Rapid target selection of object categories based on verbs: Implications for language-categorization interactions

Geertje van Bergen[1], Monique Flecken[1], and Rachel Wu[2]*

[1]Max Planck Institute for Psycholinguistics

[2]Department of Psychology, UC Riverside

*Address correspondence to:

Rachel Wu

Department of Psychology

University of California, Riverside

900 University Avenue

Riverside, CA 92521

Rachel.Wu@ucr.edu

**********************IN PRESS, *PSYCHOPHYSIOLOGY*******************

Abstract

Although much is known about how nouns facilitate object categorization, very little is known about how verbs (e.g., posture verbs such as stand or lie) facilitate object categorization. Native Dutch speakers are a unique population to investigate this issue with because the configurational categories distinguished by 'staan' (to stand) and 'liggen' (to lie) are inherent in everyday Dutch language. Using an event-related potential (N2pc), four experiments demonstrate that selection of posture verb categories is rapid (between 220-320ms). The effect was attenuated, though present, when removing the perceptual distinction between categories. A similar attenuated effect was obtained in native English speakers, where the category distinction is less familiar, and when category labels were implicit for native Dutch speakers. Our results are among the first to demonstrate that category search based on verbs is rapid, although extensive linguistic experience and explicit labels may not be necessary to facilitate categorization.

Rapid target selection of verb categories: Implications for language-categorization interactions

Noun labels facilitate processing and grouping objects with similar perceptual features, such as cars with a similar shape (Connell, 2018; Gentner & Simms, 2011; Lupyan & Lewis, 2017; Plunkett et al. 2008; Wu, Mareschal, & Rakison, 2011). By triggering a mental representation of a canonical representation (i.e., prototype) of a particular category, labels can "jump-start" visual processes (Boutonnet & Lupyan, 2015). The majority of prior research on the role of language in object perception and categorization has focused on noun labels. Although verbs have been shown to facilitate learning how one object can relate to another (e.g., a cat chasing a mouse; Gentner, 2006), it is not clear how verbs might facilitate categorization. An important difference between noun and verb labels is that the former may allow the formation of clear mental representations of canonical category members (e.g., a canonical cat), although broad, higher-level categories (e.g., food) do not afford such precise mental representations. By contrast, verbs typically describe *relational* information, such as objects placed on their natural base relative to a ground-object (i.e., standing) for posture verbs. Thus, they may evoke "fuzzier" mental representations of a category, as multiple verbs can apply to the same item in different instances, and multiple objects can apply to the same verb.

Yet, within a familiar language context, it is clear whether an object belongs in a particular verb category (e.g., whether an object is standing or lying for native Dutch speakers). Native Dutch speakers, in contrast to native English speakers, are a unique population with which to investigate object categorization based on posture verbs. The distinction between "to stand" ("staan") vs. "to lie" ("liggen") is highly familiar to native Dutch speakers, as it is engrained in the language used for everyday speech (Ameka &

Levinson, 2007; Lemmens, 2002): Posture verbs are obligatory to describe the relation between a figure and a ground (i.e., an object on a surface). For example, "de sleutels liggen op tafel" directly translates to "the keys lie on the table." By contrast, in English, object configurations are typically described with the verb ‘to be’, rather than a posture verb, such as "the keys are on the table”. The distinction between "staan" and "liggen" can be perceptual in nature, where standing objects are often taller than they are wide (a standing bottle), and lying objects are wider than they are tall (a lying bottle). However, this is not the case in all instances. For example, a standing plate is wider than it is tall, whereas a lying ball is as tall as it is wide. The distinction in Dutch posture verbs is based on properties of the object itself: if the object has a natural base, then 'staan' can be used; if not, then only 'liggen' is applicable. In addition, the specific configuration of the object in relation to the ground is relevant for the distinction: for objects with a natural base (a bottle, or a plate), 'staan' is used when the object is resting on its base, whereas 'liggen' is used in other cases. Dutch posture verbs thus do not simply encode the horizontal vs vertical orientation of a single object; rather, they mark more complex object configurations.

One particularly useful way of investigating the robustness of mental category representations is category visual search. Category search is a fundamental, everyday task, ranging from searching for your keys in different orientations (i.e., a specific category of visual percepts) to searching for something to eat (i.e., a broad category of objects). Because top-down (i.e., goal-directed) visual search requires a mental representation of the target (e.g., keeping in mind what your keys look like when searching for your keys), category-based visual search paradigms provide unique information about the effects of categorization due to stronger or weaker mental representations of a category. EEG measures, particularly the N2pc event-related potential (ERP), provides information on the timing of these processes. The N2pc is the earliest and most reliable marker of target selection, emerging approximately 200 ms after stimulus onset (Eimer, 1996; Luck & Hillyard, 1994). Prior N2pc category search studies have shown that categories based on nouns (e.g., letters, numbers, faces, kitchen items, Nako et al., 2014a, 2014b; Wu et al., 2015), and even adjective categories (e.g., healthy vs. unhealthy food; Wu et al., 2017) can elicit reliable N2pc components. These studies have demonstrated that searching for categories, especially highly familiar noun categories, can elicit N2pc components that are as large as, and emerge as quickly as, searching for a specific item, which affords a clear mental representation. When the mental representation of the category becomes "fuzzier " because not all items in the category share similar perceptual features to conform to a canonical representation, such as with healthy vs. unhealthy foods, the N2pc component, although still present, becomes attenuated (e.g., Nako et al., 2014b; Wu et al., 2016, 2018a).

The present study investigated the mental representations of categories based on posture verbs: to stand vs. to lie. If such verb labels evoke mental representations similar to those of familiar noun labels, this finding would suggest that verb labels impact categorization and subsequent attentional processes related to the categories. Although perceptual categories elicit clear mental representations based on prototypes, demonstrating a robust N2pc component even after removing the perceptual nature of the distinction would provide strong evidence that verb labels can impact mental representations of categories, extending previous aforementioned work on the role of language in vision. By reducing prior linguistic experience with the categories (via testing non-Dutch speakers, i.e., native English speakers) or by investigating differences in effects between explicit vs. implicit category labels, we can study in detail the contribution of verb labels on categorization.

The present study demonstrates that searching for categories based on the verbs *staan* vs. *liggen* can elicit N2pc components in native Dutch speakers when the categories are labelled explicitly, and the objects conform to the prototypical perceptual distinction (standing

= taller, lying = wider; Experiments 1A and 1B). A reliable N2pc (albeit attenuated) was elicited even when removing the perceptual distinction (standing and lying categories both contained mostly wider than taller objects) for the explicitly labelled categories (Experiment 2). This attenuated effect also was found in English speakers when the categories were explicit (Experiment 3), demonstrating that there may not be an obvious benefit for category search in this task related to extensive prior linguistic experience with the distinction. Finally, an attenuated (but present) N2pc was elicited in Dutch speakers when the categories were implicit (i.e., not explicitly labelled), suggesting that, at least in familiar language contexts, explicit verb labels may not be necessary to elicit the category effects obtained (Experiment 4).

Experiment 1

Using either the same or different objects, Experiments 1A and 1B investigated whether categories based on posture verbs would elicit an N2pc during category search when the objects conformed to canonical representations of the categories.

Method

Participants

Following prior N2pc category search studies (e.g., Wu et al., 2016), for Experiment 1A, data from 16 native Dutch speakers (the pre-determined stopping point; 9 female, M = 22.8 years, SD = 2.61, range = 18-28 years) were included in the analyses. An additional six participants were excluded from the final analyses due to excessive eye movements (> 50% of trials excluded due to horizontal eye movements). Experiment 1B included another 16 Dutch native speakers (10 female, M = 21.4 years, SD = 2.28, range = 18-27 years). Data from an additional three participants were excluded from analyses due to excessive eye movements.

Participants in both experiments were tested at the Max Planck Institute for Psycholinguistics and provided informed consent. All had normal or corrected-to-normal vision and were paid €18 at the end of the experiment.

Stimuli

The stimuli that were in part adapted from van Bergen and Flecken (2017) were created using Blender, an open-source animation software (www.blender.org). The stimuli consisted of grey-scale images of household objects placed on a table (eight objects for Experiment 1A and 16 objects for Experiment 1B). For Experiment 1A, between categories, each object had two configurations: one configuration was standing (Figure 1, upper panel) and the other was lying (lower panel). Using identical objects across categories ensured that each object from the standing category was visually similar (except for its orientation) to its counterpart in the lying category. In Experiment 1B, each object occurred in only one configuration, belonging to either the standing or the lying category (Figure 2). All stimuli conformed to canonical representations of the categories (standing = tall, lying = wide). The images subtended 3.6° x 3.6° and were presented at 3.03° from the central fixation point on a 24-inch 60 Hz Dell monitor.

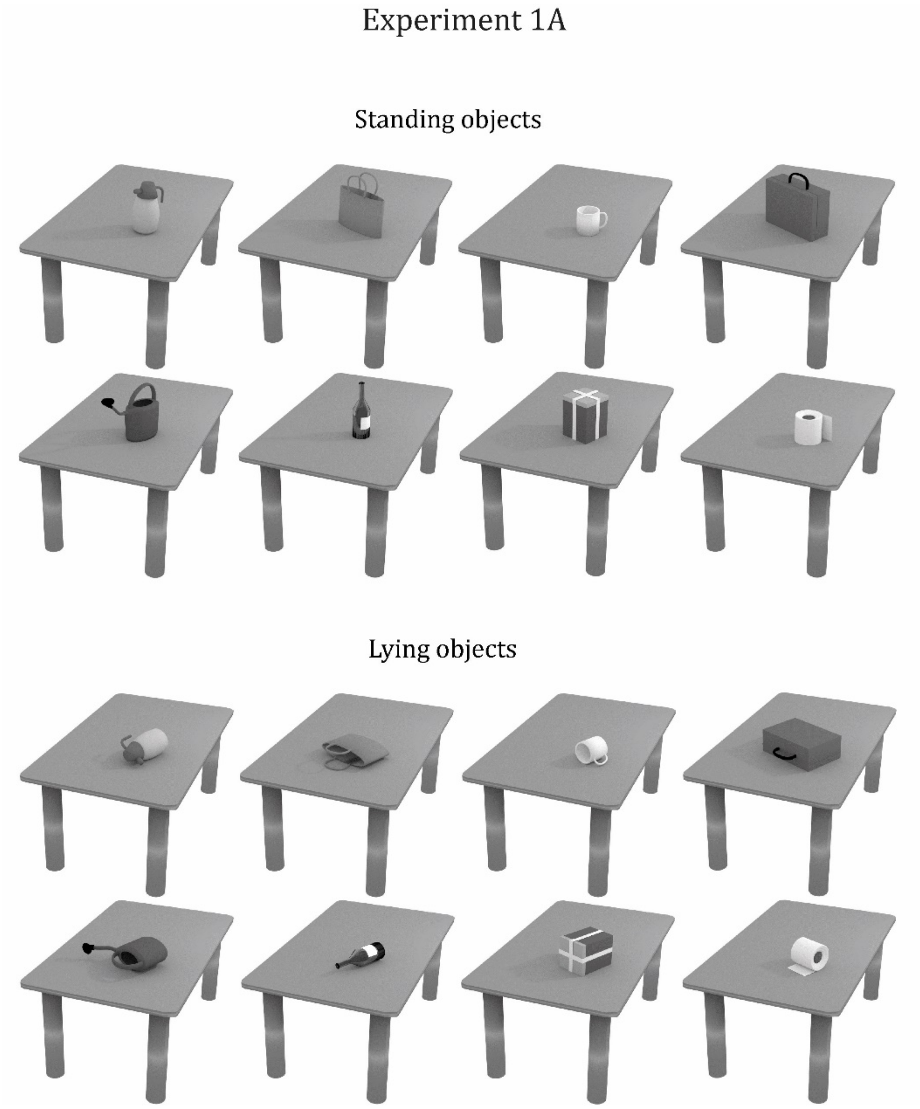

Figure 1. 'Standing' (top) and 'lying' (bottom) objects used as search stimuli in Experiment 1A.

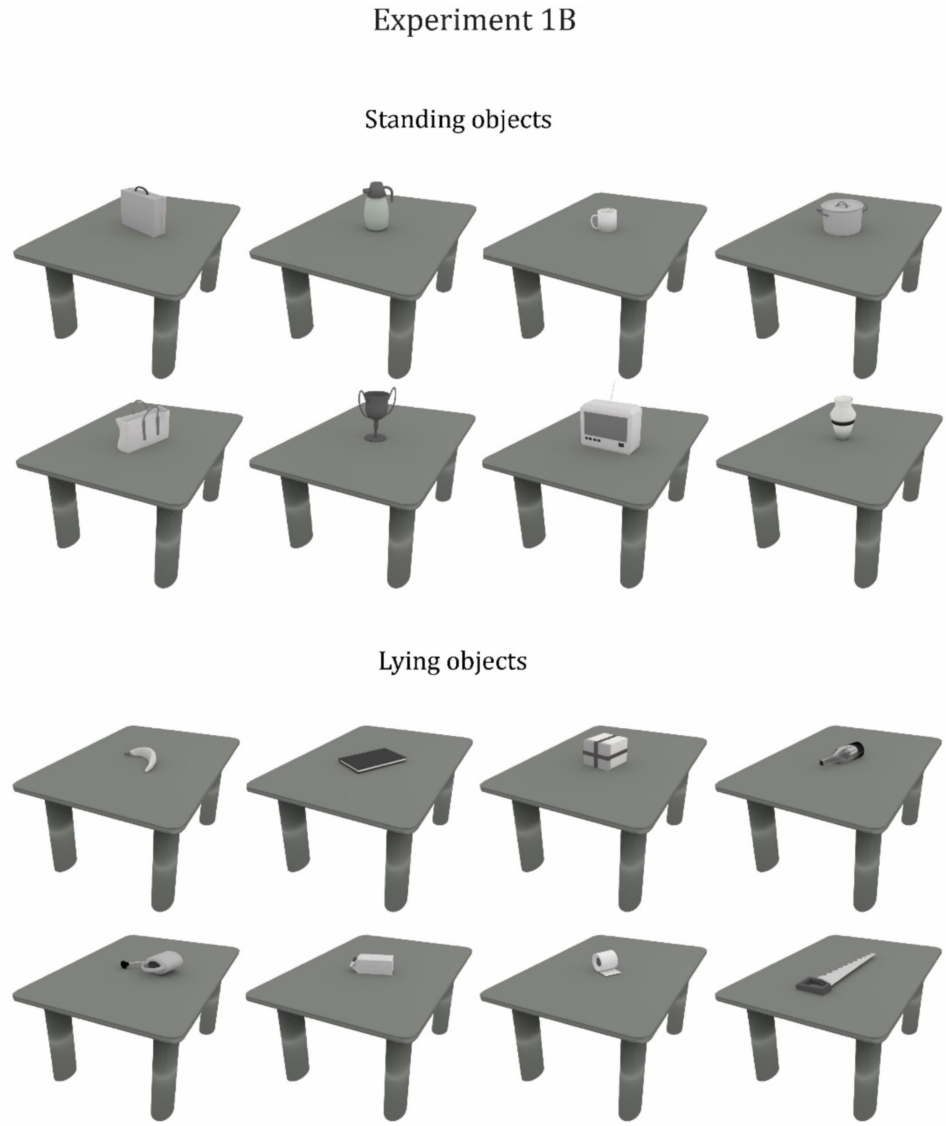

Figure 2. 'Standing' and 'lying' objects used as search stimuli in Experiment 1B.

Design and Procedure

Both Experiments included two category search tasks, one which involved searching for any standing object, and another for any lying object. In addition, an exemplar search task was included to confirm that the individual objects can elicit N2pc components (e.g., search for this specific object in this specific configuration; see Supplementary Materials). The task order for category and exemplar search was varied across participants. For example, one participant may have completed the tasks in the following order: exemplar search, standing category search, lying category search, whereas another may have completed the following order: lying category search, exemplar search, standing category search. The same stimuli were displayed in all conditions throughout the study for all participants. Therefore, only the instructions determined whether particular stimuli were the targets or distractors.

The category search tasks each included eight blocks, and each task included 224 Category Match trials and 224 No Category Match target trials. Category Match trials consisted of any object from the target category appearing on one side, while an object from the nontarget category appeared on the other side, and No Category Match trials only displayed objects from the nontarget category. Each of the eight blocks contained 28 Category Match trials and 28 No Category Match trials. At the beginning of the experimental session, participants were provided a printout with all of the objects presented simultaneously under their category label to help them complete the task. Participants were allowed to study the printouts as long as they wanted to, typically not more than a few minutes. The experimenter pointed to which item or items were the target on the printouts before each task. In addition, for the Exemplar search tasks, participants were shown the exact target image in the center of the screen before the start of each block until they advanced to the search trials. For the Category search tasks, participants were shown a black icon of a standing or a lying person before each block to indicate search for a standing or lying category, respectively. Participants searched for the same target for all eight blocks.

The search array presented two objects simultaneously on the left and right side of a black fixation dot on a white background for 200ms. This array was followed by a 1600ms inter-stimulus interval displaying only the white background and the fixation dot (Figure 3). As a constraint in N2pc studies, participants fixated the dot throughout the experiment and searched for targets via their peripheral vision. Participants also were instructed to press the right (green) button on a button box with their right index finger when a target was present, and the left (red) button with their left index finger when the target was absent. A target present response was required on half of the trials, and a target absent response on the other half. Participants had 1800 ms to respond to the array from the onset until the next array, and they were asked to respond as quickly and as accurately as possible. This design is similar to that from Wu et al. (2016, 2017, 2018a, 2018b). 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.

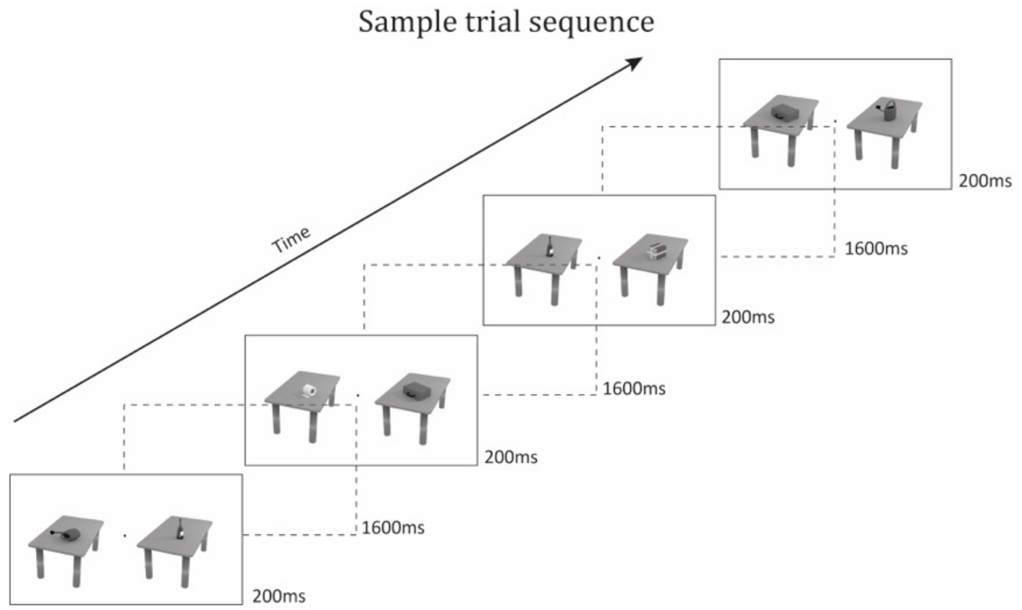

Figure 3. Example search arrays and trial sequence from the search tasks.

EEG recording, data preprocessing and analyses

The EEG data were recorded from 27 cap-mounted active electrodes placed according to the 10/20 system, and four active electrodes placed at the outer canthi of both eyes and above and below the center of the left eye to capture blinks and horizontal eye movements. Electrode impedances were kept below 20 kΩ. Signals were amplified using BrainAmp DC amplifiers with a band-pass filter between 0.01 and 150 Hz and digitized at a sampling frequency of 500 Hz. Data were recorded in reference to the left mastoid, and re-referenced offline to the average of the two mastoids. Bipolar VEOG and HEOG signals were calculated separately via the difference of the VEOG and HEOG channels, respectively.

We used a 100ms pre-stimulus baseline for epochs from -100ms to 500ms relative to the search array onset. For the artifact rejection criteria, we used horizontal EOG exceeding ± 25 μV, vertical EOG exceeding ± 60 μV, all other channels exceeding ± 80 μV. For all ERP analyses, only trials with a correct response from 100 ms to 1500 ms after stimulus onset were included. We retained 79% of all correct trials on average per participant after rejecting trials based on eye-movement artifacts. We used the 220-320 ms time window after search array onset to assess the mean N2pc amplitude at lateral posterior electrodes PO7 and PO8 (cf. Wu et al., 2015).

Results

ERP results

Only Category Match trials (rather than No Category Match) were included in the ERP analyses because the contralateral nature of the N2pc component requires identifying the location of a target in the analyses. The ERP results of all experiments are illustrated in Figures 4 and 5, and the mean N2pc amplitudes are listed in Table 1. A paired samples *t*-test between contralateral and ipsilateral mean amplitudes revealed a significant N2pc component for Category Match trials: Experiment 1A – $t(15) = 5.98$, $p < 0.001$, Experiment 1B – $t(15) = 4.98$, $p < 0.001$. Comparing the N2pc effect between Experiments 1A and 1B, an independent samples *t*-test on the difference waves did not reveal a significant difference, $t(30) = 1.06$, $p = 0.298$. The lack of a difference in the N2pc effect between Experiments 1A and 1B suggests that increased category complexity via additional objects does not affect attentional target selection. These results suggest that participants rapidly selected category targets possibly based on relational information accompanied with posture verb labels, although the obvious perceptual distinction (i.e., standing = tall, lying = wide) may have played a role.

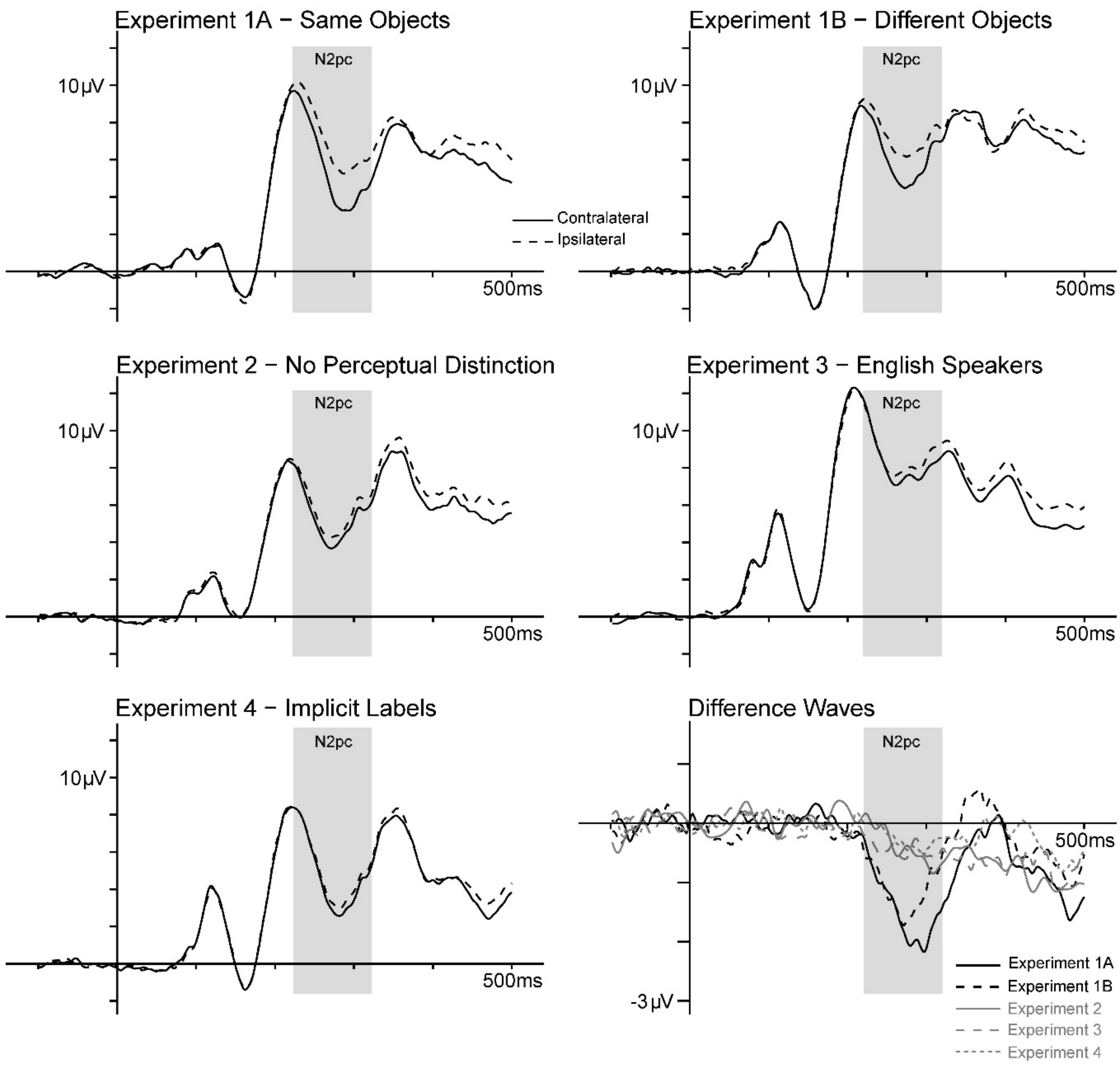

Figure 4. Grand average ERPs elicited by search arrays for category match trials at posterior electrodes PO7/8 contralateral and ipsilateral to a category target in Experiments 1 to 4. N2pc difference waveforms were obtained by subtracting ipsilateral from contralateral ERP waveforms.

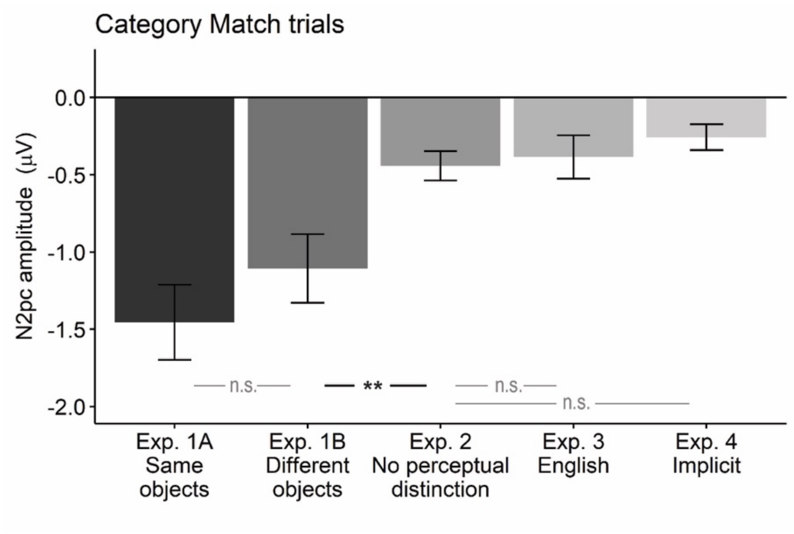

Figure 5. Mean N2pc amplitudes for Category Match trials from Experiments 1 to 4. A one-way ANOVA across experiments ($F(4,75) = 9.47$, $p < 0.001$, $\eta^2 = 0.34$) and follow-up *t*-tests revealed that Experiments 1A and 1B did not differ, and both were larger than Experiments 2, 3, and 4, which did not differ. Error bars represent ±1 SE.

Behavioral results

We compared participants' accuracy and (log-transformed) reaction times between Category Match trials and No Category Match trials separately for each experiment (Figure 6, Table 1). For Experiment 1A, accuracy for Category Match trials was significantly lower than for No Category Match trials, $t(15) = -2.59$, $p = 0.02$. By contrast, reaction times for Category Match trials were significantly faster for No Category Match trials, $t(15) = -4.46$, $p < 0.01$. For Experiment 1B, accuracy did not significantly differ between Category Match trials and No Category Match trials, $t(15) = 0.98$, $p = 0.34$, but reaction times for Category Match trials were faster than for No Category Match trials, $t(15) = -4.53$, $p < 0.01$. These results suggest that correctly selecting a category target may be largely similar in difficulty as correctly indicating that no target was present.

To investigate whether there were differences between experiments, we conducted two 2 (trial type) x 2 (experiment) omnibus ANOVAs, on for accuracy and one for reaction times.

Results from the accuracy analysis revealed no significant main effect of Experiment, $F(1,30) = 1.89$, $p = 0.18$, nor a significant interaction, $F(1,30) = 0.77$, $p = 0.39$. Similarly, for reaction times, we found no evidence for a main effect of Experiment, $F(1,30) = 1.03$, $p = 0.32$, nor for an interaction, $F(1,30) = 0.93$, $p = 0.34$. These results suggest that searching for a more varied set of objects may not be too impactful on behavioral outcomes during category search based on verb labels.

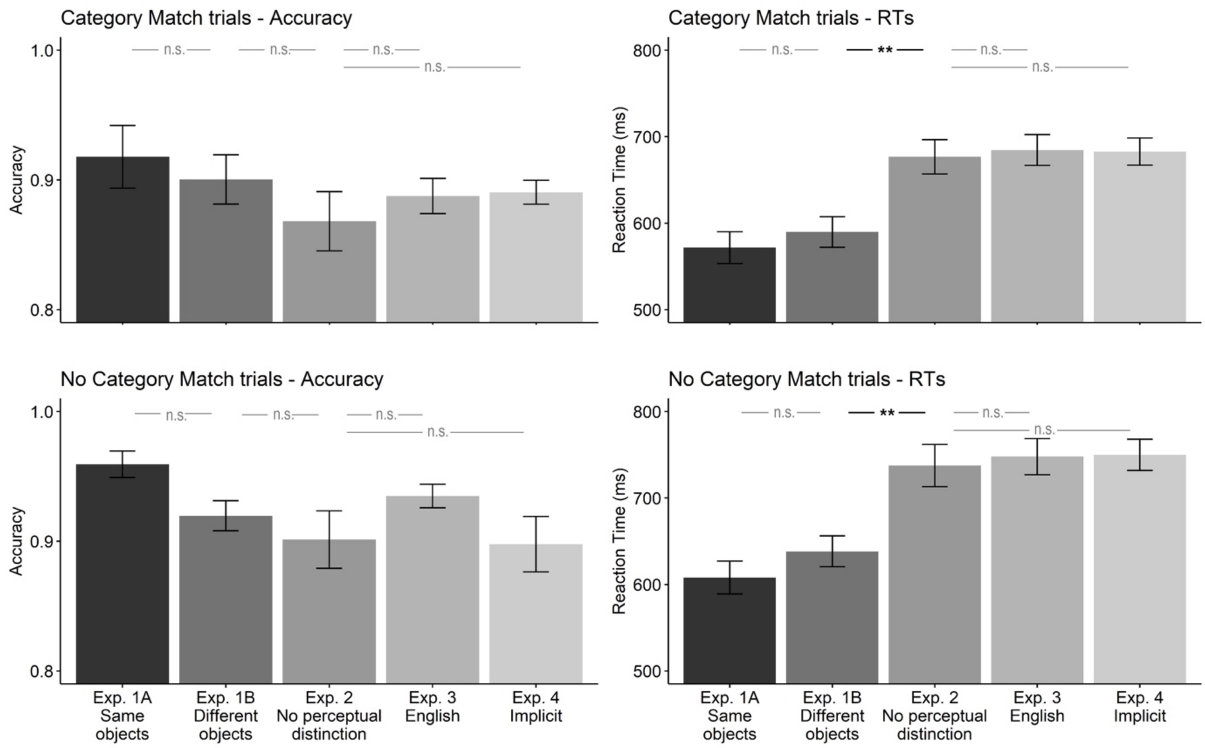

Figure 6. Accuracy (left panels) and response times (right panels) for Category Match trials (upper panels) and No Category Match trials (lower panels) in Experiments 1A-4. Error bars represent ±1 SE.

Table 1. Mean and SD values for EEG and behavioral measures.

		N2pc (µV)		**Accuracy**		**RT (ms)**	
		M	*SD*	*M*	*SD*	*M*	*SD*
Category Match	Exp. 1A	-1.45	0.97	0.92	0.10	572	74
	Exp. 1B	-1.11	0.88	0.90	0.08	589	71
	Exp. 2	-0.44	0.38	0.86	0.09	676	80
	Exp. 3	-0.38	0.56	0.89	0.05	684	71
	Exp. 4	-0.26	0.34	0.89	0.04	683	63
No Category Match	Exp. 1A	-	-	0.96	0.04	608	76
	Exp. 1B	-	-	0.92	0.05	638	71
	Exp. 2	-	-	0.90	0.09	737	98
	Exp. 3	-	-	0.94	0.04	748	83
	Exp. 4	-	-	0.90	0.09	750	72

Experiment 2

Given that the N2pc is highly affected by salience and pop-out effects, to evaluate the influence of verb concepts on the N2pc, and therefore categorization, we minimized the perceptual distinctions between the two categories. In this experiment, category membership was now determined only by the more complex concepts provided by the verbs, rather than perceptual distinctions.

Method

Participants

Sixteen Dutch native speakers (14 female, *M* = 22.1 years, *SD* = 3.97, range = 18 - 30 years) participated in Experiment 2. Participants were tested at the Max Planck Institute for Psycholinguistics. The data from an additional six participants were excluded from the final analyses due to excessive eye movements. All participants had normal or corrected-to-normal vision and were compensated €18.

Stimuli, Design, & Procedure

The stimuli consisted of 16 household objects placed on a table, eight in each category (Figure 7). The objects in both categories tended to be wider than they were tall, or similar in height and width (e.g., a mug), and each object from the standing category was visually similar to an object from the lying category (e.g., a standing printer vs. a lying suitcase). The design and procedure were identical to those from Experiments 1A and 1B, as were EEG recording and data preprocessing procedures. We retained 77% of all correct trials on average per participant.

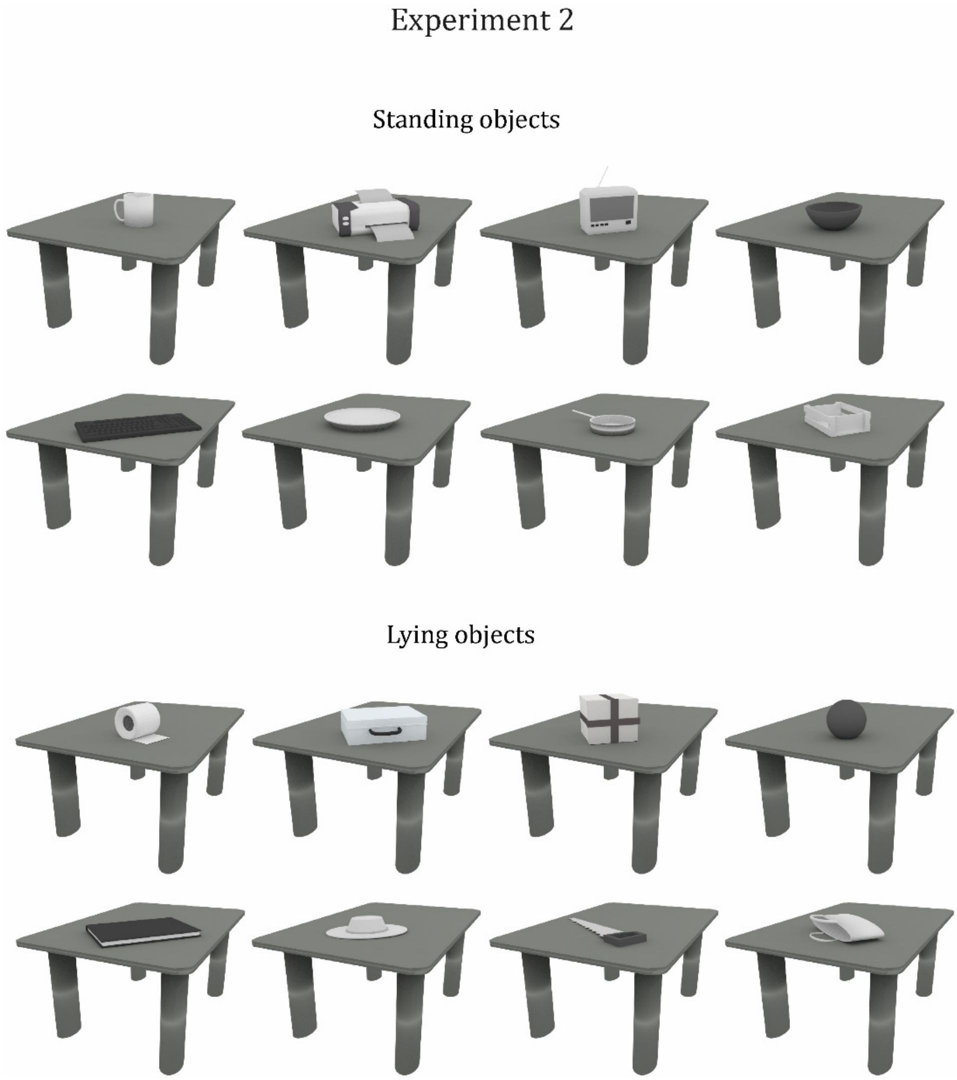

***Figure** 7*. 'Standing' and 'lying' objects used as search stimuli in Experiment 2.

Results

ERP results

A paired samples *t*-test between contralateral and ipsilateral mean amplitudes revealed a significant N2pc component for Category Match trials, $t(15) = 4.68$, $p < 0.001$. An independent samples *t*-test comparing the N2pc in Experiment 1B and Experiment 2 revealed that the N2pc was significantly smaller in Experiment 2 compared to Experiment 1B, $t(30) =$

2.75, $p = 0.01$. These data suggest that verb labels may facilitate categorization and category search, although the effect is attenuated when perceptual cues are removed.

Behavioral results

We found no difference between Category Match trials and No Category Match trials in accuracy, $t(15) = -1.80$, $p = 0.09$, but reaction times were significantly faster for Category Match trials than for No Category Match trials, $t(15) = -9.10$, $p < 0.001$ (Figure 6). An omnibus analysis showed that accuracy was similar across Experiments 1B and 2 (all $F < 1$). An omnibus analysis of reaction times revealed a significant main effect of Experiment, $F(1,30) = 11.06$, $p < 0.01$, but no significant interaction, $F(1,30) = 0.09$, $p = 0.77$, showing that reaction times were slower in Experiment 2 compared to those from Experiment 1B for both Category Match and No Category Match trials. These results demonstrate that when removing perceptual cues from a category search task, the task becomes somewhat more difficult.

Experiment 3

Experiment 3, with native English speakers, investigated whether the N2pc elicited from the first two experiments were due to the Dutch participants' extensive experience with the verbal categories. If linguistic experience does not play a crucial role, then we would find a similarly large N2pc between Experiments 2 and 3. To evaluate the explicit influence of verb concepts on categorization in native English speakers, we used the stimuli from Experiment 2, which minimized the perceptual distinctions between the two categories.

Method

Participants

Data from sixteen English native speakers (11 female, *M* = 22.2 years, *SD* = 4.68, range = 18 - 35 years) were analyzed; six participants were excluded from the final analyses due to excessive eye movements. Participants were tested at Brain and Cognitive Sciences at the University of Rochester. All had normal or corrected-to-normal vision and were paid $25 at the end of the experiment.

Stimuli, Design, & Procedure

The stimuli were identical to those from Experiment 2, as were the design and procedure. The EEG data were DC-recorded at standard positions of the extended 10–20 system (500-Hz sampling rate, 40-Hz low-pass filter) using 32 electrodes. The EEG data were re-referenced offline to the averaged earlobes. Data preprocessing and analyses were identical to the procedures from Experiments 1 and 2. We retained 72% of all correct trials on average per participant.

Results

ERP results

A paired samples *t*-test between contralateral and ipsilateral mean amplitudes revealed a significant N2pc component for Category Match trials, $t(15) = 2.74$, $p = 0.02$ (Figures 4-5). In other words, Category Match trials elicited a significant N2pc in native English native speakers, despite their lower level of experience with the linguistic categories. Comparing the N2pc for Category Match trials between Experiment 2 (Dutch participants) and Experiment 3 (English participants), a *t*-test revealed no significant difference, $t(30) = -0.34$, $p = 0.73$.

Behavioral results

Category Match trials and No Category Match trials differed significantly in accuracy, $t(15) = -4.2$, $p < 0.01$, and reaction times, $t(15) = -6.41$, $p < 0.01$, such that Category Match trials elicited faster but less accurate responses than No Category Match trials (Figure 6). Omnibus analyses showed no significant differences in accuracy or reaction times between Experiments 2 and 3 (all $F < 1.3$, $p > 0.25$). These results demonstrate that native English speakers with lower prior exposure to the verb categories were able to complete the task similarly to native Dutch speakers with a lifetime of experience.

This finding suggests that if the participant can complete the task with high accuracy, linguistic experience may not be as relevant for this task (similar to other studies demonstrating little to no effect of prior experience on category search tasks; Wu et al., 2017). However, studies have found effects on exemplar trials due to prior experience (see Supplementary Materials; e.g., Wu et al., 2017; 2018a; 2018b). Although a lifetime of experience with the linguistic categories may not impact category search, it may be the case that the English participants quickly learned the categories after they were labeled explicitly. To investigate whether category search for configurational categories with no clear perceptual distinction can still be successful *without* explicit labels, Experiment 4 included Dutch participants who were only presented with the categories from Experiment 2 without explicit verb labels (i.e., the categories were only referenced as Category A and Category B). Implicit linguistic distinctions have been shown to facilitate perceptual discrimination and categorization (e.g., color, space, Holmes & Regier, 2017; Malt & Majid, 2013; Regier & Kay, 2009). If labels played a substantial role in Experiment 2 with Dutch participants (for whom the distinction is familiar), and if they were quickly learned and applied in Experiment 3 with English participants, we would expect to see a smaller N2pc effect, if any, in Experiment 4.

Experiment 4

To evaluate the implicit influence of verb concepts on categorization, we used the stimuli from Experiment 2, which minimized the perceptual distinctions between the two categories.

Method

Participants

We analyzed data from sixteen Dutch native speakers (12 female, M = 22.1 years, SD = 3.97, range = 18 - 30 years). Two participants were excluded from final analyses due to excessive eye movements. All participants were tested at the Max Planck Institute for Psycholinguistics, had normal or corrected-to-normal vision, and were paid €18.

Stimuli, Design, & Procedure

All aspects were identical to those from Experiment 2, except that the categories were not explicitly labelled, but rather labelled as Category A and Category B. Per participant, we retained 78% of all correct trials on average.

Results

ERP results

A t-test revealed a significant N2pc component for Category Match trials, $t(15) = 3.05$, $p = 0.01$ (Figures 4 and 5). Comparing the N2pc on Category Match trials between Experiment 2 (explicit) and Experiment 4 (implicit) revealed no significant difference, $t(30) = -1.47$, $p = 0.15$, suggesting that explicit labels may not be necessary to elicit a category N2pc, at least in familiar language contexts.

Behavioral results

Category Match trials and No Category Match trials did not differ significantly in accuracy, $t(15) = -0.39$, $p = 0.70$, but responses to Category Match trials were faster than responses to No Category Match trials, $t(15) = -8.41$, $p < 0.01$ (Figure 6). Omnibus analyses showed that accuracy and reaction times were similar across both Experiments 2 and 4 (all $F < 1$). When asking participants at the end of the experimental session what strategy they used to remember which object belonged to the categories, not a single participant mentioned the objects' orientation. These results suggest that explicit labels do not provide additional benefits to category search in familiar language contexts.

General Discussion

Four experiments used event-related potential (ERP) and behavioral measures to investigate category target selection based on posture verbs (e.g., search for any standing object) in native Dutch speakers. In Experiment 1A and Experiment 1B, when participants were allowed to rely on canonical object *configurations* associated with the verbs, target selection was rapid, within 220-320 ms. This time window is similar to that for searching for a specific object (e.g., a standing vase; Nako et al., 2014b). Experiment 2 reduced participants' ability to rely on such perceptual features by including only objects that were wider than they were tall or similar in width and height in both categories. Results from Experiment 2 revealed an N2pc in the same time window, although the effect was attenuated compared to the N2pc from Experiment 1B.

Two subsequent experiments investigated the role of language experience (native Dutch speakers vs. native English speakers; Experiment 3) and explicit vs. implicit labels for Dutch speakers (Experiment 4). Due to the nature of the English language, English speakers were hypothesized to have less experience categorizing objects based on posture verbs. Results from Experiment 3 revealed an attenuated but reliable N2pc for English speakers, similar to the N2pc for Dutch speakers from Experiment 2 (using identical stimuli and tasks). Experiment 4 investigated with Dutch speakers the impact of explicit vs. implicit category verb labels. When labels were implicit, we found a small but reliable N2pc that was similar to the N2pc from Experiment 2 (where labels were explicit).

The finding that verb categories based on object *configurations* can lead to efficient category search is noteworthy (i.e., a reliable N2pc in all of the experiments), albeit the effects are stronger with perceptually reliable distinctions between categories. Thus far, event-related potential studies have only demonstrated that categories distinguished by object *types* via noun or adjective labels (rather than object relations via verb labels) can be selected by approximately 200ms (e.g., letters, numbers, faces, kitchen items, Nako et al., 2014a, 2014b; Wu et al., 2015; healthy vs. unhealthy food; Wu et al., 2017). The results from the present studies are among the first to show that configurational information can be useful for categorization and category search.

The results from Experiments 2-4 allow us to better understand the role of verb labels for categorization and category search, although some questions still remain. In Experiment 2, perceptual cues were removed, resulting in object configurations that do not conform to the canonical perceptual features associated with the verbs. The reliable N2pc in Experiment 2 may indicate that verb labels also facilitate category search for less canonical object configurations. The presence of an N2pc in Experiment 3, then, would suggest that English participants had quickly learned the categories after they were labeled explicitly. However, prior studies have shown that even advanced English learners of Dutch (in contrast with Dutch native speakers) do not use the configurational information encoded in posture verbs to guide their visual attention to potential sentence continuations in incremental language comprehension (van Bergen & Flecken, 2017). Moreover, the category N2pc in Experiment 4 would imply that the linguistic categories in Dutch participants are so deeply entrenched that they elicit an N2pc even in the absence of an explicit label. However, none of the participants in Experiment 4 reported using object configuration as a strategy for categorization during the task, suggesting that even the strategy itself was implicit or non-existent. Therefore, it remains unclear how participants completed the task in Experiments 2-4, and it may be the case that participants used different strategies to complete the task in each experiment.

It may be possible that participants performed a hybrid search task, using both visual and memory search (e.g., Wolfe, 2012) without categorizing the items. Hybrid search is typically characterized by a linear increase in search times as the number of items in a search array increases and a logarithmic increase as the number of items in memory sets increase. Given that all memory and search array set sizes were held constant throughout all four experiments, any differences we find across experiments were likely not due to hybrid search per se. Our findings are more in line with Cunningham and Wolfe (2014), who suggest that potential targets are compared in parallel to items from categories in long-term memory, although the timing may depend on the nature of the category (e.g., the specificity of the items within the category). Moreover, we have shown in prior studies that eight objects in a memory set that are not grouped into a category do not elicit an N2pc during category search, even if reaction times do not differ much between items that are grouped versus not grouped (e.g., Wu et al., 2016; Wu et al., 2018). We therefore take the presence of a reliable N2pc in the present experiments as evidence that the participants performed the task categorically.

Future studies could further investigate the role of verbs on category search by comparing search based on familiar categories and search based on unfamiliar ad-hoc categories, such as having different lying *and* standing objects in both categories. However, unfamiliar ad-hoc categories are very difficult for participants to learn in one experimental setting, especially when it is not clear how the objects are connected (e.g., Wu et al., 2013). Therefore, comparing search for an unfamiliar ad-hoc category to search for a familiar verb category might overestimate the contribution of verbs on category search. Perhaps comparing search for noun versus verb categories would be informative, but such a comparison is beyond the scope of the present studies, which investigated the role of verbs on category search.

It is possible that showing the participants all of the objects within their respective categories prior to starting the experiment may have helped them complete the task. However, to maximize the amount of clean data during the experimental session, we aimed to minimize online learning by familiarizing the participants to the objects beforehand. Future studies where participants perform a category search without prior familiarization to the stimuli would be highly informative, even if the data are noisier. A study including novel images on every trial would contribute to determining the robustness of the category effect. In such studies, thousands of objects per category are required for one experimental session. A study at this scale indeed would be informative, and is now justified given our positive findings in the present studies.

Another possible explanation for our findings is that verb labels help activate typical category members, rather than atypical category members, a view advocated by the *label-feedback hypothesis* (Lupyan & Thompson-Schill, 2012). Labels are especially useful cues for perceptual processing, precisely because their function is to generalize across individual exemplars, and activate a categorical and prototypical representation of a percept. Under this interpretation, the posture verbs accompanying the categories in Experiments 1A and 1B highlighted the canonical configuration distinction between standing and lying, and thereby enhanced N2pc effects relative to Experiments 2-4. The similar N2pc effects in Experiments 2 and 3 (native Dutch speakers versus native English speakers unfamiliar with the complexity of the linguistic categories in Dutch), as well as in Dutch participants not being presented explicitly with the verb labels (Experiment 4), suggest that categorization of atypical exemplars (object configurations without a clear perceptual distinction) on the basis of posture verbs does not afford as strong of a mental category representation compared to categorization of typical exemplars. Therefore, it is possible that the difference between the N2pc between Experiments 1A-1B and Experiments 2-4 may have been partially driven by inherent or canonical items in the natural category, in addition to the perceptual cues.

The present study adds to a growing literature demonstrating the value and constraints of categories on visual search. Although further empirical work is needed to clarify the role of verb labels on category search, our findings suggest the potential value, as well as limits, of verbs when performing a fundamental, everyday task.

Acknowledgements

We thank Gary Lupyan and Richard Aslin for providing helpful feedback on these studies, which were partly funded by the Netherlands Organisation for Scientific Research (NWO, Veni grant 275-89-022, awarded to G. van Bergen).

References

Ameka, F. K., & Levinson, S. C. (Eds.). (2007). The typology and semantics of locative predication: Posturals, positionals and other beasts [Special Issue]. *Linguistics*, *45*.

van Bergen, G., & Flecken, M. (2017). Putting things in new places: Linguistic experience modulates the predictive power of placement verb semantics. *Journal of Memory and Language,* 92, 26-42. doi:10.1016/j.jml.2016.05.003.

Boutonnet, B., & Lupyan, G. (2015). Words jump-start vision: A label advantage in object recognition. *Journal of Neuroscience, 24*, 9329-9335. doi: 10.1523/jneurosci.5111-14.2015.

Connell, L. (2018). What have labels ever done for us? The linguistic shortcut in conceptual processing. *Language, Cognition and Neuroscience*. Advanced online publication. doi: 10.1080/23273798.2018.1471512.

Cunningham, C. A., & Wolfe, J. M. (2014). The role of object categories in hybrid visual and memory search. *Journal of Experimental Psychology: General, 143*(4), 1585-1599. doi: 10.1037/a0036313

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology, 99*, 225–234.

Gentner, D. (2006). Why verbs are hard to learn. In K. Hirsh–Pasek, & R. Golinkoff, (Eds.), *Action Meets Word: How Children Learn Verbs* (pp. 544–564). Oxford: Oxford University Press.

Gentner, D., & Simms, N. (2011). Language and analogy in conceptual change. *Behavioral and Brain Sciences, 34*, 128–129. doi: 10.1017/S0140525X10002736.

Holmes, K.J., & Regier, T. (2017). Categorical perception beyond the basic level: The case of warm and cool colors. *Cognitive Science, 41,* 1135–1147. doi: 10.1111/cogs.12393.

Lemmens, M. (2002). The semantic network of Dutch posture verbs. In J. Newman (Ed.),*The linguistics of sitting, standing and lying* (pp. 103–140). Amsterdam/Philadelphia: John Benjamins.

Luck, S.J., & Hillyard, S.A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology, 31,* 291–308.

Lupyan, G., & Lewis, M. (2017). From words-as-mappings to words-as-cues: the role of language in semantic knowledge. *Language, Cognition and Neuroscience*. Advanced online publication. doi:10.1080/23273798.2017.1404114.

Lupyan, G., & Thompson–Schill, S.L. (2012). The evocative power of words: activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology: General, 141,* 170–186. doi: 10.1037/a0024904.

Malt, B.C., & Majid, A. (2013). How thought is mapped into words. *Wiley Interdisciplinary Reviews: Cognitive Science, 4*, 583–597. doi: 10.1002/wcs.1251.

Nako, R., Wu, R., & Eimer, M. (2014a). Rapid guidance of visual search by object categories. *Journal of Experimental Psychology: Human Perception and Performance, 40*, 50–60. doi: 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*, 1283–1288. doi: 10.1037/a0036885.

Regier, T. & Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends in Cognitive Science, 13,* 439–446. doi: 10.1016/j.tics.2009.07.001

Regier T., & Xu, Y. (2017). The Sapir–Whorf hypothesis and inference under uncertainty. *Wiley Interdisciplinary Reviews: Cognitive Science, 8*, e1440. doi: 10.1002/wcs.1440.

Wolfe, J. M. (2012). Saved by a log: how do humans perform hybrid visual and memory search? *Psychological Science, 23*(7), 698-703. doi: 10.1177/0956797612443968.

Wu, R., Mareschal, D., & Rakison, D.H. (2011). Attention to multiple cues during spontaneous object labeling. *Infancy, 16*, 545–556. doi: 10.1111/j.1532–7078.2010.00061.x.

Wu, R., McGee, B., Echiverri, C., & Zinszer, B.D. (2018a). Prior knowledge of category size impacts visual search. *Psychophysiology, 55,* e13075. doi: 10.1111/psyp.13075.

Wu, R., McGee, B., Rubenstein, M., Pruitt, Z., Cheung, O.S., & Aslin, R.N. (2018b). Emergence of the benefits and costs of grouping for visual search. *Psychophysiology, 55*, e13087. doi: 10.1111/psyp.13087.

Wu, R., Nako, R., Band, J., Pizzuto, J., Ghoreishi, Y., Scerif, G., & Aslin, R.N. (2015). Rapid selection of non-native stimuli despite perceptual narrowing. *Journal of Cognitive Neuroscience, 27,* 2299–2307. doi: 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, & Psychophysics, 78,* 749–776. doi: 10.3758/s13414-015-1039-6.

Wu, R., Pruitt, Z., Zinszer, B., & Cheung, O. (2017). Increased experience amplifies the activation of task-irrelevant category representations. *Attention, Perception, & Psychophysics*, 79, 522–532. https://doi.org/10.3758/s13414-016-1254-9.

Wu, R., Scerif, G., Aslin, R., 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,* 719 –729. doi: 10.1162/jocn_a_00352.

SUPPLEMENTARY MATERIALS

Exemplar search task

Design

The exemplar search task included eight blocks, and contained 224 Exemplar Match trials, 224 Foil trials, and 48 No Exemplar Match trials. Category Match trials consisted of a specific target (e.g., a standing suitcase) appearing on one side, while an object from the nontarget category (a lying object) appeared on the other side. Foil trials consisted of an object from the target's category (e.g., a standing bottle) and an object form the nontarget category (a lying object). No Exemplar Match trials displayed objects from the nontarget category only (Figure A). Each of the eight blocks contained 28 Exemplar Match trials, 28 Foil trials, and six No Exemplar Match trials.

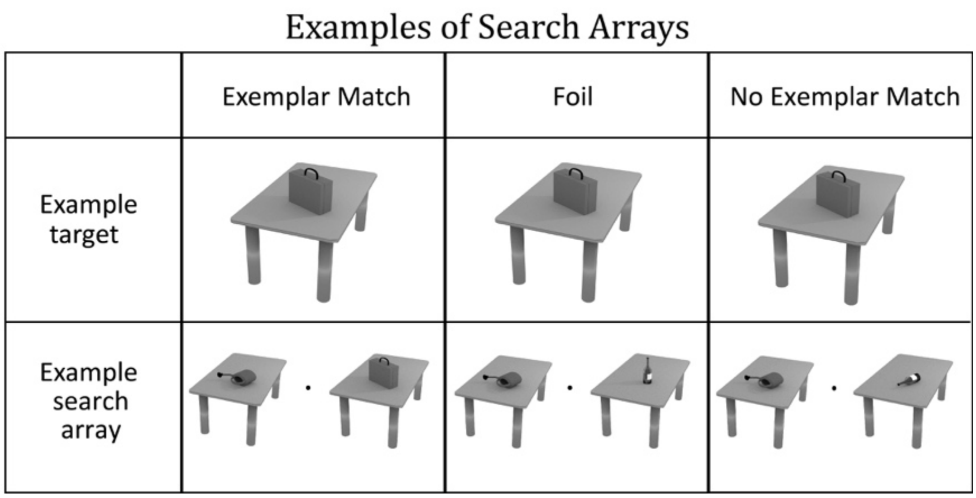

Figure A. Examples of search arrays for the Exemplar search task.

ERP results

The ERP results of the exemplar search task in all experiments are presented in Figures B-D. Assessing the presence of an N2pc component in each trial type across experiments, a 2 (laterality) × 2 (trial type) x 5 (experiment) ANOVA revealed a significant main effect of laterality, $F(1, 75) = 19.04$, $p < 0.001$, $\eta^2 = .002$, and a significant interaction between laterality and trial type, $F(1, 75) = 21.04$, $p < 0.001$, $\eta^2 = .001$. A paired samples *t*-test between contralateral and ipsilateral mean amplitudes revealed a significant N2pc component for Exemplar Match trials, $t(79) = 5.22$, $p < 0.01$, but not for Foil trials, $t(79) = 0.40$, $p = 0.69$. We found a marginally significant interaction between trial type and experiment, $F(4, 75) = 2.28$, $p = 0.07$, $\eta^2 = .002$, but no evidence for a three-way interaction, $F(4, 75) = 0.86$, $p = 0.49$, suggesting that N2pc differences between Exemplar trials and Foil trials were similar across all experiments.

Since we did not find an N2pc for Foil trials in any of the experiments, we only compared the N2pc for Exemplar Match trials across experiments. A one-way ANOVA of the difference waves revealed a marginally significant effect of Experiment, $F(4,75) = 2.21$, $p = 0.07$, $\eta^2 = .11$. Planned independent *t*-tests (corrected $\alpha = 0.0125$) between experiments showed that the N2pc for Exemplar Match trials was marginally smaller in Experiment 1B compared to Experiment 1A, $t(30) = -2.27$, $p = 0.03$, suggesting that visual search is more difficult when searching for exemplars if the experiment contains more target objects. There were no significant differences between Experiments 1B and 2, Experiments 2 and 3, or Experiments 2 and 4, $|t|(30) < 1.35$, $p > 0.18$, confirming that the other experimental manipulations did not affect attentional target selection efficiency.

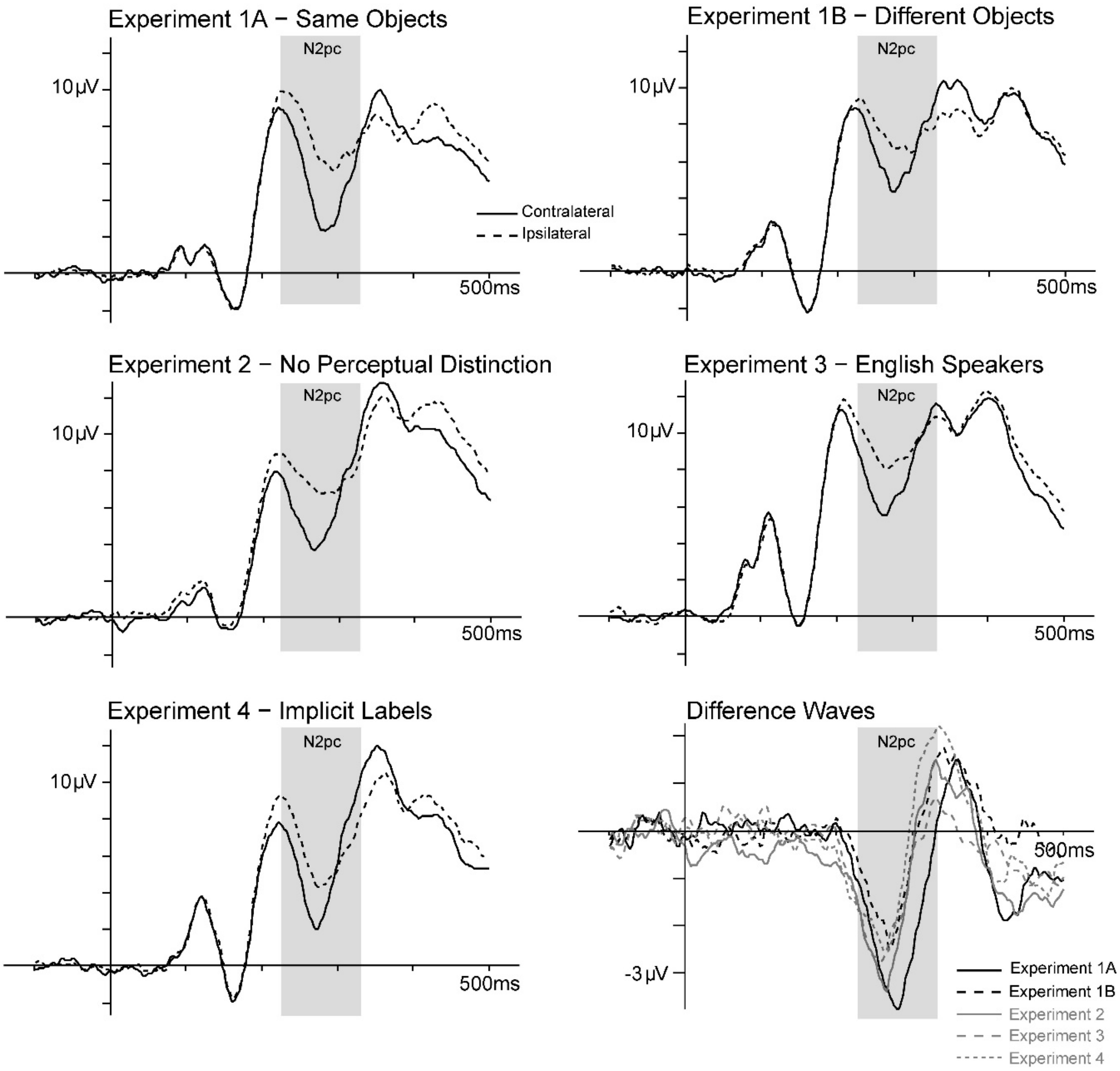

Figure B. Grand average ERPs elicited by search arrays for exemplar match trials at posterior electrodes PO7/8 contralateral and ipsilateral (and difference waves) to the exemplar target in Experiments 1 to 4.

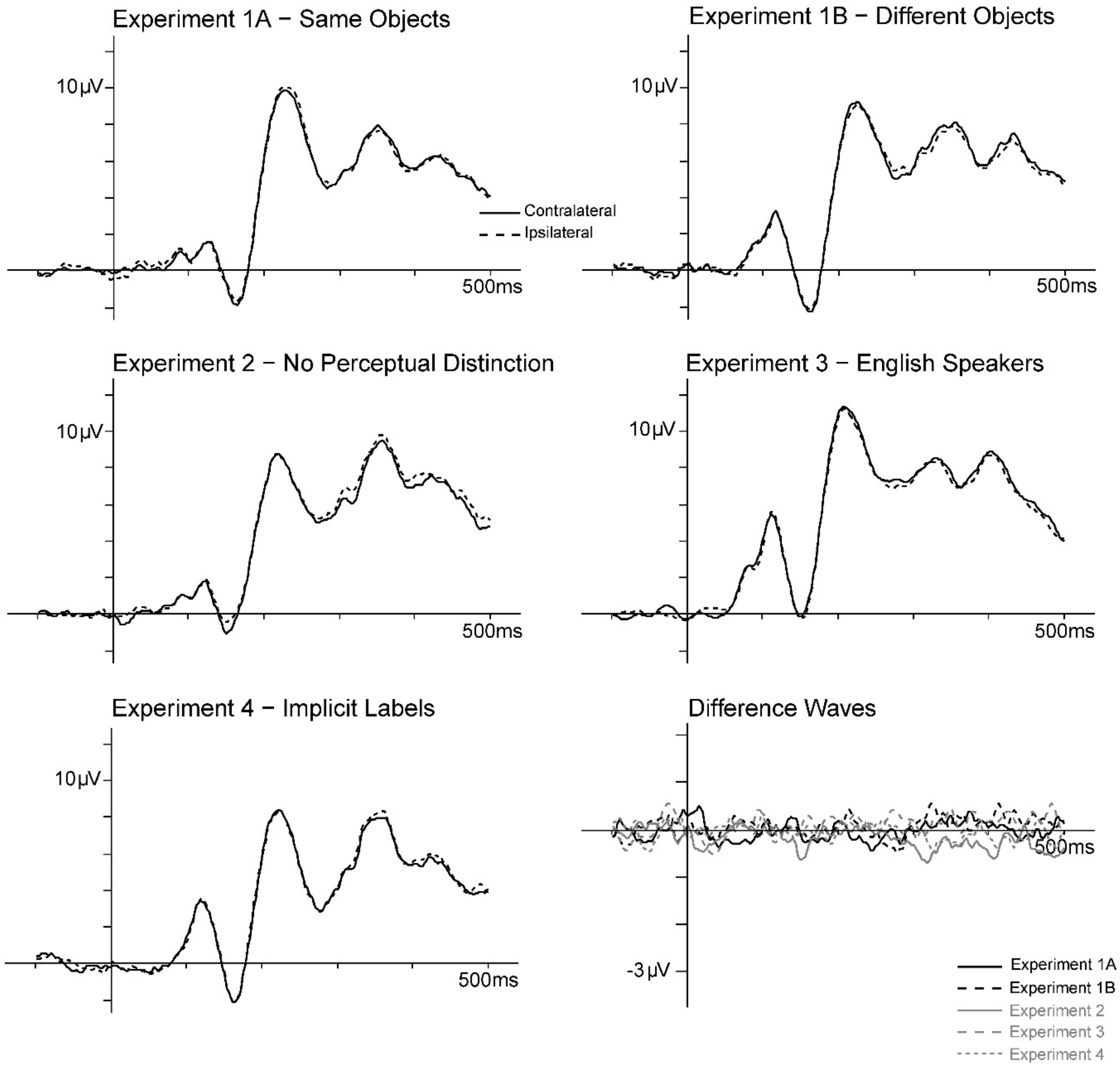

Figure C. Grand average ERPs elicited by search arrays for foil trials at posterior electrodes PO7/8 contralateral and ipsilateral (and difference waves) to an object from the target category in Experiments 1 to 4.

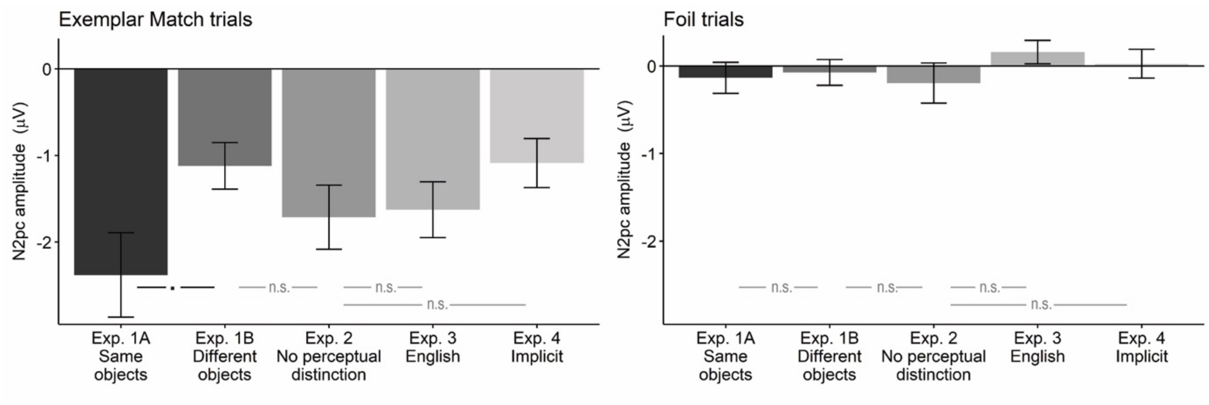

Figure D. Mean N2pc amplitudes for exemplar match trials (left panel) and foil trials (right panel) from Experiments 1 to 4. Error bars represent ±1 SE.

Behavioral results

The behavioral results of the exemplar search task are presented in Figure E. A 3 (Trial Type) x 5 (Experiment) omnibus ANOVA on accuracy revealed a main effect of trial type, $F(2,150) = 8.93$, $p < 0.001$, $\eta^2 = .06$; we found no evidence for differences based on Experiment, $F(4,75) = 0.79$, $p = 0.54$, nor a Trial Type x Experiment interaction, $F(8,150) = 0.75$, $p = 0.62$. Follow-up pairwise Bonferroni-corrected comparisons (adjusted $\alpha = 0.017$) showed that accuracy for No Exemplar Match trials was significantly higher than for Exemplar Match trials, $t(79) = 3.87$, $p < 0.001$, and for Foil trials $t(79) = 3.96$, $p < 0.001$. There was no difference in accuracy between Exemplar Match trials and Foil trials, $t(79) = 0.95$, $p = 0.34$.

An omnibus analysis on reaction times showed a small effect of Trial Type, $F(2,150) = 4.47$, $p = 0.02$, $\eta^2 = .002$; there was no significant main effect of Experiment, $F(4,75) = 0.35$, $p = 0.85$, nor an interaction between Trial Type and Experiment, $F\ (8,150) = 1.41$, $p = 0.22$. Pairwise *t*-tests (corrected $\alpha = 0.017$) revealed that response times for Exemplar Match trials were marginally faster than for No Exemplar Match trials, $t(79) = -2.42$, $p = 0.02$. There were no significant differences in reaction times between Exemplar Match trials and Foil trials $t(79) = -1.84$, $p = 0.07$, nor between Foil trials and No Exemplar trials, $t(79) = -1.43$, $p = 0.16$. Taken together, these findings confirm that our experimental manipulations did not affect participants' behavior on the exemplar search task.

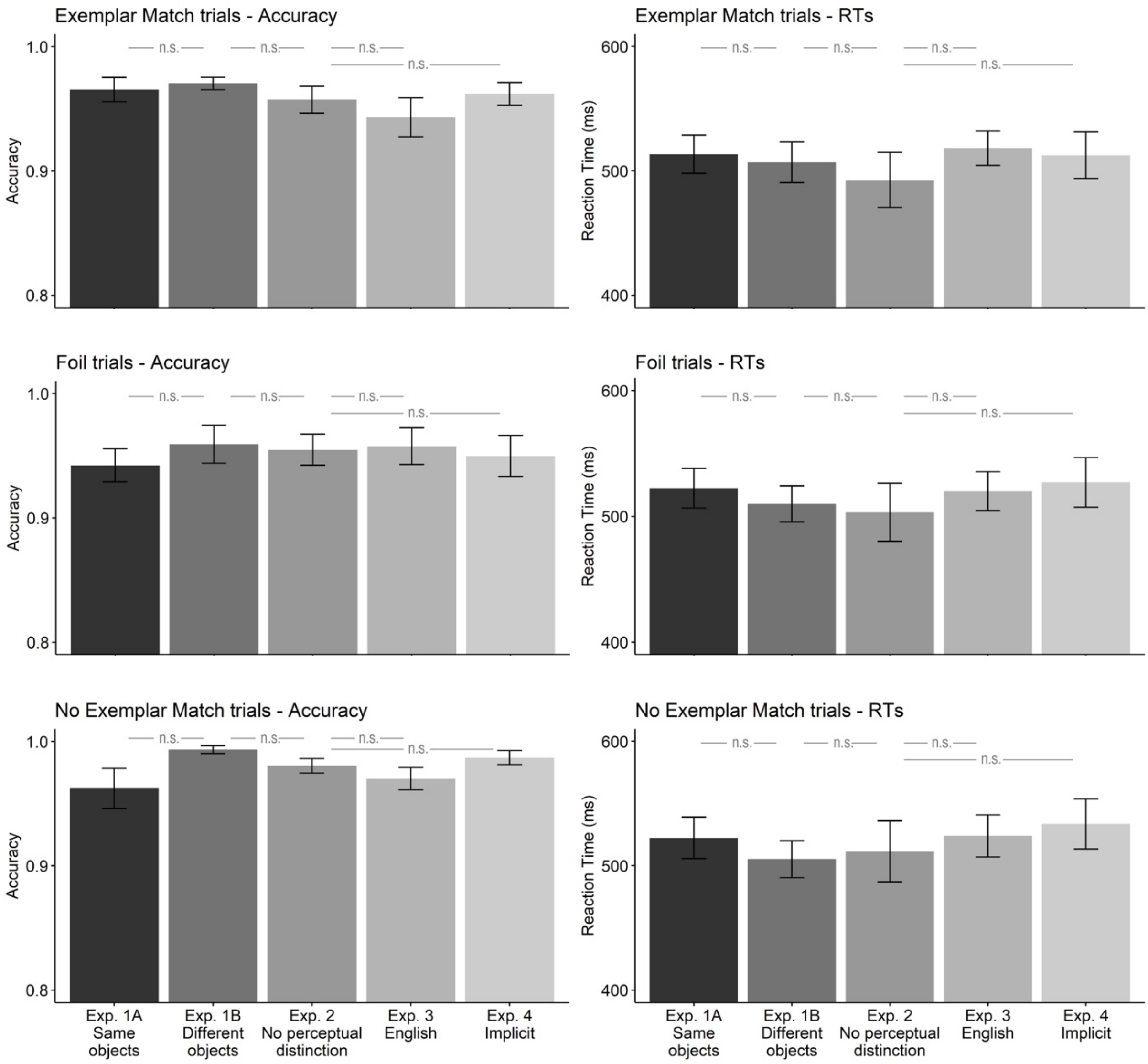

Figure E. Accuracy (left panels) and response times (right panels) to Exemplar Match trials (upper panels), Foil trials (middle panels) and No Exemplar Match trials (lower panels) in Experiments 1 to 4. Error bars represent ±1 SE.