Effects of stereoscopic disparity on early ERP components during classification of three-dimensional objects

Pegna, Alan; Darque, Alexandre; Roberts, Mark; Leek, Charles

Quarterly Journal of Experimental Psychology

DOI:
10.1080/17470218.2017.1333129

Published: 01/06/2018

Peer reviewed version

Stereo viewing modulates early evoked potentials associated with the classification of three-dimensional object shape

Journal:	*Quarterly Journal of Experimental Psychology*
Manuscript ID	QJE-STD 16-269.R1
Manuscript Type:	Standard Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Pegna, Alan; Universite de Geneve Faculte de Psychologie et des Sciences de l'Education, Psychology; University of Queensland, Psychology Darque, Alexandra; Universite de Geneve Faculte de Psychologie et des Sciences de l'Education, Psychology Roberts, Mark; Bangor University, Psychology Leek, Charles; Bangor University, Psychology
Keywords:	3D SHAPE PERCEPTION, GLOBAL CONFIGURATION, LOCAL PART STRUCTURE, EVOKED POTENTIALS, ERP

Stereo viewing modulates early evoked potentials associated with the classification of three-dimensional object shape

Alan J. Pegna[1,2], Alexandre Darque[1], Mark V. Roberts[3] & E. Charles Leek[3,4]

[1]Faculty of Psychology and Educational Sciences, University of Geneva, Geneva, Switzerland

[2]School of Psychology, University of Queensland, Brisbane, Australia

[3]Wolfson Centre for Clinical and Cognitive Neuroscience, School of Psychology, Bangor University, UK

[4]Laboratoire de Psychologie et NeuroCognition (LPNC), Université Grenoble Alpes, France

RUNNING HEAD: ERP study of 3D object classification

Address for Correspondence:
Alan J. Pegna, School of Psychology, University of Queensland, St Lucia Qld 4072, Australia. Email: a.pegna@uq.edu.au Tel: (+61) (0) 733 656 412

Charles Leek, Wolfson Centre for Clinical and Cognitive Neuroscience, School of Psychology, University of Bangor, Bangor, UK. Email. e.c.leek@bangor.ac.uk Tel: (+44) (0) 1248 382948

Keywords: 3D SHAPE PERCEPTION, GLOBAL CONFIGURATION, LOCAL PART STRUCTURE, EVOKED POTENTIALS, ERP

Word count: 5252 (without abstract and figure captions)
Figures: 7

ABSTRACT

This study investigates the effects of stereo disparity on the perception of three-dimensional (3D) object shape. We tested the hypothesis that stereo input modulates the brain activity related to perceptual analyses of 3D shape configuration during image classification. High-density (256-channel) EEG was used to record the temporal dynamics of visual shape processing under conditions of two-dimensional (2D) and three-dimensional (3D) visual presentation. On each trial, observers made image classification judgements ('Same'/'Different') to two briefly presented, multi-part, novel objects. On different-object trials, stimuli could either share volumetric parts but not the global 3D shape configuration, have different parts but the same global 3D shape configuration, or differ on both aspects. Analyses using mass univariate contrasts showed that the earliest sensitivity to 2D versus 3D viewing appeared as a negative deflection over posterior locations on the N1 component between 160ms-220ms post stimulus onset. Subsequently, ERP modulations during the N2 time window between 240ms-370ms were linked to image classification. N2 activity reflected two distinct components – an early N2 (240ms-290ms) and a late N2 (290ms-370ms) that showed different patterns of responses to 2D and 3D input, and differential sensitivity to 3D object structure. The results revealed that stereo input modulates the neural correlates of 3D object shape. We suggest that this reflects differential perceptual processing of object shape under conditions of stereo or mono input. These findings challenge current theories that attribute no functional role for stereo input during 3D shape perception.

INTRODUCTION

The human visual system is able to perceive, and rapidly classify, the shapes of complex three-dimensional (3D) objects with remarkable speed and accuracy (e.g., (Arguin & Leek, 2003; Harris, Dux, Benito, & Leek, 2008; Hummel, 2013; Leek, 1998a, 1998b; Leek, Atherton, & Thierry, 2007; Leek & Johnston, 2006; Pizlo, Sawada, Li, Kropatsch, & Steinman, 2010; Tarr & Bulthoff, 1999). One important theoretical issue is whether the perceptual processes that support the classification of 3D object shape are modulated by stereo visual input. Neurophysiological studies have shown that binocular disparity is resolved relatively early in visual cortex (e.g., DeAngelis & Newsome, 1999; Livingstone & Hubel, 1988), and that object processing areas in infero-temporal cortex can respond to shape defined solely by stereo cues (e.g., Gilaie-Dotan, Ullman, Kushnir & Malach, 2001; Tanaka, Uka, Yoshiyama, Kato & Fujita, 2001). However, it remains less clear whether stereo information modulates high-level perceptual processing of 3D object shape.

In principle, stereo disparity might facilitate perceptual analyses of 3D object shape - at least under some circumstances - by providing cues to properties such as local surface slant, global depth orientation and 3D shape configuration. But while stereo input has been shown to play an important role in our interactions with objects for tasks such as prehensile movement (e.g., Watt & Bradshaw, 2003), several current theories of object recognition attribute no particular functional significance to stereo information in the perceptual analysis of 3D shape (e.g., Biederman, 1987; Cadieu, Kouh, Pasupathy, Connor, Riesenhuber & Poggio, 2007; Pizlo et al., 2010; Riesenhuber & Poggio, 1999; Serre, Oliva & Poggio, 2007).

Current empirical evidence on this issue is inconclusive and largely confined to studies investigating the perception of shape equivalence across changes in

viewpoint (Bennett & Vuong, 2006; Burke, Taubert, & Higman, 2007; Chan, Stevenson, Li, & Pizlo, 2006; Cristino, Davitt, Hayward, & Leek, 2015; Edelman & Bulthoff, 1992; Humphrey & Khan, 1992; Lee & Saunders, 2011; Liu, Ward, & Young, 2006; Pasqualotto & Hayward, 2009; Rock & DiVita, 1987). Some studies - typically involving deformed 'paperclips' or 'amoeba' stimuli, have found stereo viewing advantages for image interpolation across viewpoint changes (e.g., Bennett & Vuong, 2006; Edelman & Bulthoff, 1992; Lee & Saunders, 2011), although stereo viewing costs have also been reported under some conditions (e.g., Pasqualotto & Hayward, 2009). A limitation of these studies is their use of stimulus types that are different from most common solid 3D objects and are not readily decomposable into a structural description of volumetric parts. Cristino et al. (2015) have recently shown - in one of the few studies to use complex, multi-part, 3D objects - that stereo presentation yields advantages in recognition when observers are required to make difficult target-distracter discriminations. One interpretation of this finding is that stereo cues may be used to constrain perceptual analyses of 3D object shape when image classification requires the computation of 3D shape configuration. That is, while stereo cues to 3D shape are computed (where available) from visual sensory input, they may only be used when object discrimination is dependent on perceptual analyses of 3D shape structure.

The aim of this study was to test this hypothesis using an image classification task contrasting 2D and 3D visual presentation of stimuli, using shapes that varied systematically along dimensions that are relevant to the analysis of 3D object shape[1]. We used complex multi-part 3D objects that (a) varied according to their global

[1]In this study, we used the term '2D' to describe non-stereo visual input (that is, where there is no disparity between visual inputs to the left and right eye), and '3D' to refer to visual input with stereo disparity.

configuration (their constituent parts being otherwise identical), or (b) varied across their local volumetric parts (their global configuration being otherwise equal) or (c) varied on both aspects (see Fig. 1A for an illustration). A further goal was to address a methodological limitation of previous studies based solely on recordings of standard behavioural responses (i.e., RTs, accuracy) which do not provide an online measure of perceptual sensitivity to stereo disparity during image classification. To address this we used event-related potentials (ERPs) in order to provide a high temporal resolution measure of cortical activation to stereo input. The task involved observers making same/different shape judgements to pairs of these sequentially presented, surface-rendered, 3D objects. Stimuli were viewed in 2D and 3D display modes in different blocks of trials by the same observers. 3D object shape similarity was factorially manipulated in terms of shared or different 3D object parts, and shared or different 3D shape configuration. These factors have previously been shown to play an important role in the perception of 3D object shape – and they are fundamental components of structure description models of recognition (e.g., Arguin & Saumier, 2004; Behrmann, Peterson, Moscovitch & Satoru, 2006; Behrmann & Kimchi, 2003; Biederman, 1987; Hummel, 2013; Hummel & Stankiewicz, 1996). This design allowed us to provide a strong test of whether stereo-defined shape information modulates the neural correlates of shape perception. We predicted that ERP responses to stimuli sharing local but not global aspects would reveal processing of local shape properties, while ERPs for stimuli sharing global but not local features would indicate global shape processing, and that moreover, stereo presentation would differentially modulate the responses associated with the discrimination of 3D objects when image classification (on "different" object trials) was dependent on specification of 3D shape configuration.

EXPERIMENTAL STUDY

Methods

Participants

Fourteen university students (7 females; mean age = 26.1; SD = 3.5) took part in the experiment and were paid for their participation. Handedness was assessed using the Edinburgh-Oldfield handedness inventory (Oldfield, 1971) and revealed that participants were all right-handed (mean laterality index: 85.7 ±22.4). None had any previous history of neurological or psychiatric disorder. All had normal or corrected-to-normal visual acuity and stereo vision. The study was approved by the local Ethics Committee. Participants signed a written informed consent form before beginning the experiment.

Stimuli

The stimulus set comprised 40 novel objects (see Figure 1A for an example), each containing a unique spatial configuration of four volumetric parts defined by variation among non-accidental properties (NAPs): Edges (Straight vs. Curved), symmetry of the cross section, tapering (co-linearity) and aspect ratio (Biederman, 1987).

The models were designed in Strata 3D CX software (Strata Inc. USA) and rendered with a smooth surface and no texture cues in a yellow mustard colour (RGB: 227, 190, 43) using a stereoscopic camera with an Inter-Pupillary Distance (IPD) of 62mm. Each rendered image pair were converted to a red-cyan anaglyph. Stimuli were scaled to have the same maximum dimensions of 7.5˚ by 7.5˚ from a viewing distance of 120cm and extended over a 12’ crossed to 6’ uncrossed disparity. Stimuli were

displayed on a white background. The stimuli could be viewed in stereo with red-cyan glasses or in mono with red-red or cyan-cyan glasses (see Design and Procedure).

The set contained 10 'target' objects (referred to as 'SS' for 'Same Part/Same Configuration – see left hand model in Figure 1A). For each of the 10 'SS' targets three distracters were created with different levels of similarity (here, similarity was defined in terms of shared local 3D parts and/or global shape configuration (Biederman, 1987): the 'DS' (Different Parts/Same Configuration) distracters shared spatial configuration but had different parts; the 'SD' (Same Parts/Different Configuration) distracters shared parts but had a different global shape configuration to the target; finally, 'DD' (Different Parts/Different Configuration) served as a baseline contrast in which distracters shared neither parts, nor global shape configuration with the corresponding 'SS' target. In order to allow us to present sequential image pairs on each trial while preventing a matching strategy based solely on pixel-to-pixel similarity, or 2D global shape outline, stimuli were presented at different viewpoints. Each stimulus was rendered at five orientations (0°, 60°, 120°, 240° and 300°) around a vertical axis relative to the stereo camera position from an arbitrary baseline viewpoint (0°). In addition to the novel object set, we also created four masks each composed of fragments from all 10 SS stimuli arranged in a random configuration. Stereo versions of the masks were made in the way as for the novel objects. The experiment was programmed using E-prime (v.1.1; www.pstnet.com/eprime).

Design and Procedure.

The experiment involved a 2 (Viewing condition: 2D, 3D) x 4 (Stimulus type: SS, SD, DS, SS) repeated measures design. The factor "Stimulus type" refers to the

kind of stimulus pairs shown on each trial. These could either be a repetition of the same shape (SS-SS – 'Same' response) or three types of 'Different' object trials SS/SD, SS/DS or SS/DD). There were four blocks of 160 trials each (N total = 640), comprising two blocks of 2D and two blocks of 3D object trials. Half of the participants viewed the blocks in the order 3D-2D-2D-3D while the others viewed the blocks in the order 2D-3D-3D-2D. For each viewing condition (2D/3D) there were 320 trials across blocks, comprising 80 trials for each type of stimulus pairing (SS/SS; SS/DS; SS/DS; SS/DD). In the 3D blocks, participants wore red/cyan lenses which allowed the anaglyphs to be perceived in 3D, while in the 2D blocks, participants wore glasses with two cyan lenses in one block, and two red lenses in the other, causing the stimuli to be seen without stereo disparity.

Each trial involved the sequential presentation of a comparison stimulus (S1), mask, and a target (S2). The S1 and S2 stimulus pairs could be either the same shape (SS – 25% trials) or one of three different types of distracters (SD; DS; DD – 75% of trials). On all trials, S1 and S2 were presented at different viewpoints to ensure that shape equivalence judgements could not be made using an image-based low-level (e.g., pixel matching) strategy. S2 was always presented at a fixed arbitrary 0° viewpoint (so that ERPs to S2 would not be affected by S2 viewpoint). S1 was presented with equal frequency (40 trials per block) at 60°, 120°, 240° or 300°. Stimulus orientation was not predictive of trial type. Trials began with a fixation cross presented at a crossed disparity of 3' that lasted 750ms. This was followed by S1 that appeared for 750ms and was followed by a mask for another 750ms. Finally, the S2 stimulus was presented for 750ms, followed by a blank screen until the participant gave a response. Feedback was then provided indicating a "correct" or "incorrect" response for 1000ms). There was a blank inter-trial interval of 1000ms. Fig. 1B

illustrates a typical experimental trial. Observers were asked to respond by pressing a key with the left or right index to indicate if they thought the S2 target was the same as S1. Half of the subjects used the right index to respond "same" and the left index to respond "different". This was reversed in the other half of the group. Participants were seated comfortably in an electrically shielded, sound attenuated, room. A chin rest was used to maintain the distance from the screen. Prior to the main task, observers completed four practice trials using two additional novel objects that were not included in the main experiment. In order to ensure correct binocular fusion of stereo images by the participants, a stereo image (a hand appearing to emerge from the screen) was viewed with anaglyphic glasses and they were asked to report whether or not they perceived it clearly in 3D. All participants confirmed that perceived the stimulus in depth.

INSERT FIGURE 1 ABOUT HERE

EEG-ERP recording. Continuous EEG was acquired at 1000 Hz using a Hydrocel Geodesic Sensor Net from 256 equally-spaced AgCl carbon-fibre coated scalp electrodes referenced to the vertex. The EEG was band-pass filtered between 0.01-100Hz and impedances were kept below 30kΩ. Epochs for correct trials, beginning 100ms before onset of S2 target stimulus and ending 800 ms afterwards were used to compute the ERPs. The 100 ms pre-stimulus period was used to establish baseline. The EEG was then visually inspected for eye movements or other sources of noise and rejected if artefacts were present. Channels that displayed frequent or continuous

noise in the individual's ERP were removed and replaced using a 3D spherical spline interpolation procedure (Perrin, Pernier, Bertrand, Giard, & Echallier, 1987). Electrodes located on the cheeks were systematically excluded such that 204 electrodes were finally retained for statistical analyses. Epochs were then filtered from 1 to 30 Hz and re-computed against the average reference.

The epochs were used to calculate the mean ERPs of the 8 conditions (2 Viewing Condition: 2D/3D) X 4 (Stimulus type: SS, SD, DS, DD) in every participant and across subjects (grand average ERPs). Analyses of the ERP data were carried out using the Cartool software (Murray, Brunet & Michel, 2008).

Event-related potentials analysis

Mass univariate analyses (e.g., Groppe, Urbach & Kutas, 2011; Guthrie & Buchwald, 1991) were used to elucidate the time course of brain activation to 2D versus 3D viewing, and the sensitivity of perceptual classification of 3D object shape to stereo viewing across conditions of shape similarity. This involved using pair wise, time-frame by time-frame, permutation tests based on repeated measures t-tests across all 204 electrodes from 0-800ms. An a priori criterion for significance testing was adopted in which a threshold of $p<.01$ (two-tailed) must be attained for at least 10 consecutive time frames in at least 5 neighbouring electrodes (Guthrie & Buchwald, 1991; Murray et al., 2008).

Component analysis. As 2D and 3D visual presentation may interact and modulate object recognition differentially, the components sensitive to viewing condition and to image classification (N1, P2, N2) were further analysed using repeated-measures analyses of variance (ANOVAs). These included Viewing condition (2D/3D), Stimulus type (SS, DS, SD, DD) and Laterality (right versus left

hemisphere electrodes) as repeated measures. The N1 was defined as the maximum negative amplitude in the 160-220ms time window, in the left and right groups of posterior electrodes (the regions of interest, or ROIs are shown in the insets in figures 1 and 2). Amplitudes and latencies for this component were established in all participants for subsequent statistical analysis. The P2, determined as the positive peak following the N1, was visible in the grand averages and peaked around 235ms. However, a maximum could not be clearly observed in the individual ERPs, thus the computations were computed using the mean amplitudes over a 30ms time window centred on the P2 (i.e., 220-250ms) in the 2 posterior ROIs of each participant. Finally, for the second negative deflection (N2), mean amplitudes were computed separately on an early N2 (240-290ms) and late N2 (290-370ms) part of the component in the same 2 posterior ROIs.

Greenhouse-Geisser corrections were applied to the analyses when appropriate. Post-hoc comparisons were applied using Tukey's Honest Significant Difference (HSD) test to compare effects when interactions were significant. For all analyses, exact two-tailed alpha values are reported (p = x) except where $p < .0001$.

RESULTS

Behavioural data

Due to technical difficulties, the behavioural results of two participants were lost and statistics were consequently computed on 12 subjects.

Mean correct responses (+/-standard deviation) for SS, DS, SD, DD respectively were for 2D: 63(8.5), 65(7.9), 77(2.4), 78(1.9) and for 3D: 62.7(9.5), 68.0(7.1), 77.6(2.3), 78.4(1.6) out of a maximum of 80. Response accuracy was entered into a 2 (Viewing condition: 2D, 3D) X 4 (Stimulus type: SS, DS, SD, DD) repeated measures ANOVA. There was no effect of Viewing condition, F (1, 11) = 1.24, p=.29; but Stimulus type was significant, F (3, 33) = 21.26, p<.0001. This was due to a significantly smaller number of correct responses for the SS and DS conditions compared to both SD and DD (Tukey HSD post-hoc: ps<. 001). There was no difference between SD and DD (p=.969) nor between SS and DS (p=.38), and no interaction, F (3, 33) = .93, p=.44.

Event-related potentials

Figure 2 and 3 show the grand average ERP traces for 2 electrodes (P9 and P10) that were part of the two ROIs used for statistical computation. These regions showed the greatest differences across conditions on visual inspection.

INSERT FIGURE 2 and 3 ABOUT HERE

1. Analysis of ERP sensitivity to stereo presentation

In the first analysis we wanted to determine the effect of stereo and non-stereo visual input on the time course of electrical activity (see figure 3 for the ERPs of stereo and non-stereo presentations). To do so, mass univariate analyses were used to

identify a temporal marker defining the earliest time point of differential ERP sensitivity to 2D versus 3D viewing. A point-wise mass univariate analysis performed on the 2D versus 3D viewing conditions showed that the earliest differences began during the N1 (figure 4). The difference affected a large group of posterior, temporo-occipital and anterior leads beginning at about 160 ms until around 220ms, encompassing the N1 component. This confirms early sensitivity to 2D and 3D visual input used in the study.

INSERT FIGURE 4 ABOUT HERE

N1

To examine this further, we conducted additional analyses on the ERP data from the N1 component. Mean N1 latencies for SS, DS, SD and DD were respectively of 174ms (±10.4), 173ms (±11.1), 175ms (±10.4) and 176ms (±11.3) for 2D targets, and 178ms (±6.7), 176ms (±10), 179ms (±8.7) and 177ms (±11.3) for 3D targets. A 2 (Viewing condition: 2D, 3D) x 4 (Stimulus type: SS, SD, DS, DD) repeated measures ANOVA showed no significant main effects or interaction.

The amplitudes values of the N1 were also entered into a 2 (Viewing condition: 2D, 3D) x 4 (Stimulus type: SS, SD, DS, DD) x 2 (Laterality) repeated measures ANOVA and revealed a significant main effect of Viewing condition, F (1, 13) = 6.9, p<.05; with 3D stimuli producing more negative N1 peaks. There were no other main effects or interactions. Thus, once again, stereo viewing affected cortical activation early in the stream of visual processing. Furthermore, sensitivity to stereo disparity was equivalent for both left and right hemisphere electrodes and was not modulated by stimulus type.

2. *Temporal marker for 3D image classification*

The aims of these analyses were to (1) identify a temporal marker defining the earliest time point at which differential responses to the stimulus types are distinguishable in the ERP – and, by hypothesis, reflect differentiation of 3D object shape during perceptual classification; and (2) to determine whether the time course of 3D object discrimination is modulated by shape similarity. We used a mass univariate approach to identify the earliest statistically significant time points for each stimulus contrast in 2D and 3D (SS-SD; SS-DS; SS-DD). The temporal distributions of these difference contrasts across all 204 electrodes for 2D and 3D viewing are shown in Figure 5A and B.

INSERT FIGURE 5A and B ABOUT HERE

This showed that, for both 2D and 3D viewing, differences were statistically significant at the .01 level from approximately 230-240ms, starting during the latter part of the P2 and extending across the N2 until around 370ms post-stimulus onset. While the onsets of these distributions are highly similar across conditions they appear to vary in amplitude (that is, frequency of significant difference contrasts). To explore this further we computed difference contrasts for 2D versus 3D viewing for SS-SD; SS-DS and SS-DD. Figure 6 shows a time series plot of the frequency distribution of significant 2D/3D difference contrasts sub-sampled into 10ms bins.

INSERT FIGURE 6 ABOUT HERE

This plot shows that during the N2 (240-380ms) the distributions of significant differences between 2D and 3D viewing varies as a function of stimulus contrast. The distribution for the SS/DS contrast is clearly bimodal with an early peak at 260ms and a later peak at 360ms. The SS/SD contrast peaks at 270ms and then rapidly declines. The SS/DD distribution is also somewhat bimodal with an early peak around 250ms and a latter peak at 320ms. These data were analysed as a non-parametric time series using the Friedman test which showed that the frequency distributions were significantly different, χ^2 N=19; d.f. = 2) = 8.84, p = .01. This provides statistical evidence that the N2 comprises potentially distinct early and late components.

Detailed analyses were consequently performed on the P2, as well as early and late N2 components.

P2

The mean amplitudes of the P2, obtained from the posterior left and right ROIs for 2D and 3D values and for SS, DS, SD, and DD, were entered into a 2 (Viewing condition: 2D, 3D) x 4 (Stimulus type: SS, DS, SD, DD) x 2 (Laterality) repeated measures ANOVA. There was a significant main effect of Viewing condition, F (1, 13) = 18.98, p=.001; Mean amplitude: 2D = 0.77μV (± 2.9); 3D = -0.19 μV (±2.2). There were no other significant main effects or interactions.

N2

Early N2 (240-290ms)

Figure 7A shows the mean amplitudes for the early N2 component as a function of viewing condition, stimulus type and laterality. This shows, similar to the N1, more negative amplitudes for 3D than 2D viewing.

INSERT FIGURE 7 ABOUT HERE

The mean early N2 amplitudes were analysed using a 2 (Viewing condition: 2D, 3D) x 4 (Stimulus type: SS, SD, DS, DD) x 2 (Laterality) repeated measures ANOVA. There were significant main effects of Viewing condition, F (1, 13) = 15.07, p = .002; and Stimulus type, F (3, 39) = 19.50, p < .0001; and a significant three-way interaction, F (3, 39) = 3.21, p = .03. To explore this further we conducted two separate 2 (Viewing condition: 2D, 3D) x 4 (Stimulus type: SS, SD, DS, DD) ANOVAs for the left and right ROIs. For the left hemisphere ROI there was a significant main effect of Viewing condition, F (1, 13) = 6.55, p = .02; and Stimulus type, F (3, 39) = 18.02, p < .0001, but no interaction. For the right hemisphere there were significant main effects of Viewing condition, F (1, 13) = 22.66, p < .0001, and Stimulus type, F (3, 39) = 9.29, p < .0001, but no interaction.

Late N2 (290-370ms)

Figure 7b shows the mean amplitudes for the late N2 component as a function of viewing condition, stimulus type and laterality. In contrast to the N1 and early N2,

this shows more negative amplitudes for 2D than 3D viewing (that is, a reversal of 2D/3D amplitude negativity).

The mean late N2 amplitudes were analysed using a 2 (Viewing condition: 2D, 3D) x 4 (Stimulus type (Stimulus type: SS, SD, DS, DD) x 2 (Laterality) repeated measures ANOVA. There were significant main effects of Viewing condition, F (1, 13) = 8.87, p = .01; and Stimulus type, F (3, 39) = 18.31, p < .0001; and a significant three-way interaction, F (3, 39) = 2.84, p = .03. To explore this further we conducted two separate 2 (Viewing condition: 2D, 3D) x 4 (Stimulus type: SS, SD, DS, DD) ANOVAs for the left and right ROIs. For the left hemisphere ROI there was a significant main effect of Stimulus type, F (3, 39) = 15.40, p < .0001, but no effect of Viewing condition, and no interaction. In contrast, for the right hemisphere there were significant main effects of Viewing condition, F (1, 13) = 12.01, p = .004, and Stimulus type, F (3, 39) = 5.62, p = .003, and a significant interaction, F (3, 39) = 3.22, p = .03. Post-hoc analyses using Tukey HSD confirmed the source of the interaction arising from a significant difference in mean amplitude for the 2D versus 3D contrast between the DS conditions ($p < .0001$), but no significant differences for the other pair wise contrasts.

Table 1 below summarises the significant results of our study.

INSERT TABLE 1 HERE

DISCUSSION

The main findings of the study can be summarised as follows; (1) the behavioural data showed that observers performed equally well with both 2D and 3D visual presentation, but they made more errors in the SS and DS conditions; (2) analyses of the ERP data showed an early sensitivity to 2D versus 3D viewing occurring during the N1 component (160ms-220ms) and (3) the neural correlates of object shape discrimination were evidenced between approximately 240ms-370ms during the N2 component with the later part showing a differential sensitivity between 2D and 3D viewing.

These findings provide new evidence about the time course underlying the perceptual analysis of 3D object shape and are consistent with our prediction that ERPs should be modulated by stereoscopic viewing when depth is required for object classification. It thus provides some of the first electrophysiological evidence underscoring the effect of stereo information on 3D shape perception.

Two main points should be highlighted. First, the evidence indicating an early sensitivity to 2D versus 3D presentation on the N1 component, which was characterised as an amplitude modulation with more negative deflections for 3D relative to 2D viewing, regardless of stimulus condition. While this result is not surprising given other neurophysiological evidence of sensitivity to binocular disparity in early visual cortex (e.g., DeAngelis & Newsome, 1999; Livingstone & Hubel, 1988), it is relevant to the interpretation of the current data because it shows that our manipulation of stereo viewing was sufficient to elicit an early perceptual response. Second, the differential ERP modulation for mono and stereo observed on the early and late N2. This was revealed by variations in the distributions of significant mass univariate difference contrasts for 2D versus 3D viewing across

conditions, and by modulations in the N2 amplitudes. In fact, two potentially distinct components were observed within the N2. The early component between 240ms-290ms showed a similar pattern of sensitivity as the N1 for 2D versus 3D viewing– that is, greater negativity for the latter relative to the former. In contrast, the later N2 component from 290ms-370ms showed a reversal of this pattern with higher negative deflections for 2D relative to the 3D viewing. Moreover, during the late N2 component this difference interacted with both laterality and stimulus condition. Notably, there was a greater amplitude difference between 2D and 3D viewing for the DS condition over right compared to the left hemisphere electrodes.

The effect of 3D processing on the N1 does not come altogether as a surprise. Indeed the role of 3D cues in visual processing has been addressed by several studies. For example, Kasai and Morotomi (2001) used dynamic random-dot stereograms in a visual attention paradigm in which participants had to selectively attend either to the shape of a stimulus (a rectangle placed vertically or horizontally) or its depth, based on visual disparity. Attention to stereo-defined depth produced a greater negative deflection over the lateral occipito-temporal regions starting from around 175 ms, whilst attention to shape produced a slightly later ERP effect that arose after about 200ms, indicating that depth was processed earlier in time. This led the authors to conclude that the two processes operated independently and prior to perceptual integration. In a subsequent study, Kasai, Morotomi, Katayama, & Kumada (2003) studied the P1 and N1 ERP responses in a 3D attentional task using stereoscopic viewing. They found that the N1 was enhanced in response to stimuli that were attended in specific spatial positions, defined both in the plane and in depth. This suggested that the N1 component is at least partly connected to spatial representation in 3D. In another ERP study using random dot stereograms (but only one occipital

electrode: Oz, referenced to Fpz), depth perception was found to be associated with a posterior negative deflection occurring at around 210ms (Akay & Celebi, 2009). This negative deflection being the first one observed after stimulus presentation, most likely reflected the N1. Its later appearance could probably be explained the more medial electrode location on which it was measured.

The N1 sensitivity to stereo-defined depth has also received some support from work carried out with 3D stimuli determined not by binocular disparity, but by perspective or depth cues. Severac-Cauquil, Trotter & Taylor (Severac Cauquil, Trotter, & Taylor, 2006) measured the ERP while subjects viewed stimuli that were flat (2D drawings and textures) or denoted a perspective implying depth (3D). In line with our results, the authors found that the N1 component was increased for 3D views, whether they were specifically attending depth or not. The study further included a LORETA source localisation analysis performed on the scalp topography at the N1 latency, which pointed to an increased activation in the right parietal lobe for 3D views. This activity was hypothesised to be due to early processing of depth cues through the dorsal parietal route which likely activated temporal and temporo-occipital areas through recurrent feedback loops (Severac Cauquil et al., 2006). More recently, Gao et al. (2015) investigated the ERP response to line drawings of 3D objects, as well as 2D renderings of textures or perspectives. The authors found that the N1 was greater for 3D objects than for textures or perspectives again suggesting that the N1 is a marker of 3D viewing. However, in this study, no 2D objects were included and comparisons were performed between objects in depth and 2D texture-like perspectives. Thus, in this case, one cannot determine with certainty whether the N1 was responsive to 3D objects or to objects more generally.

Nevertheless, the evidence available so far appears to support our findings of an early response to 3D information occurring prior to shape discrimination, which could therefore subsequently contribute to complex, multi-part object processing if available and necessary.

The differential sensitivity of the N2 for mono versus stereo viewing during 3D shape classification upholds this assumption of stereo viewing modulating high-level perceptual processes involved in the classification of 3D object shape. In our study, the effects of viewing condition were found in the amplitude data and was most pronounced in the DS condition, when observers had to make shape classification judgments between objects that shared 3D configurations. By hypothesis, these could only be differentiated by their local part structure suggesting that the differential ERP response reflected the incorporation of stereo cues to 3D global shape configuration in line with our hypothesis.

Interestingly, the N2 effect may be related to early perceptual processes supporting figure-ground segmentation (e.g., Mendola, Dale, Fischl, Liu, & Tootell, 1999; Murray, Imber, Javitt, & Foxe, 2006; Pegna, Khateb, Murray, Landis, & Michel, 2002). For example, Doniger et al. (Doniger et al., 2000; Doniger et al., 2001) investigated object recognition using fragmented line drawing of familiar objects that were either identifiable, or too degraded for recognition to occur. The ERPs for identifiable stimuli produced a greater N2 component than unidentifiable ones over lateral posterior electrodes between 230ms and roughly 400ms, following a similar pattern to our N2. Furthermore, the negativity appeared to build up when the stimuli were presented with progressively less fragmentation, culminating when the stimuli were recognised (Doniger et al., 2001). The authors suggested that this negativity reflected the neural responses linked to the processing of increasing amounts of

object-relevant information that ultimately leads to perceptual closure and recognition. This is consistent with our interpretation of the N2 data in the current study of multi-part 3D objects.

To conclude, this investigation reveals that stereo disparity initially modulates the neural correlates of perceptual analyses of three-dimensional (3D) object shape during the N1 component. Subsequent ERP modulations occur in the N2 time window that are linked to image classification and are composed of two distinct components – an early and a late N2, which show different patterns of responses to 2D and 3D input, as well as a differential sensitivity to 3D object structure. It therefore supports the view that stereo input modulates cortical activity during 3D object shape processing. More broadly, the current observations present a challenge to models of object recognition that do not posit a functional role for stereo information during the perceptual analysis of 3D object shape (e.g., Biederman, 1987; Cadieu et al., 2007; Pizlo et al., 2010; Riesenhuber & Poggio, 1999; Serre et al., 2007).

Acknowledgements

This investigation was supported by a Royal Society Grant to E.C.L. and by a Swiss National Science Foundation grant (no. 320030-144187) to A.P.

FIGURE LEGENDS

Table 1: Summary of statistically significant effects obtained in the ANOVAs. Effects of stereoscopic vs non-stereoscopic viewing and of stimulus shape on performance (number of errors made by the participants) and ERPs (N1, P2, early N2, late N2) are reported. For the N2, left and right regions of interest (ROIs) are reported separately as an interaction (Stereo X shape X ROI) was noted. For the late N2, the analysis of the right ROI further revealed a stereo X shape interaction, reason for which this effect is decomposed (ns: non-significant).

Figure 1: (A) 4 sample stimuli of the set used in the current study. SS: Same parts/Same spatial configuration; SD: Same parts/Different spatial configuration; DS: Different parts/Same spatial configuration; DD: Different parts/Different spatial configuration.
(B): Experimental procedure. The figure shows the timeline for a single trial (ISI: interstimulus interval). Here S1 is followed by a "different" S2. In this case S2 possesses different parts, but the same configuration and is therefore classified as "DS".

Figure 2: Grand average ERP traces for 2D (above) and 3D (below) conditions for each stimulus type - SS: Same parts/Same configuration; SD: Same parts/Different configuration; DS: Different parts/Same configuration; DD: Different parts/Different configuration. The inset at the centre shows the electrode placement (scalp viewed from above with the nose on top and the left ear on the left). Two representative

electrodes are shown: Electrodes P9 and P10, highlighted in blue in the inset. Electrodes highlighted in purple show the left and right posterior group of electrodes (regions of interest – ROI) that were used for the statistical computation of the N1, P2 and N2.

Figure 3: Grand average ERP traces for SS (above) and DD (below) stimuli under 2D (black) and 3D (red) viewing conditions. The inset at the centre shows the electrode placement (scalp viewed from above with the nose on top and the left ear on the left). Two representative electrodes are shown: Electrodes P9 and P10, highlighted in blue in the inset. Electrodes highlighted in purple show the left and right posterior group of electrodes (regions of interest – ROI) used for statistical computation.

Figure 4: 2D versus 3D point-wise mass univariate contrast comparing 2D and 3D presentations over time (x axis) and electrodes (y axis). All 204 electrodes are shown with right frontal leads on top, followed by the left frontal, left posterior and finally the right posterior leads. The time scale is shown below with 0 indicating the onset of stimulus presentation. Dark areas in the panel above indicate periods and electrodes significant at $p<.01$. The two arrows indicate the spatial position of the significant electrodes on the scalp at the given time instant. On the representation below, the red circles highlight the electrodes significant at $p<.01$ at the time indicated by the arrow.

Figure 5: Mass univariate contrasts showing time (x axis) and electrodes (y axis) for 2D and 3D visual presentation for each stimulus type. **A:** (a) 2D: SS/DS (b) 2D: SS/SD **B:** (c) 2D: SS/DD (d) 3D: SS/DS (e) 3D: SS/SD (f) 3D: SS/DD.

All 204 electrodes are shown from top right to bottom right with right frontal leads on top, followed by the left frontal, left posterior and finally the right posterior leads. The time scale is shown below with 0 indicating the onset of (S2) stimulus presentation. Dark areas in the panel above indicate periods and electrodes significant at $p<.01$. For each, the electrode montage shows the electrodes significant at $p<.01$ at 300ms post-stimulus onset.

Figure 6: Time series distribution showing the frequency of significant difference contrasts from the mass univariate analysis between 210ms-390ms. The contrasts shown are between 2D and 3D viewing for SS/DD (blue), SS/SD (red) and SS/DS (green).

Figure 7: Mean amplitudes (in microVolts) of (A) the early N1: 220ms-290ms and (B) the late N1: 290ms-370ms for 2D (blue) and 3D (red) presentations as a function laterality (left, right panels) and stimulus type - SS: Same parts/Same configuration; SD: Same parts/Different configuration; DS: Different parts/Same configuration; DD: Different parts/Different configuration. Bars show 95% confidence intervals.

REFERENCES

Akay, A., & Celebi, G. (2009) A brain electrophysiological correlate of depth perception. *Neurosciences*, 14, 139-142.

Arguin, M., & Saumier, D. (2004). Independent processing of parts and of their spatial organisation in complex visual objects. *Psychological Science, 15*, 629-633.

Arguin, M., & Leek, E. C. (2003). Orientation invariance in visual object priming depends on prime-target asynchrony. *Percept Psychophys, 65*(3), 469-477.

Behrmann, M., Peterson, M.A., Moscovitch, M. & Satoru, S. (2006). Independent representation of parts and the relations between them: Evidence from integrative agnosia. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 1169-1184.

Behrmann, M. & Kimchi, R. (2003). What does visual agnosia tell us about perceptual organisation and its relationship to object perception? *Journal of Experimental Psychology: Human Perception and Performance, 29*, 19-42.

Bennett, D. J., & Vuong, Q. C. (2006). A stereo advantage in generalizing over changes in viewpoint on object recognition tasks. *Percept Psychophys, 68*(7), 1082-1093.

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychol Rev, 94*(2), 115-147.

Burgess, A. P., Rehman, J., & Williams, J. D. (2003). Changes in neural complexity during the perception of 3D images using random dot stereograms. *Int J Psychophysiol, 48*(1), 35-42.

Burke, D., Taubert, J., & Higman, T. (2007). Are face representations viewpoint dependent? A stereo advantage for generalizing across different views of faces. *Vision Res, 47*(16), 2164-2169.

Cadieu, C., Kouh, M., Pasupathy, A. Connor, C., Riesenhuber, M. & Poggio, T. (2007). A mode of v4 shape selectivity and invariance. *Journal of Neurophysiology, 98*, 1733-1750.

Chan, M. W., Stevenson, A. K., Li, Y., & Pizlo, Z. (2006). Binocular shape constancy from novel views: the role of a priori constraints. *Percept Psychophys, 68*(7), 1124-1139.

Courtney, S. M., Petit, L., Haxby, J. V., & Ungerleider, L. G. (1998). The role of prefrontal cortex in working memory: examining the contents of consciousness. *Philos Trans R Soc Lond B Biol Sci, 353*(1377), 1819-1828.

Cristino, F., Davitt, L., Hayward, W. G., & Leek, E. C. (2015). Stereo disparity facilitates view generalization during shape recognition for solid multipart objects. *Q J Exp Psychol (Hove), 68*(12), 2419-2436.

Curran, T., Tanaka, J. W., & Weiskopf, D. M. (2002). An electrophysiological comparison of visual categorization and recognition memory. *Cogn Affect Behav Neurosci, 2*(1), 1-18.

DeAngelis, G.C., & Newsome, W.T. (1999). Organisation of disparity-selective neurons in macaque area MT. *Journal of Neuroscience, 19*, 1398-1415.

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

Doniger, G. M., Foxe, J. J., Schroeder, C. E., Murray, M. M., Higgins, B. A., & Javitt, D. C. (2001). Visual perceptual learning in human object recognition areas: a repetition priming study using high-density electrical mapping. *Neuroimage, 13*(2), 305-313.

Edelman, S., & Bulthoff, H. H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vision Res, 32*(12), 2385-2400.

Friedman, D., & Johnson, R., Jr. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: a selective review. *Microsc Res Tech, 51*(1), 6-28.

Gao, F., Cao, B., Cao, Y., Li, F., & Li, H. (2015). Electrophysiological evidence of separate pathways for the perception of depth and 3D objects. *Int J Psychophysiol, 96*(2), 65-73.

Gilaie-Dotan, S., Ullman, S., Kushnir, T., & Malach, R. (2001). Shape-selective stereo processing in human object-related visual areas. *Human Brain Mapping, 15*, 67-79.

Groppe, D.M., Urbach, T.P., Kutas, M. (2011) Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology. 48*, 1711-1725.

Guthrie, D., & Buchwald, J.S. (1991). Significance testing of difference potentials. *Psychophysiology, 28*, 240-244.

Harris, I. M., Dux, P. E., Benito, C. T., & Leek, E. C. (2008). Orientation sensitivity at different stages of object processing: evidence from repetition priming and naming. *PLoS One, 3*(5), e2256.

Hummel, J. E. (2013). Object Recognition. In D. Reisburg (Ed.), *Oxford Handbook of Cognitive Psycholog* (pp. 32-46). Oxford: Oxford University Press.

Hummel, J. E. & Stankiewicz, B. J., (1996). An architecture for rapid, hierarchical structural description. In T. Inui & J. McCelland (Eds.), *Attention and Performance XVI: On information integration in perception and communication* (pp.93-121). Cambridge, MA: MIT Press.

Humphrey, G. K., & Khan, S. C. (1992). Recognizing novel views of three-dimensional objects. *Can J Psychol, 46*(2), 170-190.

Kasai, T., & Morotomi, T. (2001). Event-related brain potentials during selective attention to depth and form in global stereopsis. *Vision Res, 41*(10-11), 1379-1388.

Kasai, T., Morotomi, T., Katayama, J., & Kumada, T. (2003). Attending to a location in three-dimensional space modulates early ERPs. *Brain Res Cogn Brain Res, 17*(2), 273-285.

Khateb, A., Annoni, J. M., Landis, T., Pegna, A. J., Custodi, M. C., Fonteneau, E., et al. (1999). Spatio-temporal analysis of electric brain activity during semantic and phonological word processing. *Int J Psychophysiol, 32*(3), 215-231.

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. *Vision Res, 46*(11), 1762-1776.

Klingberg, T., Kawashima, R., & Roland, P. E. (1996). Activation of multi-modal cortical areas underlies short-term memory. *Eur J Neurosci, 8*(9), 1965-1971.

Kourtzi, Z., & Connor, C. E. (2011). Neural representations for object perception: structure, category, and adaptive coding. *Annu Rev Neurosci, 34*, 45-67.

Lee, Y. L., & Saunders, J. A. (2011). Stereo improves 3D shape discrimination even when rich monocular shape cues are available. *J Vis, 11*(9).

Leek, E. C. (1998a). The analysis of viewpoint-dependent time costs in visual recognition. *Perception, 27*, 803-816.

Leek, E. C. (1998b). Effects of stimulus viewpoint on the identification of common polyoriented objects. *Psychonomic Bulletin & Review, 5*, 650-658.

Leek, E.C., Reppa, I., & Arguin, M. (2005). The structure of 3D object shape representations: Evidence from part-whole matching. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 668-684.

Leek, E. C., & Johnston, S. J. (2006). A polarity effect in misoriented object recognition: The role of polar features in the computation of orientation-invariant shape representations. *Visual Cognition, 13*, 573-600.

Leek, E.C., Atherton, C. J., Thierry, G. (2007) Computational mechanisms of object constancy for visual recognition revealed by event-related potentials. Vision Res;47(5):706-713.

Lehmann, D. (1987). *Methods of analysis of brain electrical and magnetic signals. Handbook of electroencephalography and clinical neurophysiology (Vol. 1).* . Amsterdam: Elsevier.

Lehmann, D., & Skrandies, W. (1984). Spatial analysis of evoked potentials in man: a review. *Progress in Neurobiology, 23*(3), 227-250.

Liu, C. H., Ward, J., & Young, A. W. (2006). Transfer between two- and three-dimensional representations of faces. *Visual Cognition, 13*(1), 51-64.

Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, 240, 740-749.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proc R Soc Lond B Biol Sci, 200*(1140), 269-294.

Martin, A. (2007). The representation of object concepts in the brain. *Annu Rev Psychol, 58*, 25-45.

Mecklinger, A. (2000). Interfacing mind and brain: a neurocognitive model of recognition memory. *Psychophysiology, 37*(5), 565-582.

Mendola, J. D., Dale, A. M., Fischl, B., Liu, A. K., & Tootell, R. B. (1999). The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *J Neurosci, 19*(19), 8560-8572.

Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., & Grave de Peralta, R. (2004). *EEG source imaging*. Paper presented at the Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology.

Michel, C. M., Thut, G., Morand, S., Khateb, A., Pegna, A. J., Grave de Peralta, R., et al. (2001). Electric source imaging of human brain functions. *Brain Research Reviews, 36*(2-3), 108-118.

Mouchetant-Rostaing, Y., & Giard, M. H. (2003). Electrophysiological correlates of age and gender perception on human faces. *J Cogn Neurosci, 15*(6), 900-910.

Murray, M. M., Brunet, D., & Michel, C. (2008). Topographic ERP Analyses: A Step-by-Step Tutorial Review. *Brain Topogr, 20*(4), 249-264.

Murray, M. M., Imber, M. L., Javitt, D. C., & Foxe, J. J. (2006). Boundary completion is automatic and dissociable from shape discrimination. *J Neurosci, 26*(46), 12043-12054.

Murray, M. M., Wylie, G. R., Higgins, B. A., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). The spatiotemporal dynamics of illusory contour processing: combined high-density electrical mapping, source analysis, and functional magnetic resonance imaging. *J Neurosci, 22*(12), 5055-5073.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh Inventory. *Neuropsychologia, 9*, 97-113.

Oram, M. W., & Perrett, D. I. (1992). Time course of neural responses discriminating different views of the face and head. *J Neurophysiol, 68*(1), 70-84.

Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1994). Low resolution electromagnetic tomography: a new method to localize electrical activity in the brain. *Int J Psychophysiol, 18*, 49-65.

Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1995). Segmentation of brain electrical activity into microstates: model estimation and validation. *Biomedical Engineering, IEEE Transactions on, 42*(7), 658-665.

Pasqualotto, A., & Hayward, W. G. (2009). A stereo disadvantage for recognizing rotated familiar objects. *Psychon Bull Rev, 16*(5), 832-838.

Pegna, A. J., Khateb, A., Murray, M. M., Landis, T., & Michel, C. M. (2002). Neural processing of illusory and real contours revealed by high-density ERP mapping. *Neuroreport, 13*(7), 965-968.

Pegna, A. J., Khateb, A., Spinelli, L., Seeck, M., Landis, T., & Michel, C. M. (1997). Unraveling the cerebral dynamics of mental imagery. *Hum Br Mapp, 5*, 410-421.

Perrin, F., Pernier, J., Bertrand, O., Giard, M. H., & Echallier, J. F. (1987). Mapping of scalp potentials by surface spline interpolation. *Electroencephalogr Clin Neurophysiol, 66*(1), 75-81.

Pierce, L. J., Scott, L. S., Boddington, S., Droucker, D., Curran, T., & Tanaka, J. W. (2011). The n250 brain potential to personally familiar and newly learned faces and objects. *Front Hum Neurosci, 5*, 111.

Pizlo, Z., Sawada, T., Li, Y., Kropatsch, W. G., & Steinman, R. M. (2010). New approach to the perception of 3D shape based on veridicality, complexity, symmetry and volume. *Vision Res, 50*(1), 1-11.

Riesenhuber, M & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience, 2*, 1019-1025.

Rock, I., & DiVita, J. (1987). A case of viewer-centered object perception. *Cogn Psychol, 19*(2), 280-293.

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. (2007). Neurophysiological evidence for the time course of activation of global shape, part, and local contour representations during visual object categorization and memory. *J Cogn Neurosci, 19*(5), 734-749.

Schendan, H. E., & Lucia, L. C. (2010). Object-sensitive activity reflects earlier perceptual and later cognitive processing of visual objects between 95 and 500ms. *Brain Res, 1329*, 124-141.

Schweinberger, S. R., Pickering, E. C., Jentzsch, I., Burton, A. M., & Kaufmann, J. M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Brain Res Cogn Brain Res, 14*(3), 398-409.

Scott, L. S., Tanaka, J. W., Sheinberg, D. L., & Curran, T. (2006). A reevaluation of the electrophysiological correlates of expert object processing. *J Cogn Neurosci, 18*(9), 1453-1465.

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.

Serre, T., Oliva, A., & Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proc Natl Acad Sci U S A, 104*(15), 6424-6429.

Severac Cauquil, A., Trotter, Y., & Taylor, M. J. (2006). At what stage of neural processing do perspective depth cues make a difference? *Exp Brain Res, 170*(4), 457-463.

Spinelli, L., Andino, S. G., Lantz, G., Seeck, M., & Michel, C. M. (2000). Electromagnetic inverse solutions in anatomically constrained spherical head models. *Brain Topogr, 13*(2), 115-125.

Szatkowska, I., Grabowska, A., & Szymanska, O. (2001). Evidence for the involvement of the ventro-medial prefrontal cortex in a short-term storage of visual images. *Neuroreport, 12*(6), 1187-1190.

Tanaka, J. W., Curran, T., Porterfield, A. L., & Collins, D. (2006). Activation of preexisting and acquired face representations: the N250 event-related potential as an index of face familiarity. *J Cogn Neurosci, 18*(9), 1488-1497.

Tanaka, H., Uka, T, Yoshiyama, K., Kato, M., & Fujita, I. (2001). Processing of shape defined by disparity in monkey inferior temporal cortex. *Journal of Neurophysiology, 85*, 735-744.

Tarr, M. J., & Bulthoff, H. H. (Eds.). (1999). *Object recognition in man, monkey, and machine*. Cambridge, MA: MIT Press.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*(6582), 520-522.

Tibshirani, R., & Walther, G. (2005). Cluster Validation by Prediction Strength. *Journal of Computational and Graphical Statistics, 14*(3), 511-528.

Ullman, S. (2006). Object recognition and segmentation by a fragment-based hierarchy. *Trends in Cognitive Science, 11*, 58-64.

VanRullen, R., & Thorpe, S. J. (2001). Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects. *Perception, 30*(6), 655-668.

Vogels, R. (1999). Effect of image scrambling on inferior temporal cortical responses. *Neuroreport, 10*(9), 1811-1816.

Watt, S.J. & Bradshaw, M.F. (2003). The visual control of reach and grasping: Binocular disparity and motion parallax. Journal of Experimental Psychology: Human Perception and Performance, 29, 404-415.

			Effect of Stereo vs Non-stereo	Effect of Stimulus shape
Behaviour	*Errors*		ns	SS = DS > SD = DD
N1	*Amplitude*		3D<2D	ns
	Latency		ns	ns
P2	*Mean amplitude (220-250ms)*		3D>2D	ns
N2 (early)	*Mean amplitude (240-290ms)*	*Left ROI*	3D<2D	Left ROI: SS<DS/SD/DD
		Right ROI	3D<2D	Right ROI:SS<DS/SD/DD
N2 (late)	*Amplitude (290-370ms)*	*Left ROI*	ns	SS<DS/SD/DD
		Right ROI	2D<3D	SS<DS/SD/DS
			And interaction of stereo viewing with stimulus shape: amplitude of SS in 2D = amplitude of SS in 3D amplitude of SD in 2D = amplitude of SD in 3D amplitude of DS in 2D < amplitude of DS in 2D 3D amplitude of DD in 2D = amplitude of DD in 2D 3D	

Table 1

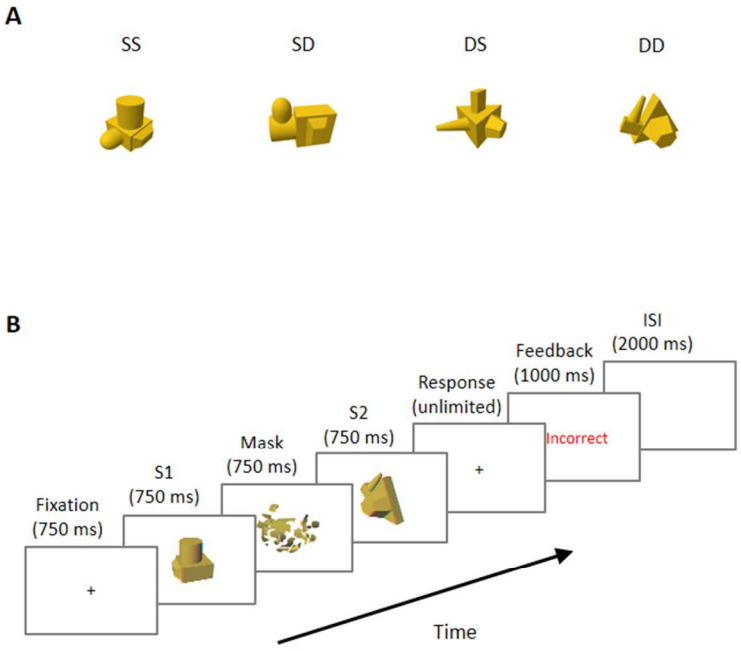

FIGURE 1

Figure 1: (A) 4 sample stimuli of the set used in the current study. SS: Same parts/Same spatial configuration; SD: Same parts/Different spatial configuration; DS: Different parts/Same spatial configuration; DD: Different parts/Different spatial configuration.
(B): Experimental procedure. The figure shows the timeline for a single trial (ISI: interstimulus interval). Here S1 is followed by a "different" S2. In this case S2 possesses different parts, but the same configuration and is therefore classified as "DS".

INSERT FIGURE 1 ABOUT HERE
190x275mm (96 x 96 DPI)

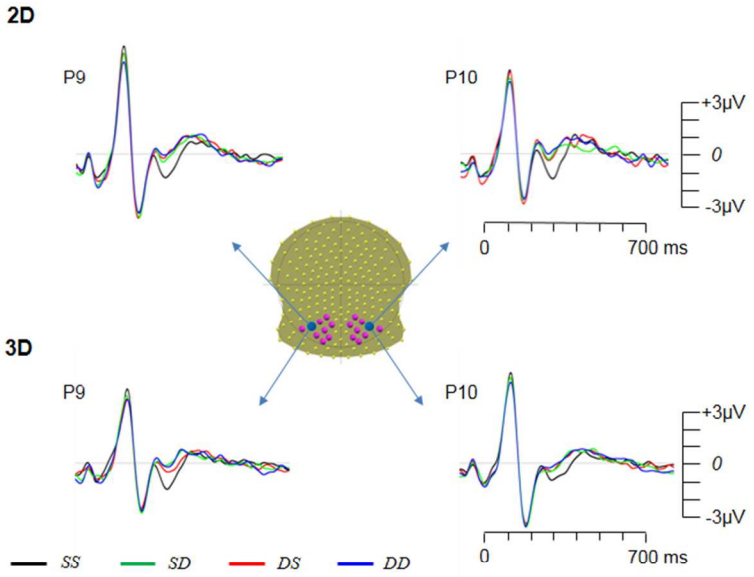

FIGURE 2

Grand average ERP traces for 2D (above) and 3D (below) conditions for each stimulus type - SS: Same parts/Same configuration; SD: Same parts/Different configuration; DS: Different parts/Same configuration; DD: Different parts/Different configuration. The inset at the centre shows the electrode placement (scalp viewed from above with the nose on top and the left ear on the left). Two representative electrodes are shown: Electrodes P9 and P10, highlighted in blue in the inset. Electrodes highlighted in purple show the left and right posterior group of electrodes (regions of interest – ROI) that were used for the statistical computation of the N1, P2 and N2.

INSERT FIGURE 2

190x275mm (96 x 96 DPI)

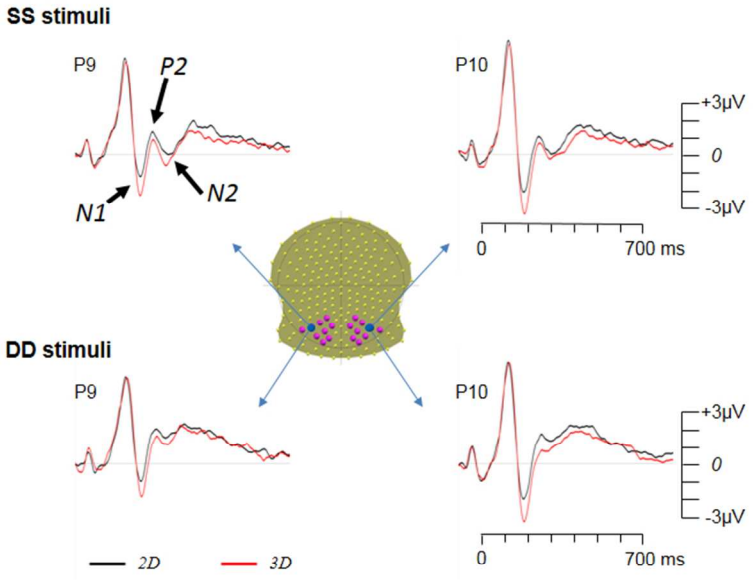

FIGURE 3

Figure 3: Grand average ERP traces for SS (above) and DD (below) stimuli under 2D (black) and 3D (red) viewing conditions. The inset at the centre shows the electrode placement (scalp viewed from above with the nose on top and the left ear on the left). Two representative electrodes are shown: Electrodes P9 and P10, highlighted in blue in the inset. Electrodes highlighted in purple show the left and right posterior group of electrodes (regions of interest – ROI) used for statistical computation.
and 3 ABOUT HERE
190x275mm (96 x 96 DPI)

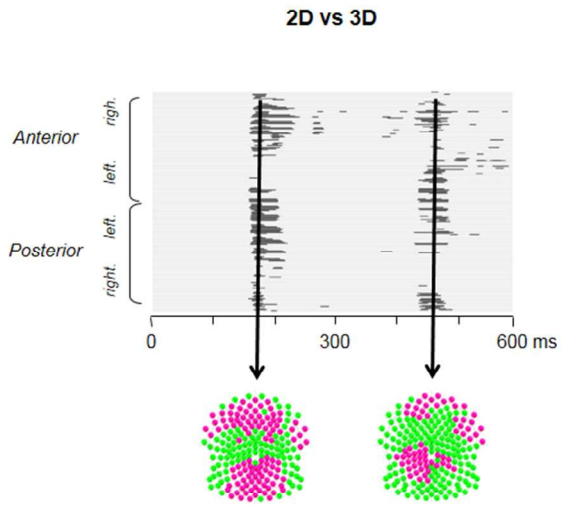

FIGURE 4

Figure 4: 2D versus 3D point-wise mass univariate contrast comparing 2D and 3D presentations over time (x axis) and electrodes (y axis). All 204 electrodes are shown with right frontal leads on top, followed by the left frontal, left posterior and finally the right posterior leads. The time scale is shown below with 0 indicating the onset of stimulus presentation. Dark areas in the panel above indicate periods and electrodes significant at $p<.01$. The two arrows indicate the spatial position of the significant electrodes on the scalp at the given time instant. On the representation below, the red circles highlight the electrodes significant at $p<.01$ at the time indicated by the arrow.

INSERT FIGURE 4 ABOUT HERE

190x275mm (96 x 96 DPI)

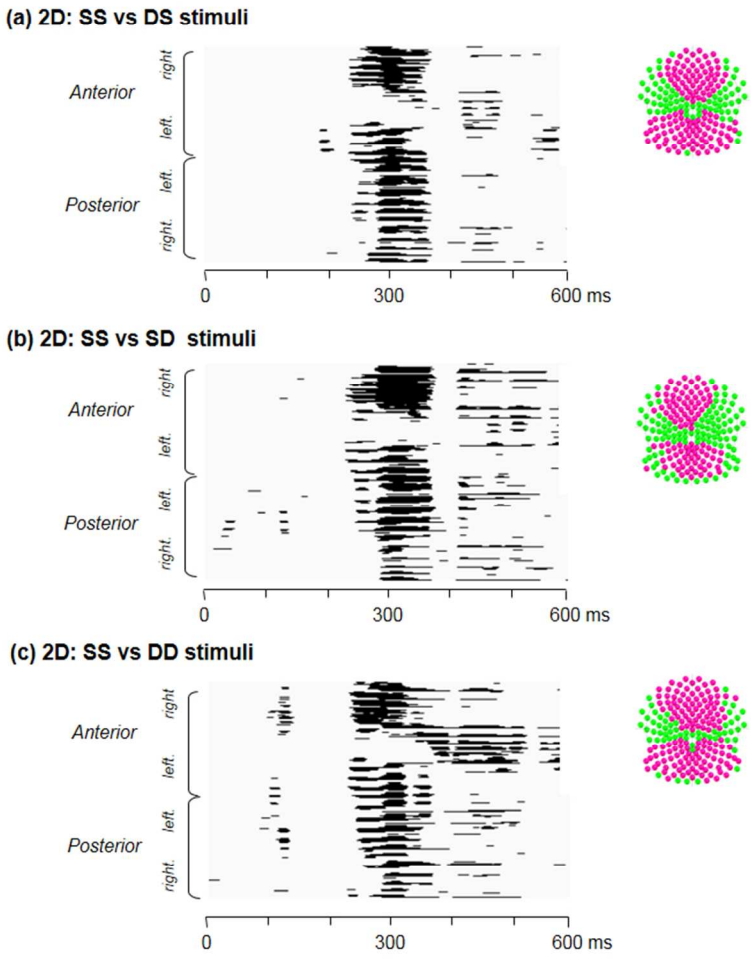

FIGURE 5A

Figure 5: Mass univariate contrasts showing time (x axis) and electrodes (y axis) for 2D and 3D visual presentation for each stimulus type. A: (a) 2D: SS/DS (b) 2D: SS/SD

INSERT FIGURE 5A

190x275mm (96 x 96 DPI)

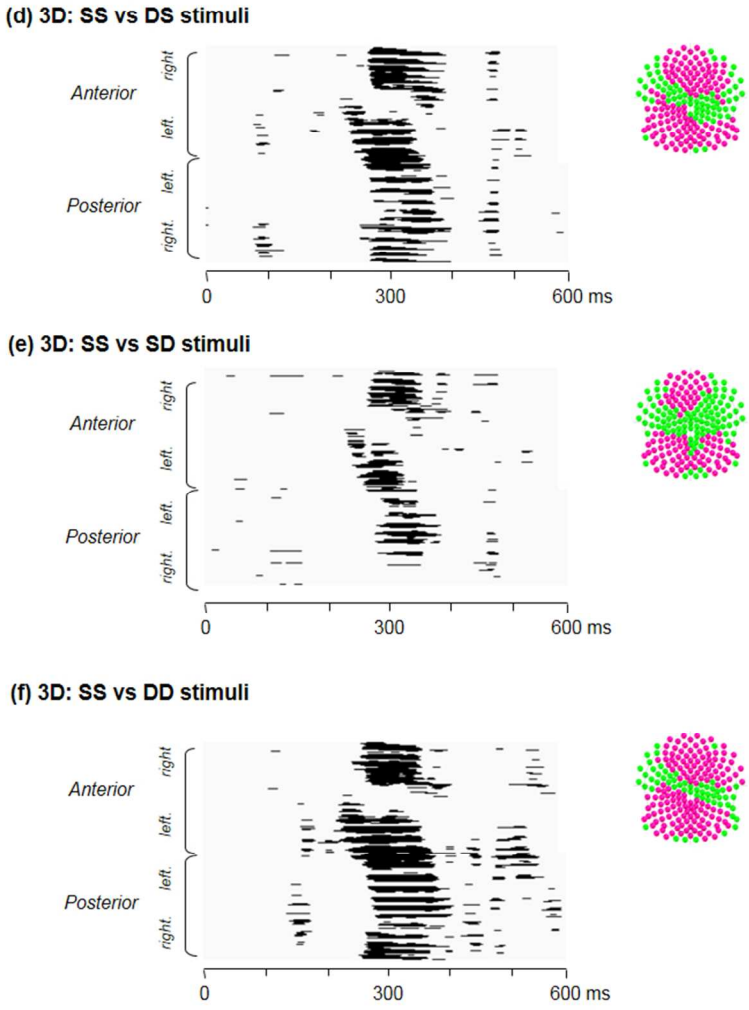

FIGURE 5B

Figure 5B: (c) 2D: SS/DD (d) 3D: SS/DS (e) 3D: SS/SD (f) 3D: SS/DD.† All 204 electrodes are shown from top right to bottom right with right frontal leads on top, followed by the left frontal, left posterior and finally the right posterior leads. The time scale is shown below with 0 indicating the onset of (S2) stimulus presentation. Dark areas in the panel above indicate periods and electrodes significant at p<.01. For each, the electrode montage shows the electrodes significant at p<.01 at 300ms post-stimulus onset.
and B ABOUT HERE
190x275mm (96 x 96 DPI)

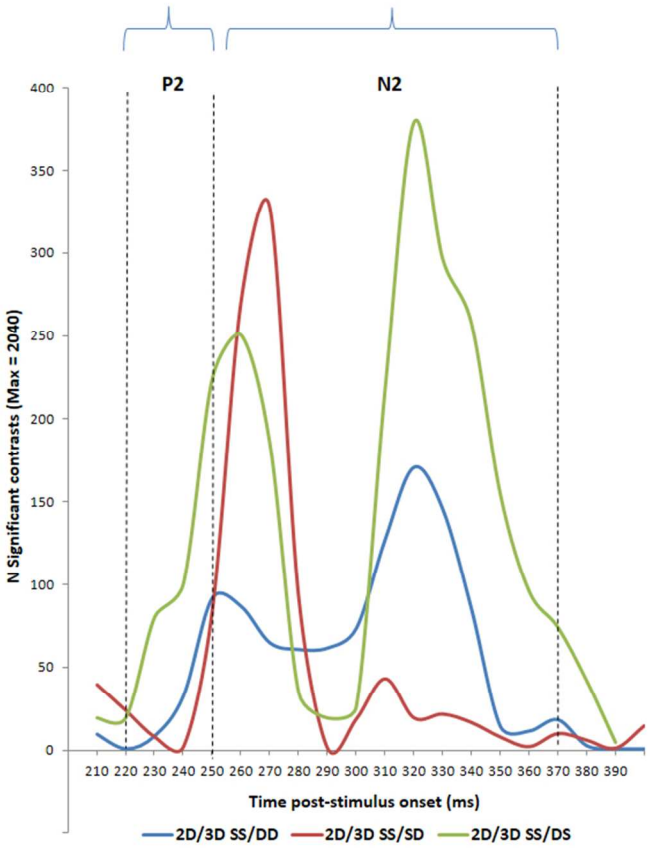

FIGURE 6

Figure 6: Time series distribution showing the frequency of significant difference contrasts from the mass univariate analysis between 210ms-390ms. The contrasts shown are between 2D and 3D viewing for SS/DD (blue), SS/SD (red) and SS/DS (green).

INSERT FIGURE 6 ABOUT HERE

190x275mm (96 x 96 DPI)

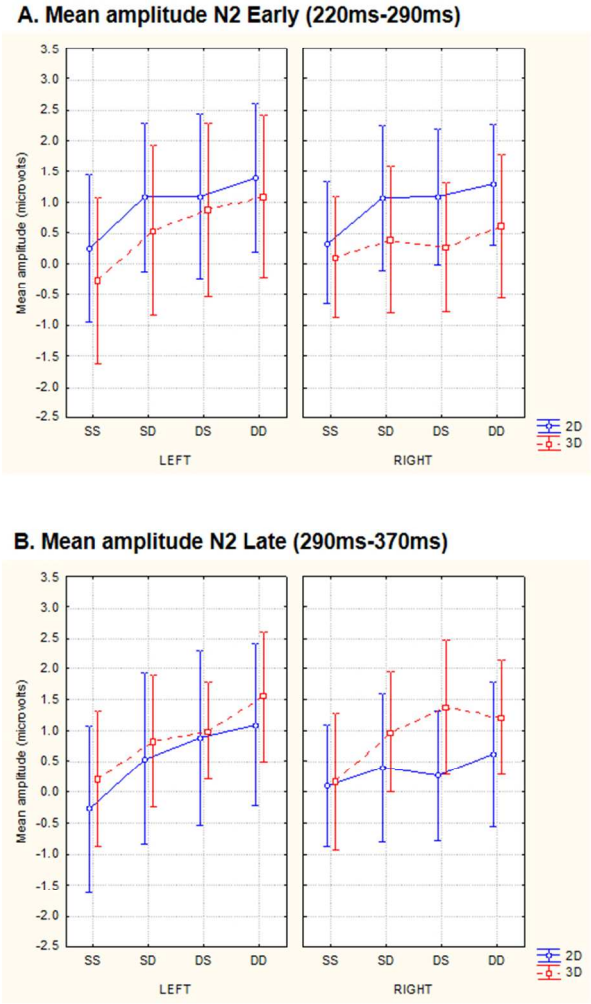

FIGURE 7

Figure 7: Mean amplitudes (in microVolts) of (A) the early N1: 220ms-290ms and (B) the late N1: 290ms-370ms for 2D (blue) and 3D (red) presentations as a function laterality (left, right panels) and stimulus type - SS: Same parts/Same configuration; SD: Same parts/Different configuration; DS: Different parts/Same configuration; DD: Different parts/Different configuration. Bars show 95% confidence intervals.
INSERT FIGURE 7 ABOUT HERE
190x275mm (96 x 96 DPI)