Neuropsychologia 117 (2018) 222–232

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Are incongruent objects harder to identify? The functional significance of the N300 component

Alyssa Truman[a], Liad Mudrik[a,b,*]

[a] Sagol School for Neuroscience, Tel-Aviv University, Ramat Aviv, POB 39040, Tel Aviv 69978, Israel
[b] School of Psychological Sciences, Tel-Aviv University, Ramat Aviv, POB 39040, Tel Aviv 69978, Israel

ARTICLE INFO

Keywords:
Object identification
Context
Object-scene relations
N300
EEG

ABSTRACT

Objects in the real world typically appear within a broader context, having relationships with the environment. Do these relations between objects and the contexts in which they appear affect object identification? Previous findings of an N300 component evoked by scene-incongruent objects were taken as evidence for such an effect, since N300 is held to reflect object identification processes. Yet this conjuncture was never directly tested, and ignores differences between the fronto-central incongruency-evoked N300 and the typically bi-polar fronto-occipital identification-related N300. Here, the possible influence of context on object identification was examined by manipulating both object-scene congruency and object identifiability. N300 effects were found both for incongruity and for identifiability, in line with previous studies. Critically, a comparison of divergence times of waveforms evoked by congruent/incongruent objects and waveforms evoked by unidentifiable objects showed that incongruent objects started to diverge from unidentifiable ones later than congruent objects did. This provides first direct evidence for the effect of scene context on object identification; arguably, rapidly extracted gist activates scene-congruent schemas which facilitate the identification of congruent objects in comparison to incongruent ones.

1. Introduction

"No man is an island", wrote the 17th century English poet John Donne, and similarly – no object in our world appears in a void. Objects in the real world are always a part of a broader context, having relationships with the environment. Do these relations between objects and the contexts in which they appear affect the processes that underlie object identification? Even after more than three decades of research, the answer to this question is still debated.

Numerous behavioral studies have shown that context indeed influences subjects' ability to identify objects: when the object is incongruent with its context (i.e., violates the probability of appearing in the scene, like a lion in the kitchen; Biederman, 1981), subjects are both slower (e.g., Davenport and Potter, 2004; Neider and Zelinsky, 2006; Palmer, 1975; Rieger et al., 2008) and less accurate (Biederman, 1972; Boyce et al., 1989; Davenport and Potter, 2004; Underwood, 2005) in identifying it. Contextual facilitation is even more pronounced when the critical object is ambiguous (Brandman and Peelen, 2017).

Yet these results cannot provide univocal support for the claim that context indeed affects object identification itself: prolonged reaction times and poorer identification could stem from post-perceptual processes, in which the observer tries to reconcile the already identified incongruent object with the scene in which it appears (and to which it does not belong). Arguably then, object identification could still be independent of scene processing and isolated from it; only after the two have been identified, their semantic relations are being processed, hereby affecting subjects' performance (De Graef et al., 1990; Hamm et al., 2002; Hollingworth and Henderson, 1998).

This interpretation, termed the *functional isolation model* (Hollingworth and Henderson, 1999), was first strengthened by an EEG study that compared neural activity in response to congruent and incongruent scenes, and only found late differences between the conditions; namely, a greater negativity in the 300–500 ms time window was evoked by incongruent scenes, compared with congruent ones (Ganis and Kutas, 2003; see also Demiral et al., 2012). This negativity was reminiscent of the widely studied N400 component, first reported for incongruent sentence endings (Holcomb, 1993; Kutas and Hillyard, 1980a, 1980b) and commonly held to reflect semantic integration attempts (for review, see Kutas and Federmeier, 2011). Indeed, N400 was also found for other stimuli which seem to require semantic integration, like unrelated word pairs (Holcomb and Neville, 1990), real-world knowledge violations (Hagoort et al., 2004) or even mathematical errors (Niedeggen and Rösler, 1999). Much like the effect reported by

---

* Corresponding author at: School of Psychological Sciences, Aviv University, Ramat Aviv, POB 39040, Tel Aviv 69978, Israel.
E-mail address: mudrikli@tau.ac.il (L. Mudrik).

https://doi.org/10.1016/j.neuropsychologia.2018.06.004
Received 7 November 2017; Received in revised form 3 June 2018; Accepted 6 June 2018

Ganis and Kutas (2003), N400-like effects were reported for other visual, non-linguistic/symbolic stimuli, such as semantically unrelated pairs of images (Barrett and Rugg, 1990; Hamm et al., 2002; McPherson and Holcomb, 1999) or video clips with incongruent endings (Sitnikova et al., 2008, 2003).

However, more recent studies reported earlier differences between congruent and incongruent scenes, either at the 200–300 ms time window (Dyck and Brodeur, 2015; Mudrik et al., 2010, 2014; Sauvé et al., 2017; Võ and Wolfe, 2013) or even sooner, between 100 and 200 ms (Coco et al., 2017). These relatively early effects lend support to the *matching models* of object identification (e.g., Bar and Ullman, 1996; Bar, 2004), which assign a crucial role to scene-object relations. According to these models, when presented with a new scene, the gist is rapidly extracted during early perceptual processing. Though this initial analysis of the scene is still crude, it is held to suffice for activating schemas of scene-congruent objects, which are being compared with incoming visual information. This search for a match between activated schemas and visual input was suggested to guide object identification. Arguably then, incongruent objects are slower and less accurately identified because they do not share the properties of the contextually-preactivated schemas (or informed guesses about objects' identity; Trapp and Bar, 2015), hereby prolonging the search for a match (Mudrik et al., 2010, 2014). These accounts are further supported by a recent MEG and fMRI study, showing contextual effects in early visual areas and in object-selective areas (Brandman and Peelen, 2017).

Indeed, the timing and polarity of the reported effects in the 200–300 ms time window fit well with those of the N300 component, which is claimed to reflect object identification processes per se. The classical N300 was originally found for fragmented or degraded objects, which are naturally harder to identify (Doniger et al., 2000; Folstein et al., 2008; Schendan and Kutas, 2002, 2003). It is accordingly considered functionally distinct from the N400 (Hamm et. al, 2002), representing a difficulty to identify the object even before semantic integration takes place (Ganis and Kutas, 2003). As an extrapolation from that, the early negativities evoked by scenes including incongruent objects were also termed N300, and were held to represent genuine difficulties to identify such objects.

Yet the conjuncture that the incongruency-evoked N300 indeed represents identification difficulties was never directly examined. It could be argued that this component actually represents another process, uniquely induced by incongruent objects, which may not directly pertain to object identification. Such a claim is supported by the different scalp distributions of the two N300 components: a typically positive-occipital topography, sometimes accompanied by negative-frontal activity for the classical object-identification-N300 (Doniger et al., 2000 (note that there the reverse polarity was found due to the use of nose reference); McPherson and Holcomb, 1999; Sehatpour et al., 2006 (again with nose reference)), and a negative frontocentral topography of the incongruency-evoked one (Dyck and Brodeur, 2015; Mudrik et al., 2010, 2014). In effect, aside from the similar timing and partially similar polarity, there are no evidence which directly tie the incongruency-N300 with the difficulty to identify the incongruent object. Thus, only a direct comparison between scene-congruent, scene-incongruent and unidentifiable objects could provide clear support for contextual effects on object identification (two previous studies examined waveforms elicited by related and unrelated objects pairs, as well as unidentifiable objects; McPherson and Holcomb (1999) found N300 in both comparisons, again with a different distribution, and Holcomb and McPherson (1994) only reported later effects. Critically, these studies cannot shed light on the possible influence of scene context on object identification, as they used isolated objects and did not conduct a latency analysis).

In this study, such a direct comparison was achieved by manipulating two variables: object-scene congruency and object identifiability. EEG waveforms were compared for congruent/incongruent objects that were either intact or scrambled, specifically targeting the N300 and N400 time windows. Our objectives were threefold. First, to examine the reproducibility of the early incongruency-evoked N300, which was found in some studies (Dyck and Brodeur, 2015; Mudrik et al., 2010, 2014; Sauvé et al., 2017; Võ and Wolfe, 2013), but not others (Ganis and Kutas, 2003; Demiral et al., 2012). Second, to directly compare the incongruency-evoked N300 effect (Dyck and Brodeur, 2015; Mudrik et al., 2010, 2014) with the classical object-identification N300 (Doniger et al., 2000; Sehatpour et al., 2006), by inspecting their waveform, timing and distribution. Third, to specifically contrast activity evoked by congruent/incongruent intact and scrambled objects: this comparison could indicate when incongruent objects are distinguishable from unidentified objects, and whether this differentiation appears later for incongruent than congruent objects. Such a difference would support the notion that the identification of incongruent objects is indeed delayed, in line with matching models. Taken together, these comparisons allow us to assess the nature of the incongruency-N300 and the classical identification-N300, and shed more light on the ongoing debate between the matching and the functional isolation models.

2. Methods

2.1. Participants

Sixteen Tel Aviv University students (6 females, 15 right handed, mean age = 25.12, SD = 3.00) with reportedly normal or corrected-to-normal sight and no psychiatric or neurological history participated in the study for payment (~$14 per hour). Five additional subjects were excluded from the analysis due to two criteria: (a) excessive eye movements, muscular artifacts, or noisy recording resulting in too few trials in each condition (fewer than 20; four subjects) and (b) accuracy lower than 3 standard deviations from the group mean (one subject). The experiment was approved by the ethics committee of Tel Aviv University, and informed consent was obtained after the experimental procedures was explained to the subjects. The experiment was pre-registered in the Open Science Framework system (OSF experiment code: osf.io/rpgek; the raw data can also be found there), including sample size, predefined EEG analyses and exclusion criterion (a). Note that exclusion criterion (b) was added post-hoc; based on previous experiments (Mudrik et al., 2010, 2014) we did not anticipate big variation in subjects' accuracy, yet in this experiment, one subject substantially deviated from the overall mean, possibly since he did not pay any attention to task instructions, and was accordingly removed.

2.2. Stimuli

Thirty-nine pairs of congruent and incongruent scenes in which the critical objects were intact and 39 pairs of congruent and incongruent scenes in which the critical objects were scrambled were used in this experiment (Fig. 1). To build the stimuli bank, we first chose a subset of 92 pairs out of a larger stimuli bank of real-life scenes which depict a person performing an action that involves either a congruent or an incongruent object (Mudrik et al., 2010, 2014). To create the original stimuli bank, images of people performing actions (e.g., a bride throwing her bouquet; see Fig. 1) were first chosen from Internet resources. Then, both the congruent and incongruent objects were replaced using Photoshop software, either by another exemplar of the same object (congruent), or by a different object which was unrelated to the scene and was similar in size to the original object (incongruent; e.g., a dynamite). Thus, incongruent scenes included a contextual violation based on the probability to appear in the scene, and not on size/color modifications (Biederman et al., 1982). The congruency of all stimuli was validated by naïve subjects, and the stimuli were further tested for low-level visual differences (in chromaticity, spatial frequency or contrast), which were not found (for details, see Mudrik et al., 2010). For each pair of scenes, versions with scrambled objects

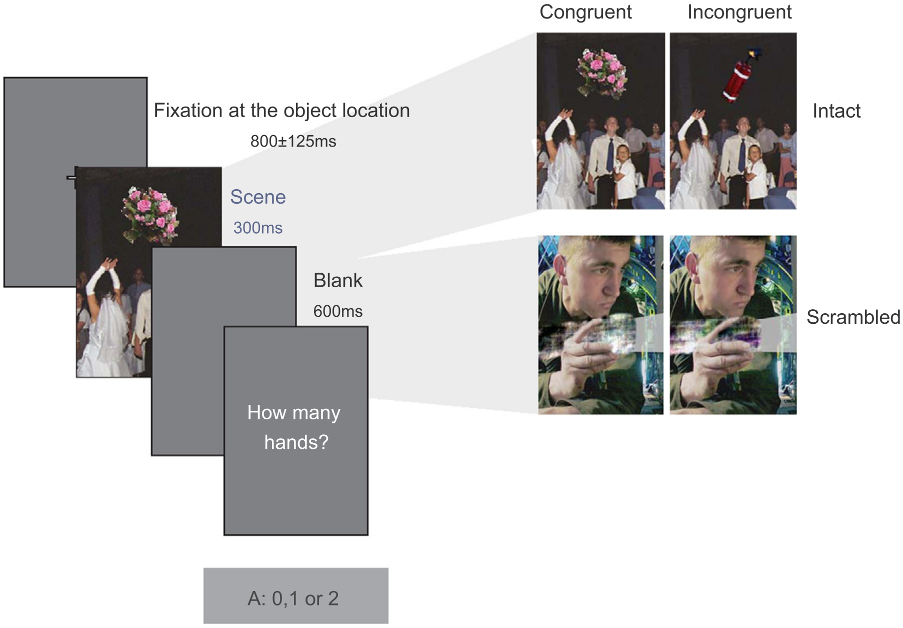

Fig. 1. Experiment timeline. Each trial started with a 800 ± 125 ms fixation cross at the location of the to-be-presented critical object in the scene. Then the scene appeared for 300 ms, followed by the question "How many hands?" which referred to the number of hands that touched the critical object in the scene. Subjects typed their responses using the keys 0, 1 and 2. Note that differnt scenes with intact and scrambled objects were used, so that subjects would not infer the identity of the scrambled object based on a previous exposure to an inteact version of it.

were created using Photoshop and a phase scrambling Matlab code: a rectangle including the relevant object (congruent/incongruent) was cut from the scene using Photoshop. It was then scrambled in Matlab, and inserted back to the scene using Photoshop to replace the original object. The object was not inserted as a rectangle; instead, it was cut in a shape that was largely similar to the shape of the original object, keeping its overall size similar, yet blurring the exact details of the shape to avoid identification of the original object (see again Fig. 1). Using this phase-scrambling method, low-level features like chromaticity and luminance were kept while object identity was rendered unidentifiable.

To make sure the scrambled objects were indeed unidentifiable, an online Qualtrics pretest ($N = 200$, 100 males, mean age $= 41.94$, $SD = 13.72$) was conducted in which either the scrambled-congruent or the scrambled-incongruent version of each scene was presented. Subjects were asked to try and identify the objects, while ignoring the context, since it cannot always provide information about the object's identity. Fifty-three pairs of images were excluded as the congruent version differed from the incongruent version in one of two ways: (1) percentage of participants that managed to recognize the actual object, (2) number of participants that did not recognize the actual object, but still gave a context-matching answer (for example, in a context of a man drinking out of scrambled bottle, responding 'cup' would be defined as a context-matching answer). For the remaining 39 pairs, subjects were unable to identify the scrambled object, with no difference between the congruent condition ($M = 22.05$, $SD = 15.75$) and the incongruent one ($M = 22.17$, $SD = 16.81$; $t(38) = 0.034$, $p = 0.972$, Cohen's $d = 0.008$). In addition, they did not provide more context-matching answers for congruent ($M = 29.10$ $SD = 14.40$) than incongruent objects ($M = 27.94$, $SD = 14.84$; $t(38) = 0.349$, $p = 0.728$, $d = 0.079$). Out of the excluded 53 pairs, 39 pairs were selected for the intact condition, so to avoid a repetition between scrambled and intact scenes which might have facilitated object identification of the scrambled objects. Thus, the final stimuli bank included 39 pairs of scenes including intact congruent/incongruent objects, and 39 other pairs of scenes including scrambled congruent/incongruent objects.

2.3. Apparatus

Subjects sat in a dimly lit room. The stimuli were presented on a LCD monitor (23" ASUS SyncMaster) with 1920 * 1080 resolution and 100 Hz refresh rate, using Matlab and psychtoolbox 3, (Brainard, 1997; Pelli, 1997). The stimuli appeared on a grey background (RGB: 128, 128, 128) at the center of the computer screen and subtend 6.67° (width) 9.52° (height) of visual angle. The screen was located 60 cm away from subjects' eyes.

2.4. Procedure

The experiment included 312 trials that differed in congruency (congruent/incongruent) and identifiability (intact/scrambled). Each subject saw all 156 images twice, presented in four different blocks, with 78 images per block. Within each block, half the images were congruent and half incongruent, half intact and half scrambled with no repetition within a pair during the block. Thus, each scene either appeared with a congruent or an incongruent object, which in turn was either intact or scrambled. Trial order was pseudo-randomly intermixed, with the constraint that object identifiability (intact/scrambled) was not repeated for four consecutive trials (and, consequently, object congruency in intact images could not be repeated for four consecutive trials). The session began with eight practice trials (two of each type, with different images than the ones which were subsequently used in the experiment), in which subjects performed the task in the experimenter's presence, to ensure that they follow the instructions correctly. Trials were self-paced.

In each trial, a fixation cross was presented for 800 ± 125 ms at the location of the to-be-presented critical object in the scene. Immediately after the cross, the scene appeared for 300 ms, followed by a blank of 600 ms, and then subjects were presented with the question: "How many hands?" which referred to the number of hands used by the person in the image, touching the critical object. The question was aimed at having subjects focus on the action performed by the person in the image (that involved the object), without explicitly asking about the congruency of the objects. Subjects typed their responses using the keys 0, 1 and 2 (Fig. 1).

2.5. EEG recording

EEG was recorded with an Active 2 system (BioSemi, the Netherlands) at 512Hrz, using a 64-electrodes cap arranged according to the extended 10–20 system. Seven external electrodes were placed on the mastoids (two electrodes), tip of the nose (one electrode) and

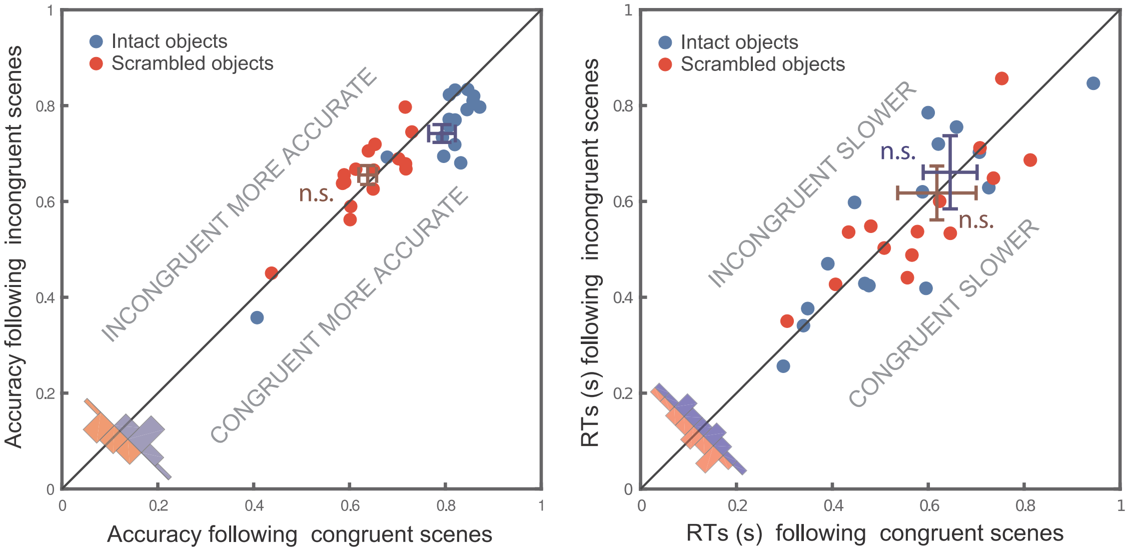

Fig. 2. Behavioral results for Accuracy (Left) and RTs (right) in the different experimental conditions. Each dot represents an individual subject, with the coordinates denoting accuracy/RT for congruent scenes (x-coordinate) and incongruent ones (y-coordinate), when the objects were either intact (blue) or scrambled (red). The diagonal line represents equal accuracy/RT for congruent and incongruent scenes. Data points below the diagonal indicate higher accuracy or longer RT for congruent scene, while data point above the diagonal indicate higher accuracy or longer RT for incongruent scene. The histograms at the left corner of each plot sum the number of the dots with respect to the diagonal line. Note that for accuracy, the blue histogram (i.e., intact objects) is asymmetrical and centered below zero, while this is not the case for the red histogram (i.e., scrambled objects).

around the eyes the EOG (four electrodes: two located at the outer canthi of the right and left eyes and two above and below the center of the right eye). All electrodes were referenced online to a common-mode signal (CMS) electrode between POz and PO3. High-pass filter of 0.05Hrz was used in the online recording.

2.6. ERP analysis

Offline analysis was conducted using the “Brain Vision Analyzer” software (version 2.1; Brain Products, Germany). The data was re-referenced to the average of both mastoids, in line with previous studies (Ganis and Kutas, 2003; Mudrik et al., 2010). It was digitally high-pass filtered at 0.1 Hz (24 dB/octave) to remove slow drifts, using a Butterworth zero-shift filter. Bipolar EOG channels, which facilitate the detection of blinks and saccades activity, were calculated by subtracting the left from the right horizontal EOG channels, and the inferior from the superior vertical EOG channels. They then served for artifact detection and removal using independent component analysis (ICA) (Jung et al., 2000). Additional detection of other artifacts included segments that had a steep gradient change of 100 μV/ms, amplitude exceeding ± 100 μV, differences beyond 100 μV within a 150 ms interval, or activity below 0.5 μV for over 100 ms, in any channel. Data including such artifacts was discarded from further analysis.

The preprocessed data was then segmented into 1000-ms long epochs starting 100 ms prior to scene onset. The segments were averaged separately for each of the four conditions (congruent intact/ congruent scrambled/ incongruent intact/ incongruent scrambled), with an average number of 65.78 segments (range: 40–80) in each condition. Segmented data for each subject in each condition was filtered with low-pass 30 Hz filter, and baseline corrected by subtracting the mean amplitude of the pre-stimulus period of each ERP from all the data points in the segment.

To do reduce the number of comparisons, the 64 electrodes were pooled into nine areas (see Mudrik et al., 2010, 2014), arranged by laterality and area, using the following division: Left Frontal [Fp1, AF3, AF7, F3, F5, F7]; Middle Frontal [Fpz, AFz, Fz, F1, F2]; Right Frontal [Fp2, AF4, AF8, F4, F6, F8]; Left Central [FC3, FC5, FT7, C3, C5, T7, CP3, CP5, TP7]; Middle Central [FCz, FC1, FC2, Cz, C1, C2, CPz, CP1, CP2]; Right Central [FC4, FC6, FT8, C4, C6, T8, CP4, CP6, TP8]; Left Parieto-Occipital [P3, P5, P7, P9, PO3, PO7, O1]; Middle Parieto-Occipital [Pz, P1, P2, POz, Oz, Iz]; Right Parieto-Occipital [P4, P6, P8, P10, PO4, PO8, O2].

EEG results were assessed in two ways. First, a four-way ANOVA[1] with Congruency (congruent, incongruent) X Identifiability (intact, scrambled) X Area (frontal, central, occipital) X Laterality (left, middle, right) was conducted on the average amplitudes in two different time windows: 200–300 ms, aimed at detecting N300, and 300–500 ms, aimed at detecting N400. Greenhouse–Geisser correction was used where appropriate. The uncorrected degrees of freedom are reported along with the Greenhouse–Geisser epsilon values (Picton et al., 2000).

A cluster-based non-parametric permutation test (Maris and Oostenveld, 2007) with 2000 permutations was applied on all 64 electrodes at all time points in the time window of interest (200–500 ms) to assess the latency and distribution of the effects, and specifically – to detect the point in time in which the waveforms induced by incongruent/congruent intact objects diverged from scrambled objects.

3. Results

3.1. Behavioral results

Subjects' responses to the hands task were analyzed by averaging mean accuracy scores as well as RTs, and conducting a two-way ANOVA with Congruency (congruent, incongruent) and Identifiability (intact, scrambled). For accuracy (Fig. 2, left), a main effect of identifiability was found ($F(1,15) = 68.05$; $p < 0.001$, $\eta^2 = 0.819$), so that subjects were more accurate in judging scenes with intact objects ($M = 76.73\%$, $SD = 11.49$) compared with scrambled ones ($M = 64.65\%$, $SD = 7.58$). A marginal main effect of Congruency was also found ($F(1,15) = 4.31$; $p = 0.056$, $\eta^2 = 0.223$). Importantly, an interaction between Congruency and Identifiability was obtained ($F(1,15) = 20.36$;

---

[1] Note that in the OSF pre-registration form, we wrote “two-way ANOVA” by mistake; our original intention was to conduct a four-way ANOVA, following the analysis in Mudrik et al., (2010, 2014).

Table 1
Summary of all ANOVA results for the N300 (light grey columns) and N400 (dark grey columns) time windows.

	N300 time window			N400 time window		
	F	p	η^2	F	p	η^2
Congruency	1.01	0.332	0.063	12.54	0.003	0.455
Intactness	0.01	0.940	0.000	4.68	0.047	0.238
Area	87.16	0.000	0.853	101.41	0.000	0.871
Laterality	36.39	0.000	0.708	9.37	0.001	0.385
Congruency *Intactness	0.45	0.515	0.029	3.50	0.081	0.189
Congruency*Area	1.91	0.166	0.113	4.45	0.044	0.229
Intactness*Area	2.89	0.096	0.162	30.06	0.000	0.667
Congruency*Intactness*Area	0.12	0.788	0.008	1.72	0.209	0.103
Congruency*Laterality	0.56	0.577	0.036	1.88	0.170	0.111
Intactness *Laterality	6.49	0.005	0.302	3.48	0.044	0.188
Congruency*Intactness*Laterality	0.37	0.623	0.024	0.137	0.873	0.009
Area*Laterality	11.77	0.000	0.440	5.88	0.007	0.281
Congruency*Area*Laterality	0.22	0.925	0.015	0.78	0.542	0.050
Intactness*Area*Laterality	7.94	0.000	0.346	7.55	0.000	0.335
Congruency*Intactness*Area*Laterality	2.80	0.065	0.157	2.15	0.085	0.125

$p < 0.001$, $\eta^2 = 0.576$): subjects were more accurate for congruent intact ($M = 79.28\%$, $SD = 11.15$) than incongruent intact scenes ($M = 74.18\%$, $SD = 11.49$; $t(15) = 4.55$, $p < 0.001$, $d = 0.450$), but not scrambled ones ($M_{cong} = 63.79\%$, $SD_{cong} = 7.43$, $M_{inc} = 65.51\%$, $SD_{inc} = 7.88$; $t(15) = 1.56$, $p = 0.139$, $d = -0.224$). No effects were found for RTs (Intactness ($F = 1.87$); Congruency ($F = 0.22$); Intactness X Congruency ($F = 0.1$)) (Fig. 2, right. Note that this was not a speeded task, so that subjects were not instructed to answer as quickly as they can).

3.2. EEG results

3.2.1. N300 time window (200–300 ms)

A repeated measures ANOVA conducted on the mean EEG waveforms in the N300 time window per region revealed no main effects (identifiability: $F = 0.006$, $p = 0.940$, $\varepsilon = 1$, $\eta^2 = 0.000$; congruency: $F = 1.006$, $p = 0.332$, $\varepsilon = 1$, $\eta^2 = 0.063$). Yet interactions were found between identifiability and laterality ($F(2,30) = 6.49$, $p = 0.005$, $\varepsilon = 0.98$, $\eta^2 = 0.302$), and between identifiability, area and laterality ($F(4,60) = 7.94$, $p < 0.0001$, $\varepsilon = 0.68$, $\eta^2 = 0.346$). A marginally significant interaction between congruency, identifiability, area and laterality was also found ($F(4,60) = 2.80$, $p = 0.065$, $\varepsilon = 0.60$, $\eta^2 = 0.157$; For a summary of all ANOVA results, see Table 1. For the mean waveforms, see Fig. 3). Post-hoc Tukey comparisons showed that intact and scrambled objects mostly differed in right electrodes ($M = 0.65$, $SD = 2.61$, $p = 0.035$, $d = 0.108$; means and SDs refer to the difference between intact and scrambled objects across the three areas in each laterality) but not medial ($M = -0.24$, $SD = 3.14$, $p = 0.831$, $d = 0.038$) and left ones ($M = -0.25$, $SD = 2.55$ $p = 0.821$, $d = 0.048$). For the congruency effect, incongruent intact objects differed from congruent intact ones in four areas (medial frontal: $M = -0.50$, $SD = 2.42$, $p = 0.023$, $d = 0.11$; right frontal: $M = -0.63$, $SD = 2.04$, $p = 0.0006$, $d = 0.133$; left central: $M = -0.57$, $SD = 1.42$, $p = 0.003$, $d = 0.207$; right central: $M = -0.52$, $SD = 1.67$, $p = 0.012$, $d = 0.152$; here, means and SDs refer to the difference between intact congruent and incongruent objects), and were undistinguished in others (left frontal: $M = -0.40$, $SD = 1.84$, $p = 0.209$, $d = 0.094$; medial central: $M = -0.43$, $SD = 2.17$, $p = 0.112$, $d = 0.11$; left occipital: $M = -0.23$, $SD = 1.49$, $p = 0.981$, $d = 0.056$; medial occipital: $M = -0.21$, $SD = 1.63$, $p = 0.996$, $d = 0.047$; right occipital: $M = 0.03$, $SD = 1.70$, $p = 1.000$, $d = 0.008$).

The finding of a marginally significant interaction with congruency was followed by two analyses. First, a non-parametric cluster-based permutation analysis that assessed the moment in time where scenes including congruent and incongruent intact objects started to elicit diverging waveforms (Fig. 4a and dark orange significance patches in Fig. 3), which was 260 ms post stimulus onset. The same analysis was conducted for scrambled vs. intact objects; this revealed earlier differences, starting at 240 ms, and a different distribution of the observed clusters (Fig. 4b and light orange significance patches in Fig. 3). The second analysis examined the functional significance of the potential N300 effect by correlating it with subjects' reaction times (i.e., correlations between difference in EEG amplitude in congruent intact vs. incongruent intact and difference in RTs in these conditions). The motivation for this analysis was to examine if the N300 effect we found indeed indexes a genuine mechanism involved in object-scene processing, especially since the ANOVA only yielded marginally significant results within this time window. We focused on frontocentral areas in which the effect was obtained. A correlation was found in all six areas, and two survived a Benjamini-Hochberg correction for multiple comparisons (Left frontal: $r = 0.55$, $p = 0.027$; ***Medial frontal:* $r = 0.63$, $p = 0.009$**; ***Right frontal:* $r = 0.65$, $p = 0.006$**; Left central: $r = 0.51$, $p = 0.045$; Medial central: $r = 0.48$, $= 0.058$; Right central: $r = 0.51$, $r = 0.045$; uncorrected values, highlighted are the correlations that survived correction; Fig. 5). Similar correlations were not found for the identification-N300 component (examined for the three occipital areas, where the effect was found; all p-values > 0.5).

Based on the cluster-based permutation analysis, we focused on the left-central area (from which the congruency cluster originated), in a post-hoc, exploratory inspection aimed at directly examining the time courses of congruency vs. identifiability (Fig. 6). There, a dissociation between the waveforms patterns during the N300 vs. N400 time windows was observed, and confirmed in the post-hoc Tukey comparisons: during the N300 time window, the congruent intact waveform ($M = -5.49$, $SD = 2.65$) differed from all other waveforms (congruent scrambled: $M = -5.97$, $SD = 2.18$, $p = 0.031$, $d = 0.200$; incongruent scrambled: $M = -6.15$, $SD = 2.14$, $p = 0.0003$, $d = 0.274$; incongruent intact: $M = -6.06$, $SD = 2.84$, $p = 0.0003$, $d = 0.207$), while the incongruent intact waveform did not differ from the ones evoked by scrambled objects (incongruent scramble, $p = 1.000$, $d = 0.036$; congruent scramble, $p = 1.000$, $d = 0.032$). In the N400 time window (see full analysis in Section 3.2.2 below), this pattern flipped: here, incongruent intact objects elicited a stronger negativity ($M = -2.86$, $SD = 3.37$) compared with congruent intact objects ($M = -1.77$, $SD = 3.38$, $p = 0.0002$, $d = -0.324$), with congruent scrambled objects ($M = -2.00$, $SD = 3.23$, $p = 0.0003$, $d = -0.262$) and also marginally when compared with incongruent scrambled ones ($M = -2.28$, $SD = 3.26$, $p = 0.071$, $d = -0.175$). No differences were found between congruent intact and both scrambled conditions in the N400 time window (congruent scramble: $p = 0.998$, $d = 0.069$; incongruent scramble: $p = 0.210$, $d = 0.155$).

Finally, to directly test the effect of incongruency on object

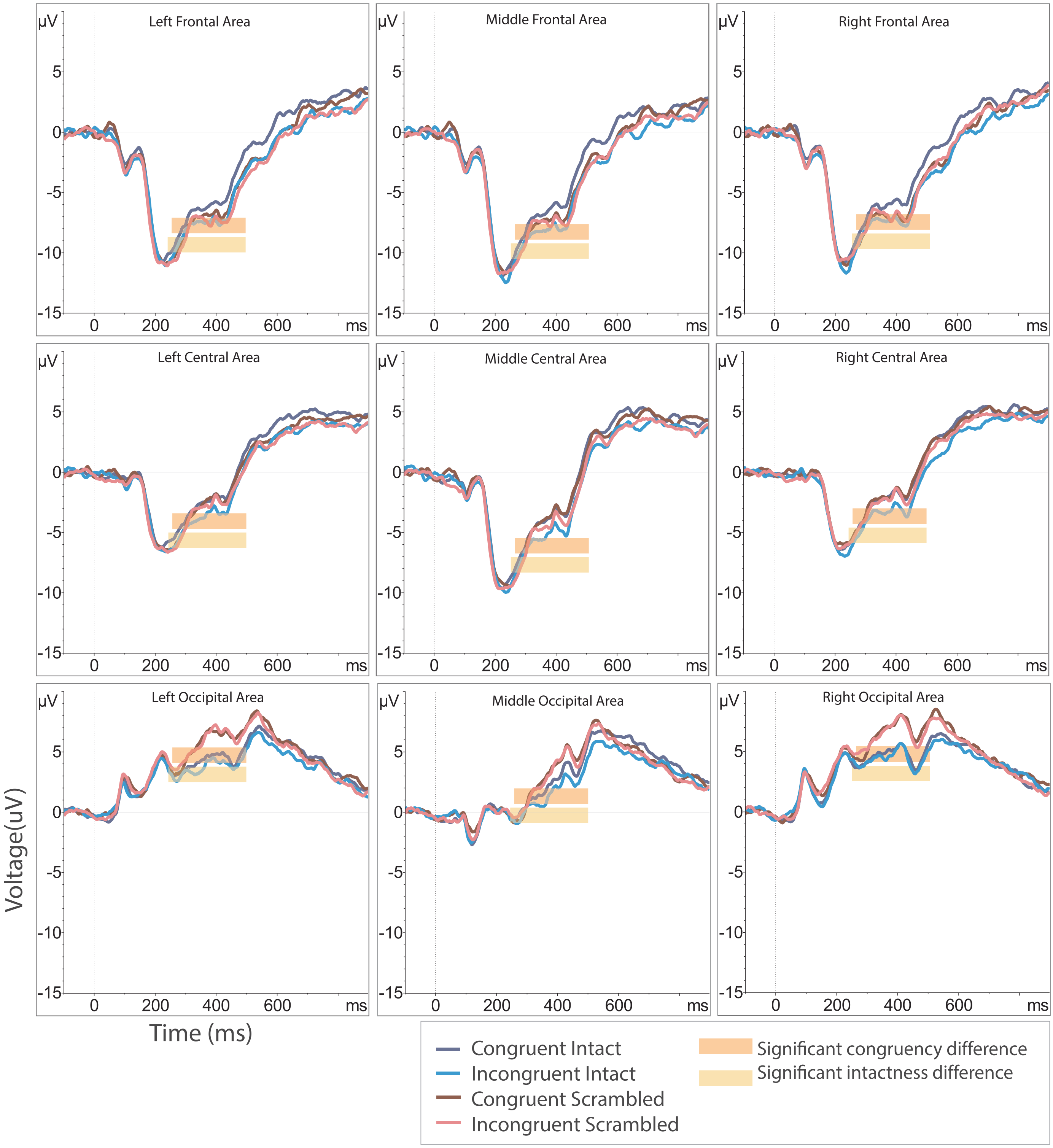

Fig. 3. ERP responses to congruent intact (dark blue), incongruent intact (light blue), congruent scrambled (dark red) and incongruent scrambled (pink) objects for each region. Zero denotes the time of scene presentation. Dark orange patches represent the time window over which a significant cluster of differential activity between congruent intact and incongruent intact scenes was detected in the non-parametric cluster-based permutation analysis (which was limited to the 200–500 ms time window). Light orange patches represent the time window over which a significant cluster of differential activity between intact and scrambled objects was found.

identification, we focused on the divergence point between congruent/incongruent intact objects and scrambled objects. If indeed incongruent objects are harder to identify, we expect them to become distinguishable from scrambled incongruent objects at a later time point, compared to the time when congruent objects become distinguishable from scrambled congruent objects. The non-parametric permutation analysis was thus run to contrast waveforms elicited by congruent intact vs. congruent scrambled objects, and then by incongruent intact vs. incongruent scrambled objects. As expected, congruent intact objects started to diverge from congruent scrambled objects earlier (cluster

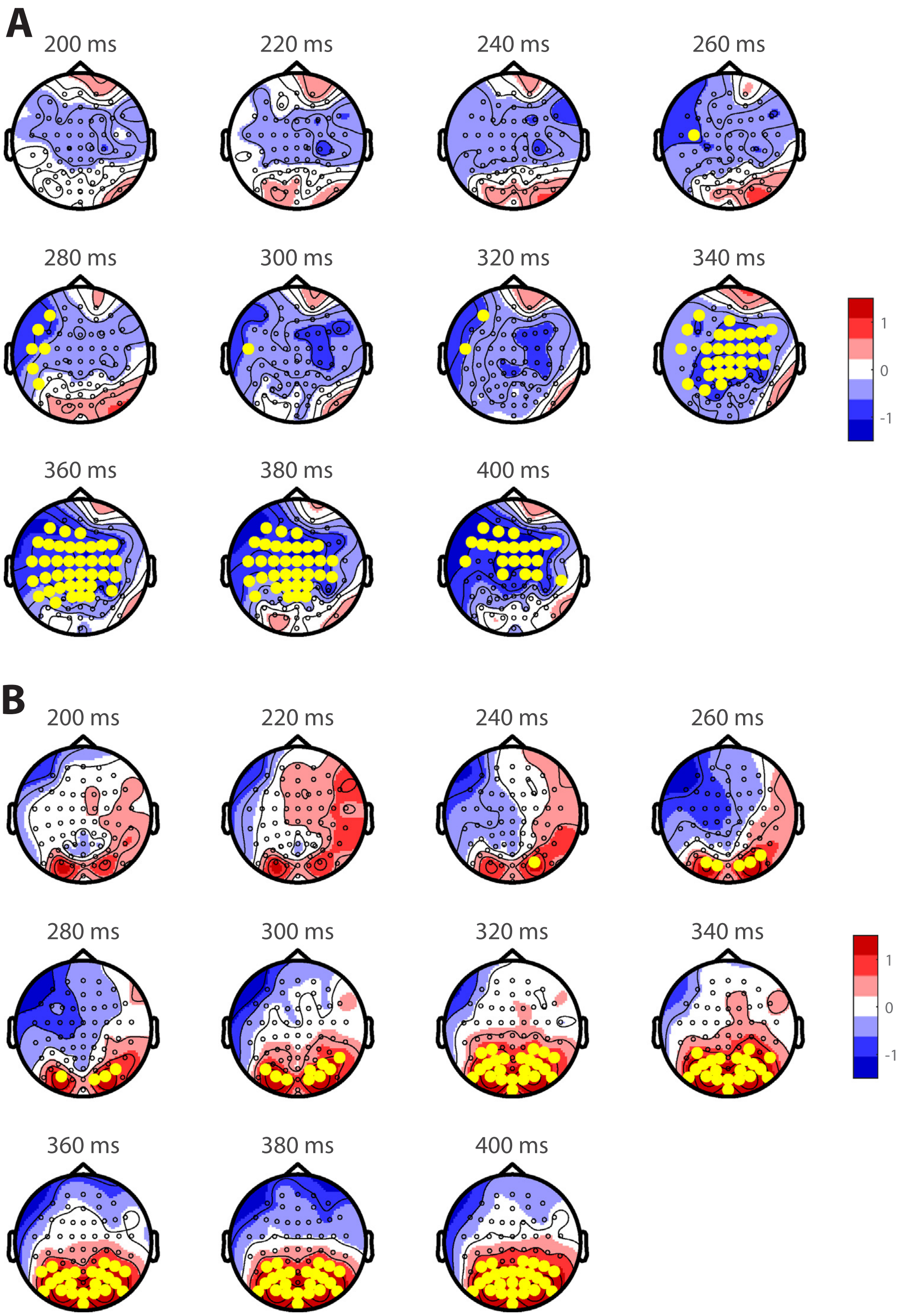

Fig. 4. A. Distribution maps of the congruency effect over time, showing a frontocentral distribution. Electrodes in which a significant effect ($p < 0.05$) was found at each time point are marked with a yellow asterisk. B. Distribution maps of the intactness effect (scrambled–intact) over time, showing a bipolar pattern with greater positivity at occipital sites.

starting point = 240 ms; the difference persisted throughout the entire analysis window, which ended at 500 ms) than incongruent intact objects started to diverge from incongruent scrambled objects (cluster starting point = 260 ms; again, the effect remained significant until the end of the analysis window).

3.2.2. N400 time window (300–500 ms)

A similar four-way ANOVA was conducted on the average amplitudes of the waveforms during the 300–500 ms time window. Main effects of congruency ($F(1,15) = 12.54$, $p = 0.003$, $\varepsilon = 1$, $\eta^2 = 0.455$) as well as an interaction between congruency and area ($F(2,30) = 4.45$, $p = 0.044$, $\varepsilon = 0.595$, $\eta^2 = 0.229$) were found (see again Fig. 3 for the waveforms). A main effect of identifiability was also found ($F(1,15) = 4.68$, $p = 0.047$, $\varepsilon = 1$, $\eta^2 = 0.238$), as well as interactions between identifiability and area ($F(2,30) = 30.06$, $p < 0.0001$, $\varepsilon = 0.613$, $\eta^2 = 0.667$), identifiability and laterality ($F(2,30) = 3.48$, $p = 0.044$, $\varepsilon = 0.849$, $\eta^2 = 0.188$), and a triple interaction of identifiability, area and laterality ($F(4,60) = 7.55$, $p < 0.0001$, $\varepsilon = 0.735$, $\eta^2 = 0.335$). Finally, a marginal interaction was found between congruency, identifiability, area and laterality ($F(4,60) = 2.15$, $p = 0.085$, $\varepsilon = 6.07$, $\eta2 = 0.125$). Post-hoc Tukey tests revealed that waveforms elicited by incongruent objects were more negative than those elicited by

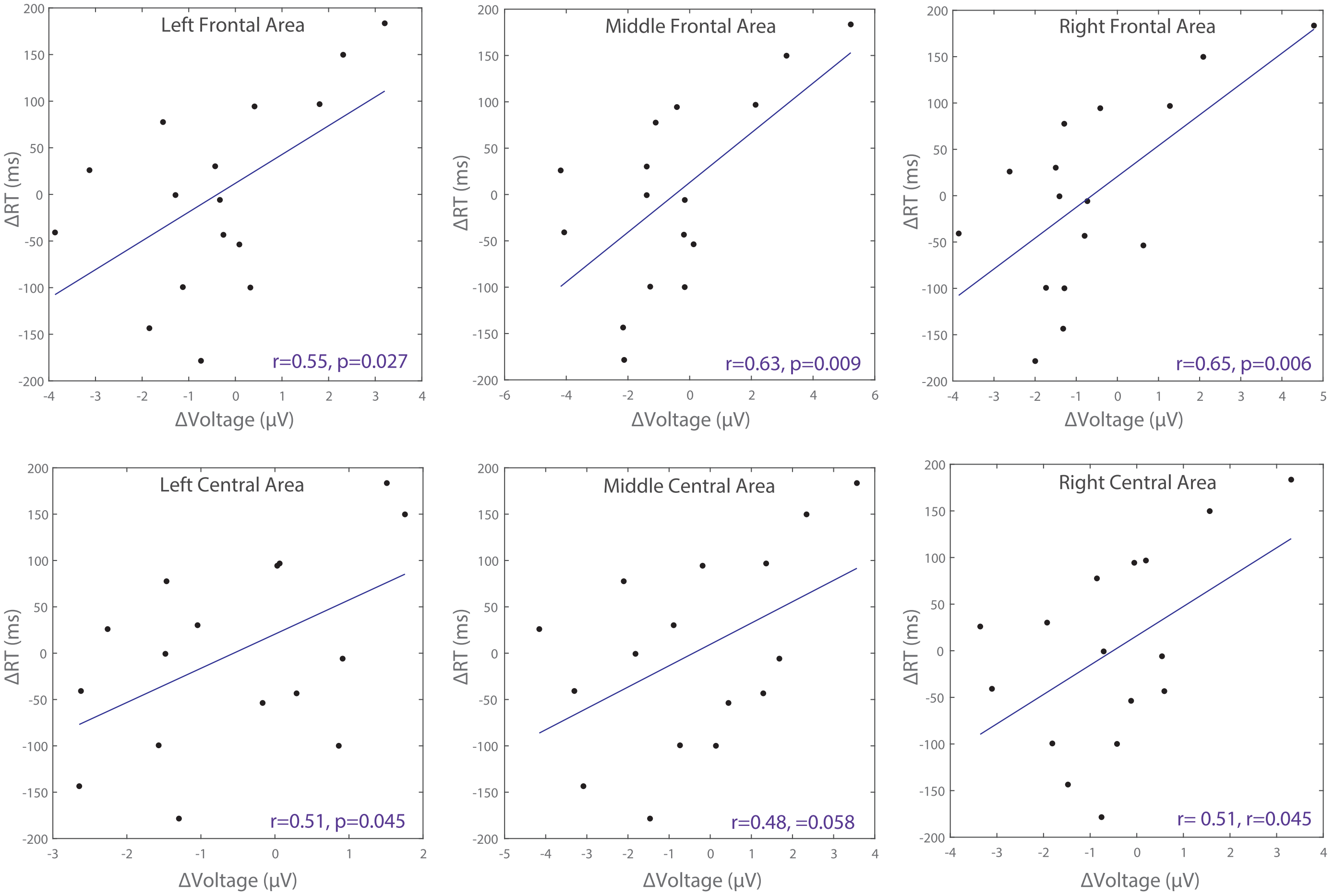

Fig. 5. Correlation between the difference in RTs (Incongruent Intact-Congruent Intact) and differences in N300 amplitude (Incongruent Intact-Congruent Intact). Black dots represent the data, while the blue line represents linear fit for the data. Correlations are presented together with uncorrected p-values.

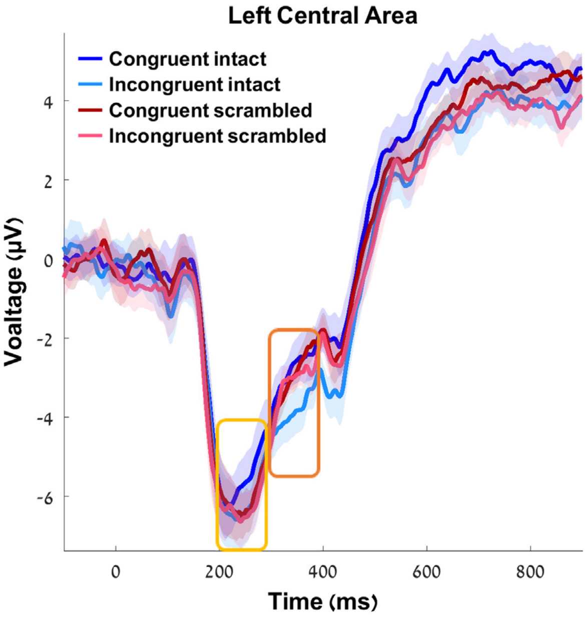

Fig. 6. Averaged waveforms in the left central area for congruent and incongruent, intact and scrambled objects. Shaded areas represent SEs across subjects. Note how the incongruent intact waveform is indistinguishable from scrambled objects waveforms in the N300 time window (highlighted by a light-orange frame; all three waveforms differ from the one evoked by congruent objects), and its divergence from them in the N400 time window (highlighted by a darker orange frame).

congruent ones in frontal and central areas (frontal: M = −1.71, SD = 2.15, p = 0.0001, d = 0.228; central: M = −1.62, SD = 1.58, p = 0.0001, d = 0.250. Means and SD refer to the difference between congruent and incongruent objects) but not occipital ones (M = −0.69, SD = 1.43, p = 0.139, d = 0.091). A difference between intact and scrambled objects was found in all frontal and all occipital areas, and in the right central area (left frontal: M = −1.15, SD = 2.37, p = 0.0002, d = 0.145; medial frontal: M = −0.69, SD = 2.62, p = 0.044, d = 0.083; right frontal: M = −0.74, SD = 2.75, p = 0.019, d = 0.103; right central: M = 1.01, SD = 1.90, p = 0.0003, d = 0.186; left occipital: M = 3.97, SD = 3.07, p = 0.0002, d = 0.478; medial occipital: M = 2.67, SD = 2.96, p = 0.0002, d = 0.317; right occipital: M = 4.24, SD = 2.98, p = 0.0002, d = 0.567; means and SD refer to the difference between intact and scrambled objects), but not in the medial central (M = 0.53, SD = 2.63, p = 0. 318, d = 0.063) or the left central areas (M = 0.35, SD = 2.01, p = 0. 903, d = 0.053). Incongruent intact objects differed from congruent intact objects in all areas but the left and right occipital ones (left frontal: M = −1.33, SD = 1.60, p = 0.0002, d = 0.317; medial frontal: M = −1.58, SD = 2.12, p = 0.0002, d = 0.366; right frontal: M = −1.40, SD = 1.85, p = 0.0002, d = 0.361; left central: M = −1.10, SD = 1.21, p = 0.0002, d = 0.324; medial central: M = −1.25, SD = 1.67, p = 0.0002, d = 0.292; right central: M = −1.08, SD = 1.32, p = 0.0002, d = 0.374; left occipital: M = −0.53, SD = 1.07, p = 0.179, d = 0.136; medial occipital: M = −0.82, SD = 1.27, p = 0.0005, d = 0.200; right occipital: M = −0.28, SD = 1.43, p = 0.994, d = 0.079; means and SD refer to the difference between intact congruent and incongruent objects). Incongruent intact objects also differed from incongruent scrambled objects in occipital sites (left occipital: M = 2.23, SD = 1.83, p = 0.0002, d = 0.516; medial occipital: M = 1.67, SD = 2.05, p = 0.0002, d = 0.367; right occipital: M = 2.15, SD = 2.03, p = p = 0.0002, d = 0.528; means and SD refer to the difference between incongruent intact and scrambled objects), in the right central area (M = 0.82, SD = 1.18, p = 0.0005, d = 0.272), and marginally in the left central area (M = 0.58, SD = 1.02, p = 0.071, d = 0.175), but not in frontal and medial central areas (left frontal: M = −0.17, SD = 1.44, p = 1.000, d = 0.038; medial frontal: M = 0.30, SD = 1.70, p = 0.984, d = 0.067; right frontal: M = 0.33, SD = 1.96, p = 0.950, d = 0.083; medial central: M = 0.53, SD = 1.57, p = 0.180, d = 0.120).

Akin to the analysis conducted on the N300 effect, we examined the correlation between the N400 magnitude and subjects' reaction times; this showed a correlation in two areas only, yet none survived a Benjamini-Hochberg correction (Left frontal: r = 0.36, p = 0.175; Medial frontal: r = 0.56, p = 0.025; Right frontal: r = 0.58, p = 0.017; Left central: r = 0.26, p = 0.328; Medial central: r = 0.22, p = 0.421; Right central: r = 0.38, r = 0.147; uncorrected values). Like in the N300 time window, no correlations were found between the amplitude of the difference waves between the intact and scrambled conditions and subjects' reaction times (all p-values > 0.26).

4. Discussion

The ongoing debate about the potential effects of context on object identification has evoked substantial scientific interest (e.g., Bar, 2003; Ganis and Gutas, 2003; Mudrik et al., 2010, 2014; Demiral et al., 2012), especially given its wider implications on the more fundamental question of whether perception is cognitively impenetrable (for a recent criticism and replies, see Firestone and Scholl, 2016). That is, is perception encapsulated from cognition (e.g., Fodor, 1983; Pylyshyn, 1980), or – alternatively – can cognitive factors, like semantic knowledge about the probability of objects to appear in specific scenes, affect perceptual processing and shape the way stimuli are perceived. The results of the current study may tip the scale in favor of cognitive penetrability, at least when applied to contextual effects on object identification: first, by replicating the finding of congruency differences in the N300 time window in a new set of subjects, and second, by directly comparing divergence times between congruent/incongruent intact objects and scrambled objects; this comparison allows us to roughly assess when the two types of intact objects were marked as different from scrambled, meaningless patches of blurred colors (i.e., scrambled objects). This moment occurred 20 ms later for incongruent objects, implying that their initial identification may indeed be delayed.

The evidence found in this study in favor of an N300 congruency effect rest on several findings. First, the results of the cluster-based permutation analysis, showing initial differences between congruent and incongruent objects 260 ms after the scenes were presented. Interestingly, the cluster we found here was similar in timing and spatial course to our previous findings (Mudrik et al., 2010, 2014). This was corroborated to some extent by the ANOVA results, which only yielded a marginally significant interaction between congruency, identifiability, area and laterality. Yet post-hoc contrasts further confirmed that congruent and incongruent intact objects differed in four frontocentral areas. Finally, additional support for the existence of the congruency-N300 effect was found in correlations between the amplitude of this component and subjects' reaction times for congruent vs. incongruent intact objects. Taken together, this suggests that contextual effects indeed occur already in the N300 time window, hereby providing further evidence for early contextual effects on scene processing. This finding, which aligns with some studies (Dyck and Brodeur, 2015; Mudrik et al., 2010, 2014; Võ and Wolfe, 2013) but not all (Ganis and Kutas, 2003; see also Demiral et al., 2012), was previously interpreted as strengthening matching models (Bar and Ullman, 1996; Bar, 2004). According to these models, objects' identification is heavily influenced by the context of the scene: rapid gist extraction triggers the activation of gist-congruent schemas. Then, matching procedures – arguably indexed by the N300 components – take place, so that these activated schemas are compared with upcoming visual information. When a match is found, object identification is achieved. Accordingly, incongruent objects are harder to identify given the discrepancy between the activated schemas and the upcoming visual information, prolonging matching procedures and evoking a greater N300. The findings thus go against the functional isolation model (Hollingworth and Henderson, 1999), which claimed that context only affects later stages of scene processing, after both scene and objects have been identified.

Critically however, our findings clearly show that the incongruency N300 and the classical object-identification N300 differ with respect to latency, polarity and distribution (see also McPherson and Holcomb, 1999). While the classical N300 started at 240 ms, arguably marking the time when objects are first marked as different from scrambled visual input, the incongruency N300 started a bit later, at 260 ms, indicating when object-scene relations start to affect object processing (note however that this is an indirect comparison, relying to two separate analyses, each one targeting the time at which each object type diverged from scrambled objects). The relative proximity in timing is interesting, as one might expect the classical N300 to start even earlier given the visual differences between intact and scrambled objects. Indeed, the object-identification N300 has a more occipital distribution (in line with previous studies; Doniger et al., 2000; Sehatpour et al., 2006), suggesting that it may stem from visual occipital areas that are involved in object recognition (Grill-Spector et al., 1998; Malach et al., 1995; Moscovitch et al., 1995) as well as from earlier striate responses to the difference in contours between the conditions (Kovács, 1996). These lower-level differences in contours and spatial frequency may drive at least some of the identification-N300 component found here; in this study, identifiability was manipulated by scrambling the objects, hereby creating stimuli that are very visually different from the intact, identifiable ones. And so, with the current design, these two factors cannot be disentangled. Note that previous studies reporting the identification-N300 also suffered from a similar limitation to some extent (Doniger et al., 2000; Sehatpour et al., 2006; though less so in McPherson and Holcomb, 1999). Future studies should aim at manipulating identifiability while minimizing low-different levels. For example, by using meaningless objects rather than scrambled ones.

Given the different distributions of the identification-N300 and the congruency-N300, one could claim that the temporal overlap between the two components does not suffice for claiming that the incongruency-N300 reflects difficulties in identifying incongruent objects. Importantly however, our design does allow us to test this question more directly than was done before: as opposed to previous studies, we could compare not only congruent and incongruent objects, but also to track the differences between these two types of objects and the scrambled, unidentifiable ones. This enabled us to assess the moment in time when objects start to be identified as objects, or as different from a non-informative blur of colors which bears no semantic information. Although this contrast does not isolate identifiability per se, given the low-level differences between the intact and the scrambled conditions, it does allow a comparison between congruent and incongruent objects: because low-level differences exist for both object types, any difference in divergence times (i.e., congruent intact vs. congruent scrambled, compared with incongruent intact vs. incongruent scrambled) cannot be explained simply by these lower-level differences. Accordingly, if indeed incongruent objects do not pose any identification difficulty, we should have found no differences in divergence latency of such objects from scrambled ones, compared with the divergence times of congruent objects: both classes should have been identified as objects at the same time, and only later would their relations with the scene should have affected processing. This critical comparison suggested that this was not the case: activity evoked by incongruent objects diverged later, compared with congruent objects, from activity evoked by scrambled ones (though notably, only by 20 ms, and derived from an indirect comparison between each object type and the corresponding scrambled condition). This difference in divergence time, serving here as a rough approximation of their identification time as intact objects, implies that their identification was indeed delayed. This finding is further strengthened by the activity pattern found in the left central area, where the difference between congruent and incongruent scenes begun; there, incongruent scenes were indistinguishable from scrambled ones in the N300 time window, while congruent scenes were.

Interestingly, in the N400 time window, an opposite effect was found, so that now it was the incongruent intact waveform that was the-odd-one-out, differing from all other conditions. This suggests a functional dissociation between the N300 and N400 components: during the N300 time window, the congruent waveform diverged from all others, possibly because the congruent objects were the only ones identified; incongruent intact objects, as well as both scrambled ones, elicited greater – and undistinguished – negativity, presumably indexing the difficulty to identify them. During the N400 time window, on the other hand, the semantic incongruency of the already identified incongruent objects had to be reconciled (e.g., one had to comprehend why a dynamite is found in a wedding scene), so that now this condition elicited greater negativity. Because such semantic incongruency is not evoked by scrambled objects, in this time window they were undistinguished from the congruent intact objects. Under such an account, then, N300 indeed taps identification processes and N400 reflects semantic integration processes (Brown and Hagoort, 1993; Friederici et al., 1999; Kutas and Van Petten, 1994). Note however that intact and scrambled incongruent objects did not differ in frontal sites as well as in the medial central area, suggesting that at least some of the N400 effects might be explained by a general violation of expectation (e.g., Summerfield and Egner, 2009) or a prediction error signal (Friston and Kiebel, 2009; Trapp and Bar, 2015) which is evoked both by intact incongruent objects, and scrambled ones (since both are not expected to be found in the scene).

Taken together, our results support a functional dissociation between the N300 and the N400 component, and confirm the relations between processes taking place in the N300 time window and object identification. Notably however, we also show that the incongruency-N300 is different from the classical identification-N300 component, both with respect to topography and to onset time. Most critically, we provide first direct evidence for the difficulty to identify incongruent objects. Going back to the debate about cognitive penetrability, our findings thus argue against perceptual encapsulation and in favor of cognitive penetration. This goes hand in hand with claims that higher-level cognitive processes and information, like subjective (Balcetis and Dunning, 2010), or emotional (Song et al., 2012) value, as well as semantic and linguistic features (Dils and Boroditsky, 2010; Meier et al., 2007) may indeed affect the manner in which stimuli are perceived. Arguably, such cognitive states do not merely affect the interpretation of the stimuli, but penetrate and modify early perceptual processing itself. These findings were recently challenged based on methodological considerations (Firestone and Scholl, 2016; but see some of the replies: Clore and Proffitt, 2016; Desseilles and Phillips, 2016; Emberson, 2016; Hackel et al., 2016; Levin et al., 2016; O'Callaghan et al., 2016; Vinson et al., 2016), reviving the claims that perception, and specifically visual processing, is independent from semantic knowledge or expectations (Fodor, 1983; Pylyshyn, 1980). Critically, these criticisms only pertained to behavioral findings, and cannot explain our results. Thus, when it comes to object identification, this study provides direct evidence for semantic effects on perception (cf. Brandman and Peelen, 2017). Top-down processes in which the context limits the horizon of possibilities among which object identity lies, direct and affect object identification. Arguably then, objects relations are not only a characteristic of the actual world, but also a core feature of our ability to perceive it.

Acknowledgements

The study was funded by the Israel Science Foundation (grant No. 1847/16). We thank Oren Matar, Leemor Zucker and Uri Korisky for their help in establishing the stimuli bank and designing the experiment.

References

Balcetis, E., Dunning, D., 2010. Wishful seeing: more desired objects are seen ascloser. Psychol. Sci. 21, 147–152.

Bar, M., 2003. A cortical mechanism for triggering top-down facilitation in visual object recognition. J. Cogn. Neurosci. 15 (4), 600–609.

Bar, M., 2004. Visual objects in context. Nat. Rev. Neurosci. 5 (8), 617–629.

Bar, M., Ullman, S., 1996. Spatial context in recognition. Perception 25 (3), 343–352.

Barrett, S.E., Rugg, M.D., 1990. Event-related potentials and the semantic matching of pictures. Brain Cogn. 14 (2), 201–212.

Biederman, I., 1972. Perceiving real-world scenes. Science 177 (43), 77–80.

Biederman, I., 1981. On the semantics of a glance at a scene. In: M. Kubovy, & J.R.

Biederman, I., Mezzanotte, R.J., Rabinowitz, J.C., 1982. Scene perception: detecting and judging objects undergoing relational violations. Cogn. Psychol. 14 (2), 143–177.

Boyce, S.J., Pollatsek, A., Rayner, K., 1989. Effect of background information on object identification. J. Exp. Psychol.-Hum. Percept. Perform. 15 (3), 556–566.

Brainard, D.H., 1997. The psychophysics toolbox. Spat. Vision. 10, 433–436.

Brandman, T., Peelen, M.V., 2017. Interaction between scene and object processing revealed by human fMRI and MEG decoding. J. Neurosci. 37 (32), 7700–7710.

Brown, C., Hagoort, P., 1993. The processing nature of the N400: evidence from masked priming. J. Cogn. Neurosci. 5 (1), 34–44.

Clore, G.L., Proffitt, D.R., 2016. The myth of pure perception. Behav. Brain Sci. 39.

Coco, M.I., Araujo, S., Petersson, K.M., 2017. Disentangling stimulus plausibility and contextual congruency: electro-physiological evidence for differential cognitive dynamics. Neuropsychologia 96, 150–163.

Davenport, J.L., Potter, M.C., 2004. Scene consistency in object and background perception. Psychol. Sci. 15 (8), 559–564.

De Graef, P., Christiaens, D., Dydewalle, G., 1990. Perceptual effects of scene con-text on object identification. Psychol. Res.- Psychol. Forsch. 52 (4), 317–329.

Demiral, Ş.B., Malcolm, G.L., Henderson, J.M., 2012. ERP correlates of spatially incongruent object identification during scene viewing: contextual expectancy versus simultaneous processing. Neuropsychologia 50 (7), 1271–1285.

Desseilles, M., Phillips, C., 2016. How cognition affects perception: brain activity modelling to unravel top-down dynamics. Behav. Brain Sci. 39.

Dils, A.T., Boroditsky, L., 2010a. Processing unrelated language can change what you see. Psychon. Bull. Rev. 17, 882–888.

Doniger, G.M., Foxe, J.J., Murray, M.M., Higgins, B.A., Snodgrass, J.G., Schroeder, C.E., et al., 2000. Activation time course of ventral visual stream object-recognition areas: high density electrical mapping of perceptual closure processes. J. Cogn. Neurosci. 12 (4), 615–621.

Dyck, M., Brodeur, M.B., 2015. ERP evidence for the influence of scene context on the recognition of ambiguous and unambiguous objects. Neuropsychologia 72, 43–51.

Emberson, L.L., 2016. Gaining knowledge mediates changes in perception (without differences in attention): a case for perceptual learning. Behav. Brain Sci. 39.

Firestone, C., Scholl, B.J., 2016. Cognition does not affect perception: Evaluating the evidence for" top-down" effects. Behav. Brain Sci. 39.

Fodor, J.A., 1983. Modularity of Mind. MIT Press, Cambridge, MA.

Folstein, J.R., Van Petten, C., Rose, S.A., 2008. Novelty and conflict in the categorization of complex stimuli. Psychophysiology 45 (3), 467–479.

Friederici, A.D., Steinhauer, K., Frisch, S., 1999. Lexical integration: Sequential effects of syntactic and semantic information. Mem. Cogn. 27 (3), 438–453.

Friston, K., Kiebel, S., 2009. Predictive coding under the free-energy principle. Philos. Trans. R. Soc. B: Biol. Sci. 364 (1521), 1211–1221.

Ganis, G., Kutas, M., 2003. An electrophysiological study of scene effects on object identification. Cogn. Brain Res. 16 (2), 123–144.

Grill-Spector, K., Kushnir, T., Edelman, S., Itzchak, Y., Malach, R., 1998. Cue-invariant activation in object-related areas of the human occipital lobe. Neuron 21 (191 ± 202).

Hackel, L.M., Larson, G.M., Bowen, J.D., Ehrlich, G.A., Mann, T.C., Middlewood, B., Garrido, C.O., 2016. On the neural implausibility of the modular mind: evidence for distributed construction dissolves boundaries between perception, cognition, and emotion. Behav. Brain Sci. 39.

Hagoort, P., Hald, L., Bastiaansen, M., Petersson, K.M., 2004. Integration of word meaning and world knowledge in language comprehension. Science 304 (5669), 438–441.

Hamm, J.P., Johnson, B.W., Kirk, I.J., 2002. Comparison of the N300 and N400 ERPs to picture stimuli in congruent and incongruent contexts. Clin. Neuro-Physiol. 113 (8), 1339–1350.

Holcomb, P.J., 1993. Semantic priming and stimulus degradation—Implications for the role of the N400 in language processing. Psychophysiology 30 (1), 47–61.

Holcomb, P.J., Neville, H.J., 1990. Auditory and visual semantic priming in lexical decision: a comparison using event-related brain potentials. Lang. Cogn. Process. 5 (4), 281–312.

Holcomb, P.J., McPherson, W.B., 1994. Event-related brain potentials reflect semantic priming in an object decision task. Brain Cogn. 24 (2), 259–276.

Hollingworth, A., Henderson, J.M., 1998. Does consistent scene context facilitate object perception? J. Exp. Psychol.-Gen. 127 (4), 398–415.

Hollingworth, A., Henderson, J.M., 1999. Object identification is isolated from scene semantic constraint: evidence from object type and token discrimination. Acta Psychol. 102 (2–3), 319–343.

Jung, T.P., Makeig, S., Humphries, C., Lee, T.W., McKeown, M.J., Iragui, V., et al., 2000. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 37 (2), 163–178.

Kovács, I., 1996. Gestalten of today: early processing of visual contours and surfaces. Behav. Brain Res. 82 (1), 1–11.

Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62, 621–647.

Kutas, M., Hillyard, S.A., 1980a. Event-related brain potentials to semantically inappropriate and surprisingly large words. Biol. Psychol. 11 (2), 99–116.

Kutas, M., Hillyard, S.A., 1980b. Reading between the lines: event-related brain potentials during natural sentence processing. Brain Lang. 11 (2), 354–373.

Kutas, M., Van Petten, C., 1994. Psycholinguistics electrified. In: Gernsbacher, M.A. (Ed.), Handbook of Psycholinguistics, pp. 83–143.

Levin, D.T., Baker, L.J., Banaji, M.R., 2016. Cognition can affect perception: restating the evidence of a top-down effect. Behav. Brain Sci. 39.

Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P. J., Brady, T.J., Rosen, B.R., Tootell, R.B., 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc. Natl. Acad. Sci. USA, 92, 8135 ± 8139.

Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG-and MEG-data. J. Neurosci. Methods 164 (1), 177–190.

McPherson, W.B., Holcomb, P.J., 1999. An electrophysiological investigation of semantic priming with pictures of real objects. Psychophysiology 36 (1), 53–65.

Meier, B.P., Robinson, M.D., Crawford, L.E., Ahlvers, W.J., 2007. When “light” and “dark” thoughts become light and dark responses: affect biases brightness judgments. Emotion 7 (2), 366–376.

Moscovitch, C., Kapur, S., Kohler, S., Houle, S., 1995. Distinct neural correlates of visual long-term memory for spatial location and object identity: a positron emission tomography study in humans. Proc. Natl. Acad. Sci. USA, 92, 3721 ± 3725.

Mudrik, L., Lamy, D., Deouell, L.Y., 2010. ERP evidence for context congruity effects during simultaneous object–scene processing. Neuropsychologia 48 (2), 507–517.

Mudrik, L., Shalgi, S., Lamy, D., Deouell, L.Y., 2014. Synchronous contextual irregularities affect early scene processing: replication and extension. Neuropsychologia 56, 447–458.

Neider, M.B., Zelinsky, G.J., 2006. Scene context guides eye movements during visual search. Vis. Res. 46 (5), 614–621.

Niedeggen, M., Rösler, F., 1999. N400 effects reflect activation spread during retrieval of arithmetic facts. Psychol. Sci. 10 (3), 271–276.

O'Callaghan, C., Kveraga, K., Shine, J.M., Adams, R.B., Bar, M., 2016. Convergent evidence for top-down effects from the “predictive brain”. Behav. Brain Sci. 39.

Palmer, S.E., 1975. Effects of contextual scenes on identification of objects. Mem. Cogn. 3 (5), 519–526.

Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spat. Vis. 10 (4), 437–442.

Picton, T.W., Bentin, S., Berg, P., Donchin, E., Hillyard, S.A., Johnson, R., Taylor, M.J., 2000. Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. Psychophysiology 37 (2), 127–152.

Pylyshyn, Z., 1980. Computation and cognition: issues in the foundations of cognitive science. Behav. Brain Sci. 3, 111–132.

Rieger, J.W., Kochy, N., Schalk, F., Gruschow, M., Heinze, H.J., 2008. Speedlimits: orientation and semantic context interactions constrain natural scenediscrimination dynamics. J. Exp. Psychol.-Hum. Percept. Perform. 34 (1), 56–76.

Sauvé, G., Harmand, M., Vanni, L., Brodeur, M.B., 2017. The probability of object–scene co-occurrence influences object identification processes. Exp. Brain Res. 1–13.

Schendan, H.E., Kutas, M., 2002. Neurophysiological evidence for two processing times for visual object identification. Neuropsychologia 40 (7), 931–945.

Schendan, H.E., Kutas, M., 2003. Time course of processes and representations supporting visual object identification and memory. J. Cogn. Neurosci. 15 (1), 111–135.

Sehatpour, P., Molholm, S., Javitt, D.C., Foxe, J.J., 2006. Spatiotemporal dynamics of human object recognition processing: an integrated high-density electrical mapping and functional imaging study of “closure” processes. Neuroimage 29 (2), 605–618.

Sitnikova, T., Holcomb, P.J., Kiyonaga, K.A., Kuperberg, G.R., 2008. Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. J. Cogn. Neurosci. 20 (11), 2037–2057.

Song, H., Vonasch, A.J., Meier, B.P., Bargh, J.A., 2012. Brighten up: Smiles facilitate perceptual judgment of facial lightness. J. Exp. Soc. Psychol. 48, 450–452.

Sitnikova, T., Kuperberg, G., Holcomb, P.J., 2003. Semantic integration in videos of real-world events: an electrophysiological investigation. Psychophysiology 40 (1), 160–164.

Summerfield, C., Egner, T., 2009. Expectation (and attention) in visual cognition. Trends Cogn. Sci. 13, 403–409.

Trapp, S., Bar, M., 2015. Prediction, context, and competition in visual recognition. Ann. N. Y. Acad. Sci. 1339 (1), 190–198.

Underwood, G., 2005. Eye fixations on pictures of natural scenes: getting the gist and identifying the components. In: Underwood, G. (Ed.), Cognitive Processes in Eye Guidance. Oxford University Press, Oxford, pp. 163–187.

Vinson, D.W., Abney, D.H., Amso, D., Chemero, A., Cutting, J.E., Dale, R., Jordan, J.S., 2016. Perception, as you make it. Behav. Brain Sci. 39.

Võ, M.L.H., Wolfe, J.M., 2013. Differential electrophysiological signatures of semantic and syntactic scene processing. Psychol. Sci. 24 (9), 1816–1823.