairs spanned inferior frontal and temporal sites (F7/8, Ft7/8, Tp7/8, T5/6). The vertical electrooculogram (EOG) was recorded from an electrode below the right eye. Scalp and vertical EOG electrodes were referenced to the left mastoid during the recording, and digitally re-referenced to the mean of the right and left mastoids offline. Electrodes lateral to the external canthi of the two eyes were used to record the horizontal EOG. The EEG was amplified by a Grass Model 12 polygraph (Grass, West Warwick, RI) with a gain of 50,000 and half-amplitude cutoffs of 0.01 and 100 Hz, digitized online at a sampling rate of 250 hz, and stored on CD along with stimulus codes for subsequent averaging. Impedances for scalp channels were less than or equal to 5K Ohms; EOG channels were less than 10K Ohms. Trials contaminated by eye movements (identified as high-amplitude activity on the horizontal EOG channel), blinks (identified as high-amplitude polarity inversions between the lower eye and Fpz channels), amplifier saturation (identified as flat lines), or movement artifact (identified as activity exceeding  $\sim 100 \mu V$  on any channel) were rejected prior to averaging. ERPs were averaged for 700 ms post-stimulus, with a 200 ms pre-stimulus baseline. After excluding trials with incorrect behavioral responses and electrical artifacts, an average of 104, 53, and 52 trials per subject remained in the Typical, Conflicting, and Novel conditions, respectively.

### **Measurement and analysis**

Data were analyzed via ANOVAs with factors of Condition (3 levels), anterior-to-posterior scalp location (4 levels), and left vs. midline vs. right location (3 levels). When these omnibus ANOVAs yielded significant effects of condition, they were followed by comparisons between pairs of conditions. The Huynh-Feldt correction for non-sphericity of variance was used for all F-ratios with more than one degree of freedom in the numerator.

N2 was defined as the mean amplitude from 190 to 370 ms post-stimulus onset at twelve prefrontal, frontal and central scalp sites: Fp1, Fpz, Fp2, F3, Fz, F4, Fc3, Fcz, Fc4, C3', Cz, and C4'. Because one recent report includes N2 effects that are larger in averages timelocked to the behavioral responses rather than the stimuli (Nieuwenhuis et al., 2003), we also examined N2

amplitude in response-locked averages, but no significant effects were observed, so that we omit this null aspect of the results (which is reported in Folstein, 2007).

For each participant, P3 peak amplitude was defined as the most positive data point between 300 and 700 ms at centroparietal and parietal scalp sites Cp3, Cpz, Cp4, P3, Pz, and P4, after applying a 10 Hz low-pass filter; P3 peak latency was defined as the latency of that same data point. We chose to quantify the P3 using peak rather than mean amplitude measures due to differing latencies across conditions (described in Results). P3 measures were analyzed via ANOVAs as above, but with two levels of the anterior-to-posterior spatial factor.

## Results

**Accuracy and reaction time.** Table 2 shows that categorization accuracy was very high overall, but small differences among conditions led to a main effect of condition ( $F(2,34) = 8.29$ ,  $p < .001$ ,  $\epsilon = 1.00$ ). Followup analyses showed that participants were slightly more accurate in categorizing Typical stimuli than Novel or Conflicting stimuli ( $ts(17) > 2.90$ ,  $ps < .01$ ), which did not differ from one another ( $t(17) = 1.05$ ). The omnibus ANOVA for RT yielded a main effect of condition ( $F(2,34) = 20.7$ ,  $p < .0005$ ,  $\epsilon = 1.00$ ). Followup analyses showed that responses were faster to Typical than to Novel and Conflicting stimuli, ( $ts(17) > 4.69$ ,  $ps < .001$ ), which did not differ from one another ( $t(17) = 0.52$ ).

---

Insert Table 2 about here

---

**Event Related Potentials.** Table 2 and Figure 2 show that there were no differences in N2 or P3 amplitude among conditions ( $F_s < 1.22$  for main effects,  $F_s < 1.68$  for interactions with scalp site). The figure and table show that P3 latency was shorter for the Typical condition than the other two, leading to a main effect of condition in the omnibus ANOVA ( $F(2,34) = 5.94$ ,  $p < .01$ ,  $\epsilon = 1.0$ ), and main effects in pairwise ANOVAs comparing Typical to the other conditions ( $F_s(1,17) > 9.61$ ,  $ps < .01$ ), which did not differ from each other ( $F < 1$ ).

## Discussion

During the test phase, categorization decisions were slower for stimulus classes that had not been encountered during the training phase. This was equally true of Novel stimuli whose global body shape was newly introduced during the test phase, and of Conflicting stimuli that were comprised of familiar body shapes but a rule-relevant feature that specified the opposite category than that of the training phase. P3 latencies paralleled the RT results in showing a latency delay for stimuli other than the Typical exemplars repeated from the training phase. These effects indicate that participants did note the rule-irrelevant perceptual features of the stimuli. However, the pattern of results suggests only a beneficial impact of familiarity – speeded responses in the presence of a frequent combination of frequent features in the Typical condition. There was no suggestion that the reversed-from-usual combination of rule-relevant and rule-irrelevant features in the Conflicting condition led to a unique delay in assigning a category.

The ERP results contained no hint that the N2 was modulated by either response conflict or perceptual novelty as implemented here. We had hypothesized that, by associating rule-irrelevant features with one category assignment, and using a feature verification task to call attention to these associations, irrelevant features in the Conflicting condition would cue the retrieval of exemplars or prototypes from the incorrect category, prompting incorrect response

preparation and triggering an anterior N2 when the opposing response was activated or selected. However, as described below, some prior studies show that categorization performance can be insensitive to rule-irrelevant dimensions when perfect performance can be achieved via attention to a single dimension.

Behavioral studies examining the impact of rule-irrelevant features when a single dimension perfectly predicts category membership indicate that correlations between rule-irrelevant features and the single relevant dimension influence categorization behavior in only limited circumstances. Garner and Felfoldy (1970) instructed participants to sort circles by size or by the orientation of a radial line. In addition to a baseline set of stimuli with no variation in the irrelevant dimension, participants categorized sets in which the irrelevant dimension was perfectly correlated with the relevant dimension, or variable but uncorrelated. As compared to the baseline condition, no benefit from the correlation of rule-relevant and rule-irrelevant features was observed, and very little cost of irrelevant variability. In a study inspired by the phenomenon of blocking in classical conditioning, Williams and colleagues have similarly shown that when participants are instructed that a single element can cause an outcome, and have experience in this condition, they assign little predictive value to a subsequently correlated element (Williams, Sagness, & McPhee, 1994). In a single-phase categorization task, Nosofsky (1991) asked participants to categorize lines of various lengths as long or short, but varied the probability that short lines had a particular (rule-irrelevant) orientation. The probability manipulation had no significant effect on categorization speed or accuracy, or on subsequent judgments of categorical typicality when stimuli were judged in isolation.

There is one prominent exception to the general finding that correlations between rule-relevant and rule-irrelevant features do not influence categorization when a unidimensional rule is in place. In the same card-sorting task described above, Garner and Felfoldy (1970) found a speed benefit when the brightness and saturation of color patches were correlated, although only one dimension was rule-relevant. This benefit disappeared when brightness and saturation were co-varied, but in spatially-separated color patches appearing on the same card. More generally, it has been suggested that a significant influence of correlated features during unidimensional rule use constitutes evidence that the two dimensions in question are perceptually integral, whereas null effects indicate that the dimensions are separable (Nosofsky & Palmeri, 1997).

In Exp. 1, participants using a unidimensional rule were not entirely insensitive to the frequency of features and their combinations, as they showed lower accuracies, slower RTs, and longer P3 latencies when exemplars were comprised of rare combinations of familiar features (Conflicting) and when they were comprised of rare features (Novel), as compared to familiar combinations of familiar features (Typical). However, the lack of difference between the Conflicting and Novel conditions suggests that, although unfamiliarity slows stimulus evaluation, features that were irrelevant to a unidimensional rule were not associated with a category despite their objective correlation with that category. In Garner's (1970) terms, the global body shape and body markings proved to be separable dimensions, although they occupied the same spatial locations. The absence of a large N2 in the Conflicting condition suggests that potentially misleading information occurring in a dimension separate from that governing rule application does not result in response conflict, at least when a unidimensional rule is used.

The novelty manipulation implemented in Exp. 1 was also somewhat different than those of previous ERP studies showing enhanced N2s. In one type of previous paradigm, the novel stimuli that elicited large frontocentral N2s were nontargets that demanded no response from the

participants, but were unique in the context of the experiment (Courchesne et al., 1975; Czigler & Balázs, 2005; Daffner et al., 2000). In a second type of paradigm, the stimuli eliciting larger N2s mismatched an immediately prior stimulus in one or more visual features, and participants were asked to make same/different judgments about the sequentially presented pairs of stimuli (e.g. Wang, Cui, Wang, Tian, & Zhang, 2004; Wang, Tian, Wang, Cui, Zhang, & Zhang, 2003). In contrast to both of these general sorts of paradigms, the Novel stimuli of Exp. 1 contained task-relevant features that were not unusual or unfamiliar, namely the body markings that determined category membership. The results suggest that, when a unitary task-relevant feature is present, unusual but task-irrelevant features have little impact on the N2.

## EXPERIMENT 2

The design of Experiment 2 includes conflicting information regarding category assignment, and a manipulation of perceptual novelty, but both are implemented within perceptual dimensions that are relevant to the assigned categorization task. Participants were instructed on a two-dimensional rule that mandated attention to both global body shape and body markings. Conflicting trials are those in which the two features occupying these dimensions are diagnostic of different categories, and participants must resolve the tie by responding “Other”. Novel trials are those in which the two features occupying the rule-relevant dimensions are only rarely encountered, and also require a response of “Other”. We predicted that both response conflict and low feature-probability would have more potent effects on reaction time and brain activity than the variations in irrelevant features used in Exp. 1, because both stimulus dimensions are relevant to the assigned categorization task.

## Methods

### Participants

Eighteen healthy young adults (eight men, ten women, mean age of 27 years) were paid for their participation after providing informed consent. All were free of neurological or psychiatric disorders by self-report, reported no medications likely to affect the central nervous system, and had not participated in any other categorization experiments in our laboratory. They had a mean of 17 years of formal education. Seventeen participants were right-handed, one was left-handed. An additional five individuals participated, but generated unusable data: four due to an excessive (greater than 50%) number of trials contaminated by eyeblink or movement artifact, and one due to computer failure.

### Materials

Stimulus examples are shown in Figure 3. Global body shapes, or “species”, were humanoid or horse-like or fish-like, and were filled with spots or stripes or a brick pattern. Category membership was determined by a conjunction of species and body markings, such that for the examples in the figure, spotted horses are Mogs, striped humanoids are Nibs, and anything else is an Other (the mapping between species and category membership was rotated across participants). As described in greater detail below, “Others” formed a perceptually heterogeneous group of stimuli, but a single response category, such that completely accurate responding in the test phase would yield one-third “Mog” responses, one-third “Nib” responses, and one-third “Other” responses.

1) *Typical Mogs* and *Typical Nibs* contained the two features diagnostic of their category. As in Exp. 1, there were multiple variants of the four features contributing to global body shape

(or species): eight different “horse” heads, eight different humanoid arms, etc. An additional rule-irrelevant dimension present in all exemplars was color (eight distinct colors).

Combinations of different heads, legs, colors, etc were used to create 32 individually unique Mogs and 32 individually unique Nibs. Typical Mogs and Nibs are exemplars that were repeated in identical form from the training phase to the test phase.

2) *Similar Mogs* and *Similar Nibs* (32 individual exemplars in each category) also contained the two features diagnostic of their category, but differed by two rule-irrelevant features (e.g. head and tail) from the most similar Typical stimulus encountered during the training phase. These were included to test whether participants were sensitive to repetition of individual body parts from the training phase or if, instead, they were sensitive primarily to global body shape.

3) *Conflicting Others* were identical to the Typical Mogs and Nibs, except that their body markings were switched across category. Thus, Conflicting stimuli always had one Mog feature and one Nib feature, precluding them from membership in either category.

4) *Novel Others* had global body shape (species) and body-markings that were unassociated with either the Mog or Nib categories (see Figure 3). As a consequence, Novel species and markings were presented less frequently than those of Mogs and Nibs, as shown in Table 3.

5) *Training Others* had one Mog or Nib property (either markings or species with equal probability) and one that belonged to neither category. The purpose of the Training Others was to motivate participants to attend to both relevant dimensions. If a participant attended to only one dimension during training, his or her responses would be correct for all the Mogs and Nibs, but incorrect for half of the Training Others.

---

Insert Table 3 and Figure 3 about here

---

## Procedure

Because both the conflict and novelty manipulations were implemented within the rule-relevant dimensions of global body shape and body-markings, the preliminary feature verification task was no longer necessary to induce attention to these features. Participants were immediately instructed how to categorize the aliens into Mogs, Nibs, and Others using the two dimensional conjunctive rule described above. They were shown eight members of each category in random order for a duration of five sec each, followed by the appearance of the word “Mog”, “Nib”, or “Other”, corresponding to the alien’s category. After this initial training, participants pressed keys to categorize 64 Mogs, 64 Nibs, and 64 Training Others, receiving feedback on their accuracy after each trial. In order to maximize response conflict in the later Conflicting condition, the Mog and Nib keys were always on the same hand while the Other button was always on the index finger of opposite hand (right versus left hands for Other were counterbalanced across participants).

During the test phase without feedback, participants categorized 128 Typical aliens, 64 Similar aliens, 48 Conflicting Others, and 48 Novel Others for a total of 288 stimuli divided into eight trial blocks. This list composition ensured that aliens requiring “Mog”, “Nib”, and “Other” responses occurred with equal probability (96 each).

Electrophysiological recording, measurement and analysis methods were the same as in Exp. 1, except that ERPs were averaged in 900 ms epochs with respect to a 100 ms pre-stimulus baseline. After excluding trials with incorrect behavioral responses and electrical artifacts, an

average of 98, 47, 36, and 36 trials per subject remained in the Typical, Similar, Conflicting, and Novel conditions, respectively.

## Results and Discussion

**Accuracy and reaction time.** Table 4 shows that neither accuracy nor RT differed between the Typical and Similar conditions ( $ts(17) < 1.6$ ), which were collapsed into a single *Categorical* condition for subsequent analyses. Omnibus ANOVAs comparing the Categorical, Conflicting and Novel conditions yielded main effects of condition for both accuracy ( $F(2,34) = 31.3, p < .001, \epsilon = .59$ ) and RT ( $F(2,34) = 79.9, p < .0001, \epsilon = .62$ ). Accuracy in the Conflicting condition was 6 to 7 percent lower ( $ts(17) > 4.5, ps < .0005$ ), and RTs more than 200 ms slower than either the Categorical or Novel conditions ( $ts(17) > 8.0, ps < .0005$ ). Novel stimuli received slightly more accurate ( $t(17) = 2.8, p < .05$ ), and slightly faster responses than Categorical ( $t(17) = 2.4, p < .05$ )

---

Insert Table 4 about here

---

Figure 4 shows ERPs in the four conditions at all electrode sites. The Typical and Similar conditions elicited indistinguishable waveforms, so that these were collapsed into a single condition of *Categorical* in statistical analyses.

---

Insert Figure 4 about here

---

**N2.** The omnibus ANOVA showed a main effect of condition ( $F(2,34) = 5.19, p < .01, \epsilon = 1.00$ ). N2 amplitudes did not differ between the Conflicting and Categorical conditions ( $F < 1$ ). Unlike Exp. 1, Novel stimuli elicited a larger N2 than the other conditions (Novel vs. Categorical:  $F(1,17) = 8.65, p < .01$ ; Novel vs. Conflicting:  $F(1,17) = 5.92, p < .05$ ). The novelty effect interacted with the left vs. midline vs. right location of scalp sites in both comparisons (Novel vs. Categorical: ( $F(2,34) = 3.56, p < .05, \epsilon = .75$ ; Novel vs. Conflicting: ( $F(2,34) = 3.80, p < .05, \epsilon = .78$ ). A followup analysis excluding the midline scalp sites showed that the novelty effect was slightly larger over the right than the left (Novel vs. Categorical:  $F(1,17) = 4.10, p = .06$ ; Novel vs. Conflicting:  $F(1,17) = 4.28, p = .05$ ).

### P3 amplitude

The omnibus ANOVA showed a main effect of condition, ( $F(2,34) = 6.93, p < .01, \epsilon = 1.00$ ). Table 4 and Figure 4 show that the P3 to Novel stimuli was larger than both the Categorical ( $F(1,17) = 19.3, p < .0005$ ) and Conflicting conditions ( $F(1,17) = 4.47, p < .05$ ), which did not differ from each other ( $Fs < 1$ ). Interactions between novelty and left/midline/right scalp site were also significant. Followup analyses excluding the midline sites confirmed that the impact of novelty on the P3 was larger over the left than the right (Novel vs. Categorical: ( $F(1,17) = 7.52, p < .05$ ; Novel vs. Conflicting: ( $F(1,17) = 4.01, p = .06$ ).

The P3 in the Conflicting condition was distributed more frontally than the Categorical and Novel conditions and this resulted in significant conflict by centroparietal/parietal interactions ( $Fs(1,17) > 4.80, ps < .05$ ). The frontal distribution of the Conflicting P3 is analyzed in more detail below.

### P3 latency

The omnibus ANOVA showed a main effect of condition, ( $F(2,34) = 5.61, p < .01, \epsilon = 0.80$ ). P3 latency in the Conflicting condition was delayed relative to the Categorical condition ( $F(1,17) = 10.2, p < .01$ ). Although Figure 4 suggests that the peak of the P3 in the Conflicting condition was also later than the Novel condition, this difference was not significant ( $F(1,17) = 1.85$ ). As compared to the Categorical condition, Novel stimuli also elicited a later P3 ( $F(1,17) = 7.31, p < .05$ ), despite the fact that Novel stimuli elicited faster RTs. In comparisons between the Categorical condition and the other two conditions, the three-way interaction between condition, centroparietal/parietal, and left/midline/right electrode site was also significant. Followup analyses excluding the midline electrodes showed that the interaction was driven by somewhat later latencies in the Novel and Conflicting conditions at left centroparietal electrode CP3 ( $F_s(1,17) > 4.90, p_s < .05$ ).

Comparison of the Novel and Categorical conditions yielded slower P3 latencies combined with faster RT, a dissociation between the ordering of RTs and P3 latencies across conditions that is rarely observed (Kutas et al., 1977). We suggest that responses to Categorical stimuli were slower than responses to Novels because Categorical stimuli were evaluated as **candidates** to be Conflicting stimuli and therefore treated with more caution. This interpretation suggests that participants processed both relevant features of the Categorical stimuli before selecting a response, but were able to initiate a response to the Novels after processing only global body shape. In contrast, P3 latency may reflect more complete evaluation of the Novels that lagged behind response selection.

### Positivities over prefrontal and frontal scalp

Figure 4 shows that the Conflicting condition not only elicited a later P3 at the parietal scalp sites where this component is most prominent, but also late positive waves at prefrontal and frontal scalp sites where a pronounced positive peak is small to absent in the other conditions. This visual impression was evaluated by measuring peak amplitudes at the midline and dorsal lateral sites spanning the anterior-posterior extent of the electrode array from prefrontal to occipital sites, normalizing these to eliminate amplitude differences between conditions (McCarthy & Wood, 1985), and entering the values into ANOVAs with factors of condition (two levels), anterior-to-posterior location (AP, seven levels), and laterality (left/midline/right). Significant interactions between condition and the AP factor confirmed that the Conflicting condition elicited more frontally-distributed positive potentials than both the Categorical ( $F(6,102) = 7.06, p < .005, \epsilon = .37$ ) or Novel ( $F(6,102) = 4.92, p < .005, \epsilon = .45$ ) conditions.

Our previous work has suggested that prefrontal ERPs are particularly sensitive to categorization strategy (Folstein & Van Petten, 2004), and inspection of Figure 4 suggests that the larger prefrontal positivity in the Conflicting condition is not simply the anterior residue of the P3. The Conflicting condition becomes more positive than the other conditions at prefrontal sites earlier than at frontal sites, and the difference does not come to a peak and resolve to baseline as it does posterior to the prefrontal sites. We thus conducted additional analyses of the prefrontal sites (Fp1, Fpz, Fp2) alone: mean amplitudes in latency windows of 400 to 650, and 650 to 900 ms. For the earlier window, ERPs to the Conflicting stimuli were more positive than both Categorical and Novel conditions ( $F_s(1,17) > 7.30, p_s < .05$ ). For the later epoch, the Conflicting condition was significantly more positive than the Categorical condition ( $F(1,17) = 6.82, p < .05$ ), but only marginally more positive than the Novel condition ( $F(1,17) = 3.69, p = .07$ ). The Categorical and Novel conditions did not differ from each other during either epoch.

## GENERAL DISCUSSION

### **Conflicting information during unidimensional versus multidimensional rule use**

Experiment 2 used a two-dimensional categorization rule to examine the effects of conflicting information and perceptual novelty when both conflicting and novel features were relevant to the categorization rule. In contrast to Exp. 1, conflicting information led to a dramatic delay in categorization decisions. The conflict effect on reaction times cannot be attributed to the selection of a low-probability response, as Novel stimuli that shared the same assigned keypress elicited much faster responses. The prolonged RTs in the Conflicting condition also cannot be attributed to the mere presentation of unfamiliar stimuli, as the features of the Novel stimuli were rare in both the training and test phases. Instead, the substantial conflict effect in Exp. 2 as compared to Exp. 1 must be attributed to the use of a multidimensional rather than unidimensional rule.

In the current designs, the dimensionality of the assigned rule, and the presence of conflicting information in rule-relevant versus rule-irrelevant dimensions were deliberately confounded. When participants followed a one-dimensional rule in Exp. 1, the global body shape associated with the opposite category was necessarily rule-irrelevant. In Exp. 2, features associated with two different categories occurred in dimensions that had to be analyzed in order to follow the two-dimensional rule. There are thus two possible accounts for the significant conflict effect in Exp. 2 as compared to Exp. 1: 1) that only relationships between rule-relevant features are learned, or 2) that following a multidimensional rule necessarily leads to a broader focus of attention such that **all** feature correlations are spontaneously encoded and can lead to subsequent interference when the correlation structure is altered. Some prior results favor the latter account. As described in the Introduction, Allen and Brooks (1991) observed poorer performance when exemplars resembled the wrong category due to the presence of correlated, but rule-irrelevant features. However, the conflict effects observed in that study hinged -- by design -- on memorization of the individual exemplars and the combination of the relevant and irrelevant features to form a distinctive gestalt (Regehr & Brooks, 1993). In the present experiments, many more individual exemplars were presented, and the indistinguishable results for items occurring in the training phase and those that were similar but new in the test phase (Typical vs. Similar conditions in Exp. 2) suggest that participants remembered the general patterns of what they had seen, but not individual exemplars (see Folstein & Van Petten, 2004 for a similar comparison, also with null effects). It would be possible, and of some interest, to combine design elements of Experiments 1 and 2 so that rule-irrelevant rather than rule-relevant features were correlated with category membership in the context of a multidimensional rule. Delayed RTs for conflicting stimuli in this design would suggest that all feature correlations are encoded when following a multidimensional rule, even when individual exemplars are too numerous to memorize. While this manipulation might produce a large conflict effect on reaction time, we can be confident based on the results of Experiment 2 that this effect would not extend to the N2: although large conflict effects were observed in RT and P3 latency, there was no impact on the N2.

### **The anterior N2, perceptual novelty, and conflict**

As described in the Introduction, substantial prior literatures link a frontocentral N2 with both cognitive control and perceptual novelty, but these have rarely been manipulated within a single study (see Folstein & Van Petten, in press for more extensive review). In the current

experiments, novel features embedded within multi-feature stimuli did not influence the N2 when participants used other features to classify the stimuli (Exp. 1), but resulted in large N2s when the novel features were relevant to the assigned categorization rule (Exp. 2). At first blush, the lack of N2 modulation to novel stimuli in Exp. 1 might appear inconsistent with experiments showing that perceptually novel stimuli elicit large N2s when they serve as nontargets requiring no overt response (Courchesne et al., 1975; Czigler & Balázs, 2005; Daffner et al., 2000; Suwazono et al., 2000; Thomas & Nelson, 1996). However, this inconsistency is only apparent, as the novel stimuli in those prior studies needed to be attended and compared to the mental template of a target in order to determine that they warranted no response (a type of categorization decision). Via the use of multidimensional stimuli, the current series of experiments show that it is necessary for novel features to occur in a task-relevant dimension in order to obtain substantial enhancements of the frontocentral N2. This conclusion is consistent with that of Wang and colleagues, who have observed larger N2s when an eliciting stimulus differs from a prior stimulus than when it is the same, but also that this “mismatch effect” is larger when it occurs in a dimension that governs a same/different judgment than when it occurs in a task-irrelevant dimension (Wang et al., 2003; Wang et al., 2004).

Despite the substantial impact of conflicting information on RTs and P3 latencies in Exp. 2, no increase in N2 amplitude was observed. This result places limits on the generality of the N2’s sensitivity to conflicting information during decision making. Below, we describe some prior studies that have observed modulation of N2 amplitude when participants discriminated perceptually similar stimuli calling for different responses, as well as some studies that have not. We then suggest that the former paradigms are likely those that encouraged early response preparation based on preliminary stimulus processing, whereas paradigms that encourage a more cautious strategy of complete stimulus analysis prior to response preparation are less likely to elicit N2 effects.

Two recent studies show that cross-category similarity can enhance the N2. Nieuwenhuis and colleagues manipulated auditory perceptual overlap between frequent “go” and rare “no-go” stimuli. Typically, no-go trials elicit larger N2s than go trials (Pfefferbaum et al., 1985). When participants discriminated similar-sounding letters “F” and “S”, this No-go N2 effect was larger than when they discriminated the dissimilar letters “F” and “T” (Nieuwenhuis, Yeung, & Cohen, 2004). A second study employed a more complex design in which participants performed a choice RT task on complex geometric shapes (Azizian, Freitas, Parvaz, & Squires, 2006). One of the shapes was in a rare “target” category (probability 10%) while the frequent “non-target” category (90%) included four stimulus types: rare (10%) “target-similar” stimuli that were very similar to the target, two rare (10%) “target-dissimilar” stimulus types with very different shapes from the target, and a frequent (60%) stimulus that was also visually distinct from the target. The target-similar stimuli elicited slower reaction times and a larger N2 than the rare target-dissimilar stimuli despite identical stimulus and response probabilities. These results indicate that stimuli that are near-matches to other items demanding a different response can drive N2 amplitude.

Two other studies that shared many design elements with the ones above failed to observe a conflict N2 effect. In a different study, Azizian et al. had participants perform a counting task on schematic cartoon faces with eight possible features: two eyes, two eyebrows, two ears, a nose, and a mouth (Azizian, Freitas, Watson, & Squires, 2006). In the first experiment, the counted target had all eight features and the non-targets had anywhere from one to all eight of those features deleted. The most similar non-targets thus shared all but one feature

with the eight-feature target, while the least similar non-target was an empty circle. The second experiment was the mirror image: the target was an empty circle while the non-targets had one to eight features added. In both experiments, as non-targets became increasingly similar to the target, they elicited larger and larger P3s but not larger N2s. A salient aspect of this design is the very gradual similarity gradient between the target and non-target categories (seven gradations of similarity between target and the most dissimilar non-target). We suspect that something about this gradient encouraged a strategy in which participants were unlikely to prepare a response before completion of stimulus evaluation. Previous studies have shown that encouraging a cautious strategy – for instance, instructions to favor accuracy over speed -- also reduce the amplitude of the no-go N2 (Jodo & Kayama, 1992). In Azizian et al.'s (2006) experiment, the category structure itself might have pushed participants towards a cautious strategy.

A similar explanation is likely for a second failure to observe an N2 enhancement in the face of conflicting information about a category assignment. Fox and colleagues presented equiprobable colored letters requiring an overt response (go) or not (go-go; Fox, Michie, Wynne, & Maybery, 2000). The letters in the Go category were “N”, “J”, or “W” in blue, and “O” in red, green, or yellow; No-go stimuli were pink “E”s or blue “O”s. The pink “E” was called an *elemental No-go* stimulus because it was comprised of two features never associated with a go response. This elemental no-go condition elicited a large and significant frontocentral N2 relative to go trials, as predicted. The blue “O” was called a *configural no-go* stimulus because it was comprised of features from the Go category. *A priori*, the configural No-go condition would have seemed perfectly situated to elicit a large N2, given the addition of conflicting information to a no-go response requirement, but yielded no significant difference from the go condition. A possible reason for the lack of a conflict effect in this experiment is that participants used a cautious response strategy due to the complex response rule necessary to perform this task correctly. The rule for the go category has the logical form of *exclusive or*: “go if (blue or O) and not (blue and O)”. Use of this rule might have encouraged serial processing of the individual features, and/or double-checking before responding, thus discouraging fast response preparation.

We suspect that enhancements of the N2 due to conflicting information occur when participants begin to prepare a motor response before evaluation of a stimulus is complete, and not when they adopt a more deliberate strategy of postponing response preparation until evaluation is complete. This hypothesis suggests that the “conflict” in the “conflict N2” is response conflict, rather than a more general perception that a stimulus might be compatible with more than one decision option. Under this hypothesis, the participants in Exp. 2 adopted a strategy of evaluating both rule-relevant features and their fit with the instructed rule instead of beginning to prepare a response that might prove incorrect.

Although lacking an enhanced N2, the Conflicting condition of Exp. 2 did elicit different ERPs than the Typical or Novel conditions. In addition to a delayed, more frontally distributed P3, the Conflicting stimuli yielded a prolonged prefrontal positivity that appeared distinct from the frontal aspect of the P3. We previously observed a similar wave that was larger in participants who used multiple features as opposed to single features in their categorization strategy (Folstein & Van Petten, 2004). All of the stimuli in that study contained some conflicting information – distinct features suggestive of different category assignments. We therefore suspect that conflicting information in complex categorization tasks recruits additional

executive processing resources and that these resources are reflected in prefrontal positivities that occur later in time than the latency range of the N2.

In summary, conflicting and novel irrelevant information both slowed stimulus evaluation time but did not engage cognitive control processes in a manner similar to the Eriksen flanker task, or tasks that seem to demand explicit response inhibition such as No-go conditions. The fact that novelty and not conflict drove the N2 in Experiment 2 suggests that the N2-like potential observed in our previous study of categorization (Folstein & Van Petten, 2004) may have been driven by stimulus atypicality rather than response inhibition as originally hypothesized. Future studies should further investigate the effect of category structure on the processing of conflicting information, focusing on the effects of rule complexity and gradations in typicality.

## REFERENCES

- Allen, S. W., & Brooks, L. R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General*, *120*, 3-19.
- Azizian, A., Freitas, A. L., Parvaz, M. A., & Squires, N. K. (2006). Beware misleading cues: perceptual similarity modulates the N2/P3 complex. *Psychophysiology*, *43*, 253-260.
- Azizian, A., Freitas, A. L., Watson, T. D., & Squires, N. K. (2006). Electrophysiological correlates of categorization: P300 amplitude as index of target similarity. *Biological Psychology*, *71*, 278-288.
- Breton, F., Ritter, W., Simson, R., & Vaughan, H. G., Jr. (1988). The N2 component elicited by stimulus matches and multiple targets. *Biological Psychology*, *27*, 23-44.
- Courchesne, E., Hillyard, S. A., & Galambos, R. (1975). Stimulus novelty, task relevance and the visual evoked potential in man. *Electroencephalography & Clinical Neurophysiology*, *39*, 131-143.
- Czigler, I., & Balázs, L. (2005). Age-related effects of novel visual stimuli in a letter-matching task: an event-related potential study. *Biological Psychology*, *69*, 229-242.
- Daffner, K. R., Mesulam, M. M., Scinto, L. F., Calvo, V., Faust, R., & Holcomb, P. J. (2000). An electrophysiological index of stimulus unfamiliarity. *Psychophysiology*, *37*, 737-747.
- Folstein, J.R. (2007). *On the category's edge: Event related potential correlates of novelty and conflicting information in rule-based categorization*. Unpublished doctoral dissertation, University of Arizona, Tucson, Arizona.
- Folstein, J. R., & Van Petten, C. (2004). Multidimensional rule, unidimensional rule, and similarity strategies in categorization: event-related potential correlates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 1026-1044.
- Folstein, J. R., & Van Petten, C. (in press). Influence of cognitive control and mismatch on the N200: a review. *Psychophysiology*.
- Fox, A. M., Michie, P. T., Wynne, C. D., & Maybery, M. T. (2000). ERP correlates of response inhibition to elemental and configural stimuli in a negative patterning task. *Clinical Neurophysiology*, *111*, 1045-1053.
- Garner, W. R., & Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, *Vol. 1*, 225-241.
- Gehring, W. J., Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Probability effects on stimulus evaluation and response processes. *Journal of Experimental Psychology: Human Perception & Performance*, *18*, 198-216.
- Heil, M., Osman, A., Wiegmann, J., Rolke, B., & Hennighausen, E. (2000). N200 in the Eriksen-task: inhibitory executive processes? *Journal of Psychophysiology*, *14*, 218-225.
- Heinze, H.-J., Münte, T.-F., & Kutas, M. (1998). Context effects in a category verification task as assessed by event-related brain potential (ERP) measures. *Biological Psychology*, *47*, 121-135.
- Holroyd, C. B. (2004). A Note on the Oddball N200 and the Feedback ERN. In M. Ullsperger & M. Falkenstein (Eds.), *Errors, conflicts, and the brain: Current opinions on performance monitoring*. Leipzig: MPI of Cognitive Neuroscience.
- Jodo, E., & Kayama, Y. (1992). Relation of negative ERP component to response inhibition in a go/no-go task. *Electroencephalography & Clinical Neurophysiology*, *82*, 477-482.
- Kopp, B., Rist, F., & Mattler, U. (1996). N200 in the flanker task as a neurobehavioral tool for investigating executive control. *Psychophysiology*, *33*, 282-294.

- Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science*, *197*, 792-795.
- McCarthy, G., & Donchin E. (1981). A metric for thought: a comparison of P300 latency and reaction time. *Science*, *211*, 77-80.
- McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. *Electroencephalography & Clinical Neurophysiology*, *62*, 203-208.
- McCloskey, M., & Glucksberg, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology*, *11*, 1-37.
- Nieuwenhuis, S., Yeung, N., & Cohen, J. D. (2004). Stimulus modality, perceptual overlap, and the go/no-go N2. *Psychophysiology*, *41*, 157-160.
- Nieuwenhuis, S., Yeung, N., van den Wildenberg, W., & Ridderinkhof, K. R. (2003). Electrophysiological correlates of anterior cingulate function in a go/no-go task: effects of response conflict and trial type frequency. *Cognitive, Affective & Behavioral Neuroscience*, *3*, 17-26.
- Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39-57.
- Nosofsky, R. M. (1991). Typicality in logically defined categories: Exemplar-similarity versus rule instantiation. *Memory & Cognition*, *19*, 131-150.
- Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, *104*, 266-300.
- Patalano, A. L., Smith, E. E., Jonides, J., & Koeppel, R. A. (2001). PET evidence for multiple strategies of categorization. *Cognitive, Affective & Behavioral Neuroscience*, *1*, 360-370.
- Pfefferbaum, A., Ford, J. M., Weller, B. J., & Kopell, B. S. (1985). ERPs to response production and inhibition. *Electroencephalography & Clinical Neurophysiology*, *60*, 423-434.
- Pritchard, W. S., Shappell, S. A., & Brandt, M. E. (1991). Psychophysiology of N200/N400: A review and classification scheme. In J. R. Jennings & P. K. Ackles (Eds.), *Advances in psychophysiology: A research annual*. (Vol. 4, pp. 43-106): Jessica Kingsley Publishers, Ltd.
- Regehr, G., & Brooks, L. R. (1993). Perceptual manifestations of an analytic structure: The priority of holistic individuation. *Journal of Experimental Psychology: General*, *122*, 92-114.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573-605.
- Simson, R., Vaughn, H.G., Jr., & Ritter, W. (1977). The scalp topography of potentials in auditory and visual discrimination tasks. *Electroencephalography & Clinical Neurophysiology*, *42*, 528-535.
- Smith, E. E., Patalano, A. L., & Jonides, J. (1998). Alternative strategies of categorization. *Cognition*, *65*, 167-196.
- Squires, N.K., Donchin, E., Squires, K.C., & Grossberg, S. (1977). Bisensory stimulation: inferring decision-related processes from P300 component. *Journal of Experimental Psychology: Human Perception & Performance*, *3*, 299-315.
- Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, *150*, 1187-1188.

- Suwazono, S., Machado, L., & Knight, R. T. (2000). Predictive value of novel stimuli modifies visual event-related potentials and behavior. *Clinical Neurophysiology*, *111*, 29-39.
- Thomas, K. M., & Nelson, C. A. (1996). Age-related changes in the electrophysiological response to visual stimulus novelty: a topographical approach. *Electroencephalography & Clinical Neurophysiology*, *98*, 294-308.
- Verleger, R. (1997). On the utility of P3 latency as an index of mental chronometry. *Psychophysiology*, *34*, 131-156.
- Wang, Y., Cui, L., Wang, H., Tian, S., & Zhang, X. (2004). The sequential processing of visual feature conjunction mismatches in the human brain. *Psychophysiology*, *41*, 21-29.
- Wang, Y., Tian, S., Wang, H., Cui, L., Zhang, Y., & Zhang, X. (2003). Event-related potentials evoked by multi-feature conflict under different attentive conditions. *Experimental Brain Research*, *148*, 451-457.
- Williams, D. A., Sagness, K. E., & McPhee, J. E. (1994). Configural and elemental strategies in predictive learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 694-709.

### **ACKNOWLEDGEMENTS**

Financial support was provided by the National Institute of Mental Health (MH073703). We are grateful to Dianne K. Patterson, Jonathan Forster, and Paul Krewski for technical support. The experiments formed part of a dissertation submitted in partial requirement for a doctorate in Psychology from the University of Arizona. We are grateful to Ken Forster, Rebecca Gomez, and Lee Ryan for serving on the committee, and two anonymous reviewers for their comments on a previous version.

**Table 1. Average frequency of rule-relevant and rule-irrelevant features in Experiment 1**

	TEST PHASE			ENTIRE EXPERIMENT		
	Relevant body marking	Irrelevant global shape	Irrelevant individual body part	Relevant body marking	Irrelevant global shape	Irrelevant individual body part
Typical	50.0	37.5	23.4	50.0	42.8	26.8
Conflicting	50.0	37.5	23.4	50.0	42.8	26.8
Novel	50.0	25.0	3.1	50.0	14.2	1.8

---

**Note.** The frequency of a feature in the test phase is the percentage of test trials in which that feature occurs; feature frequency is also shown calculated across all trials (feature verification, categorization training, and test phases). In Exp. 1, the only rule-relevant features were spots and stripes (type of body marking). Spots occurred in half of the trials, and stripes in the other half of the trials, so that average frequency of these two relevant features was 50%. In the test phase, one of the global body shapes occurred on three-eighths of the trials (spread across the Typical and Conflicting conditions), another global body shape on three-eighths of the trials (spread across the Typical and Conflicting conditions), and the third body shape on two-eighths of the trials (all of which were in the Novel condition only). Finally, there were multiple individual features comprising each global body shape, and more individual variants were used in the Novel condition, leading to lower frequencies for the specific irrelevant features overall, but particularly in the Novel condition.

**Table 2. Dependent measures in Experiment 1 (mean and standard error)**

<b>Condition</b>	<b>Accuracy (%)</b>	<b>Reaction time (ms)</b>	<b>N2 amplitude (<i>uV</i>)</b>	<b>P3 amplitude (<i>uV</i>)</b>	<b>P3 latency (ms)</b>
Typical	99.2 (0.2)	507 (17)	1.78 (0.31)	13.7 (0.4)	478 (9)
Conflicting	96.9 (0.5)	542 (19)	1.69 (0.29)	13.8 (0.4)	529 (10)
Novel	97.6 (0.6)	545 (19)	1.94 (0.28)	14.4 (0.5)	519 (8)

**Table 3. Average frequency of rule-relevant and rule-irrelevant features in Experiment 2**

	TEST PHASE		ENTIRE EXPERIMENT	
	Relevant (global shape and markings)	Irrelevant (individual body part)	Relevant (global shape and markings)	Irrelevant (individual body part)
Typical	41.7	5.4	41.7	5.5
Similar	41.7	5.0	41.7	4.9
Conflicting	41.7	5.4	41.7	5.5
Novel	16.7	2.1	16.7	2.1

---

**Note.** The frequency of a feature in the test phase is the percentage of test trials in which that feature occurs. In Exp. 2, rule-relevant features were humanoid, horse, and fish global body shapes, and spots, stripes, and brick-pattern body markings. For a given subject, humanoid shapes would occur in 5 out of 12 trials and horse shapes on 5 out of 12 trials on average, yielding an average frequency of 5/12 (41.7%) for these rule-relevant shapes – which were distributed across the Typical, Similar, and Conflicting conditions. Fish shapes would occur in the remaining 2 out of 12 trials (16.7%), and these trials would occur only in the Novel condition. The calculations for body markings are identical, such that a Typical stimulus would be comprised of some body shape with a 41.7% frequency, and some marking with a 41.7% frequency, for an average frequency of 41.7%. Global shapes were comprised of different variants of each body part across stimuli, such that average feature frequencies for any specific body part (e.g., a type of head) are lower. Average feature frequency is also shown calculated across all trials (training plus test phases).

**Table 4. Dependent measures in Experiment 2 (mean and standard error)**

<b>Condition</b>	<b>Accuracy (%)</b>	<b>Reaction time (ms)</b>	<b>N2 amplitude (<math>\mu</math>V)</b>	<b>P3 amplitude (<math>\mu</math>V)</b>	<b>P3 latency (ms)</b>
Typical	98.7 (0.4)	634 (25)	1.57 (0.30)	14.9 (0.4)	459 (9)
Similar	99.2 (0.4)	639 (24)	1.76 (0.29)	14.7 (0.4)	477 (8)
Conflicting	92.6 (1.2)	876 (41)	1.62 (0.31)	15.1 (0.5)	550 (13)
Novel	99.9 (0.1)	610 (20)	0.51 (0.31)	17.3 (0.5)	512 (5)

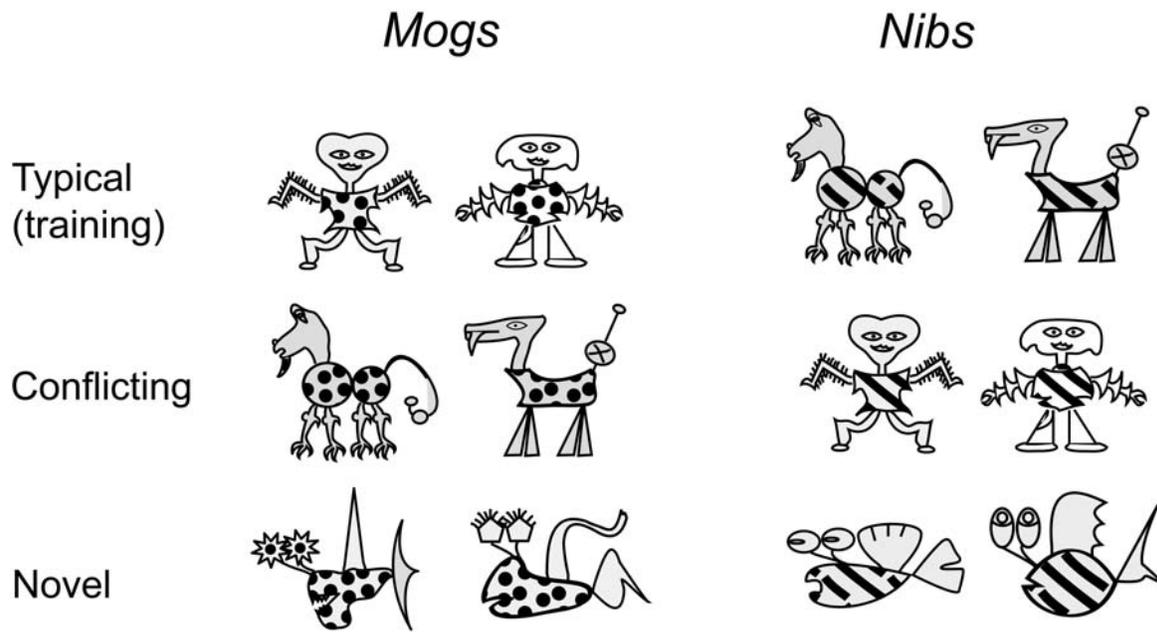
## FIGURE CAPTIONS

**Figure 1.** Examples of the stimuli in Experiment 1. The body markings – spots or stripes – determined whether they were “Mogs” or “Nibs” according to the instructed rule. Trained exemplars were the same as those in Typical condition during test, and shared a global body shape (humanoid, horse-like, fish-like) with members of their own category. Stimuli in the Conflicting condition had a global body shape that was correlated with the opposite category in the prior training phase. Novel stimuli had a global body shape not present in the training phase, and of lower frequency than the other two in the test phase.

**Figure 2.** Grand average event-related potentials from the twelve prefrontal, frontal, frontocentral, and central sites used to measure N2 amplitude, and the six centroparietal and parietal sites used to measure P3 amplitude and latency. Left scalp sites are shown in the left column, midline sites in the middle column, and right scalp sites in the right column.

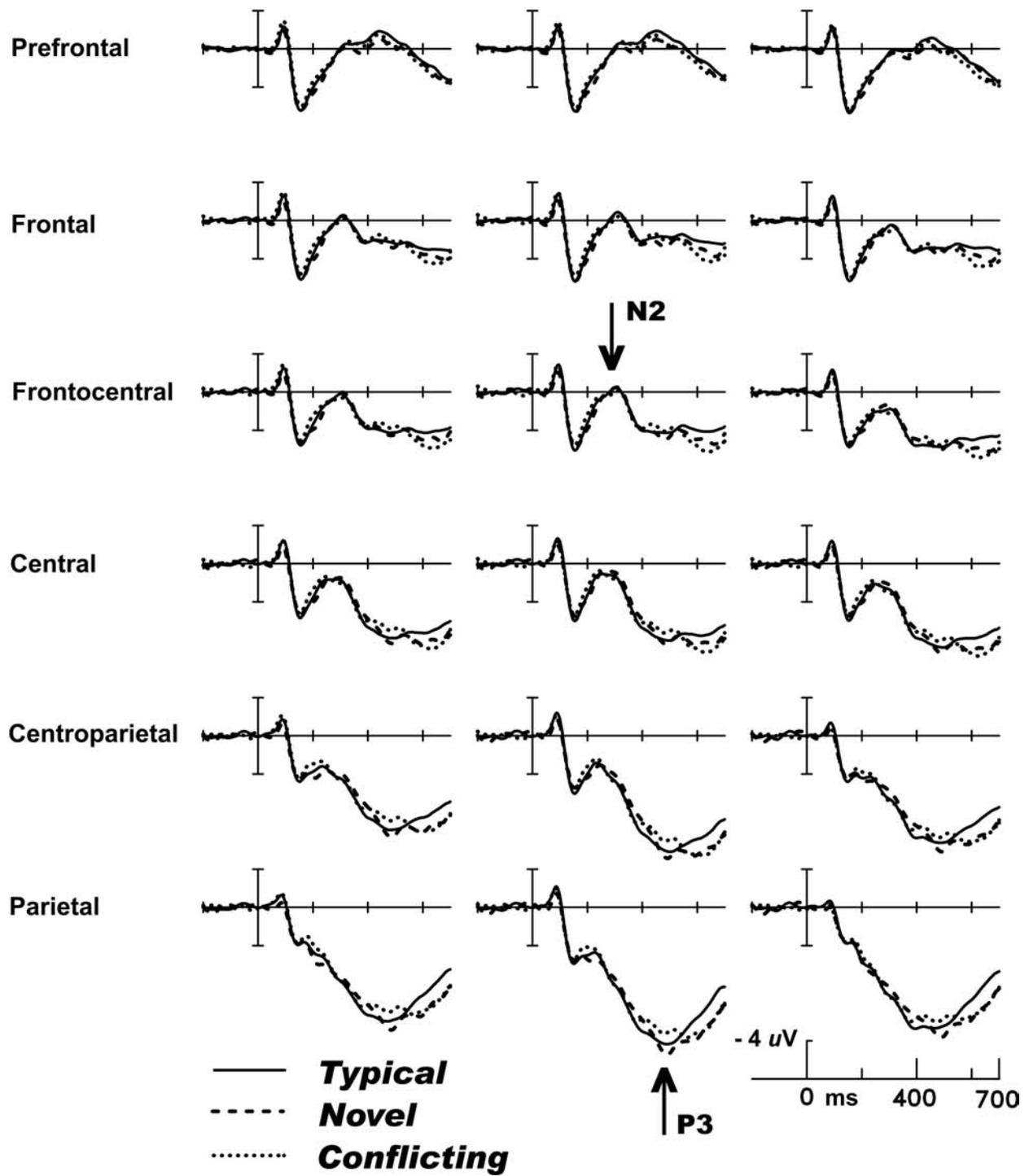
**Figure 3.** Examples of the stimuli presented during the training and test phases of Experiment 2, together with the correct responses according to a rule for which both body markings (spots, stripes or brick pattern) and global body shape (humanoid, horse-like, fish-like) determined category membership.

**Figure 4.** Grand average event-related potentials from all scalp sites in Experiment 2. Electrode sites are arranged from anterior (top row) to posterior (bottom row) and from left to right, roughly corresponding to head location.

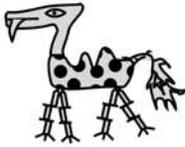
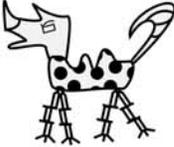
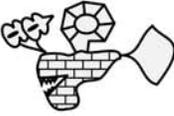
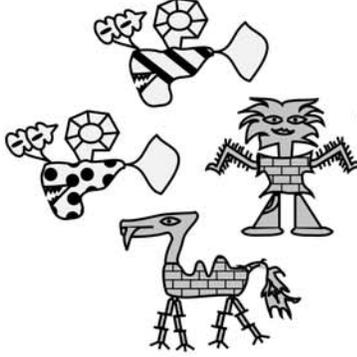


**Figure 1**

**Experiment 1**



**Figure 2**

Condition	Stimulus	Correct response	Viewed in
<b>Typical</b> (2 Mog features)		MOG	Training & Test
<b>Typical</b> (2 Nib features)		NIB	Training & Test
<b>Similar</b> (2 Mog features)		MOG	Test only
<b>Similar</b> (2 Nib features)		NIB	Test only
<b>Conflicting</b> (1 Mog, 1 Nib feature)		OTHER	Test only
<b>Conflicting</b> (1 Nib, 1 Mog feature)		OTHER	Test only
<b>Novel</b> (2 Other features)		OTHER	Test only
<b>Training Others</b> (1 Mog or Nib feature, 1 Other feature)		OTHER	Training only

**Figure 3**



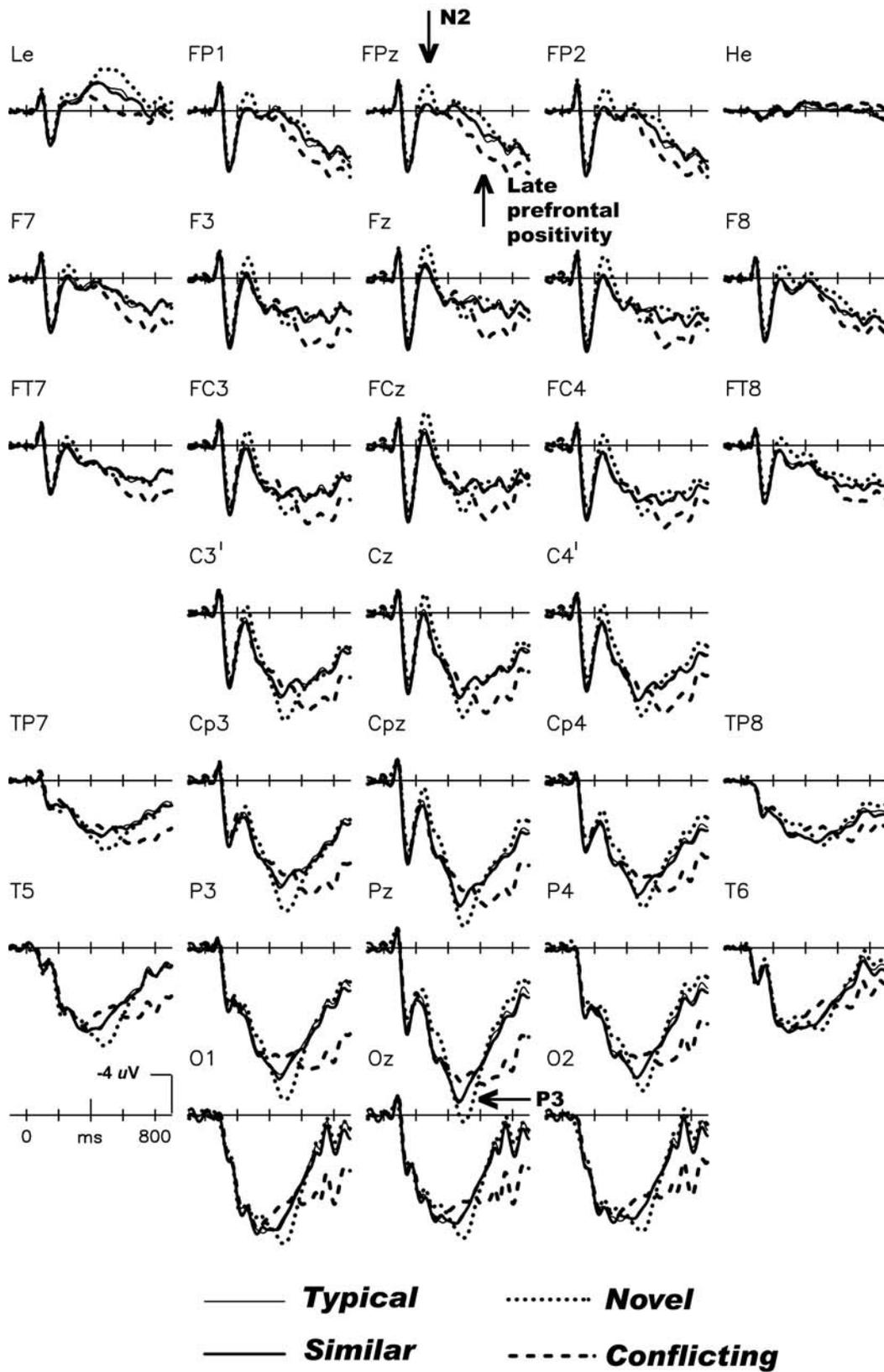


Figure 4