Event-Related Potentials of the Semantically Informed Perception of Unfamiliar Objects

Alexander Enge[1,2], Franziska Süß[3], & Rasha Abdel Rahman[1]

[1] Humboldt-Universität zu Berlin

[2] Max Planck Institute for Human Cognitive and Brain Sciences

[3] Fachhochschule des Mittelstands

Author Note

Alexander Enge https://orcid.org/0000-0003-0100-2297

Rasha Abdel Rahman https://orcid.org/0000-0002-8438-1570

The preprocessed data and code for this study are openly available at https://osf.io/uksbc/.

We have no conflict of interest to disclose.

Correspondence concerning this article should be addressed to Rasha Abdel Rahman, Rudower Chaussee 18, 12489 Berlin. E-mail: rasha.abdel.rahman@hu-berlin.de

Abstract

Does our perception of an object change as soon as we discover what function it serves? This question is relevant not only for our everyday lives, where we may encounter novel tools and gadgets as parts of our dynamic working and private environments; it also pertains to the long-standing debate around the (im)penetrability of perception by higher cognitive capacities. In this experiment, we showed participants ($n = 24$) pictures of 120 unfamiliar objects either together with matching information about their function—leading to semantically informed perception—or together with non-matching information—resulting in naive perception. We measured event-related potentials (ERPs) to investigate at which stages in the visual processing hierarchy these two types of object perception differed from one another. We found that semantically informed as compared to naive perception was associated with larger amplitudes in the N170 component and reduced amplitudes in the N400 component. When the same objects were presented once more (without any information), the N400 effect persisted and we now also observed enlarged amplitudes in the P1 component in response to objects for which semantically informed perception had taken place. We replicated these novel findings in an independent sample ($n = 24$). Consistent with previous work, these results suggest that obtaining semantic information about previously unfamiliar objects alters aspects of their lower-level visual perception (P1 component), higher-level visual perception (e.g., holistic perception; N170 component), and semantic processing (N400 component).

Keywords: objects, semantic knowledge, visual perception, event-related potentials

Event-Related Potentials of the Semantically Informed Perception of Unfamiliar Objects

How does our perception of an object change as soon as we discover what function it serves? In this study, we presented participants with unfamiliar objects that were preceded either by matching information about their function, leading to semantically informed perception, or by non-matching information, leading to naive perception. We expected semantically informed perception to differ from naive perception within the object processing hierarchy and we capitalized on the high temporal resoluation of event-related potentials (ERPs) to find out at which stage (or stages) this was indeed the case. It allowed us to show for the first time that knowledge about the function of a previously unfamiliar object alters its early cortical processing in the same trial as it is being discovered and within less than 200 ms after the presentation of the object.

The claim that higher-level cognitive capacities such as knowledge or language can modulate lower-level perceptual processes has been the source of a long-running debate. This debate revolves around the question of whether or not our perception of the (visual) world is "cognitively impenetrable" (Pylyshyn, 1999) or, in other words, determined solely by the visual input reaching our retina. According to one view, higher-level mental states and functions are not able to penetrate perception and kick in only after lower-level perceptual analysis of the visual input has taken pace (e.g., Firestone & Scholl, 2016; Fodor, 1983; Pylyshyn, 1999). Visual perception itself is treated as an encapsulated module that processes the retinal input in a feed-forward fashion, progressing from lower areas with smaller receptive field sizes to areas representing increasingly complex shapes and, eventually, whole objects (DiCarlo et al., 2012; Marr, 1982).

This position is challenged by the alternative view that perceptual processing dynamically interacts with other aspects of cognition from early on (e.g., Churchland et al., 1994; Lupyan, 2015; Teufel & Nanay, 2017). The myriad feedback connections from areas higher up the visual hierarchy (e.g., the middle temporal and lateral occipital cortices) to early visual areas (e.g., V1 and V2; Bullier, 2001; Gilbert & Li, 2013)—as well as behavioral evidence—have inspired theories of visual perception that emphasize the top-down influence of cognitive processes. For instance, the reverse hierarchy theory (Ahissar & Hochstein, 2004; Hochstein & Ahissar, 2002) assumes that conscious visual perception occurs initially at the level of whole objects or object categories. Only after this high-level interpretation of the stimulus has been obtained, its more for fine grained visual details—if relevant for the current task—are being accessed from lower-level areas via top-down connections. Along similar lines, the label feedback hypothesis (Lupyan, 2012) posits that the activation of an object's name (i.e., a high level property) transiently warps perceptual space so that its diagnostic visual features are being processed preferentially. Finally, an active role of top-down influences is also promoted by theories of vision in the framework of Bayesian inference and predictive coding (e.g., Clark, 2013; Lupyan, 2015; Panichello et al., 2013; Yuille & Kersten, 2006).

Empirical findings from experimental psychology have recently added additional support to the view that perception can indeed be penetrated by other aspects of cognition (for review, see Collins & Olson, 2014; Vetter & Newen, 2014). These aspects of cognition may include transient states such as emotions (e.g., Bocanegra & Zeelenberg, 2009; Phelps et al., 2016) or intentions (e.g., Balcetis & Dunning, 2010; Cole et al., 2012) as well as learned capacities such as language (e.g., Boutonnet & Lupyan, 2015; Maier & Abdel Rahman, 2018; Mo et al., 2011) or declarative knowledge about faces (e.g., Abdel Rahman, 2011; Eiserbeck & Abdel Rahman, 2020; Suess et al., 2015) and objects (e.g., Abdel Rahman & Sommer, 2008; Gauthier, James, et al., 2003; Weller et al., 2019). Despite this wealth of empirical evidence, critics of the idea that cognitive functions can penetrate perception have remained skeptical and pointed out important methodological shortcomings in large parts of this literature (Firestone & Scholl, 2016; Machery, 2015). They argue, for example, that most behavioral paradigms have not been able to discern between perceptual and post-perceptual (e.g., memory-related) effects. Furthermore, cognitive influences in some studies were confounded with differences in the low-level visual input itself. We will revisit these and other potential pitfalls with regards to the present study in the Discussion section once we have laid out our methodology and results.

In general, however, it is important to note that one productive way to circumvent at least some of these concerns is to use neurophysiological recording methods such as ERPs (Luck, 2014): Their high temporal resolution allows for a direct test of which processing stages—perceptual and/or post-perceptual—are influenced by higher-level capacities such as semantic knowledge. This is made possible by the fact that when participants are presented with visual stimuli, their averaged brain responses show characteristic deflections (ERP components), each of them with its typical polarity (positive or negative), peak latency, spatial scalp distribution, and functional role(s).

The visual P1 component of the ERP refers to a positive deflection peaking as early as 100–150 ms after stimulus onset at occipital channels. It is generated by multiple sources in the extrastriate cortices of the middle occipital and fusiform gyri (Di Russo et al., 2001; Mangun, 1995). The P1 is thought to reflect the processing of low-level visual characteristics of the stimulus (e.g., size, luminance, and contrast; Di Russo et al., 2001; Johannes et al., 1995; Luck, 2014), which is why it being modulated by higher-level cognitive capacities would challenge a purely bottom-up, encapsulated view of visual perception. It is well-established that the P1 is enhanced when participants pay attention to the spatial location of a stimulus (Luck et al., 2000; Mangun, 1995; Mangun & Hillyard, 1991) but there is also more recent evidence for higher-level cognitive influences that go beyond spatial attention (see below).

The visual N170 (or N1) component refers to the negative deflection following the P1 and is maximal around 150–200 ms after stimulus onset at occipito-temporal channels. Just as the P1, it is influenced by visual parameters of the stimulus (Johannes et al., 1995) and by selective attention, especially when participants are required to discriminate between different visual stimuli (Mangun & Hillyard, 1991; Vogel & Luck, 2000). The N170 is enlarged (i.e., more negative) in response to faces (Bentin et al., 1996; Rossion & Jacques, 2011) and other visual stimuli for which one happens to be an expert (e.g., birds, dogs, or cars; Gauthier, Curran, et al., 2003; Tanaka & Curran, 2001). This category selectivity seems to reflect both the processing of individual visual features of the respective objects (including faces) as well as the holistic configuration of these features (Eimer et al., 2011; Jacques & Rossion, 2010; Rossion et al., 1999; Sagiv & Bentin, 2001). Thus, a modulation of the N170 by cognitive capacities such as knowledge would suggest these capacities having an influence on higher-level, holistic perception of the visual stimulus.

Finally, the N400 component refers to more negative ERP amplitudes when stimuli require more as compared to less resources for semantic processing or integration (Kutas & Federmeier, 2011; Lau et al., 2008; Rabovsky et al., 2018). It begins approximately 300 ms after stimulus onset and is most pronounced at centro-parietal channels. A modulation of the N400 component by cognitive capacities such as knowledge can be taken as evidence for a post-perceptual influence on stimulus processing.

Previous studies have focused on these and other ERP components to investigate which processing stages are indeed influenced by knowledge about visual objects (Abdel Rahman & Sommer, 2008; Gratton et al., 2009; Maier et al., 2014; Maier & Abdel Rahman, 2019; Rossion et al., 2004, 2002; Samaha et al., 2018; Tanaka & Curran, 2001; Weller et al., 2019). Together, they suggest that not only higher-level ERPs like the semantic N400 component are influenced by acquiring information about objects, but also earlier components that typically reflect bottom-up visual processing. This was the case, for instance, when participants were presented with a range of unfamiliar objects, receiving in-depth verbal descriptions about their function for half of the objects and irrelevant verbal information (i.e., cooking recipes) for the other half of the objects (Abdel Rahman & Sommer, 2008, Experiment 1). After this learning phase, the same objects were presented in three different ERP tasks which did not require explicit access to any of the learned semantic information. In all three tasks, ERP amplitudes in response to objects for which in-depth knowledge had been acquired differed from those in response to objects for which this had not been the case. Crucially, these differences occurred not just in the N400 component, indexing a modulation of post-perceptual semantic processing, but also in the P1 component, suggesting a top-down modulation of lower-level perceptual processing. Interestingly, the ERPs in the in-depth knowledge condition even were qualitatively similar to those for untrained but well-known objects. The modulation of the P1 component has recently been replicated when the same objects were presented under circumstances of limited attentional resources in an attentional blink paradigm (Weller et al., 2019). In this study, the (neurophysiological) differences in P1 amplitudes between objects with in-depth versus minimal knowledge were correlated with (behavioral) differences in their detection rate during the attentional blink. This can be taken as further evidence for an influence of semantic knowledge on the early visual perception of unfamiliar objects.

A complementary approach has been taken in a recent study investigating the ERPs in response to familiar (rather than unfamiliar) objects which were rendered difficult to recognize by converting them to two-tone, "Mooney" images (Samaha et al., 2018, Experiment 4). In this experiment, participants were trained with meaningful verbal cues, telling them what kind of objects they should look for in the images, or with a non-meaningful perceptual task, familiarizing them with the images but not with their semantic content. When the EEG was measured in a delayed matching task later on, the two types of training led to differential effects, again suggesting a modulation of lower-level visual perception: Over left posterior electrodes, P1 amplitudes were larger and alpha-band power was higher for meaning-trained images than for perceptually-trained images. As before, these effects of perceiving objects in a semantically informed way seem to be behaviorally relevant, as participants tended respond faster and more accurately to meaning-trained as compared to perceptually-trained images in the matching task. This alternative approach further corroborates an early influence of semantic information on the visual perception of objects.

All the previously mentioned studies have in common that they investigated the effects of semantic knowledge on the visual perception of objects only *after* an extensive learning phase had taken place (Abdel Rahman & Sommer, 2008; Gauthier, James, et al., 2003; Maier et al., 2014; Maier & Abdel Rahman, 2019; Rossion et al., 2004, 2002; Samaha et al., 2018; Weller et al., 2019). Participants performed one or multiple training sessions during which they encountered each object together with its respective label or description multiple times. The EEG was usually not measured—or at least not analyzed and reported—during these training sessions, which sometimes took place on a separate day (Abdel Rahman & Sommer, 2008; Maier et al., 2014; Maier & Abdel Rahman, 2019) or were spread out across multiple days or weeks (Rossion et al., 2004, 2002). While experimental designs with such substantive training maximize the chances of detecting even subtle top-down effects of the knowledge acquired, they leave at least three conceptual questions unresolved.

The first question is if we can detect any electrophysiological correlates of *semantic insight*, that is, the critical presentation during which an understanding of the previously unfamiliar object is happening (instead of asking if this understanding leads to differential effects in an orthogonal task later on). The second question is how much learning is actually necessary before reliable top-down effects of knowledge on ERP amplitudes can be obtained: Does it actually take dozens of repetitions per object or may a single exposure to the object together with the relevant semantic information be enough? The third and closely related question is about the nature of the effects of knowledge on perception: Are they reflecting genuine top-down effects of the semantic system altering perception online (while the object is being perceived), or do they merely reflect the (re-)activation of stored representations of the objects which have been altered over the course of the learning process? By measuring the ERPs before, during, and after participants received semantic hints about unfamiliar objects, we were able to provide tentative answers to all three of these questions.

In the two experiments subsequently reported, participants were presented with real-word objects that were presumed to be unfamiliar to most of them. They first viewed each of these objects without any semantic information. This first part served as a naive baseline to rule out that ERPs would differ in response to the objects based solely on low-level visual differences. Next, participants viewed each unfamiliar object for a second time, now preceded by verbal keywords. These keywords could either be matching the typical function of the object, thus making it possible for participants to understand what kind of object they were viewing, or they could be non-matching by describing the function of a different object, thus keeping the perception of the object semantically naive. During this second part, we were able to measure the online influence of semantic insight on object-evoked ERPs as it happened. Finally, the objects were presented for a third time (without keywords, as in the first part) to investigate downstream effects of having acquired semantic knowledge about them—thereby mimicking previous studies (e.g., Abdel Rahman & Sommer, 2008; Samaha et al., 2018; Weller et al., 2019). In all three parts, we examined the influence of semantic information on ERPs associated with lower-level visual perception (P1 component, 100–150 ms), higher-level visual perception (N170 component, 150–200 ms), and semantic processing (N400 component, 400–700 ms).

Experiment 1

Method

Participants

Participants for Experiment 1 were 24 German native speakers (13 female, 11 male) with a mean age of 24 years (range 18 to 31) and no history of psychological disorder or treatment. No a priori power analysis was carried out and the sample size was chosen in line with previous EEG studies on object processing in our lab. All participants were right-handed according to the Edinburgh inventory (Oldfield, 1971) and reported normal or corrected-to-normal vision. They gave written informed consent before starting the experiment and received a compensation of €8 per hour for participating.

Materials

Stimuli for Experiments 1 and 2 consisted of 240 grayscale photographs of real-world objects, 120 of which were well-known everyday objects (e.g., a bicycle, a toothbrush), serving as filler stimuli of no interest, whereas the other 120 were rare objects presumed to be unfamiliar to the majority of participants (e.g., a galvanometer, an udu drum). A list of these unfamiliar objects can be found in Appendix A. All stimuli were presented on a light blue background with a size of 207 $\times$ 207 pixels on a 19-inch LCD monitor with a resolution of 1,280 $\times$ 1,024 pixels and a refresh rate of 75 Hz. At a standardized viewing distance of 90 cm, the images of the objects subtended approximately 3.9 degrees of participants' horizontal and vertical visual angle.

For each unfamiliar object, a pair of keywords—a noun and a verb—was selected, describing the typical function or use of the object in a way that could typically be related to its visual features and their configuration (e.g., voltage, measuring; clay pot, drumming). As our central experimental manipulation, the presentation of half of the objects was preceded by keywords that correctly matched their respective function, whereas the presentation of the other half of the objects was preceded by non-matching keywords belonging to one of the other objects. The matching keywords were expected to induce semantically informed perception (i.e., participants suddenly understanding what kind of object they were viewing), whereas the non-matching keywords were expected hamper such an understanding and keep the perception of the object semantically naive. All participants saw each unfamiliar object with only one type of keywords (matching or non-matching). This assignment of keywords to objects was counterbalanced across participants so that each object was presented with matching keywords (leading to semantically informed perception) and non-matching keywords (leading to naive perception) to an equal number of participants. The experiment was programmed and displayed using Presentation® software (Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com).

Procedure

Each experimental session consisted of three parts (see Figure 1A). In the *pre-insight* part, after written informed consent had been obtained and the EEG had been prepared, all 240 familiar and unfamiliar objects were presented once in random order and without any keywords. Each trial consisted of a fixation cross presented in the middle of the screen for 0.5 s, followed by the presentation of the object until participants made a response or until a time out after 3 s. The inter-trial interval until the presentation of the next fixation cross was 0.5 s and participants took a self-timed break after each block of 60 objects. The task, which was kept the same across all three parts, was to classify each object using one of four response alternatives: (a) "I know what this is or have a strong assumption," (b) "I have an assumption what this is," (c) "I have rather no assumption what this is," or (d) "I don't know what this is and have no assumption." Participants were asked to respond as quickly and as accurately as possible by pressing one out of four buttons with the index or middle finger of their left or right hand, respectively. The mapping of the rating scale to the four buttons (left to right or right to left) was counterbalanced across participants.

In the *insight* part, the 120 unfamiliar objects were presented for a second time, now preceded either by matching keywords (leading to semantically informed perception) or by non-matching keywords (leading to naive perception). Each trial consisted of a fixation cross presented for 0.5 s, followed by the presentation of the keywords for 2.5 s. Then, an asterisk was presented in the middle of the screen for another 0.5 s, followed by the presentation of the object until a response was made or until a time out after 3 s. The objects were presented in blocks of 30 trials so that within each block (a) there were 15 objects from each of the two experimental conditions and (b) objects were heterogeneous in terms of their shape, visual complexity, and functional category (e.g., medical devices, musical instruments).

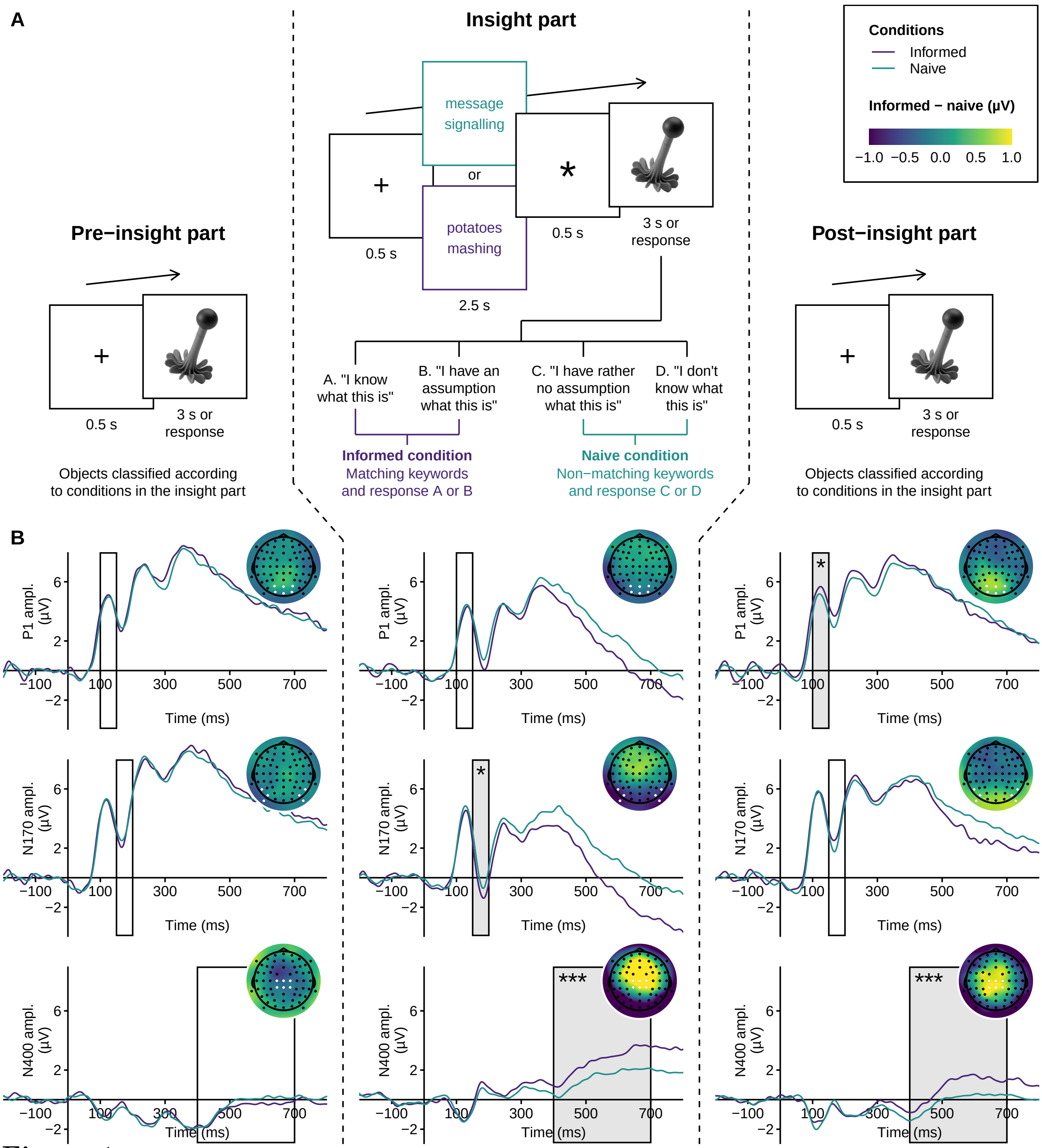

Figure 1

Procedure and Results of Experiment 1

Finally, in the *post-insight* part, the unfamiliar objects were presented for a third time with an identical trial structure as in the pre-insight part, that is, without any keywords. Note that the insight and post-insight parts were presented in an interleaved fashion so that after the presentation of one block of 30 objects in the insight part (with keywords), participants took a self-timed break and continued with the same block of 30 objects in the post-insight part (without keywords) before moving on to the next block consisting of 30 different objects. They continued like this until all four blocks were completed in both parts. In total, the experiment consisted of 480 trials (120 familiar objects in the pre-insight part and 120 unfamiliar objects in the pre-insight, insight, and post-insight parts). It took participants approximately 35 minutes to complete.

EEG Recording and Preprocessing

The continuous EEG was recorded from 62 Ag/AgCl scalp electrodes placed according to the extended 10–20 system (American Electroencephalographic Society, 1991) and referenced online to an external electrode placed on the left mastoid (M1). Two additional external electrodes were placed on the right mastoid (M2) and below the left eye (IO1), respectively. During the recording, electrode impedance was kept below 5 kΩ. An online band-pass filter with a high-pass time-constant of 10 s (0.016 Hz) and a low-pass cutoff frequency of 1000 Hz was applied before digitizing the signal at a sampling rate of 500 Hz.

Offline, the data were preprocessed using the MNE software (Version 0.21.0; Gramfort et al., 2013) in Python (Version 3.8.5; van Rossum & Drake, 2009). First, all scalp electrodes were re-referenced to the common average. Next, artifacts resulting from blinks and eye movements were removed using independent component analysis (ICA). The first 15 components were extracted by the FastICA algorithm (Hyvärinen, 1999) after temporarily low-pass filtering the data at 1 Hz. Any components showing substantive correlations with either of two virtual EOG channels (VEOG: IO1 - Fp1, HEOG: F9 - F10) were removed automatically using the *find_bads_eog* function. After artifact correction, a zero-phase, non-causal FIR filter with a lower pass-band edge at

0.1 Hz (transition bandwidth: 0.1 Hz) and an upper pass-band edge at 30 Hz (transition bandwidth: 7.5 Hz) was applied. Next, the continuous EEG was epoched into segments of 2 s, starting 500 ms before the onset of the visual presentation of each unfamiliar object. The epochs were baseline-corrected by subtracting the average voltage during the 200 ms before stimulus onset. Epochs containing artifacts despite ICA, defined as peak-to-peak amplitudes exceeding 200 µV, were removed from further analysis. This led to the exclusion of an average of 2.4 trials per participant (= 0.7%; range 0 to 24 trials). Single-trial event-related potentials were computed as the mean amplitude across time windows and regions of interests (ROIs) defined a priori, namely 100–150 ms after object onset at electrodes PO3, PO4, POz, O1, O2, and Oz for the P1 component, 150–200 ms after object onset at electrodes P7, P8, PO7, PO8, PO9, and PO10 for the N170 component, and 400–700 ms after object onset at electrodes C1, C2, Cz, CP1, CP2, and CPz for the N400 component (for the positions of these electrodes on the scalp, see the topographies in Figure 1B).

Statistical Analysis

First, because we were interested in the effects of knowledge on perceiving *unfamiliar* objects only, we excluded objects from all further analyses when participants classified them as being known in the pre-insight part (i.e., before any keywords were presented). This led to the exclusion of an average of 7.5 objects per participant (= 6.3% of all unfamiliar objects). Next, to clearly delineate semantically informed and naive perception, the assignment of all other objects to one of these two conditions for statistical analysis was co-determined by our experimental manipulation (matching versus non-matching keywords in the insight part) and the behavioral responses of the participants themselves (see Figure 1A). Objects were assigned to the semantically informed condition only if they were presented with matching keywords *and* if participants indicated knowing what the object was or having an assumption. This was the case for an average of 27.0 objects per participant (= 44.9% of objects presented with matching keywords). Complementarily, objects were assigned to the naive condition only if they were presented with non-matching keywords *and* if participants indicated not knowing what the object was or having rather no assumption. This was the case for an average of 48.8 objects per participant (= 81.2% of objects presented with non-matching keywords). Although this assignment was based on the manipulation and responses in the second part—when insight was thought to occur—the same assignment was used to analyze the data from the other two parts. This allowed us to test, on the one hand, if the objects from both conditions differed in important aspects even before any keywords were presented (pre-insight part) and, on the other hand, if the semantic understanding acquired in the insight part had any down-stream effects on the subsequent perception of the objects (post-insight part).

The event-related potentials in response to objects from both conditions and all three parts were analyzed on the single trial level using linear mixed-effects regression models (Baayen et al., 2008; Frömer et al., 2018). For the purpose of the present study, these models have at least three desirable properties compared to traditional approaches such as analyses of variance (ANOVAs) performed on by-participant grand averages. First, they can leverage all of the information that is present in the data set on the single trial level, some of which would be lost when computing averages. Second, they can account simultaneously for the non-independence of data points coming from the same participant and/or from the same item. In contrast, the neglect of the item as a random variable in by-participant ANOVAs leads to anti-conservative test statistics and strictly does not allow for inferences beyond the stimulus set under study (Bürki et al., 2018; Judd et al., 2012). Finally, mixed-effects models can flexibly deal with unbalanced designs in which the number of trials differs across (combinations of) conditions. Such a situation is inevitable when the assignment of trials to conditions is co-determined by the experimental manipulation and the behavioral responses of the participants to this manipulation (e.g., Fröber et al., 2017).

Three separate models were computed predicting P1, N170, and N400 mean amplitudes, respectively. All models included three fixed effects: (a) the part of the experiment, coded as a repeated contrast (i.e., subtracting the first from the second part and the second from the third part, the intercept being the grand mean across all three parts), (b) the condition of the object, coded as a scaled sum contrast (i.e., subtracting the naive condition from the semantically informed condition, the intercept being the grand mean across both conditions), and (c) the two-way interaction of part and condition. For details on these and other contrast coding schemes in linear (mixed-effects) models, please refer to Schad et al. (2020). To determine the random effects structure, we always started with a maximal model containing by-participant and by-item random intercepts and random slopes for all fixed effects (Barr et al., 2013). We then performed a model selection algorithm as proposed by Matuschek et al. (2017) in order to increase statistical power and avoid overparameterization: Iteratively, each random effect was removed and the resulting, more parsimonious model was compared to the previous, more complex model by means of a likelihood ratio test. Only if the parsimonious model explained the data equally well as the complex model (determined by $p > .20$; Matuschek et al., 2017) did we leave the random effect out, otherwise it was kept in the final model (see Appendix B for the formulas after model selection). All models were computed in R (Version 4.0.2; R Core Team, 2020) using the lme4 package (Version 1.1.23; Bates et al., 2015). The optimizer function *bobyqa* with 20,000 iterations was used for maximum likelihood estimation. The model selection algorithm via likelihood ratio tests was performed using the buildmer package (Version 1.7.1; Voeten, 2020).

Finally, to answer our research question of whether or not semantically informed perception had an influence on the ERP components within each part, planned follow-up comparisons were calculated, contrasting the semantically informed against the naive condition within the pre-insight, insight, and post-insight parts. This was achieved using the emmeans package (Version 1.5.0; Lenth, 2020). All p-values were computed by approximating the relevant denominator degrees of freedom using Satterthwaite's method as implemented in the lmerTest package (Version 3.1.2; Kuznetsova et al., 2017).

The materials, single trial behavioral and ERP data, and all code for data analysis can be accessed via the Open Science Framework (https://osf.io/uksbc/). Participants' lack of informed consent to do so prevents us from openly sharing the raw EEG data.

Results

Single-trial ERPs were analyzed in response to unfamiliar objects before (pre-insight part), while (insight part), and after (post-insight part) participants obtained relevant semantic information about their function. In the insight part, half of the objects were preceded by matching keywords, fostering semantically informed perception. The other half were preceded by non-matching keywords, keeping the perception of the object semantically naive. The objects were analyzed according to this manipulation in combination with participants' self-report in the insight part (see Figure 1A), thereby making sure that semantically informed and naive perception did indeed occur. The analysis focused on differences between these two conditions in the P1 component (100–150 ms) as an index of lower-level visual perception, the N170 component (150–200 ms) as index of higher-level visual perception, and the N400 component (400–700 ms) as an index of semantic processing.

Averaged across conditions, P1, N170, and N400 amplitudes differed as a function of the part of the experiment, all Fs > 10.89, all ps $< .001$ (see Table 1). In addition, N400 amplitudes differed between the informed and the naive condition averaged across the three parts of the experiment, $F(1, 24.4) = 13.00$, $p = .001$. Crucially, the part × condition interaction was significant in the N170 component, $F(2, 4878.2) = 4.84$, $p = .008$, and in the N400 component, $F(2, 5120.9) = 10.91$, $p < .001$, while also showing a weak statistical trend in the P1 component, $F(2, 4605.1) = 2.30$, $p = .100$. To answer our main research question, we decomposed these interactions into the differences between the semantically informed condition and the naive condition within the three different parts of the experiment.

Table 1

Results of Linear Mixed-Effects Regression Models for Experiment 1

	P1		N170		N400	
Fixed effects	F (df)	p	F (df)	p	F (df)	p
Part	10.89 (2, 24.5)	$<$ .001	14.18 (2, 24.9)	$<$ .001	32.85 (2, 25.6)	$<$ .001
Condition	0.60 (1, 5343.5)	.438	0.30 (1, 24.2)	.586	13.00 (1, 24.4)	.001
Pt. $\times$ con.	2.30 (2, 4605.1)	.100	4.84 (2, 4878.2)	.008	10.91 (2, 5120.9)	$<$ .001
Informed - naive	Est. [CI]; t (df)	p	Est. [CI]; t (df)	p	Est. [CI]; t (df)	p
Pre-insight part	-0.03 [-0.53, 0.47] -0.11 (4996.8)	.913	-0.12 [-0.68, 0.44] -0.43 (88.8)	.665	-0.20 [-0.65, 0.26] -0.85 (102.7)	.400
Insight part	-0.18 [-0.68, 0.32] -0.69 (5310.6)	.490	-0.64 [-1.20, -0.08] -2.26 (92.4)	.026	0.93 [0.48, 1.39] 4.04 (103.7)	$<$ .001
Post-insight part	0.55 [0.05, 1.05] 2.16 (4717.7)	.031	0.43 [-0.14, 0.99] 1.51 (91.8)	.136	1.00 [0.54, 1.46] 4.31 (105.0)	$<$ .001
logLik (df)	-16555.1 (13)		-16289.6 (17)		-15350.3 (17)	
RMSE	5.07		4.79		4.05	
Conditional R^2	0.289		0.379		0.207	
Marginal R^2	0.007		0.053		0.048	

Note. Pt. = part, con. = condition, est. = estimate, CI = confidence interval.

ERPs Before Insight Was Occurring

In the pre-insight part, when objects were unfamiliar to participants and presented without keywords, no differences emerged between the semantically informed and the naive condition in the P1, N170, or N400 component, all ps $\geq$ .400 (see Table 1 and Figure 1B). On the one hand, this was to be expected given that the critical presentation of the keywords (leading to semantically informed vs. naive perception) had not yet taken place. On the other hand, the absence of reliable differences in this part can be taken as evidence—with the usual caveats when interpreting null effects—that any subsequent effect of the semantic information in the other two parts cannot be accounted for by visual differences between the objects in the two conditions. Although the presentation of matching or non-matching keywords for each object was counterbalanced across participants, the fact that different numbers of objects were assigned to the two conditions based on participants' self report would have made it possible for such visual differences to emerge as a confounding factor. If they did, however, one would expect to detect these differences even before any keywords were presented, which was apparently not the case.

ERPs While Insight Was Occurring

In the insight part, half of the unfamiliar objects were presented with matching keywords (for forming the semantically informed condition) and the other half were presented with non-matching keywords (for forming the naive condition). When semantic information informed the perception of the object, the amplitude of the N170 component was significantly enlarged (i.e., more negative), b = -0.64 µV, p = .026, and the amplitude of the N400 component was significantly reduced (i.e., less negative), b = 0.93 µV, p < .001, compared to when the object was viewed naively without relevant semantic information. As in the pre-insight part, there were no reliable differences in the P1 component, p = .490.

ERPs After Insight Had Occurred

In the post-insight part, the unfamiliar objects were presented for a third time, again without the keywords (as in the pre-insight part). This allowed us to test if the semantic information had any lasting effects on the processing of the objects. As in the insight part, the N400 component remained significantly reduced during semantically informed as compared to naive perception, b = 1.00 µV, p < .001, whereas the effect in the N170 component did not reoccur, p = .136. Instead, we now observed an even earlier modulation in the P1 component which was significantly enlarged (i.e., more positive) in response to objects for which semantically informed perception had taken place, b = 0.55 µV, p = .031.

Discussion

In Experiment 1, we measured the ERPs of participants viewing unfamiliar objects before (pre-insight part), while (insight part) and after (post-insight part) they were able to understand what kind of object they were seeing. To induce this semantically informed perception, half of the objects in the insight part were preceded by matching verbal keywords about the object's typical function or use, whereas the other half were preceded by non-matching keywords, serving as a naive baseline condition.

In the insight part, we found that semantically informed perception was associated with enlarged amplitudes in the N170 component (150–200 ms after object onset) and reduced amplitudes in the N400 component (400–700 ms). When the same objects were presented once more in the post-insight part, the reduction of the N400 component reoccurred and we also observed a modulation of the P1 component (100–150 ms), which was significantly larger for objects for which semantically informed perception had taken place.

Because of the novelty of our experimental paradigm—measuring the ERPs while rather than after participants received semantic information about unfamiliar objects—and because of the novelty of our findings, we ran a replication study to assess the robustness of these effects with another sample of participants.

Experiment 2

Method

Participants

Participants for Experiment 2 were 24 German native speakers (15 female, 9 male) with a mean age of 26 years (range 19 to 29 years) who had not participated in Experiment 1. They had no history of psychological disorders or treatment, were right-handed, and reported normal or corrected-to-normal vision. They gave written informed consent before starting the experiment and received a compensation of €8 per hour for participating.

Materials, Procedure, and Analysis

All materials, procedures, EEG-related methods, and statistical analyses were identical to Experiment 1. An average of 7.0 objects per participant (= 5.8% of all unfamiliar objects) was classified as being known in the pre-insight part and excluded from all further analyses. Based on participants' responses in the insight part, an average of 26.4 objects was assigned to the semantically informed condition (= 44.0% of objects presented with matching keywords) and an average of 49.2 objects was assigned to the naive condition (= 82.0% of objects presented with non-matching keywords). Automatic rejection of EEG epochs containing artifacts led to the exclusion of 11.0 trials per participant (= 3.0%; range 0 to 85 trials).

Results

As in Experiment 1, P1, N170, and N400 amplitudes differed between the three different parts of the experiments, all Fs > 16.57, all ps $< .001$ (see Table 2). Also as in Experiment 1, N400 amplitudes differed between the informed and the naive condition averaged across parts, $F(1, 5242.5) = 19.95$, $p < .001$. The part × condition interaction was significant in the P1 component, $F(2, 5267.1) = 3.60$, $p = .027$, and in the N400 component, $F(2, 4714.9) = 7.92$, $p < .001$, but not in the N170 component, $F(2, 4756.5) = 1.25$, $p = .287$.

ERPs Before Insight Was Occurring

As in Experiment 1, no differences between objects in the semantically informed and the naive condition emerged in the P1, N170, or N400 component, all ps $> .582$ (see Table 2 and Figure 2).

Table 2

Results of Linear Mixed-Effects Regression Models for Experiment 2

	P1		**N170**		**N400**	
Fixed effects	F (df)	p	F (df)	p	F (df)	p
Part	17.45 (2, 5267.1)	< .001	16.57 (2, 25.1)	< .001	70.12 (2, 25.2)	< .001
Condition	0.02 (1, 5276.9)	.900	0.98 (1, 5228.5)	.322	19.95 (1, 5242.5)	< .001
Pt. × con.	3.60 (2, 5267.1)	.027	1.25 (2, 4756.5)	.287	7.92 (2, 4714.9)	< .001
Informed - naive	Est. [CI]; t (df)	p	Est. [CI]; t (df)	p	Est. [CI]; t (df)	p
Pre-insight part	-0.16 [-0.71, 0.40] -0.55 (5270.9)	.582	-0.03 [-0.53, 0.48] -0.10 (4772.1)	.920	0.04 [-0.43, 0.50] 0.15 (5177.9)	.877
Insight part	-0.48 [-1.03, 0.08] -1.67 (5270.7)	.094	-0.48 [-0.99, 0.03] -1.85 (5241.8)	.064	1.33 [0.87, 1.79] 5.68 (4781.9)	< .001
Post-insight part	0.57 [0.01, 1.13] 1.99 (5270.8)	.047	0.06 [-0.45, 0.57] 0.23 (5035.9)	.818	0.47 [0.01, 0.93] 1.98 (5067.3)	.047
logLik (df)	-16707.1 (8)		-16174.3 (13)		-15648.0 (13)	
RMSE	5.62		5.03		4.58	
Conditional R^2	0.242		0.399		0.174	
Marginal R^2	0.005		0.037		0.068	

Note. Pt. = part, con. = condition, est. = estimate, CI = confidence interval.

ERPs While Insight Was Occurring

As in Experiment 1, semantically informed as compared to naive perception (induced by matching vs. non-matching keywords) was associated with an enhancement of the N170 component, although this effect marginally failed to reach statistical significance, b = -0.48 µV, p = .064. The reduction of the N400 component during semantically informed perception remained significant, b = 1.33 µV, p < .001.

ERPs After Insight Had Occurred

As in Experiment 1, the presentation of the same unfamiliar objects for a third time (without keywords, as in the pre-insight part) again led to significantly reduced amplitudes in the N400 component in response to objects for which semantically

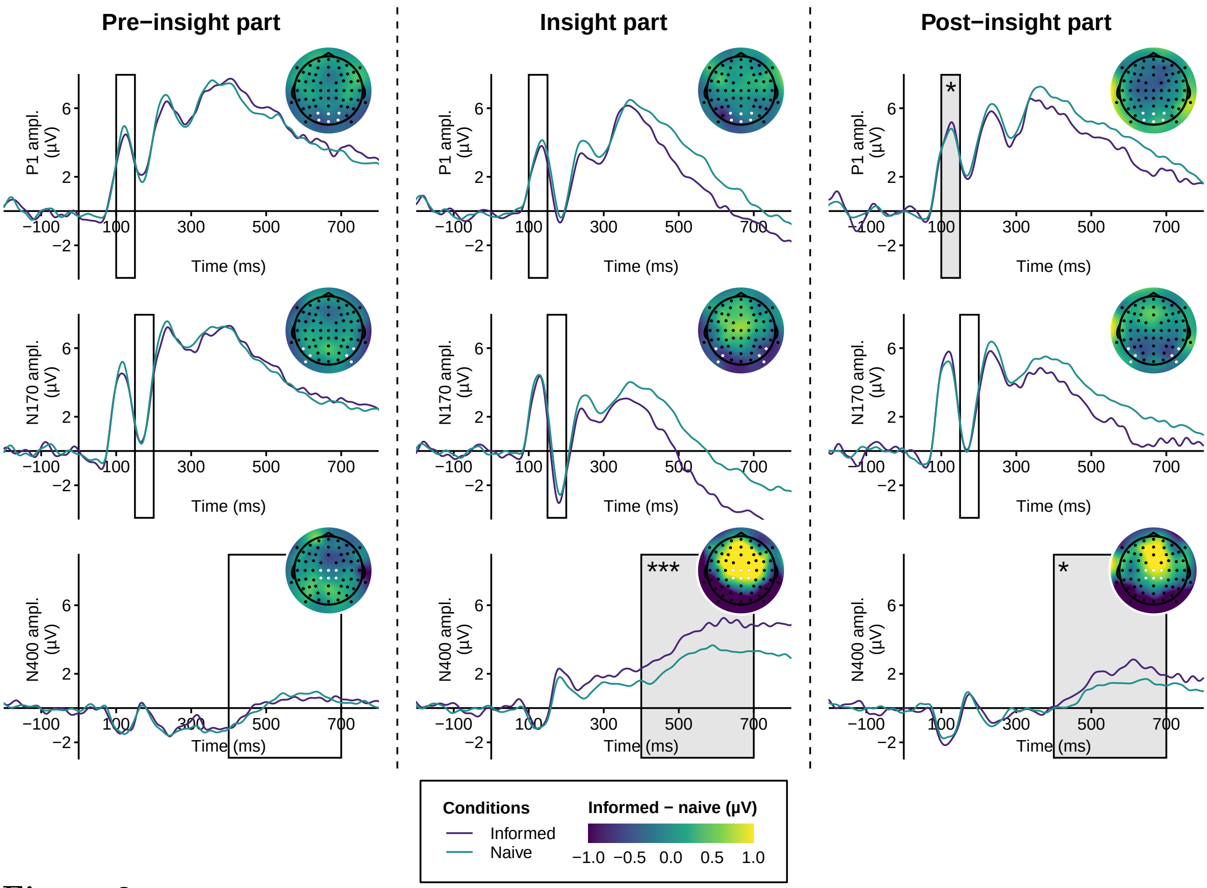

Figure 2

Results of Experiment 2

informed perception had occurred, $b = 0.47$ µV, $p = .047$. Also, P1 amplitudes in response these objects were significantly enlarged, $b = 0.57$ µV, $p = .047$.

Joint Analysis of Experiments 1 and 2

In an attempt to maximize statistical power, we combined the ERP data sets from Experiments 1 and 2. This allowed us to determine (a) if the above effects—including the marginally significant ones—were reliable when tested in a larger sample, and (b) if there were significant differences in the ERP amplitudes between Experiments 1 and 2. Methods for statistical analysis were kept unchanged apart from the addition of a new factor denoting the experiment, coded as a scaled sum contrast (i.e., subtracting Experiment 1 from Experiment 2, the intercept being the grand mean across both experiments; Schad et al., 2020). This factor and its interactions with part, condition, and part × condition were included in the linear mixed-effects regression models as fixed effects and as potential by-item random slopes. They were not included as by-participant random slopes because different participants took part in Experiments 1 and 2. Note that, as above, random effects were eventually included only if their omission led to a significant decline in model fit [Matuschek et al. (2017); Voeten (2020); see Appendix B].

As shown in Table 3, the main effect of part was significant in the P1, N170, and N400 component, all Fs > 12.34, all ps < .001, as was the main effect of condition in the N400, $F(1, 10593.6) = 41.38$, $p < .001$. Furthermore, the part × condition interaction was now observed reliably in all three components, all Fs > 3.90, all ps < .020. There was a main effect of experiment in the N400, $F(1, 48.4) = 4.96$, $p = .031$, with more negative amplitudes in Experiment 1 as compared to Experiment 2, $b =$ -1.06. However, the absence of any significant interactions of experiment with part or condition indicated that the effects of our experimental manipulations did not differ between Experiments 1 and 2.

Based on the part × condition interaction, we again computed follow-up comparisons between the semantically informed and the naive condition within each part, now collapsed across the data from both experiments (see Table 3 and Figure 3). This confirmed the absence of any reliable differences between the two conditions in the pre-insight part, all ps > .577, the significant enhancement of the N170 component in the insight part, while the semantic information was obtained, b = -0.57 µV, p = .004, the significant reduction of the N400 component in the insight part, while the semantic information was obtained, b = 1.10 µV, p < .001, and in the post-insight part, after the information had been obtained, b = 0.73 µV, p < .001, as well as the significant enhancement of the P1 component in the post-insight part, after the information had been obtained, b = 0.47 µV, p = .014.

Control Analysis

One may raise concerns whether the modulation of the N170 component in the insight part genuinely reflects the semantically informed perception of the objects in the

Table 3

Results of Linear Mixed-Effects Regression Models for Experiments 1 and 2 Combined

	P1		**N170**		**N400**	
Fixed effects	F (df)	p	F (df)	p	F (df)	p
Part	12.34 (2, 49.1)	$<$ .001	30.08 (2, 50.3)	$<$ .001	94.61 (2, 50.6)	$<$ .001
Condition	0.15 (1, 10573.1)	.699	1.16 (1, 45.2)	.287	41.38 (1, 10593.6)	$<$ .001
Experiment	0.33 (1, 48.2)	.567	2.50 (1, 48.0)	.121	4.96 (1, 48.4)	.031
Pt. × con.	3.90 (2, 9004.9)	.020	5.26 (2, 9904.1)	.005	15.34 (2, 10091.0)	$<$ .001
Pt. × exp.	0.04 (2, 49.1)	.966	0.58 (2, 50.3)	.565	2.55 (2, 50.6)	.088
Ins. × exp.	0.43 (1, 10573.1)	.514	0.06 (1, 45.2)	.805	0.07 (1, 10593.6)	.785
Pt. × con. × exp.	0.03 (2, 9004.9)	.968	0.58 (2, 9904.1)	.559	2.54 (2, 10091.0)	.079
Informed - naive	Est. [CI]; t (df)	p	Est. [CI]; t (df)	p	Est. [CI]; t (df)	p
Pre-insight part	-0.10 [-0.47, 0.28] -0.50 (9534.6)	.616	-0.10 [-0.48, 0.29] -0.51 (208.1)	.611	-0.09 [-0.39, 0.22] -0.56 (10459.4)	.577
Insight part	-0.24 [-0.62, 0.13] -1.28 (10420.2)	.202	-0.57 [-0.96, -0.18] -2.90 (214.4)	.004	1.10 [0.80, 1.41] 7.09 (10436.6)	$<$ .001
Post-insight part	0.47 [0.09, 0.84] 2.45 (9881.6)	.014	0.24 [-0.15, 0.63] 1.22 (211.4)	.225	0.73 [0.42, 1.04] 4.67 (10339.5)	$<$ .001
logLik (df)	-33272.6 (19)		-32474.6 (23)		-31045.5 (19)	
RMSE	5.32		4.91		4.32	
Conditional R^2	0.271		0.400		0.198	
Marginal R^2	0.008		0.06		0.07	

Note. Pt. = part, con. = condition, exp. = experiment, est. = estimate, CI = confidence interval.

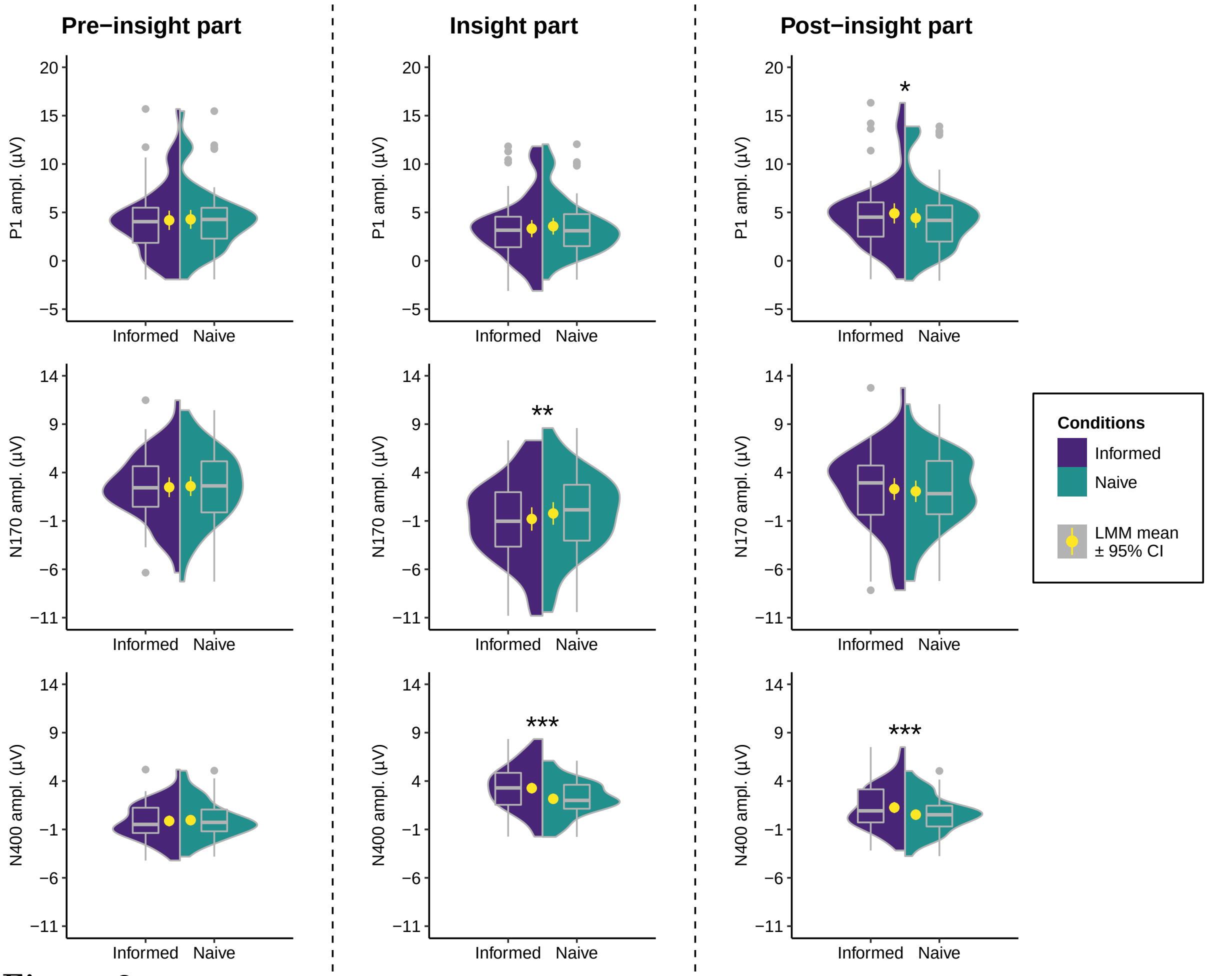

Figure 3

Measured and Modeled ERP Amplitudes for Experiments 1 and 2 Combined

respective condition, or—as an alternative explanation—whether it may be driven by the objects in the other, semantically naive condition. Remember that these objects were preceded by non-matching keywords which were picked so that they could not be related to the visual features of the object and their configuration. Thus, the modulation of the ERP components in the insight part could be a mismatch response to those objects—reflecting, for example, the fact that the visual features of the object shown in Figure 1A cannot be reconciled with the function of signalling messages. To preclude this alternative explanation, we repeated our analysis using a different baseline against which the objects in the semantically informed condition were contrasted. Instead of the naive condition (where objects were presented with non-matching keywords), we now used those objects which were presented with matching keywords (as in the informed condition), but which were excluded from the main analysis because participants indicated behaviorally that they did not understand the object they were seeing. Across both experiments, this was the case for 49.6% of objects presented with matching keywords, as compared to 44.4% of objects which did indeed lead to semantically informed perception.[1] Just as above, this control analysis revealed a robust N170 effect in the insight part, b = -0.99 µV, $p < .001$. Thus, this enhanced negativity seems to be a genuine marker of semantically informed perception, no matter if compared to objects presented with non-matching keywords or compared to objects presented with matching keywords on which participants failed to capitalize. Note that the reduction of the N400 component in the insight part also remained robust in this control analysis, b = 1.19 µV, $p < .001$.

Discussion

Experiment 2, which was an exact replication of Experiment 1, confirmed the effects of obtaining semantic information about previously unfamiliar objects on ERPs associated with lower-level visual perception (P1), higher-level visual perception (N170), and semantic processing (N400). Objects for which participants experienced semantically informed perception via matching keywords elicited enlarged N170 amplitudes and reduced N400 amplitudes. Note, however, that the effect in the N170 component marginally failed to reach the conventional level for statistical significance in this replication experiment. When the objects were presented once more without any semantic information, they again elicited reduced N400 amplitudes as well as enlarged P1 amplitudes.

General Discussion

Here we investigated if obtaining a semantic understanding of previously unfamiliar objects has an influence on how we perceive them. To this end, we measured

[1] Note that these percentages do not add up to 100% because some objects were excluded from all analyses because participants indicated knowing these objects in the pre-insight part, before any keywords had been presented (see Method).

ERPs while participants viewed unfamiliar objects before, while, and after receiving semantic information about them. For half of the objects, this information was matching the object, thus leading to semantically informed perception, whereas for the other half of the objects, the information was non-matching, thus keeping the perception of the object semantically naive. We found semantically informed perception to be accompanied by enlarged (i.e., more negative) N170 amplitudes and reduced (i.e., less negative) N400 amplitudes. When the same objects were presented again without the semantic information, the N400 component remained significantly reduced and we also observed a modulation of the P1 component, which was enlarged (i.e., more positive) in response to objects that had previously triggered semantically informed perception. We will discuss each of these effects in turn, starting with the latest (i.e., post-perceptual) effect and moving backward in time to the earlier (i.e., more perceptual) effects.

The reduction of the N400 component (400–700 ms after object onset) during semantically informed perception was the numerically largest and most robust effect. Perhaps least controversially, this effect indicates that acquiring an understanding of the objects (in the insight part) lessened participants' demand for effortful semantic processing in comparison to the naive condition (Kutas & Federmeier, 2011). This replicates previous work showing that N400 amplitudes are larger in response to pictures when they are either difficult to understand in and of themselves (e.g., Abdel Rahman & Sommer, 2008; Supp et al., 2005) or difficult to integrate into the preceding context (e.g., Barrett & Rugg, 1990; Ganis et al., 1996; Hirschfeld et al., 2011). The time course of this effect and the computational role of the N400 (Lau et al., 2008; Rabovsky et al., 2018) suggest that it has a post-perceptual locus.

In contrast to the N400, the N170 component (150–200 ms after object onset) was modulated while but not after the objects were presented together with the relevant semantic information. As such, it can be seen as an online marker of semantic insight, that is, participants suddenly understanding the visual objects in the light of the additional information provided by the verbal keywords. The N170 is typically associated with the holistic perception of faces (Eimer et al., 2011; Sagiv & Bentin, 2001) and other stimuli of visual expertise (Rossion et al., 2002; Tanaka & Curran, 2001). It being enlarged during semantically informed perception may therefore reflect that the additional semantic information made participants experience the configuration of the visual features of the objects in a new and meaningful way. This interpretation is supported by previous findings with a similar experimental paradigm in the domain of face perception (Bentin & Golland, 2002): Here, participants showed a face-like (i.e., enlarged) N170 response to a scrambled version of a schematic face after (but not before) they were primed with the intact version of the same face. This effect was absent when a visual control stimulus (a non-face object) was shown between the two presentations of the scrambled face. In the domain of non-face stimuli, enlarged N170 amplitudes have also been observed when participants were asked to discriminate between composite line drawings of meaningful objects as compared to composite line drawings of non-objects (Beaucousin et al., 2011). Of note, this effect was present only when participants were asked to decide based on the global shape of the object and it was reversed in polarity when they were asked to decide based on the constituent parts of the object. Together with the present study, these findings suggest an online impact of meaningfulness on the higher-level (i.e., holistic) perception of visual objects, integrating across their visual features.

The P1 component (100–150 ms after object onset), unlike the N400 and N170 components, was modulated by semantic information only one trial after the information had been obtained. This is consistent with previous studies showing modulations of the P1 component when participants learned meaningful information about previously unfamiliar objects (Abdel Rahman & Sommer, 2008; Maier & Abdel Rahman, 2019, 2018; Weller et al., 2019) or about familiar objects that were rendered difficult to recognize (Samaha et al., 2018). What the present study adds to these findings is that the P1 effect does not take an extensive learning phase to develop (with multiple presentations of the objects together with the respective information). Instead, it can be observed as soon as one trial after semantic insight has happened. Because the

P1 is typically associated with lower-level sensory processing (e.g., Johannes et al., 1995; Luck, 2014; Pratt, 2011), we take its susceptibility to semantic information as an indicator that knowledge about the function of an object can change how we perceive it visually.

Both the N170 and the P1 components therefore seem to be sensitive to the semantic meaningfulness of visual objects. However, the finding that these two components were modulated in different parts of our experimental design suggests that they reflect different aspects of top-down processing with different time courses and neuroanatomical implementations. It has been pointed out that the time course of the N170 component is consistent with a top-down influence of (non-visual) areas in the prefrontal and parietal cortices on visual areas, whereas modulations of the P1 component seem reflect recurrent processing *within* the visual system (Wyatte et al., 2014). Here we could show that the former pathway seems to be able to convey semantic information instantaneously (i.e., within the same trial), whereas the latter seems to take at least one—but apparently also not more than one—additional presentation of the visual object to emerge.

While the limited spatial resolution of the EEG precludes a precise localization of these effects within the ventral stream of object recognition, there is converging evidence coming from fMRI showing that semantic information can feed back into the earliest of visual areas. Hsieh et al. (2010) showed participants indiscernible two-tone ("Mooney") versions of images before and after showing them the original versions. They found that the brain responses to the original image were correlated more strongly with the second presentation of the Mooney image (after insight had taken place) than with the first presentation of the Mooney image (before insight had taken place). This increase in representational similarity was not just observed in higher-level object-sensitive areas in the lateral occipital cortex (LOC), but also in early retinotopic cortex (areas V1, V2, and V3). Both of these cortical regions are consistent with the neural generators of the N170 and P1 components of the ERP which we have found to be sensitive to the semantically informed perception of previously unfamiliar objects.

On a theoretical level, the top-down modulation of these visual ERPs by semantic information challenges a modular view of visual perception (Firestone & Scholl, 2016; Fodor, 1983; Pylyshyn, 1999; but see Clarke, 2020). However, proponents of such a modular view have pointed out important shortcomings of previous studies that claimed to demonstrate top-down effects of cognition on perception (Firestone & Scholl, 2016; Machery, 2015). We took care to address as many of these shortcomings as possible: First, we showed that no effect had been present before any semantic information was being presented (in the pre-insight part). Second, we used an objective and time-resolved measure (ERPs) to disentangle effects with a perceptual locus from those with a post-perceptual locus. Third, we reduced response and demand biases by keeping the manipulation (i.e., matching or non-matching keywords) obscure to the participants and by including well-known objects as filler stimuli. Fourth, we precluded low-level visual differences between conditions by counterbalancing the assignment of objects to conditions across participants. Fifth, we reduced priming and attentional effects by presenting all objects in a randomized order and at the same location. Sixth, we reduced memory effects by using only unfamiliar objects and by measuring online ERPs rather than delayed behavioral responses. We hope that these procedures have effectively ruled out some of the most important alternative explanations for the top-down effects that we have observed, thus making a more compelling case against the cognitive impenetrability of perception.

An interactive view of object vision with an abundance of top-down feedback also challenges the predominantly feed-forward models in computer vision (e.g., Marr, 1982; Serre et al., 2007). In fact, the lack of a semantic knowledge base that dynamically interacts with the processing of lower-level visual features may be one key reason why even state-of-the-art deep-learning algorithms need orders of magnitude more training examples to achieve human-level performance in object recognition. For these network models, single-trial learning of previously unfamiliar objects, as was observed on the behavioral and on the neurophysiological level in the present study, seems to be out of reach until they overcome this “barrier of meaning” (Mitchell, 2020). Drawing inspiration from cognitive psychology and human neuroscientific data may help to make these models more biologically plausible and, at the same time, more data efficient.

A theoretical framework that would explicitly predict or explain the observed P1 and N170 effects in our study is lacking at present. The effects are consistent, however, with the reverse hierarchy theory (Ahissar & Hochstein, 2004; Hochstein & Ahissar, 2002) which posits that objects first enter our visual consciousness at an abstract, conceptual level. Once this initial “vision at a glance” has taken place, feedback connections to earlier layers of the visual system are being accessed to extract the relevant lower-level features (“vision with scrutiny”). This reverse trajectory down the visual hierarchy may explain (a) the semantically induced changes to the fMRI signal in LOC and retinotopic cortex (Hsieh et al., 2010) as well as (b) the modulations of early visual ERP components observed in the present study and others (Abdel Rahman & Sommer, 2008; Maier et al., 2014; Maier & Abdel Rahman, 2019; Samaha et al., 2018; Weller et al., 2019). Besides this specific theory, an important role of top-down mechanisms for vision or, more specifically, object recognition is also posited by the family of predictive coding and Bayesian inference theories (e.g., Clark, 2013; Lupyan, 2015; Panichello et al., 2013; Yuille & Kersten, 2006). Despite the theoretical advances, the mechanistic details of these top-down effects at the algorithmic and implementational level (Marr, 1982) remain to be clarified.

The lack of mechanistic insight into the top-down effects that we have observed is one limitation of the present study. Another one is our reliance on rare and highly specialized objects (see Appendix A). This was necessary to induce the experience of semantic insight in a population of undergraduate students but it may not necessarily generalize to the way in which we learn about everyday objects.[2] One way of addressing

[2]Note, however, the converging evidence coming from the complementary approach of rendering images of everyday objects difficult to recognize (Samaha et al., 2018).

this issue would be to adapt the present paradigm to younger participants, using everyday objects which they have not yet learned about. Finally, the number participants ($n = 24$ per experiment) and trials ($k \approx 30$ per part in the insight condition and $k \approx 50$ per part in the naive condition) can be considered small by today's standards (Baker et al., 2020). Our analysis may therefore not have been particularly sensitive, especially given that linear mixed-effects regression models tend to have limited statistical power under a wide range of circumstances (see, e.g., the simulations by Matuschek et al., 2017). This lack of power may also explain why one of the effects that had been observed in the first experiment (i.e., the enlargement of the N170 component, $p = .026$) failed to reach statisticial signficance in the replication experiment ($p = .064$). To discern if this was due to a lack of statistical power or due to the actual absence of an effect, it would take yet another (more highly powered) replication study and/or a different data-analytic approach where it is possible to estimate the evidence both for the alternative hypothesis and for the null hypothesis (e.g., Bayesian linear mixed models). It should be noted, however, that the joint analysis of both data sets yielded a very robust N170 effect ($p = .004$). This makes it less likely that the effect observed in Experiment 1 was due to a false positive and instead suggests that statistical power in Experiment 2 was insufficient to render the effect significant. Also note that the effect in the P1 component was statistically significant in both experiments individually, making an even stronger case for the susceptibility of early cortical processing to newly acquired knowledge about objects.

Taken together, this study provides preliminary evidence that whenever we receive meaningful semantic information about a previously unfamiliar object, this information has an immediate influence on the processing of the object. This influence is immediate in at least two different ways: First, it does not require extensive training and can instead be observed within the same trial in which the information is being presented (and/or one trial later). Second, the time course of this influence suggests that it manifests itself not only at later, post-perceptual stages—typically associated with semantic processing—but also at earlier stages—typically associated with visual perception itself and happenning within less than a fifth of a second after the object is presented to us.

References

Abdel Rahman, R. (2011). Facing good and evil: Early brain signatures of affective biographical knowledge in face recognition. *Emotion*, *11*(6), 1397–1405. https://doi.org/10.1037/a0024717

Abdel Rahman, R., & Sommer, W. (2008). Seeing what we know and understand: How knowledge shapes perception. *Psychonomic Bulletin & Review*, *15*(6), 1055–1063. https://doi.org/10.3758/PBR.15.6.1055

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, *8*(10), 457–464. https://doi.org/10.1016/j.tics.2004.08.011

American Electroencephalographic Society. (1991). American Electroencephalographic Society guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, *8*(2), 200–202.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baker, D. H., Vilidaite, G., Lygo, F. A., Smith, A. K., Flack, T. R., Gouws, A. D., & Andrews, T. J. (2020). Power contours: Optimising sample size and precision in experimental psychology and human neuroscience. *Psychological Methods*. https://doi.org/10.1037/met0000337

Balcetis, E., & Dunning, D. (2010). Wishful seeing: More desired objects are seen as closer. *Psychological Science*, *21*(1), 147–152. https://doi.org/10.1177/0956797609356283

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barrett, S. E., & Rugg, M. D. (1990). Event-related potentials and the semantic matching of pictures. *Brain and Cognition*, *14*(2), 201–212. https://doi.org/10.1016/0278-2626(90)90029-N

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Beaucousin, V., Cassotti, M., Simon, G., Pineau, A., Kostova, M., Houdé, O., & Poirel, N. (2011). ERP evidence of a meaningfulness impact on visual global/local processing: When meaning captures attention. *Neuropsychologia*, *49*(5), 1258–1266. https://doi.org/10.1016/j.neuropsychologia.2011.01.039

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, *8*(6), 551–565. https://doi.org/10.1162/jocn.1996.8.6.551

Bentin, S., & Golland, Y. (2002). Meaningful processing of meaningless stimuli: The influence of perceptual experience on early visual processing of faces. *Cognition*, *86*(1), B1–B14. https://doi.org/10.1016/S0010-0277(02)00124-5

Bocanegra, B. R., & Zeelenberg, R. (2009). Emotion improves and impairs early vision. *Psychological Science*, *20*(6), 707–713. https://doi.org/10.1111/j.1467-9280.2009.02354.x

Boutonnet, B., & Lupyan, G. (2015). Words jump-start vision: A label advantage in object recognition. *Journal of Neuroscience*, *35*(25), 9329–9335. https://doi.org/10.1523/JNEUROSCI.5111-14.2015

Bullier, J. (2001). Integrated model of visual processing. *Brain Research Reviews*, *36*(2), 96–107. https://doi.org/10.1016/S0165-0173(01)00085-6

Bürki, A., Frossard, J., & Renaud, O. (2018). Accounting for stimulus and participant effects in event-related potential analyses to increase the replicability of studies. *Journal of Neuroscience Methods*, *309*, 218–227. https://doi.org/10.1016/j.jneumeth.2018.09.016

Churchland, P. S., Ramachandran, V. S., & Sejnowski, T. J. (1994). A critique of pure vision. In C. Koch & J. L. Davis (Eds.), *Computational neuroscience. Large-scale neuronal theories of the brain* (pp. 23–60). MIT Press.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Clarke, S. (2020). Cognitive penetration and informational encapsulation: Have we been failing the module? *Philosophical Studies*. https://doi.org/10.1007/s11098-020-01565-1

Cole, S., Balcetis, E., & Dunning, D. (2012). Affective signals of threat increase perceived proximity. *Psychological Science*. https://doi.org/10.1177/0956797612446953

Collins, J. A., & Olson, I. R. (2014). Knowledge is power: How conceptual knowledge transforms visual cognition. *Psychonomic Bulletin & Review*, *21*(4), 843–860. https://doi.org/10.3758/s13423-013-0564-3

Di Russo, F., Martínez, A., Sereno, M. I., Pitzalis, S., & Hillyard, S. A. (2001). Cortical sources of the early components of the visual evoked potential. *Human Brain Mapping*, *15*(2), 95–111. https://doi.org/10.1002/hbm.10010

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron*, *73*(3), 415–434. https://doi.org/10.1016/j.neuron.2012.01.010

Eimer, M., Gosling, A., Nicholas, S., & Kiss, M. (2011). The N170 component and its links to configural face processing: A rapid neural adaptation study. *Brain Research*, *1376*, 76–87. https://doi.org/10.1016/j.brainres.2010.12.046

Eiserbeck, A., & Abdel Rahman, R. (2020). Visual consciousness of faces in the attentional blink: Knowledge-based effects of trustworthiness dominate over appearance-based impressions. *Consciousness and Cognition*, *83*, 102977. https://doi.org/10.1016/j.concog.2020.102977

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences*, *39*. https://doi.org/10.1017/S0140525X15000965

Fodor, J. A. (1983). *The modularity of mind*. MIT Press.

Fröber, K., Stürmer, B., Frömer, R., & Dreisbach, G. (2017). The role of affective evaluation in conflict adaptation: An LRP study. *Brain and Cognition*, *116*, 9–16. https://doi.org/10.1016/j.bandc.2017.05.003

Frömer, R., Maier, M., & Abdel Rahman, R. (2018). Group-level EEG-processing pipeline for flexible single trial-based analyses including linear mixed models. *Frontiers in Neuroscience*, *12*. https://doi.org/10.3389/fnins.2018.00048

Ganis, G., Kutas, M., & Sereno, M. I. (1996). The search for "common sense": An electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience*, *8*(2), 89–106. https://doi.org/10.1162/jocn.1996.8.2.89

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience*, *6*(4), 428–432. https://doi.org/10.1038/nn1029

Gauthier, I., James, T. W., Curby, K. M., & Tarr, M. J. (2003). The influence of conceptual knowledge on visual discrimination. *Cognitive Neuropsychology*, *20*(3-6), 507–523. https://doi.org/10.1080/02643290244000275

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience*, *14*(5), 350–363. https://doi.org/10.1038/nrn3476

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & al., et. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*. https://doi.org/10.3389/fnins.2013.00267

Gratton, C., Evans, K. M., & Federmeier, K. D. (2009). See what I mean? An ERP study of the effect of background knowledge on novel object processing. *Memory & Cognition*, *37*(3), 277–291. https://doi.org/10.3758/MC.37.3.277

Hirschfeld, G., Zwitserlood, P., & Dobel, C. (2011). Effects of language comprehension on visual processing – MEG dissociates early perceptual and late N400 effects. *Brain and Language*, *116*(2), 91–96. https://doi.org/10.1016/j.bandl.2010.07.002

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, *36*(5), 791–804. https://doi.org/10.1016/S0896-6273(02)01091-7

Hsieh, P.-J., Vul, E., & Kanwisher, N. (2010). Recognition alters the spatial pattern of fMRI activation in early retinotopic cortex. *Journal of Neurophysiology*, *103*(3), 1501–1507. https://doi.org/10.1152/jn.00812.2009

Hyvärinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, *10*(3), 626–634. https://doi.org/10.1109/72.761722

Jacques, C., & Rossion, B. (2010). Misaligning face halves increases and delays the N170 specifically for upright faces: Implications for the nature of early face representations. *Brain Research*, *1318*, 96–109. https://doi.org/10.1016/j.brainres.2009.12.070

Johannes, S., Münte, T. F., Heinze, H. J., & Mangun, G. R. (1995). Luminance and spatial attention effects on early visual processing. *Cognitive Brain Research*, *2*(3), 189–205. https://doi.org/10.1016/0926-6410(95)90008-X

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*(1), 54–69. https://doi.org/10.1037/a0028347

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. *Nature Reviews Neuroscience*, *9*(12), 920–933. https://doi.org/10.1038/nrn2532

Lenth, R. (2020). *Emmeans: Estimated marginal means, aka least-squares means.* https://CRAN.R-project.org/package=emmeans

Luck, S. J. (2014). Overview of common ERP components. In *An introduction to the event-related potential technique* (2nd ed., pp. 71–118). MIT Press.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences*, *4*(11), 432–440. https://doi.org/10.1016/S1364-6613(00)01545-X

Lupyan, G. (2015). Cognitive penetrability of perception in the age of prediction: Predictive systems are penetrable systems. *Review of Philosophy and Psychology*, *6*(4), 547–569. https://doi.org/10.1007/s13164-015-0253-4

Lupyan, G. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00054

Machery, E. (2015). Cognitive penetrability: A no-progress report. In J. Zeimbekis & A. Raftopoulos (Eds.), *The cognitive penetrability of perception: New philosophical perspectives*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198738916.003.0002

Maier, M., & Abdel Rahman, R. (2019). No matter how: Top-down effects of verbal and semantic category knowledge on early visual perception. *Cognitive, Affective, & Behavioral Neuroscience*, *19*(4), 859–876. https://doi.org/10.3758/s13415-018-00679-8

Maier, M., & Abdel Rahman, R. (2018). Native language promotes access to visual consciousness. *Psychological Science*, *29*(11), 1757–1772. https://doi.org/10.1177/0956797618782181

Maier, M., Glage, P., Hohlfeld, A., & Abdel Rahman, R. (2014). Does the semantic content of verbal categories influence categorical perception? An ERP study. *Brain and Cognition*, *91*, 1–10. https://doi.org/10.1016/j.bandc.2014.07.008

Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology*, *32*(1), 4–18. https://doi.org/10.1111/j.1469-8986.1995.tb03400.x

Mangun, G. R., & Hillyard, S. A. (1991). Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming. *Journal of Experimental Psychology. Human Perception and Performance*, *17*(4), 1057–1074. https://doi.org/10.1037//0096-1523.17.4.1057

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. W.H. Freeman.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Mitchell, M. (2020). On crashing the barrier of meaning in artificial intelligence. *AI Magazine*, *41*(2), 86–92. https://doi.org/10.1609/aimag.v41i2.5259

Mo, L., Xu, G., Kay, P., & Tan, L.-H. (2011). Electrophysiological evidence for the left-lateralized effect of language on preattentive categorical perception of color. *Proceedings of the National Academy of Sciences*, *108*(34), 14026–14030. https://doi.org/10.1073/pnas.1111860108

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Panichello, M. F., Cheung, O. S., & Bar, M. (2013). Predictive feedback and conscious visual experience. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00620

Phelps, E. A., Ling, S., & Carrasco, M. (2016). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological Science*. https://journals.sagepub.com/doi/10.1111/j.1467-9280.2006.01701.x

Pratt, H. (2011). Sensory ERP components. In E. S. Kappenman & S. J. Luck (Eds.), *The Oxford handbook of event-related potential components* (pp. 89–114). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0050

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, *22*(3), 341–365. https://doi.org/10.1017/S0140525X99002022

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rabovsky, M., Hansen, S. S., & McClelland, J. L. (2018). Modelling the N400 brain potential as change in a probabilistic representation of meaning. *Nature Human Behaviour*, *2*(9), 693–705. https://doi.org/10.1038/s41562-018-0406-4

Rossion, B., Delvenne, J.-F., Debatisse, D., Goffaux, V., Bruyer, R., Crommelinck, M., & Guérit, J.-M. (1999). Spatio-temporal localization of the face inversion effect: An event-related potentials study. *Biological Psychology*, *50*(3), 173–189. https://doi.org/10.1016/S0301-0511(99)00013-7

Rossion, B., Gauthier, I., Goffaux, V., Tarr, M. J., & Crommelinck, M. (2002). Expertise training with novel objects leads to left-lateralized facelike electrophysiological responses. *Psychological Science*. https://doi.org/10.1111/1467-9280.00446

Rossion, B., & Jacques, C. (2011). The N170: Understanding the time course of face perception in the human brain. In E. S. Kappenman & S. J. Luck (Eds.), *The Oxford handbook of event-related potential components* (pp. 115–142). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.001.0001

Rossion, B., Kung, C.-C., & Tarr, M. J. (2004). Visual expertise with nonface objects leads to competition with the early perceptual processing of faces in the human occipitotemporal cortex. *Proceedings of the National Academy of Sciences*, *101*(40), 14521–14526. https://doi.org/10.1073/pnas.0405613101

Sagiv, N., & Bentin, S. (2001). Structural encoding of human and schematic faces: Holistic and part-based processes. *Journal of Cognitive Neuroscience*, *13*(7), 937–951. https://doi.org/10.1162/089892901753165854

Samaha, J., Boutonnet, B., Postle, B. R., & Lupyan, G. (2018). Effects of meaningfulness on perception: Alpha-band oscillations carry perceptual expectations and influence early visual responses. *Scientific Reports*, *8*(1), 1–14. https://doi.org/10.1038/s41598-018-25093-5

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, *110*, 104038. https://doi.org/10.1016/j.jml.2019.104038

Serre, T., Oliva, A., & Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proceedings of the National Academy of Sciences*, *104*(15), 6424–6429. https://doi.org/10.1073/pnas.0700622104

Suess, F., Rabovsky, M., & Abdel Rahman, R. (2015). Perceiving emotions in neutral faces: Expression processing is biased by affective person knowledge. *Social Cognitive and Affective Neuroscience*, *10*(4), 531–536. https://doi.org/10.1093/scan/nsu088

Supp, G. G., Schlögl, A., Fiebach, C. J., Gunter, T. C., Vigliocco, G., Pfurtscheller, G., & Petsche, H. (2005). Semantic memory retrieval: Cortical couplings in object recognition in the N400 window. *European Journal of Neuroscience*, *21*(4), 1139–1143. https://doi.org/https://doi.org/10.1111/j.1460-9568.2005.03906.x

Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological Science*, *12*(1), 43–47. https://doi.org/10.1111/1467-9280.00308

Teufel, C., & Nanay, B. (2017). How to (and how not to) think about top-down influences on visual perception. *Consciousness and Cognition*, *47*, 17–25. https://doi.org/10.1016/j.concog.2016.05.008

van Rossum, G., & Drake, F. L. (2009). *Python 3 reference manual*. CreateSpace.

Vetter, P., & Newen, A. (2014). Varieties of cognitive penetration in visual perception. *Consciousness and Cognition*, *27*, 62–75. https://doi.org/10.1016/j.concog.2014.04.007

Voeten, C. C. (2020). *Buildmer: Stepwise elimination and term reordering for mixed-effects regression*. https://CRAN.R-project.org/package=buildmer

Vogel, E. K., & Luck, S. J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology*, *37*(2), 190–203. https://doi.org/10.1111/1469-8986.3720190

Weller, P. D., Rabovsky, M., & Abdel Rahman, R. (2019). Semantic knowledge enhances conscious awareness of visual objects. *Journal of Cognitive Neuroscience*, *31*(8), 1216–1226. https://doi.org/10.1162/jocn_a_01404

Wyatte, D., Jilk, D. J., & O'Reilly, R. C. (2014). Early recurrent feedback facilitates visual object recognition under challenging conditions. *Frontiers in Psychology*, *5*, 674. https://doi.org/10.3389/fpsyg.2014.00674

Yuille, A., & Kersten, D. (2006). Vision as bayesian inference: Analysis by synthesis? *Trends in Cognitive Sciences*, *10*(7), 301–308. https://doi.org/10.1016/j.tics.2006.05.002