

Running Head: CATEGORIZATION AND ATTENTION

Emergence of the benefits and costs of grouping for visual search

Rachel Wu^{1*}, Brianna McGee¹, Madelyn Rubenstein², Zoe Pruitt³, Olivia S. Cheung⁴, and
Richard N. Aslin³

¹Department of Psychology, UC Riverside

²Brain and Cognitive Sciences, University of Rochester

³Haskins Laboratories

⁴Science Division, Psychology, NYU Abu Dhabi

*Address correspondence to:

Rachel Wu
Department of Psychology
University of California, Riverside
900 University Avenue
Riverside, CA 92521
Rachel.Wu@ucr.edu

Abstract

The present study investigated how grouping related items leads to the emergence of benefits (facilitation when related items are search targets) and costs (interference when related items are distractors) in visual search. Participants integrated different views ("related items") of a novel Lego[®] object via 1) assembling the object, 2) disassembling the object, or 3) sitting quietly without explicit instructions. An omnibus ANOVA revealed that neural responses (N2pc event-related potential) for attentional selection increased between pretest to posttest regardless of the training condition when a specific target view appeared (benefit) and when a nontarget view from the same object as the target view appeared (cost). Bonferroni-corrected planned comparisons revealed that assembling the object (but not disassembling the object or no training) had a significant impact from pretest to posttest, although the ANOVA did not reveal any interaction effects, suggesting that the effects might not differ across training conditions. This study is one of the first to demonstrate the emergence of the costs and benefits of grouping novel targets on visual search efficiency.

Keywords: Categorization; N2pc; Visual search

1. Introduction

Prior knowledge about targets and distractors influences visual search in meaningful ways (see Wu & Zhao, 2017). In particular, there are costs and benefits to prior knowledge on visual search. In terms of the benefits, recent visual search studies using ERP measures have shown that prior knowledge of grouping features into objects and objects into categories increases search efficiency, as reflected in the N2pc event-related potential (ERP) component (e.g., Nako et al., 2014a; 2014b; Wu et al., 2015; Wu et al., 2016). The N2pc ERP is the established marker of target selection in visual search studies, emerging 200 ms after stimulus onset in retinotopic occipito-temporal areas contralateral to the target hemifield (Eimer, 1996; Luck & Hillyard, 1994). Whereas the N2pc amplitude decreases linearly with an increasing number of unrelated search targets, the N2pc amplitude when searching for many items within a category (i.e., a group of related objects) remains stable for well-defined, familiar categories (e.g., letters and numbers; Nako et al., 2014a; clothing, Nako et al., 2014b; human faces, Wu et al., 2015). Thus, prior knowledge of a category increases search efficiency for multiple objects within that category. The cognitive benefits from grouping features and objects also have been documented in studies on categorization, chunking, associative learning, and statistical learning, which show that grouping leads to increased efficiency in attention, learning, and working memory from infancy to adulthood (e.g., Blair et al., 2009; Brady et al., 2009; Chun & Jiang, 1998; Gobet et al., 2011; Nako et al., 2014; Wolfe & Horowitz, 2004; Woodman et al., 2003; Wu et al., 2011; Wu et al., 2016; Xu & Chun, 2007; see Kruschke, 2001).

Besides the benefits of prior knowledge of grouping on visual search, there are also costs to grouping. In particular, once grouped into familiar categories, objects become difficult to "ungroup", and nontarget category members can capture attention in an obligatory manner (e.g.,

Visual search (4)

Nako et al., 2014a, Wu et al., 2017a; see also Castro & Wasserman, 2016). Nako et al. (2014a) showed that nontargets that are members of a familiar target category elicit obligatory attentional capture. For example, when instructed to search for the letter R, the letter C captured attention among a set of numbers, because C and R are both letters. This “foil effect” is similar to other findings on capture by semantically related objects (De Groot et al., 2016; Telling et al., 2010), objects of expertise in particular categories (e.g., cars, McGugin et al., 2011), and objects with similar known features (e.g., the color red for a stop sign target, even when the stop sign is in grayscale during the experiment, Olivers et al., 2011). Interestingly, visual search studies have shown that novel objects and categories do not tend to elicit foil effects during a one-hour experimental session (e.g., novel Chinese characters, Wu et al., 2013; novel alien families, Wu et al., 2016).

Although there are clear search benefits and costs to grouping features into objects and objects into categories, it is still unclear how the benefits and costs in search efficiency as measured in neural responses may emerge over a single experimental session with a short training procedure. It is important to investigate this issue because understanding when and how the costs and benefits of prior knowledge emerge in visual search eventually may inform research on how observers could maximize the benefits while minimizing the costs. The N2pc, in particular, is a useful marker indicating the costs and benefits of grouping on visual search. Two N2pc studies have used novel categories (e.g., Chinese characters, Wu et al., 2013; alien families, Wu et al., 2016) in which the training procedure was the search task itself (Wu et al., 2013) or the category was obvious enough to verbalize as a rule (same versus different shapes; Wu et al., 2016). In both cases, no training procedure separate from the visual search task itself was used. In both studies, the benefits of grouping were sustained throughout the entire

experimental session, and there were no costs observed (i.e., foil effects). A recent study (Wu et al., 2017a), however, suggests that level of real world experience with grouping particular categories (healthy vs. unhealthy food based on dieting experience) may be related to the magnitude of the foil effect (i.e., the N2pc amplitude on foil trials). Therefore, the present study investigated whether the costs and benefits of grouping would emerge in the search task after a brief training procedure in a controlled environment.

1.1 Justification for object view integration training

To investigate how the benefits and costs of grouping emerge over a single experimental session with a brief training procedure, we trained participants to integrate different viewpoints of one novel object. We chose this training procedure for both theoretical and methodological reasons. Integrating different viewpoints of objects allows learners to form stable mental representations that are critical for functional interactions and inferences. The ability to integrate across different viewpoints is a skill that develops in the first half year of life (e.g., Soska et al., 2008; see also Fagan, 1976) and presumably becomes more robust and automatic with increasing expertise. Soska et al. (2008) suggests that by 6 months, infants can predict untrained views of trained three-dimensional objects, a skill that may be related to motor exploration (Soska et al., 2010). With adults, N2pc ERP recordings reveal that familiar objects in different viewpoints can be selected as targets by 200ms, which is the same time window as selecting a target in the same viewpoint (Nako et al., 2014b). Results from Nako et al. provide converging evidence for studies showing that object representations may include multiple views of the same object in early stages of processing (e.g., Harris and Dux, 2005). In sum, integrating different views of an object is an acquired ability that involves grouping and is critical for successful object interactions.

Visual search (6)

Methodologically, we used novel objects to investigate the *emergence* of the benefits and costs on search efficiency, which would not have been possible with familiar objects that already elicit benefits or costs of grouping. We used two objects with a finite number of views to allow the categories to be learnable within a 5-minute training session. To prompt top-down effects from training while avoiding pop-out effects from differences in stimulus characteristics between categories, we used perceptually similar targets and distractors.

1.2 Goal of the present study

The present study used a visual search paradigm to investigate the benefits and costs of grouping multiple views of a novel object as they emerged over one experimental session via one 5-minute training period. We trained one group of participants to integrate different views of a multi-part Lego object by assembling it using seven Lego pieces (assemble condition). A second group of participants was trained in the disassemble condition, where they took apart the already constructed Lego object. We reasoned that assembling the object from its parts would encourage participants to group different views of the same object into one "category" because the participants had to understand how the parts fit together into a whole object during the assembly process. By contrast, the disassembling process did not necessarily require the creation of such a "category" of different views, but merely the removal of each part from the fully assembled object. A control condition (no training condition) allowed us to investigate whether participants could integrate the different object views in the absence of explicit training (i.e., either assembling or disassembling Lego pieces). Participants in this condition sat quietly for 5 minutes between completing visual search trials during pretest and posttest. In other words, the passive control condition assessed whether merely completing the visual search trials would induce

improvements from pretest to posttest. Before and after training (i.e., pretest and posttest), participants searched for both a specific view of the target object or any view of that object (i.e., treating the different views of the object as a category). Both pretest and posttest included four trial types: exemplar match, category match, foil (nontarget view of the same object as the exemplar view), and no target.

If training matters, we would expect an increase from pretest to posttest in the N2pc amplitude in the category match and foil trials for the assemble condition, but not the other two training conditions, based on the idea that assembling an object may lead to increased grouping of object views. If there is a general improvement from pretest to posttest across all training conditions, then we would expect to see only a main effect from pretest to posttest. In addition, an increase in the N2pc amplitude in category match trials from pretest to posttest would indicate an increase in task *relevant* categorization, whereas an increase in the N2pc amplitude in foil trials would indicate an increase in task *irrelevant* categorization (i.e., obligatory attentional selection). An increase in the exemplar match trials also could be possible in the assemble condition because a more robust representation of an object also may enhance representations of specific views of that object. If the other two training conditions reveal enhancements from pretest to posttest, in particular the no training condition, perhaps simply completing the visual search task twice (i.e., for pretest and posttest) could contribute to the training effect.

2. Method

2.1 Participants

The final sample consisted of 51 adults ($M = 21.43$ years, $SD = 3.02$, range: 17-32 years, $N = 4$ with missing age data, 31 females, 20 males): 18 adults participated in the assemble

Visual search (8)

condition (putting a Lego[®] object together), 16 participated in the disassemble condition (taking a Lego[®] object apart), and 17 participated in the no training condition (sitting quietly for 5 minutes). All participants had normal or corrected-to-normal vision. An additional 32 participants were excluded from the study across all three conditions due to excessive eye movements ($N = 19$, more than 50% of trials excluded), low accuracy ($N = 4$, less than 60% overall accuracy), experimenter error ($N = 3$, running the wrong experiment and failing to record), participant requesting to terminate ($N = 1$), and equipment failure ($N = 5$, hardware and software failure). This moderate attrition rate was primarily due to a difficult experimental task and training new undergraduate research assistants to conduct the study. This attrition rate is consistent with our prior studies in which difficult tasks were employed and undergraduates were trained to test participants (e.g., Wu et al., 2015; 2016; 2017). Participants in the assemble and disassemble conditions were paid \$25 after finishing the experiment at University of Rochester, whereas participants in the no training condition were provided course credit at University of California, Riverside. Informed consent was obtained for all participants prior to the start of the experiment.

2.2 Stimuli, Design, Procedure

Two physical objects were constructed using the same set of seven red Lego[®] bricks. These two objects were then rendered as 2-D images to create 8 different views of each (every 45°) using Lego[®] Digital Designer (<http://ldd.lego.com/en-us/>) (Figure 1, top panels). Thus, there were sixteen images that formed the entire stimulus set in the visual search task. Each search display presented on an LCD computer screen consisted of two objects, one of which was sometimes the target. Each object subtended 3.82° to the left and right of a central fixation point

that the participant was told to focus on for the entire experiment (Figure 2). The images were presented on a gray background (RGB 126,126,126) using Psychtoolbox and E-Prime. Each trial was presented for 200 ms, and the interstimulus interval was 1600 ms, where only the fixation dot was presented. Participants were expected to respond to the stimuli that were just presented within the 1800 ms time window. However, it is possible that occasionally they responded to stimuli presented on prior trials. Incorrect trials were excluded from ERP and reaction time analyses. We chose not to jitter the stimuli because pilot studies in our lab using a similar paradigm suggested that participants were distracted by the jittered interstimulus interval and had lower accuracy and longer reaction times. Because the present task was already very difficult, we did not want to add any distracting elements.

INSERT FIGURES 1 & 2 ABOUT HERE

This mixed design experiment had pretest and posttest phases (within-subjects factor) with a short training session in between (between-subjects factor). The pretest and posttest phases contained the same two tasks: exemplar search (i.e., search for a specific view of a particular object) and category search (i.e., search for any view of a particular object). There was one specific target for the exemplar search task (with 7 foils and 8 possible distractors), and 8 possible targets for the category search task (with 8 possible distractors). Foils never appeared with the target on the same trial during the exemplar search task. The target and search task type were specified at the beginning of every task, and participants were reminded of the target at the

Visual search (10)

beginning of every block. For the exemplar search task, the specific target (i.e., one particular view) was pseudo-randomly assigned and maintained throughout the entire experiment for each participant. Approximately half of the participants in each condition were tested and trained on object A, while the other half on object B. The order of the tasks was counterbalanced between pretest and posttest phases. In other words, half of the participants received the exemplar task first and the category task second in the pretest phase and the opposite order during the posttest, while the other half received the category task first in the pretest phase and the opposite order during the posttest. We presented 24 blocks total, with 6 continuous blocks per task (exemplar and category) for each phase (pretest and posttest).

Across the exemplar and category search tasks, we presented four trial types: exemplar match, category match, foil, and no target trials (Figure 1, bottom panel). Each of the 6 continuous blocks of the exemplar search task contained 28 exemplar match trials (specific target view appeared in the search array), 28 foil trials (nontarget view of the target object), and 6 no target trials (views from the nontarget object). Each of the 6 blocks from the category search task contained 28 category match trials (any view of the target object) and 28 no target trials (only views from the nontarget object). We included foil trials in the exemplar search task to measure task-irrelevant view integration, as well as to ensure that participants performed an identity search for an exact item, rather than a category search. The category search task required that all foils become targets, and therefore foil trials were not coded separately in that task. Target prevalence was similar between exemplar and category search trials: A target was present for 45% of the exemplar search trials and for 50% of the category search trials. The 5% difference was due to the inclusion of no target trials in the exemplar search task, while including the same number of exemplar match and foil trials. These no target trials were necessary to include, as all

of the trials requiring a target absent response in this task could not only include foil trials. It is unlikely that a 5% difference significantly impacted the results, especially given that the participants were not explicitly made aware of these target prevalence rates. The participants completed a total of 1416 trials, which lasted approximately 1.5 hours. Participants were instructed to press the left arrow key to indicate 'target present' and the right arrow key to indicate 'target absent'.

The training session for the participants in the assemble condition required them to construct either object A or object B (Figure 1). They were given seven individual Lego[®] pieces along with the 8 images of the different views of their object to help them construct it manually. They were given at most ten minutes to construct the object twice, and were required to make the object at least once without help from the experimenter. If the participant was unable to construct the object at first, the experimenter would provide the images of the 8 different views and highlight the pieces that were correctly or incorrectly assembled. The vast majority of participants finished within 5 minutes. All participants in the assemble condition were able to construct the object (double checked by the experimenter) at least once during training. Participants in the disassemble condition were instructed to take apart an already constructed object A or object B twice (the object was constructed by the experimenter). If the participants were confused by the instructions, the experimenter explained by saying, "break the object into individual blocks." Participants in the no training condition were instructed to sit quietly for 5 minutes between pretest and posttest sessions, to account for any training effects that would occur only via completing the visual search tasks. Participants in this condition had the opportunity to look at the 8 different views of the target image as often as they wanted to at the beginning of every block (similar to the participants in the two training conditions), but sat

quietly without engaging in activities related to the experiment during the 5-minute "training" period. Throughout the entire experiment, all participants were provided with a printout of the 16 different object views (8 views per object) as often as they needed them.

2.3 EEG Recording and Data Analysis

The EEG data were DC-recorded using 32 scalp electrodes at standard positions of the extended 10/20 system. The data were sampled at 500 Hz, and a 40 Hz Butterworth zero phase IIR low-pass filter (48 dB/octave) and a 0.1 Hz high-pass filter (12 dB/octave), in addition to a 60 Hz notch filter, were applied offline after re-referencing to averaged earlobes. A 100 ms pretest stimulus baseline was applied to epochs from -100 ms to 500 ms relative to the onset of the search target. Artifact rejection (-100 ms to 300 ms relative to stimulus onset) included criteria consisting of horizontal EOG exceeding $\pm 25 \mu\text{V}$ and vertical EOG exceeding $\pm 60 \mu\text{V}$, as well as all other channels exceeding $\pm 80 \mu\text{V}$. The average percentage of correct trials (i.e., correct hits and correct rejections) that was retained per participant after artifact rejection was 76.5% ($SD = 12.68$, range: 50.6-97.4%). The mean N2pc amplitudes were expressed as the contralateral minus ipsilateral difference in the 200-320 ms window after stimulus onset from lateral posterior electrode sites PO7/8 (Wu et al., 2015; 2017a).

3. Results

3.1 EEG results

3.1.1 Omnibus ANOVAs.

Using the mean amplitudes (Figure 3) from the N2pc difference waves (Figures 4-9), we conducted a 2 (Test: Pretest vs. Posttest) \times 3 (Training condition: assemble, disassemble, no

training) \times 3 (Trial type: exemplar match, category match, foil) mixed ANOVA. No target trials were not included in the analyses because N2pc analyses require either a target or foil to calculate contralateral vs. ipsilateral difference waves. We found a main effect of test on the mean amplitude of the N2pc, with an increase in N2pc amplitude from pretest to posttest, $F(1,48) = 15.80, p < .001, \eta^2 = .25$, and a main effect of trial type, $F(1,48) = 62.16, p < .001, \eta^2 = .56$. There were no other main effects or interactions, $F < 1.14$.

INSERT FIGURES 3-9 ABOUT HERE

Given our a priori hypotheses on task-relevant selection and task-irrelevant attentional capture, although the interaction between trial type and test was not significant, $F < 0.082$, we compared pretest and posttest effects based on trial type by conducting separate mixed ANOVAs for each trial type: exemplar match, category match, and foil trials. We found a main effect of test on the mean amplitude of the N2pc, with an increase in N2pc amplitude from pretest to posttest, for the exemplar match trials, $F(1,48) = 5.54, p = .023, \eta^2 = .10$, and critically for the foil trials, $F(1,48) = 12.51, p = .001, \eta^2 = .21$, but only a trend towards significance for category match trials, $F(1,48) = 3.13, p = .083, \eta^2 = .06$. There were no other main effects or significant interactions, $F < 2.33$.

To evaluate the effects of training based on our a priori hypotheses, although there was no interaction between training condition and test for all three trial types, Bonferroni-corrected planned comparisons were conducted for each condition across the three test trials (adjusted $\alpha=.017$ for each of the three trial types). In the assemble condition, there was an increase in the

Visual search (14)

N2pc amplitude from pretest and posttest for the exemplar match trials, $t(17) = 2.65, p = .017$, and there was a trend towards significance (using the Bonferroni correction) for the foil trials, $t(17) = 2.44, p = .026$, but not for the category match trials, $t(17) = 2.10, p = .051$. These results suggest that in the assemble condition, search efficiency increased from pretest to posttest for exemplar match trials, and there was a marginal effect of increased attentional capture to nontarget views of the same object as the target view. In the disassemble condition and no training condition, there were no significant differences between pretest and posttest N2pc amplitudes, all $|t|s < 1.91, ps > .076$, although numerically, the amplitudes increased from pretest to posttest.

3.1.2 Presence of the N2pc.

In addition to analyzing the change in the N2pc amplitude, determining whether an N2pc component was present is important for confirming a change in N2pc amplitude. Paired t -tests (adjusted $\alpha = .017$) between contralateral and ipsilateral mean amplitudes demonstrated that, collapsed across conditions, there was a significant N2pc during pretest and posttests for exemplar match and category match trials, $t(50) > 3.84, p < .001$. For the foil trials, there was a marginally significant N2pc across the three conditions during posttest, $t(50) = 2.25, p = .029$, but not during pretest, $t(50) = 1.57, p = .122$.

To determine whether an N2pc component was present in each of the three trial types (adjusted $\alpha = .017$) for each condition, Bonferroni-corrected planned paired t -tests were used for the subsequent analyses. The analyses were conducted separately for pretest and posttest trials in each of the three conditions. In the assemble condition, there was a significant N2pc component for exemplar match and category match trials during pretest and posttest, $|t|s > 3.04, ps < .007$,

but no significant component for foil trials during pretest and posttest, $|t|s < 1.69$, $ps > .110$. In the disassemble condition, there was a significant N2pc component for exemplar match and category match trials pretest and posttest, $|t|s > 2.91$, $ps < .011$, but no significant component for foil trials during pretest and posttest, $|t|s < 1.02$, $ps > .325$. In the no training condition, there was a significant N2pc component in the exemplar match posttest trials, $t(16) = 4.09$, $p = .001$, a trend towards significance for the N2pc component in the exemplar match pretest trials, $t(16) = 2.54$, $p = .022$, but not for the category match posttest trials, $t(16) = -2.12$, $p = .050$, or other trials, $|t|s < 1.28$, $ps > .218$. Note that for the foil trials, although the presence of the N2pc component was not statistically significant for individual training conditions, our previous analysis nonetheless demonstrated that there was a significant increase in the N2pc amplitude between pretest and posttest collapsed across all training conditions.

3.2 Behavioral results

We conducted a 2 (Test: pretest vs. posttest) \times 4 (Trial type: exemplar, foil, category, no target) \times 3 (Training condition: assemble, disassemble, no training) mixed ANOVA on accuracy and reaction time (Figure 10). We found a main effect of test for accuracy, $F(1,48) = 26.47$, $p < .001$, $\eta^2 = .36$ (i.e., accuracy increased from pretest to posttest) and reaction time, $F(1,48) = 54.05$, $p < .001$, $\eta^2 = .53$ (i.e., reaction time decreased from pretest to posttest). We also found a main effect of trial type for accuracy, $F(3,144) = 48.83$, $p < .001$, $\eta^2 = .50$, and reaction time, $F(3,144) = 46.94$, $p < .001$, $\eta^2 = .49$. Although there was no main effect of training condition for accuracy, $F(2,48) = .73$, $p = .493$, there was one for reaction time, $F(2,48) = 7.66$, $p = .001$, $\eta^2 = .24$. Bonferroni-corrected planned comparisons (adjusted $\alpha = .017$) revealed that participants in the assemble condition responded significantly faster than those in the no training condition,

Visual search (16)

$t(33) = 3.82, p = .001$, and marginally faster than those in the disassemble condition, $t(32) = 2.36, p = .025$, whereas the other two conditions did not differ from each other, $t(31) = 1.23, p = .227$.

 INSERT FIGURE 10 ABOUT HERE

To investigate the main effect of trial type for accuracy, we collapsed the trials across pretest and posttest phases and compared the two target present trials (exemplar match vs. category match) and the two target absent trials (foil vs. no target), (adjusted $\alpha = .025$). Consistent with previous results (e.g., Wu et al., 2016), for target present trials, exemplar match had higher accuracy than category match trials, $t(50) = 7.78, p < .001$. Moreover, foil trials had higher accuracy than no target trials, $t(50) = 8.64, p < .001$. To investigate the main effect of trial type for reaction time, we collapsed the trials across pretest and posttest phases and compared the two target present trials (exemplar match vs. category match) and the two target absent trials (foil vs. no target), (adjusted $\alpha = .025$). For target present trials, exemplar match had faster reaction times than category match trials, $t(50) = -6.55, p < .001$, and foil trials had faster reaction times than no target trials, $t(50) = -8.61, p < .001$.

From the omnibus ANOVA, although there were no interactions for accuracy, $F < 2.19$, there was a marginal interaction between test and training condition for reaction time, $F(2,48) = 3.13, p = .053, \eta^2 = .12$. To investigate this marginal interaction for reaction time, we collapsed across trial types and conducted pairwise comparisons for pretest and posttest reaction time for each training condition (adjusted $\alpha = .017$). Reaction time decreased from pretest to posttest for

the assemble condition, $t(17) = 4.91, p < .001$, and no training condition, $t(16) = 7.52, p < .001$, but did not significantly decrease for the disassemble condition, $t(15) = 1.81, p = .090$.

It is possible that completing a category search task prior to an exemplar search task may have inflated the foil effect. We conducted a t -test collapsed across conditions to examine this possibility. Although numerically the foil N2pcs were larger when participants completed the category task first ($M_{\text{exemplarFirst}} = .04, SD_{\text{exemplarFirst}} = .55, M_{\text{categoryFirst}} = .26, SD_{\text{categoryFirst}} = .62$), this difference was not significant ($t(49) = 1.32, p = .194$). Therefore, although it is plausible that a category search task may inflate later foil effects, it does not seem to play a significant role in the present study.

4. Discussion

The present study investigated the emergence of the benefits and costs of grouping target items on visual search efficiency. Participants were trained to integrate different views of an object by assembling it out of seven Lego pieces (assemble condition), taking apart an already assembled Lego object (disassemble condition), or sitting quietly for 5 minutes without explicit instructions (no training condition). Before and after the training procedure, a visual search task required participants to search for a specific object view (i.e., exemplar search) or any view of the trained object (i.e., category search). During exemplar search, foil trials were included that displayed the nontarget views of the same object as the target view. Task-relevant selection was measured on trials when a target was displayed (either a specific view in exemplar search, or any views of the trained object in category search). Task-irrelevant attentional capture was measured via foil trials that displayed nontarget views of the target object when the participant was searching for a specific view.

Visual search (18)

Overall, we found that exposure to a visual search task involving the integration of object views enhanced search efficiency for selecting exemplars (and possibly category targets), but also increased obligatory attentional capture to nontarget items related to the target. The omnibus ANOVA on the ERP results (N2pc component) revealed that task-relevant attentional selection of a specific object view (exemplar match trials) increased in efficiency regardless of training condition. Similarly, selection for any object views (category match trials) increased as well, but this effect was only marginal. Although there was only a marginally significant foil N2pc component (i.e., task-irrelevant attentional capture) during the posttest trials collapsed across all three training conditions, there was a main effect, from pretest to posttest, of an increase in foil N2pc amplitude regardless of training condition. The main effect, supporting the notion of an emergence of a foil N2pc component from pretest to posttest, is in line with prior findings. Previous studies that have observed foil effects have used highly familiar stimuli (e.g., letters, numbers, clothing, kitchen items; Nako et al., 2014a; 2014b) or items that partially match the target (e.g., large red bar when the target is a small red bar; Kiss et al., 2013). Other studies with similar designs to the present study that do not observe foil effects have used novel stimuli, such as Chinese characters (Wu et al., 2013) and alien stimuli (Wu et al., 2016). A recent study also showed that foil amplitude may depend on the level of experience with the target categories (Wu et al., 2017a). The behavioral results largely mirrored the pattern from the ERP results. Overall, these results suggest that practice at a visual search task involving grouping object views into categories enhances search efficiency for selecting exemplars (and perhaps category targets), but also increases attentional capture to related nontarget category members.

Interestingly, according to the results from the omnibus ANOVA, even the training group that received no hands-on training (i.e., sitting quietly for 5 minutes, and only completing the

visual search tasks) displayed similar patterns of improvement to the other groups who received hands-on training. This result suggests that our 5-minute hands-on training procedure might not be better than (or perhaps was overshadowed by) the repeated testing that the participants received when searching for specific or category targets. There is a possibility that completing more than 700 visual search trials during the pretest and having access to the two categories of images on a printed sheet throughout the experiment enhanced pretest to posttest effects in all three conditions, thereby attenuating differences among the three conditions. We had to administer the training halfway through the experimental session to obtain pretest and posttest measures. Moreover, the pretest was lengthy (a typical N2pc experimental session) so that we could obtain enough useable trials for analyses, especially given that accuracy rates were lower during pretest. Future studies could administer an extended training and only the posttest phase to draw stronger conclusions based on the training effects.

There are a number of further avenues for future research based on the results of the present study. First, although the omnibus ANOVA did not support the hypothesis that only the assemble condition would show an effect from pretest to posttest, the Bonferroni-corrected comparisons largely supported the hypothesis. Therefore, it is unclear whether assembling the target object, rather than disassembling the object or sitting quietly, might have increased search efficiency. Future work could investigate different types of training procedures, including training length and intensity, to better understand the emergence of the costs and benefits of prior knowledge on attentional selection. Second, whereas prior work has used categories consisting of discrete objects, the current study used different views of one object as the “category”. This novel way of conceptualizing a category extends our understanding of how searching for different views of the same object can rely on similar search processes compared to searching for

Visual search (20)

more abstract categories. Future work could extend the current findings to categories with discrete objects. Moreover, because only eight items were included in each category in the present study, future studies could measure differences between searching for such small categories versus larger ones. Our prior studies (e.g., Wu et al., 2013; 2015; 2016; 2017; 2018) have demonstrated similar effects using eight items in each category, but varying the breadth of the category (e.g., letters, animal faces, healthy food), although broader categories (e.g., healthy food) tend to elicit attenuated N2pc components compared to smaller categories (e.g., letters). Finally, increased task difficulty in an already difficult task seems to attenuate the N2pc amplitude (see Wu et al., 2013, which uses a similar paradigm as the present study). Future studies could minimize task difficulty perhaps via using simpler objects to maximize the likelihood of finding subtle effects, such as those suggested in the present study.

A significant limitation of the present study is that there is a noticeable difference in the N2pc component from pretest to posttest for category match trials in the no training condition, which does not seem to appear in the assemble or disassemble conditions. This difference may have appeared due to different sampling environments, including participant demographics (private university in upstate NY versus a public university in southern California), different compensation (money in the NY lab versus course credit in the CA lab), as well as different experimenters, although the first author trained all of the experimenters. Many other aspects of the testing environments were nearly identical, namely the EEG system used, the experimental procedure, and the analysis pipeline. Interestingly, we were expecting the pretest results from the category match trials from all of the conditions to resemble those from the no training condition (i.e., no category N2pc at pretest). However, we did not find this result for the first two groups. This discrepancy calls into question whether the three groups are comparable. However, the

difference among the pretest trials across conditions was not validated in the omnibus ANOVA nor in the planned Bonferroni-corrected comparisons. Moreover, other N2pc results from the exemplar trials and the behavioral results are numerically similar across all conditions from pretest to posttest. To investigate this potential discrepancy in the category match trials, future work could investigate individual differences in the ability to group novel images and objects, such as how individuals may generalize from differential expertise to novel objects, as well as the efficacy of various training procedures on different individuals. Perhaps some individuals may be able to integrate particular sets of novel images better than other sets due to overlap with prior object knowledge that conforms to similar experimental norms.

The present study found that practice at a visual search task involving grouping object views into categories enhances task-relevant attentional selection, but also may increase task-irrelevant attentional capture. This study on grouping novel stimuli informs a larger research question regarding situations when knowledge is beneficial (e.g., visual search of familiar items, generalization) and when knowledge is disadvantageous (e.g., Stroop effect, inattention blindness). The present study focuses on how the costs and benefits of knowledge on search processes develop over a single experimental session. This paradigm also may provide markers of learning (e.g., categorization) for studies with infants and children to determine how their level of experience with particular stimuli affects attentional capacities. With a better understanding of the development of the costs and benefits of knowledge on visual search, the manner in which knowledge can be leveraged to facilitate new learning and the contexts in which previous knowledge could be detrimental to new learning may be identified (Wu & Zhao, 2017; Wu et al., 2017b).

Acknowledgments

We thank Megan Runkle and other Aslin and CALLA Lab RAs for their help with data collection and Rebecca Nako for useful discussions on the study design and data interpretations. We also thank Brain Vision LLC and Brain Products for their technical support.

Funding

This research was funded by an NRSA (grant number F32HD070537) from NICHD to R.W., and an NIH grant (HD-037082) to R.N.A.

References

- Blair, M. R., Watson, M. R., & Meier, K. M. (2009). Errors, efficiency, and the interplay between attention and category learning. *Cognition*, *112*(2), 330–336.
<https://doi.org/10.1016/j.cognition.2009.04.008>
- Brady, T., Konkle, T., & Alvarez, G. (2009). Compression in visual working memory: Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General*, *138*(4), 487–502. <https://doi.org/10.1037/a0016797>
- Castro, L., & Wasserman, E. A. (2016). Attentional shifts in categorization learning: Perseveration but not learned irrelevance. *Behavioral Processes*, *123*, 63–73.
<https://doi.org/10.1016/j.beproc.2015.11.001>
- Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*, 28–71.
<https://doi.org/10.1006/cogp.1998.0681>
- Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology*, *99*(3), 225–234.
[https://doi.org/10.1016/S0921-884X\(96\)95711-2](https://doi.org/10.1016/S0921-884X(96)95711-2)
- Fagan, J. F. (1976). Infants' recognition of invariant features of faces. *Child Development*, *4*(3), 627–638. <https://doi.org/10.1111/j.1467-8624.1976.tb02226.x>
- Freedman, D. J., Riesenhuber, M., Poggio, T., Miller, & E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science*, *291* (5502), 312–316.
<https://doi.org/10.1126/science.291.5502.312>

Visual search (24)

- Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: Exploring mechanisms for face recognition. *Vision Research*, 37, 1673–1682. [https://doi.org/10.1016/S0042-6989\(96\)00286-6](https://doi.org/10.1016/S0042-6989(96)00286-6)
- Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C. H., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences*, 5(6), 236–243. [https://doi.org/10.1016/S1364-6613\(00\)01662-4](https://doi.org/10.1016/S1364-6613(00)01662-4)
- Kiss, M., Grubert, A., & Eimer, M. (2013). Top-down task sets for combined features: Behavioural and electrophysiological evidence for two stages in attentional object selection. *Attention, Perception, & Psychophysics*, 75, 216–228. <https://doi.org/10.3758/s13414-012-0391-z>
- Kruschke, J. K. (2001). Toward a unified model of attention in associative learning. *Journal of Mathematical Psychology*, 45, 812–863. <https://doi.org/10.1006/jmps.2000.1354>
- Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, 31, 291–308. <https://doi.org/10.1111/j.1469-8986.1994.tb02218.x>
- McGugin, R. W., McKeef, T. J., Tong, F., & Gauthier, I. (2011). Irrelevant objects of expertise compete with faces during visual search. *Attention, Perception, and Psychophysics*, 73, 309–317. <https://doi.org/10.3758/s13414-010-0006-5>
- Nako, R., Wu, R., & Eimer, M. (2014a). Rapid guidance of visual search by object categories. *Journal of Experimental Psychology: Human Perception and Performance*, 40(1), 50–60. <https://doi.org/10.1037/a0033228>
- Nako, R., Wu, R., Smith, T. J., & Eimer, M. (2014b). Item and category-based attentional control during search for real-world objects: Can you find the pants among the pans? *Journal of*

- Experimental Psychology: Human Perception and Performance*, 40(4), 1283–1288.
<https://doi.org/10.1037/a0036885>
- Soska, K. C., & Johnson, S. P. (2008). Development of three-dimensional object completion in infancy. *Child Development*, 79, 1230–1236. <https://doi.org/10.1111/j.1467-8624.2008.01185.x>
- Soska, K. C., Adolph, K. E., & Johnson, S. P. (2010). Systems in development: Motor skill acquisition facilitates three-dimensional object completion. *Developmental Psychology*, 46, 129–138. <https://doi.org/10.1037/a0014618>
- Telling, A., Kumar, S., Meyer, A., & Humphreys, G. (2010). Electrophysiological Evidence of Semantic Interference in Visual Search. *Journal of Cognitive Neuroscience*, 22(10), 2212–2225. <https://doi.org/10.1162/jocn.2009.21348>
- Wolfe, J. M. Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews, Neuroscience*, 5, 495–501.
<https://doi.org/10.1038/nrn1411>
- Woodman, G., Vecera, S., & Luck, S. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review*, 10(1), 80–87.
<https://doi.org/10.3758/BF03196470>
- Wu, R., Gopnik, A., Richardson, D. C., & Kirkham, N. Z. (2011). Infants learn about objects from statistics and people. *Developmental Psychology*, 47(5), 1220–1229.
<https://doi.org/10.1037/a0024023>
- Wu, R., McGee, B., Echiverri, C., & Zinszer, B. (2018). Prior knowledge of category size impacts visual search. *Psychophysiology*, e13075.

Visual search (26)

- Wu, R., Nako, R., Band, J., Pizzuto, J., Shadravan, Y., Scerif, G., & Aslin, R. N. (2015). Rapid selection of non-native stimuli despite perceptual narrowing. *Journal of Cognitive Neuroscience*, 27(11), 2299–2307. https://doi.org/10.1162/jocn_a_00857
- Wu, R., Pruitt, Z., Runkle, M., Scerif, G., & Aslin, R. N. (2016). A neural signature of rapid category-based target selection as a function of intra-item perceptual similarity despite inter-item dissimilarity. *Attention, Perception, and Psychophysics*, 78(3), 749–776. <https://doi.org/10.3758/s13414-015-1039-6>
- Wu, R., Pruitt, Z., Zinszer, B., & Cheung, O. (2017a). Increased experience amplifies the activation of task-irrelevant category representations. *Attention, Perception, and Psychophysics*, 79(2), 522–532. <https://doi.org/10.3758/s13414-016-1254-9>
- Wu, R., Rebok, G. W., & Lin, F. V. (2017b). A novel theoretical life course framework for triggering cognitive development across the lifespan. *Human Development*, 56(6), 342–365. <https://doi.org/10.1159/000458720>
- Wu, R., Scerif, G., Aslin, R. N., Smith, T. J., Nako, R., & Eimer, M. (2013). Searching for something familiar or novel: Top-down attentional selection of specific items or object categories. *Journal of Cognitive Neuroscience*, 25(5), 719–729. https://doi.org/10.1162/jocn_a_00352
- Wu, R., Tummeltshammer, K. S., Gliga, T., & Kirkham, N. Z. (2014). Ostensive signals support learning from novel attention cues during infancy. *Frontiers in Psychology*, 5:251. <https://doi.org/10.3389/fpsyg.2014.00251>
- Wu, R., & Zhao, J. (2017). Prior knowledge of object associations shapes attentional templates and information acquisition. *Frontiers in Psychology*, 8(843), 1–6. <https://doi.org/10.3389/fpsyg.2017.00843>

Xu, Y., & Chun, M. M. (2007). Visual grouping in human parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 18766–18771.
<https://doi.org/10.1073/pnas.0705618104>

Figure captions

Figure 1. Stimuli and search arrays for the experiment. The eight different views of each object (object A and object B) are displayed in the top panel. The bottom panel displays the example search arrays from target trials (exemplar or category match), foil trials, and no target trials. In the example search arrays for the exemplar match, foil, and category match trials, the item on the left is either the target or foil view from object A, and the item on the right is a view from object B.

Figure 2. Sample trial sequence. The sample trial sequence depicts a target trial (target on the right), no target trial, foil trial, and another target trial (target on the left), during exemplar search. For the category search task, this same sequence would be a target trial, no target trial, another target trial, and another target trial.

Figure 3. Graph of the N2pc mean amplitudes. Graph of the N2pc mean amplitudes for exemplar, foil, and category trials for both pretest and posttest trials for all three training conditions.

Figure 4. Grand-averaged ERPs from exemplar, category, and foil pretest trials in the assemble condition. The ERPs were recorded at posterior electrodes PO7/8 contralateral and ipsilateral to a target item. N2pc difference waveforms were obtained by subtracting ipsilateral from contralateral ERP waveforms at PO7/8.

Figure 5. Grand-averaged ERPs from exemplar, category, and foil posttest trials in the assemble condition. The ERPs were recorded at posterior electrodes PO7/8 contralateral and ipsilateral to a target item. N2pc difference waveforms were obtained by subtracting ipsilateral from contralateral ERP waveforms at PO7/8.

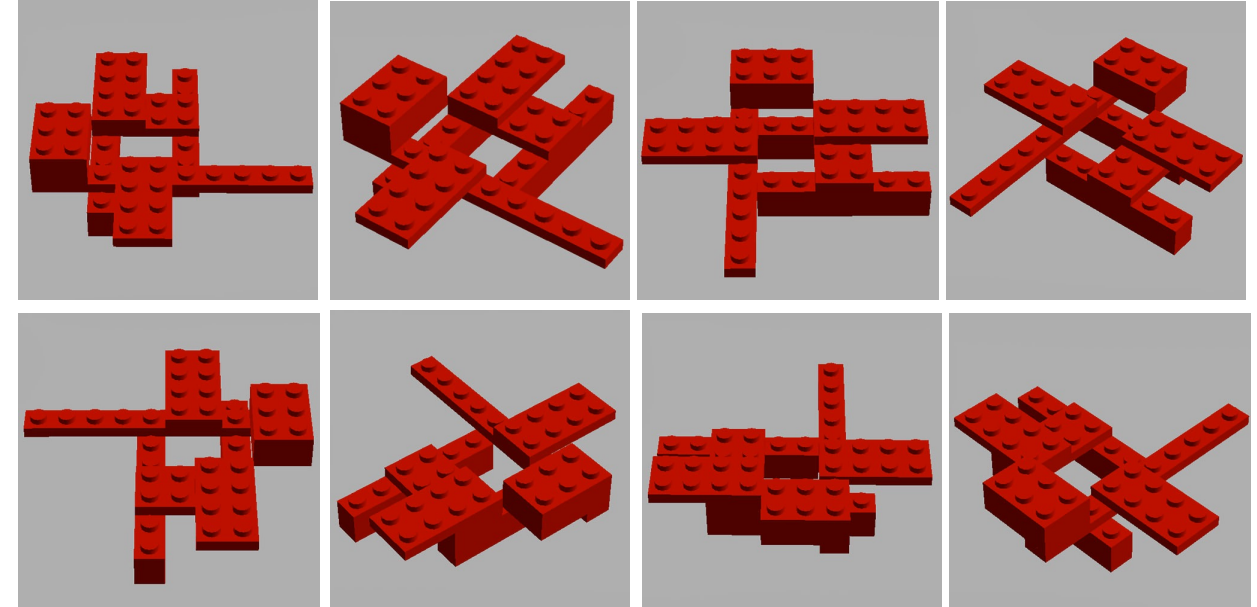
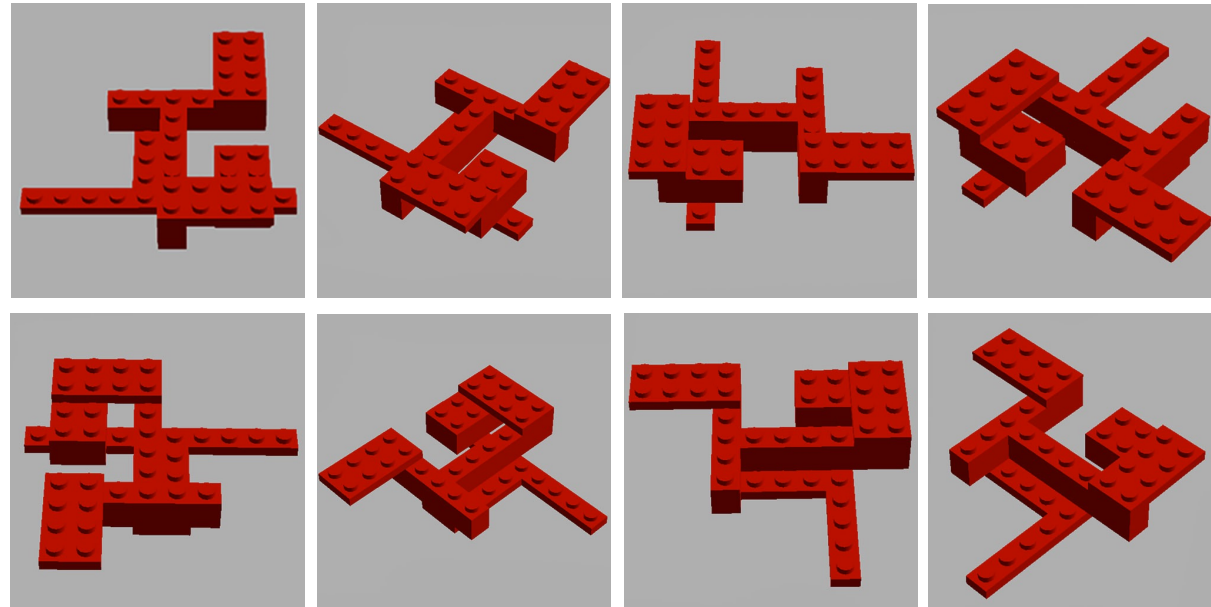
Figure 6. Grand-averaged ERPs from exemplar, category, and foil pretest trials in the disassemble condition. The ERPs were recorded at posterior electrodes PO7/8 contralateral and ipsilateral to a target item. N2pc difference waveforms were obtained by subtracting ipsilateral from contralateral ERP waveforms at PO7/8.

Figure 7. Grand-averaged ERPs from exemplar, category, and foil posttest trials in the disassemble condition. The ERPs were recorded at posterior electrodes PO7/8 contralateral and ipsilateral to a target item. N2pc difference waveforms were obtained by subtracting ipsilateral from contralateral ERP waveforms at PO7/8.

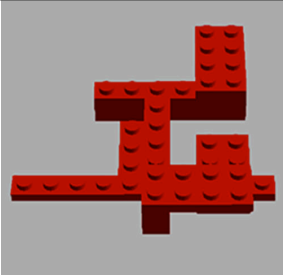
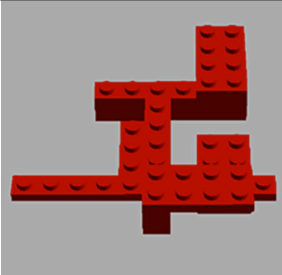
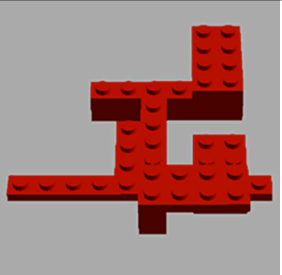
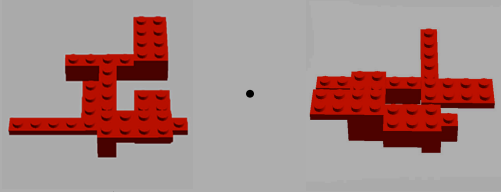
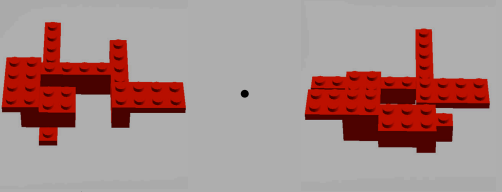
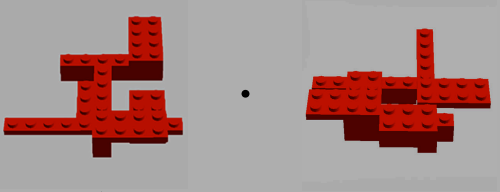
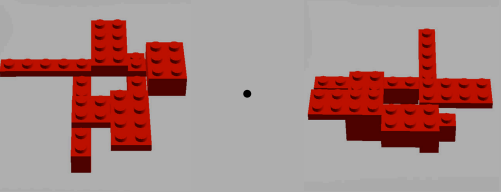
Figure 8. Grand-averaged ERPs from exemplar, category, and foil pretest trials in the no training condition. The ERPs were recorded at posterior electrodes PO7/8 contralateral and ipsilateral to a target item. N2pc difference waveforms were obtained by subtracting ipsilateral from contralateral ERP waveforms at PO7/8.

Figure 9. Grand-averaged ERPs from exemplar, category, and foil posttest trials in the no training condition. The ERPs were recorded at posterior electrodes PO7/8 contralateral and ipsilateral to a target item. N2pc difference waveforms were obtained by subtracting ipsilateral from contralateral ERP waveforms at PO7/8.

Figure 10. Behavioral results. Reaction time and accuracy for all trial types split by pretest and posttest for all three training conditions (assemble, disassemble, and no training conditions). Error bars represent 1 standard error.

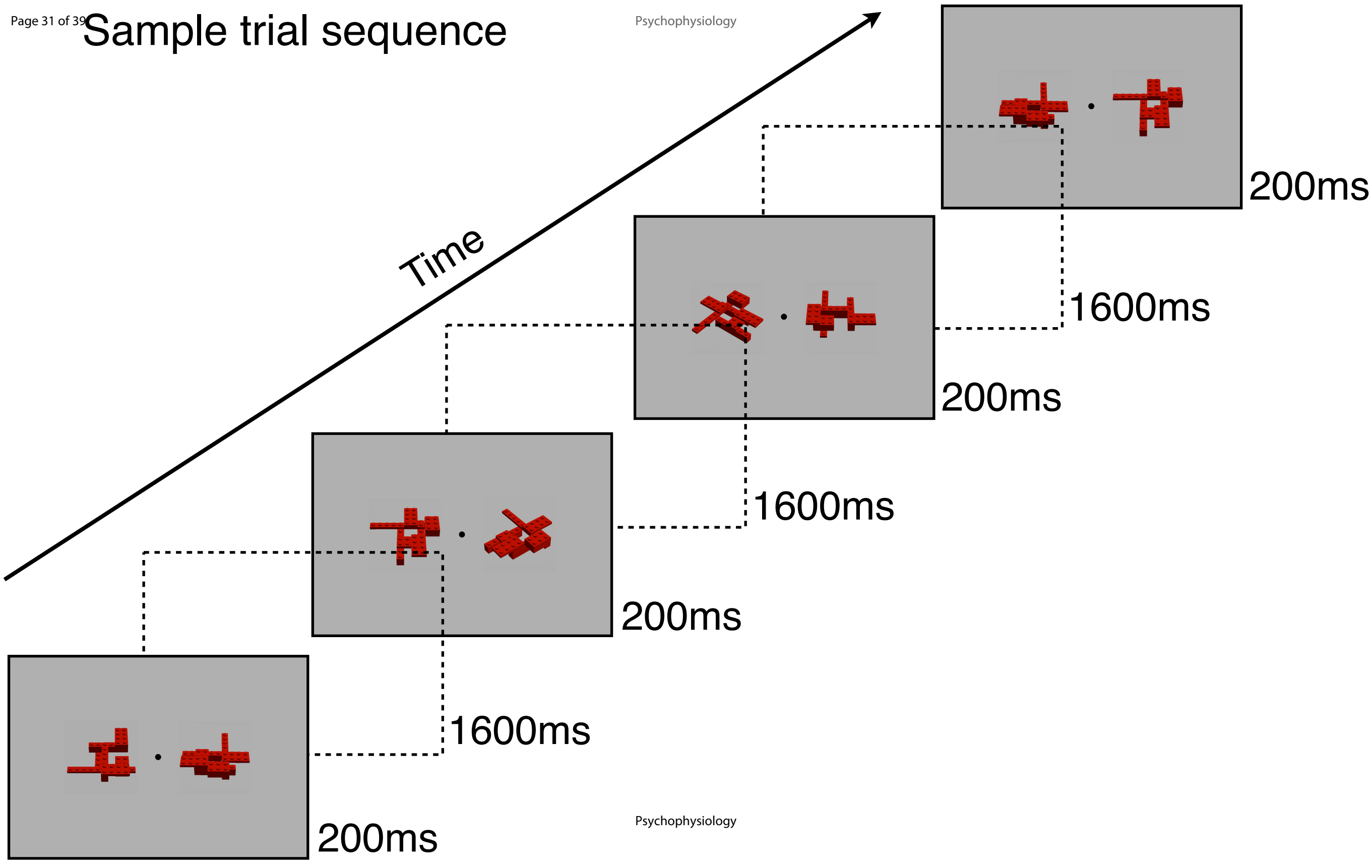


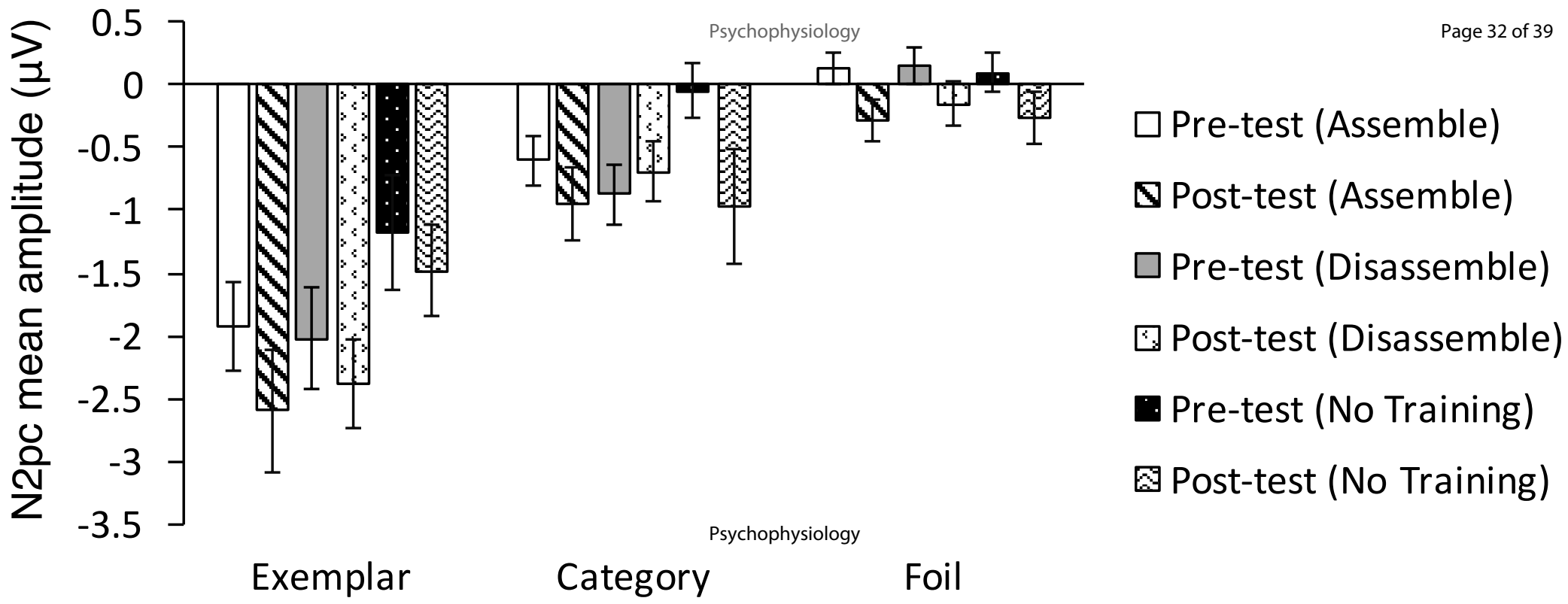
Examples of Search Arrays

	Exemplar Match (specific view)	Foil	Category Match (any view of object)	No Target
Example targets			Any view of Object A	
Example search array				
	Target	Foil	Target	No Target

Sample trial sequence

Psychophysiology

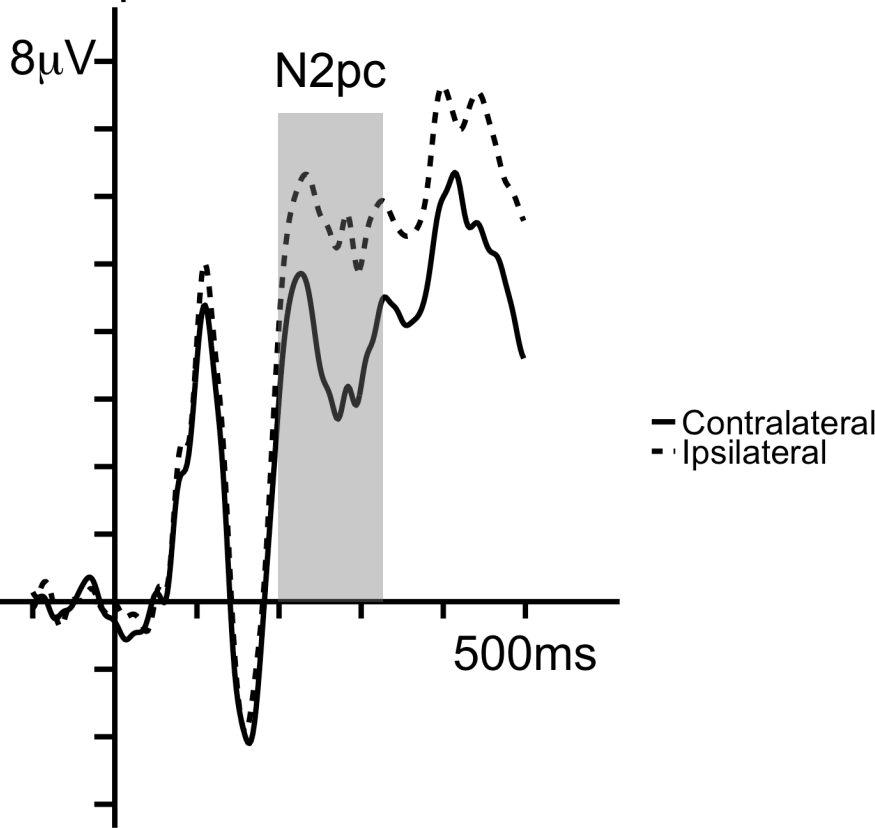




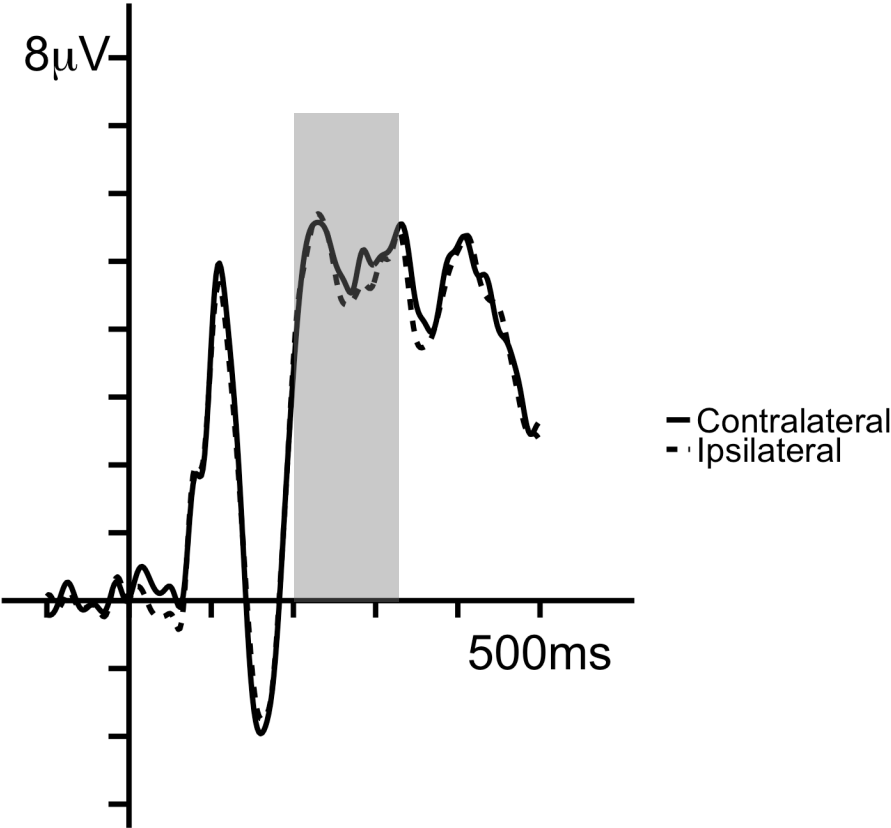
Assemble Condition: Pre-test

Psychophysiology

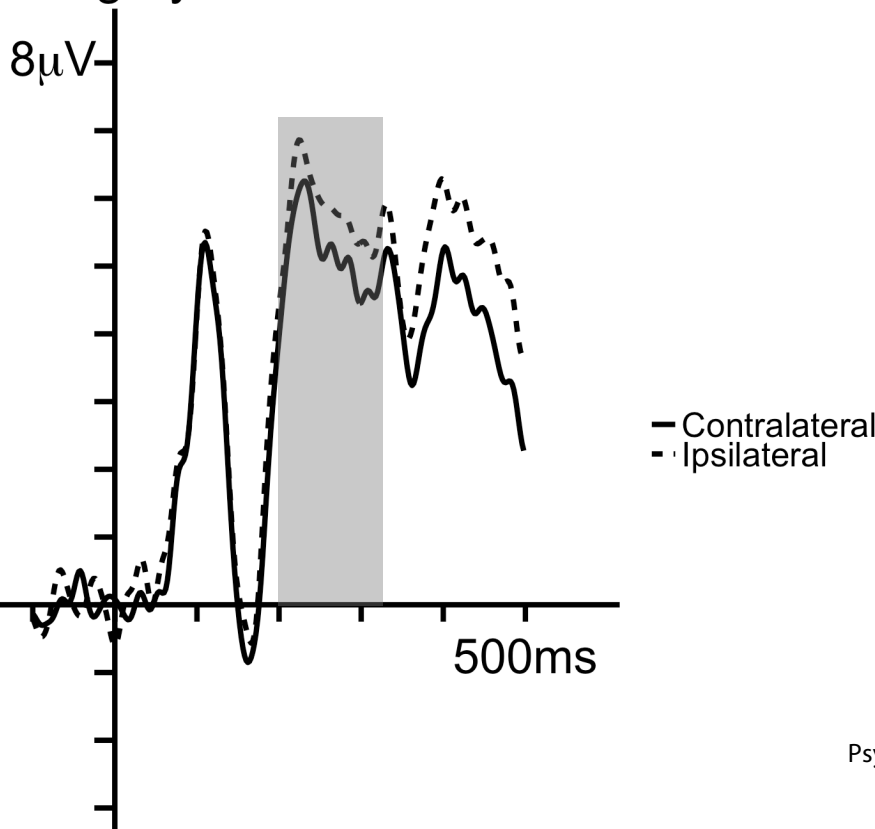
Exemplar Match trials



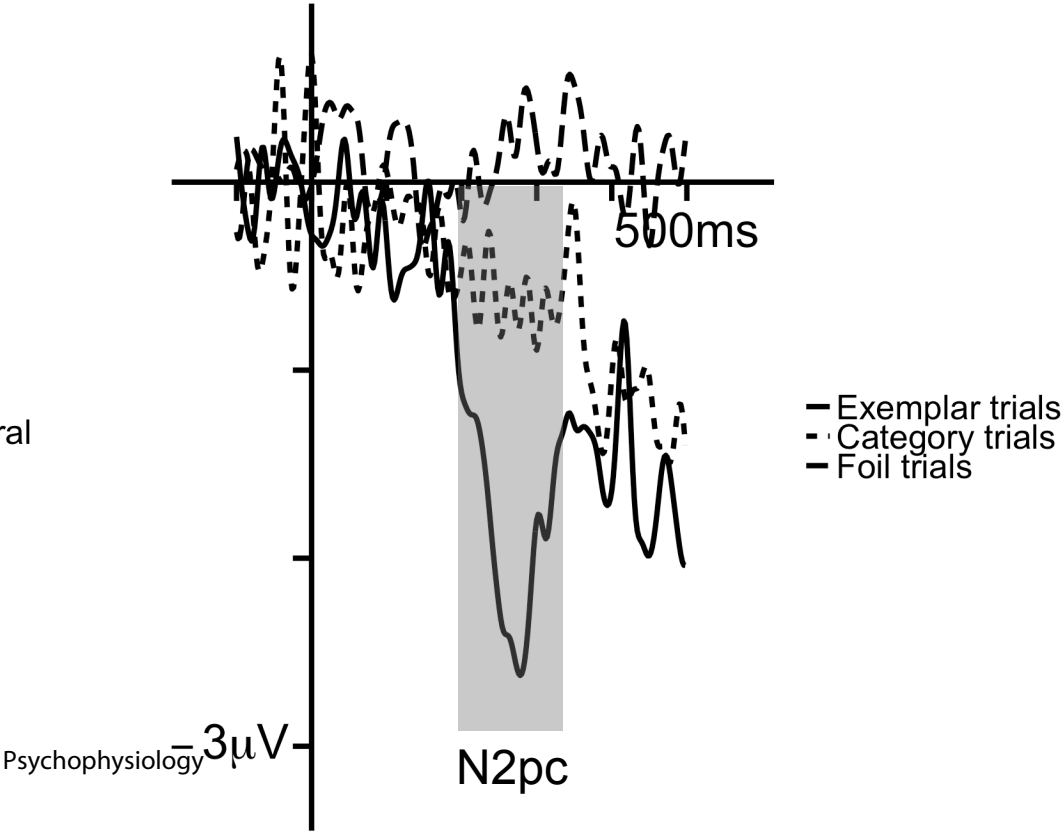
Foil trials



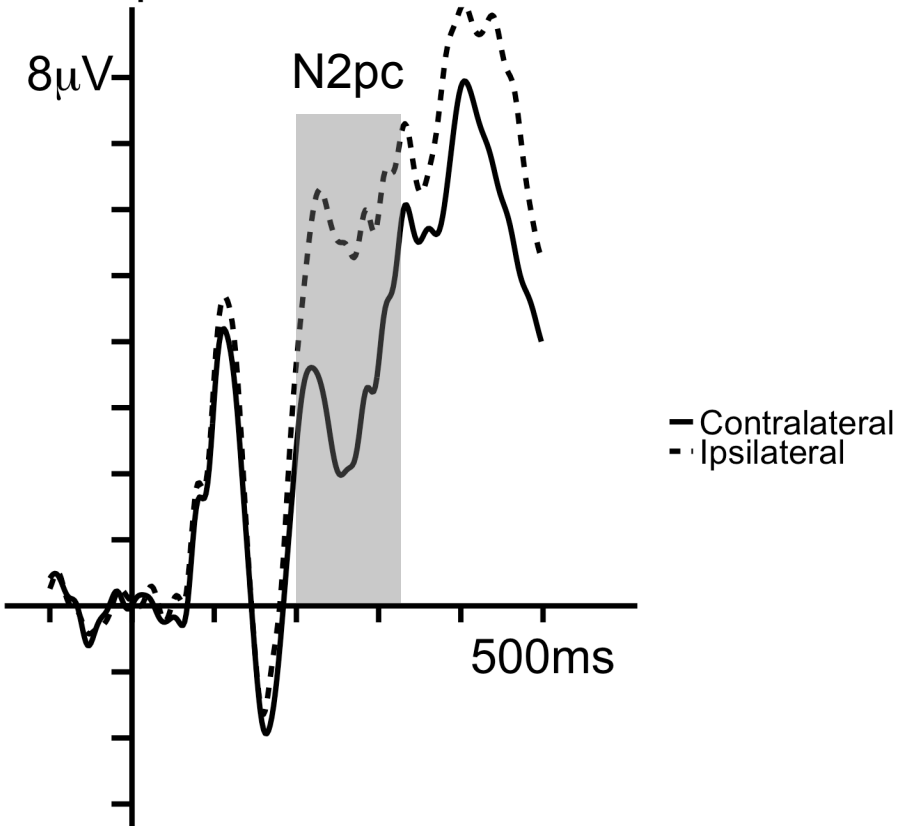
Category Match trials



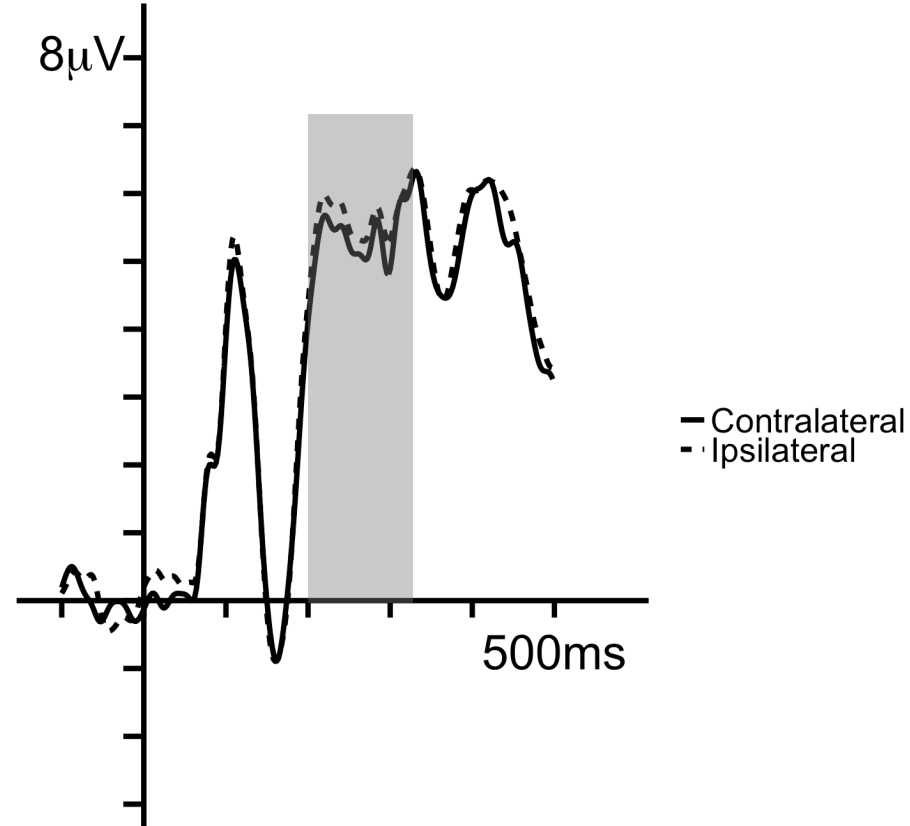
Difference waves



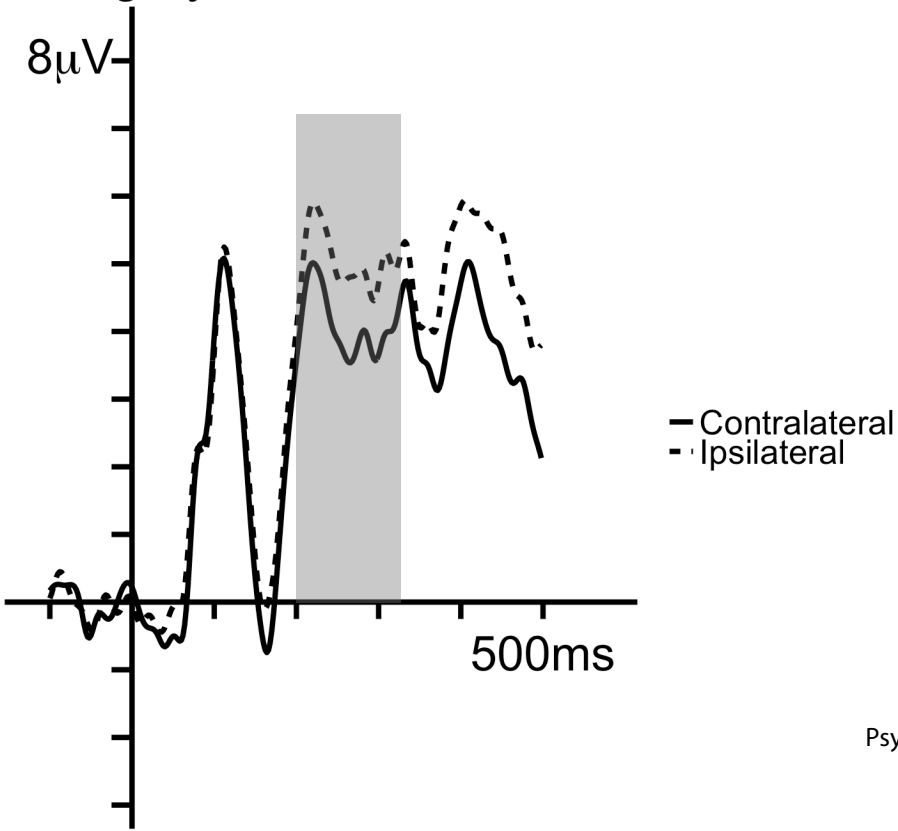
Exemplar Match trials



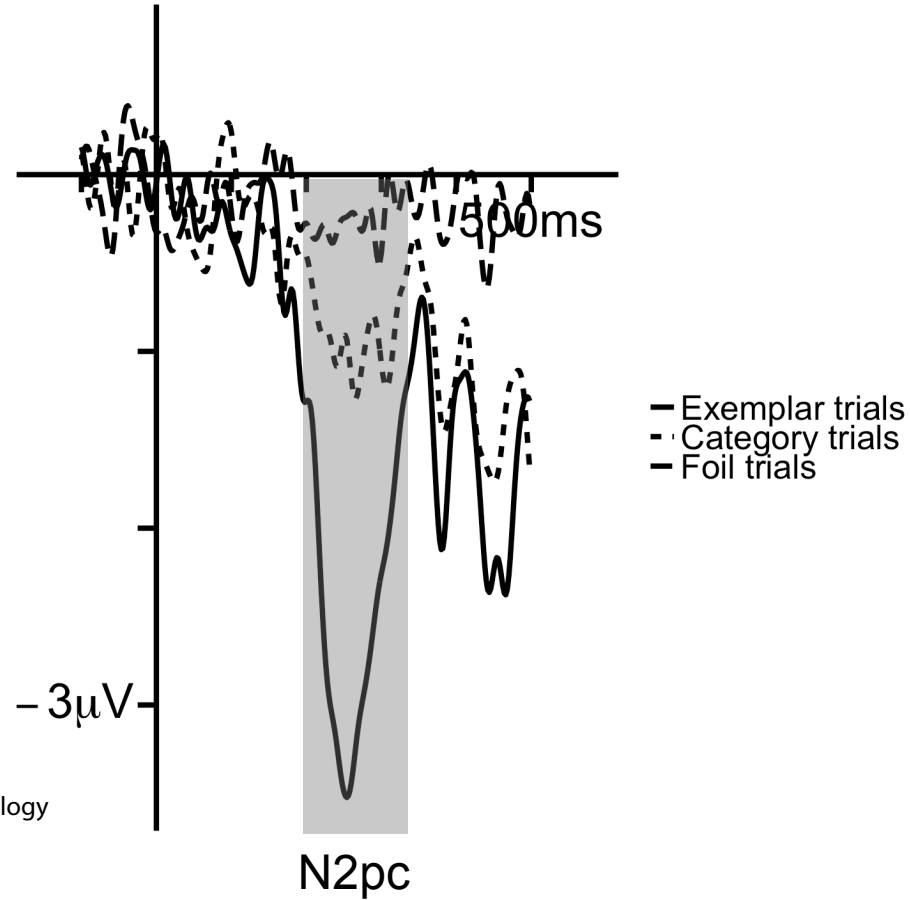
Foil trials



Category Match trials



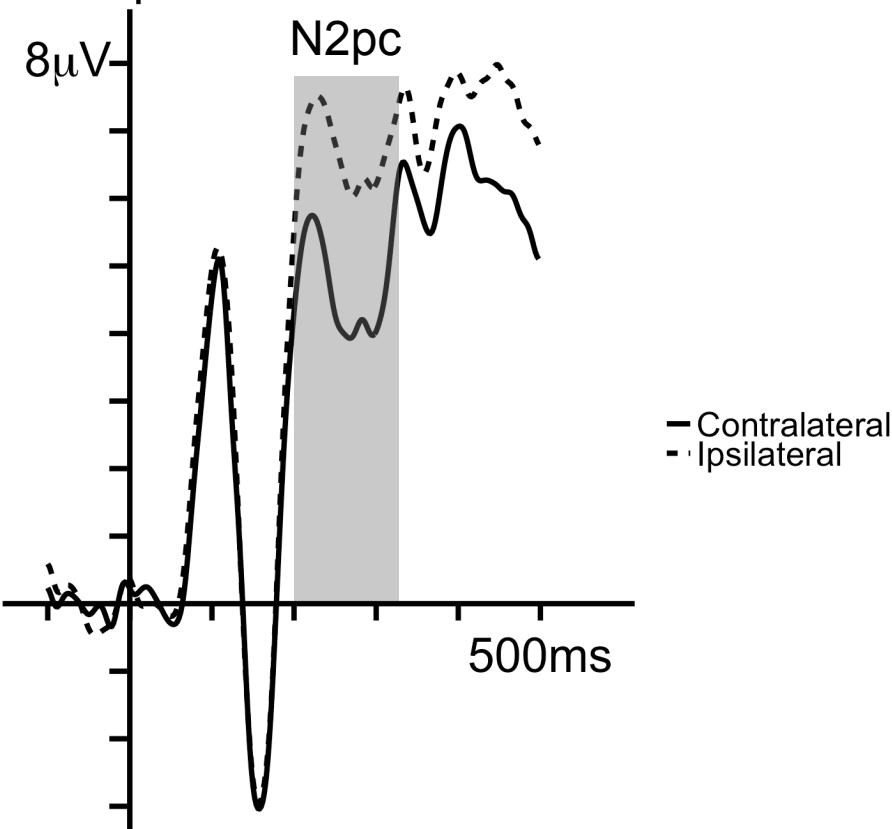
Difference waves



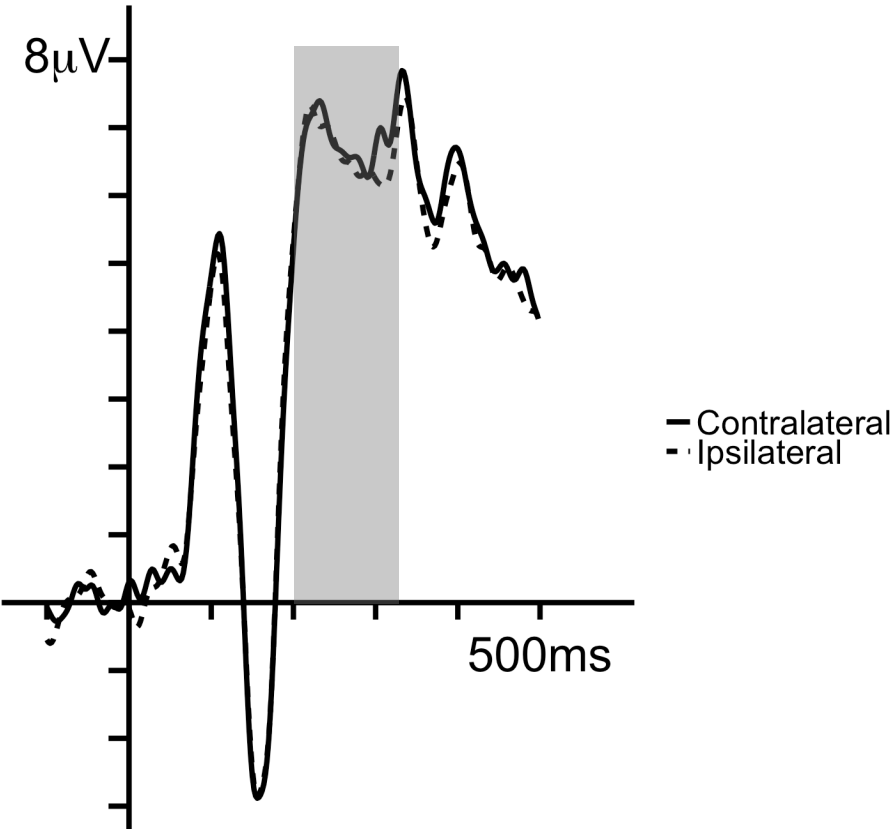
Disassemble Condition: Pre-test

Psychophysiology

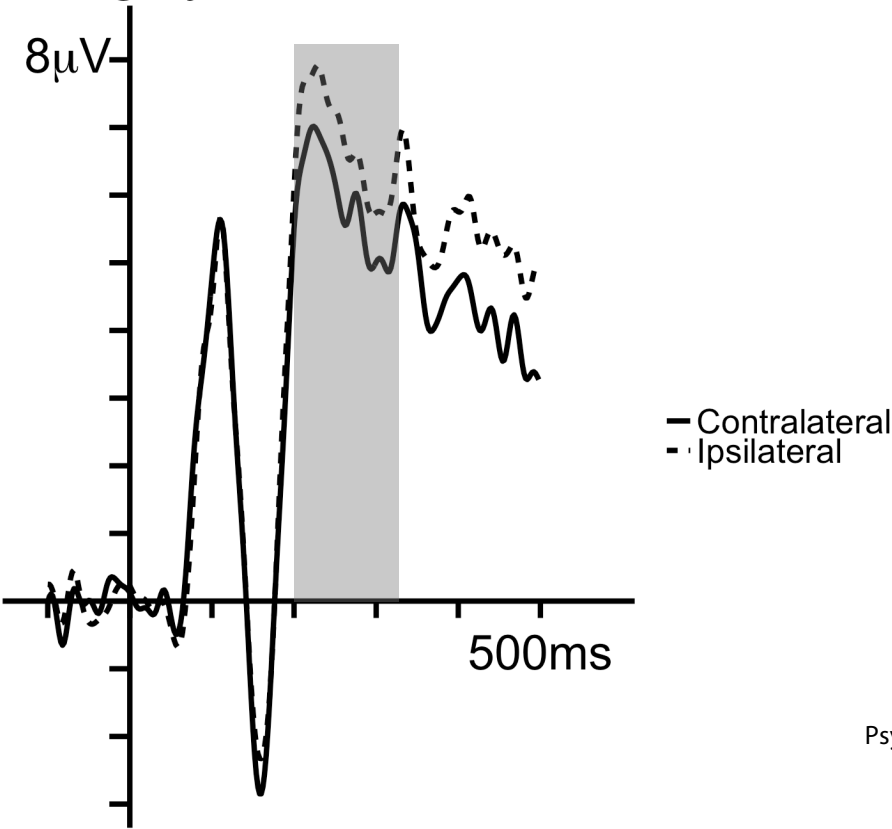
Exemplar Match trials



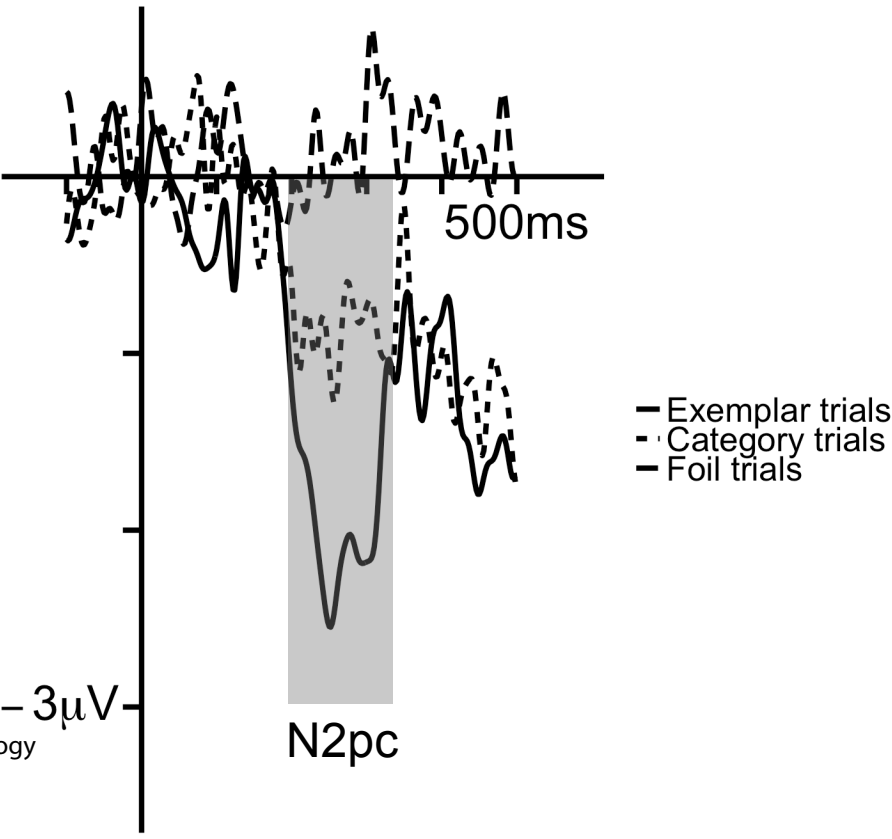
Foil trials



Category Match trials

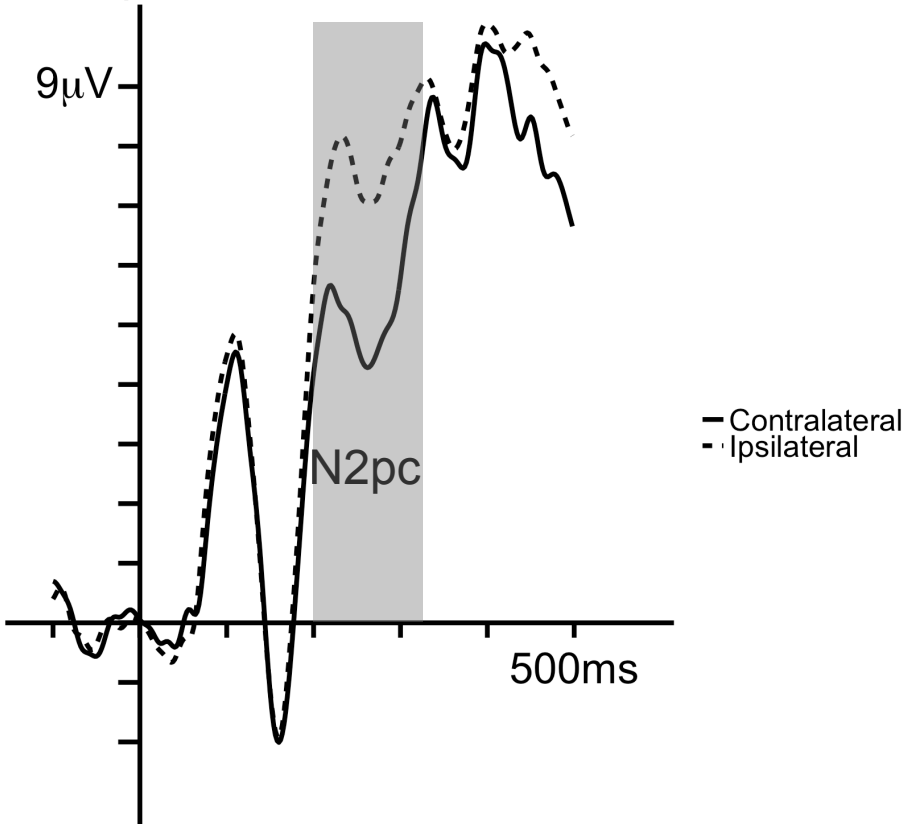


Difference waves

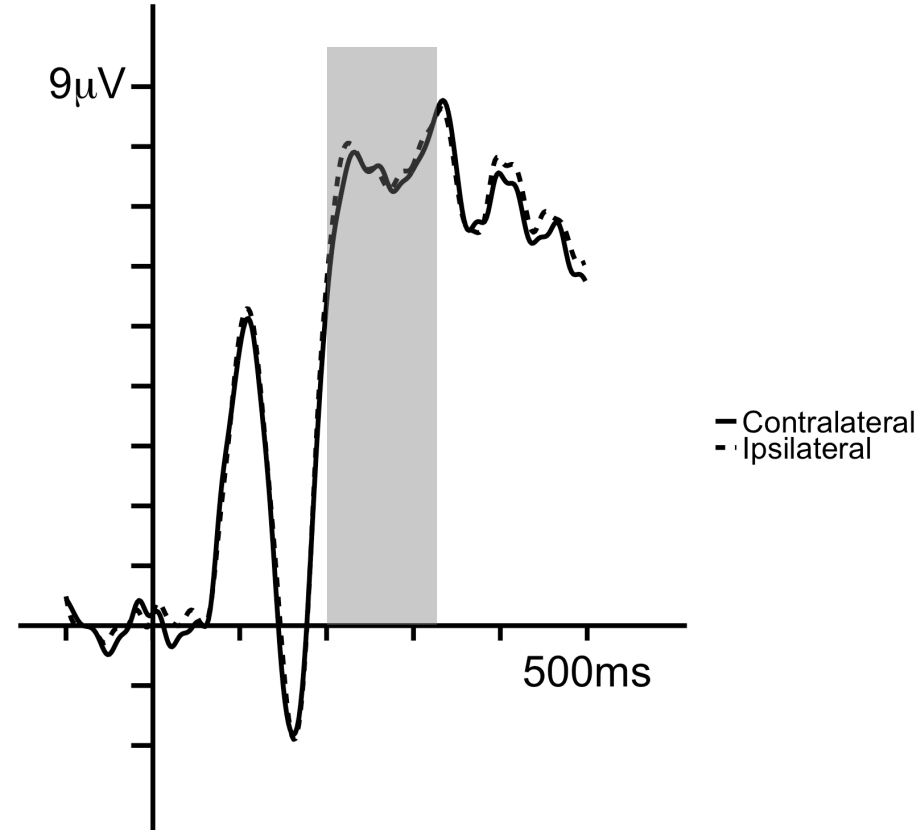


Disassemble Condition: Post-test

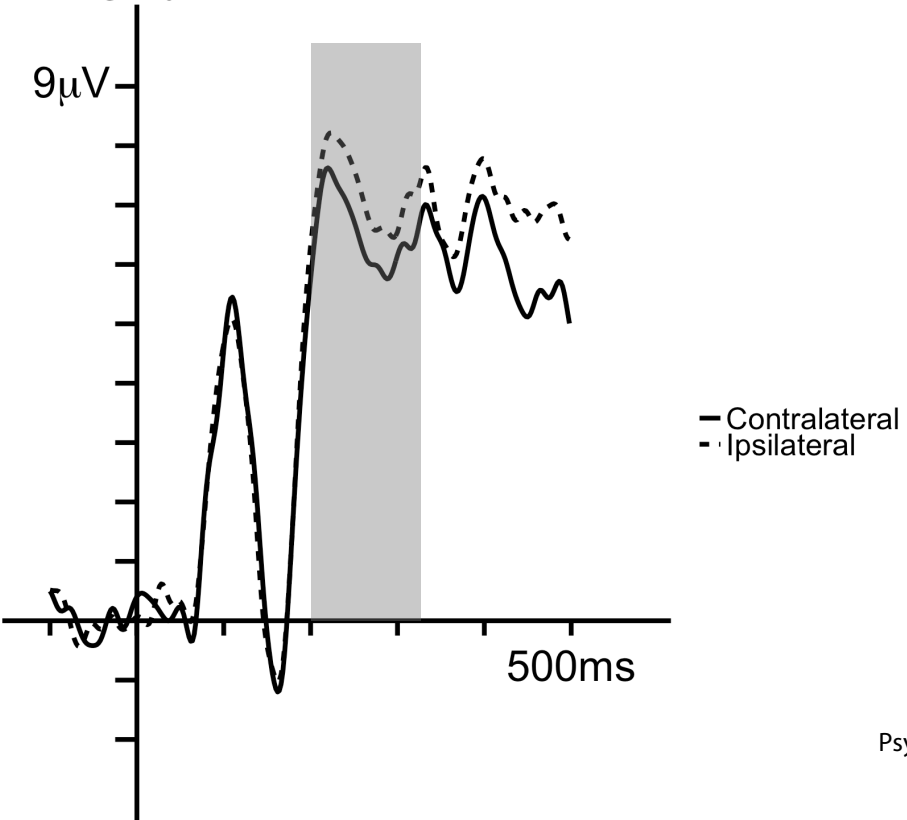
Exemplar Match trials



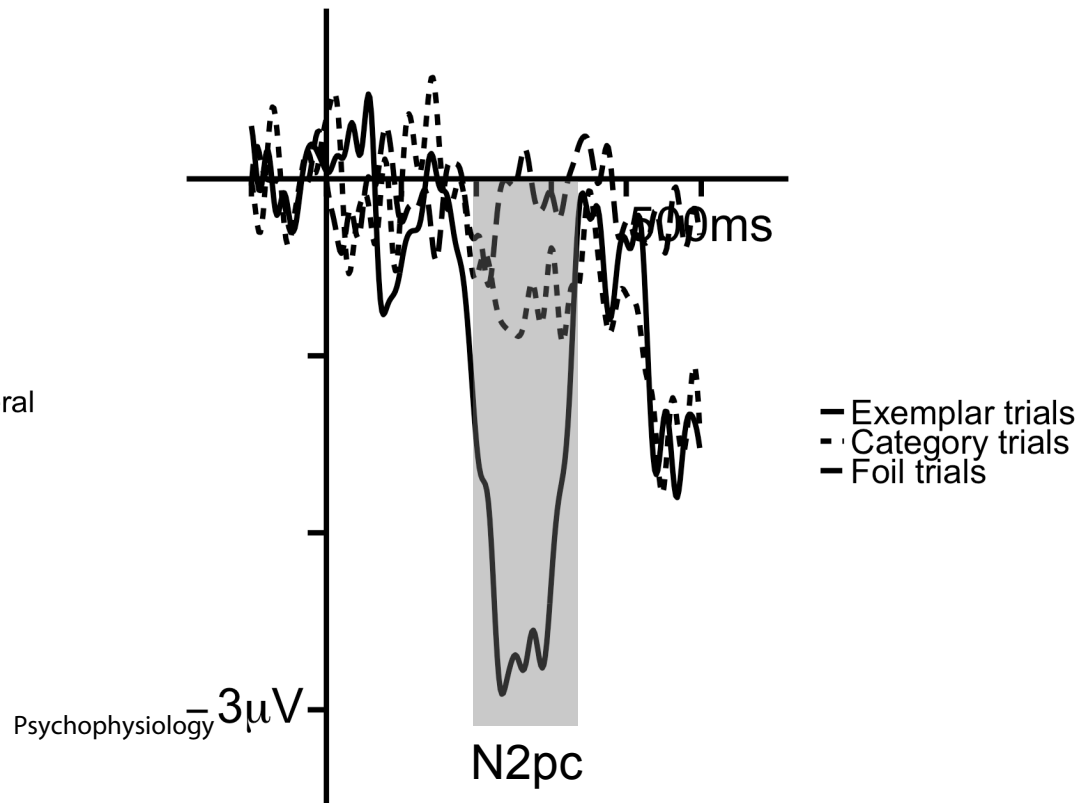
Foil Match trials



Category Match trials

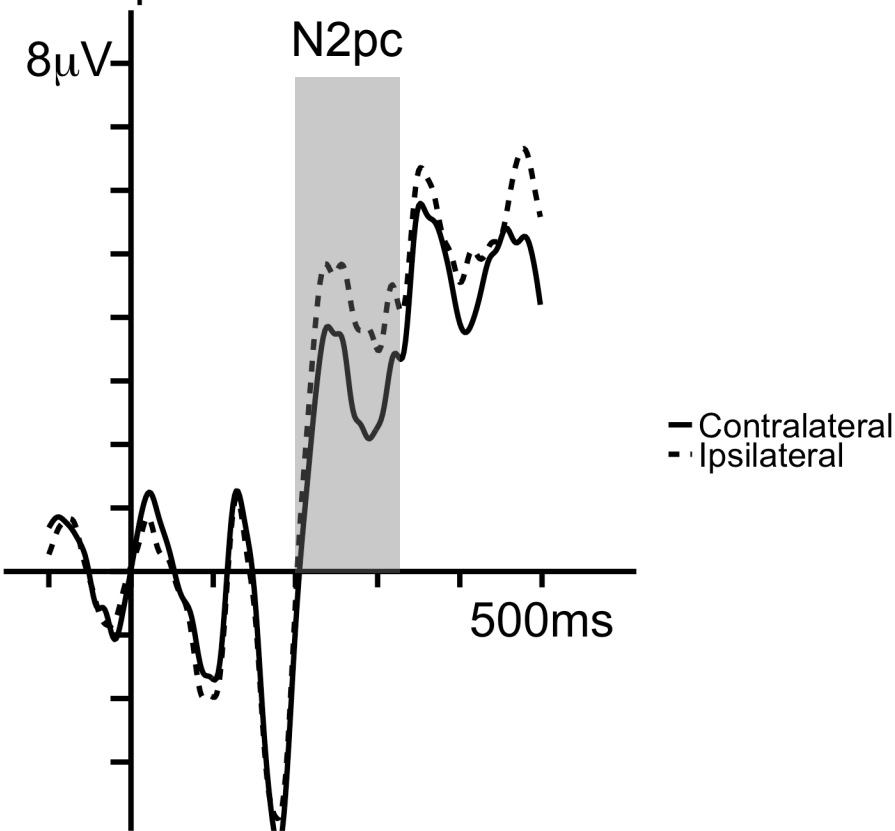


Difference waves

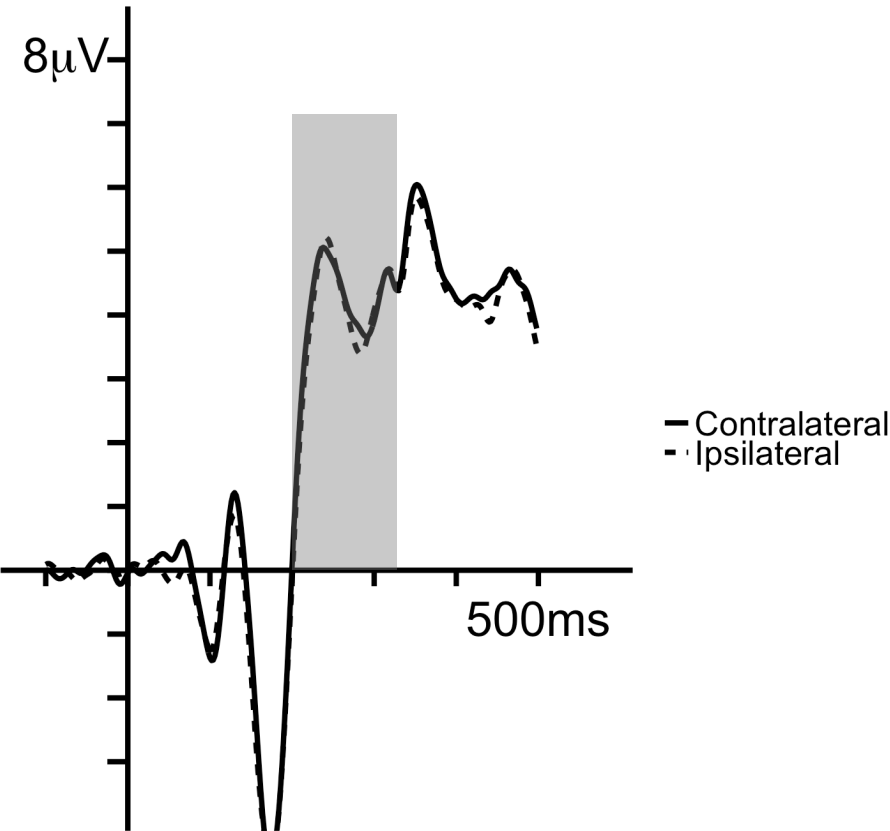


No Training Condition: Pre-test

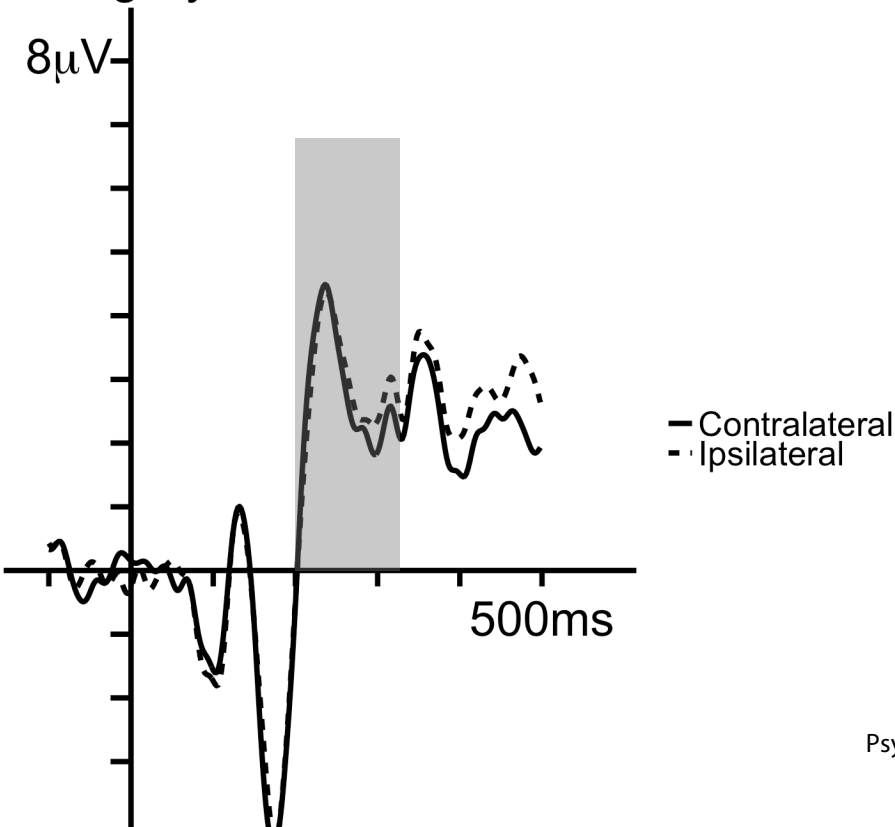
Exemplar Match trials



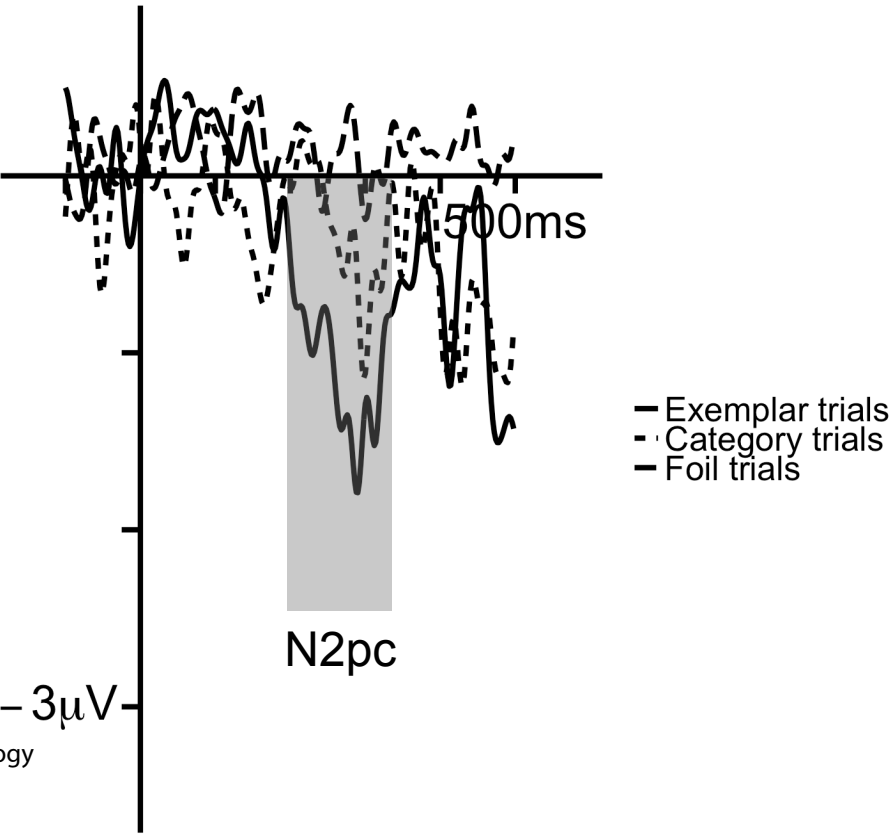
Foil trials



Category Match trials

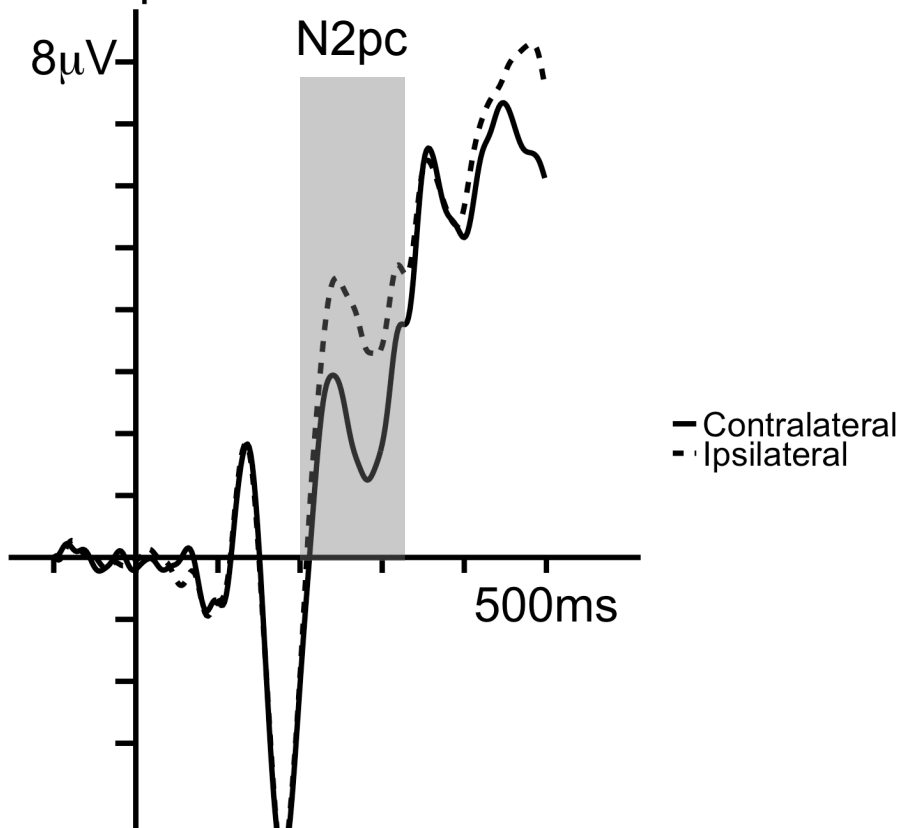


Difference waves

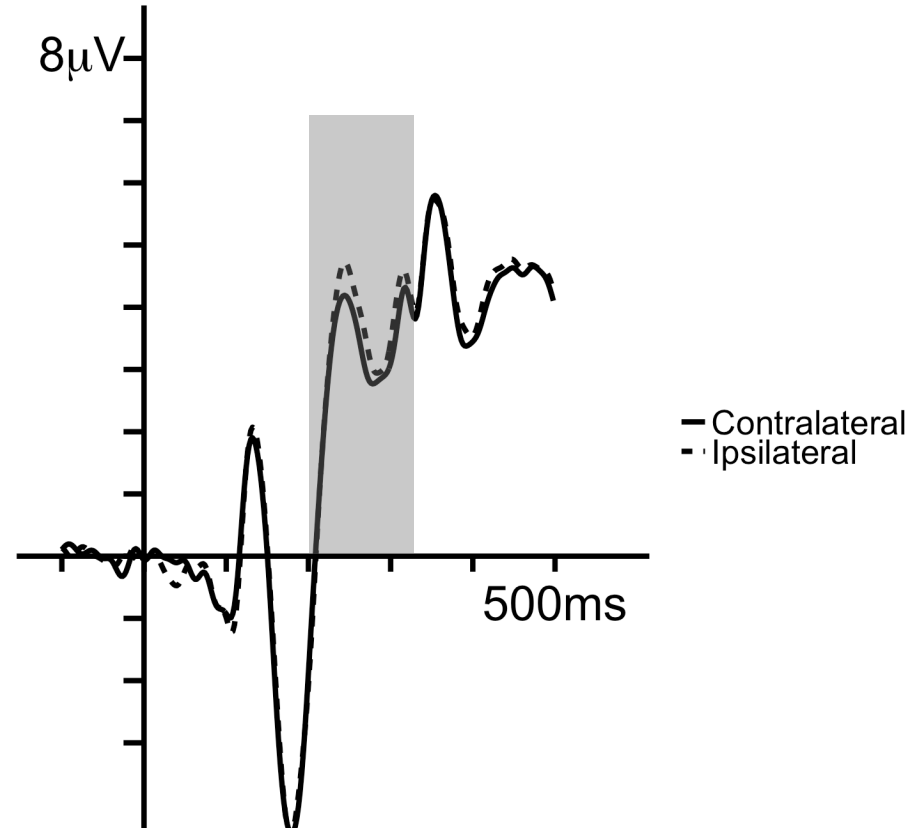


No Training Condition: Post-test

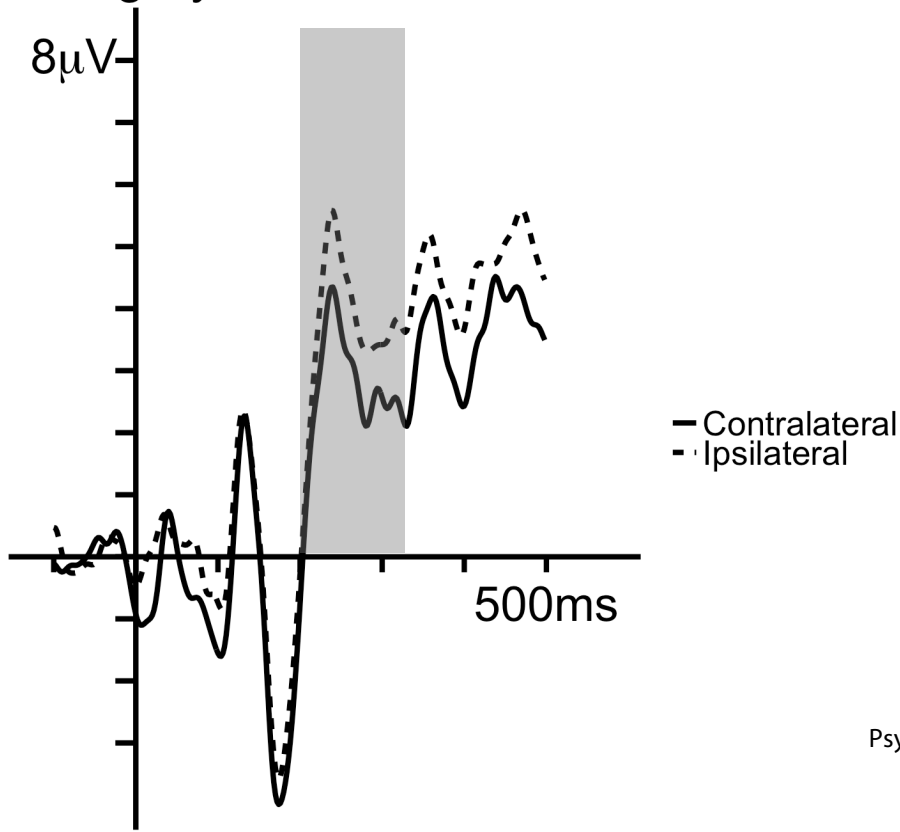
Exemplar Match trials



Foil trials



Category Match trials



Difference waves

