Modulations of depth responses in the human brain by object context: Does biological relevance matter?

https://doi.org/10.1523/ENEURO.0039-21.2021

Cite as: eNeuro 2021; 10.1523/ENEURO.0039-21.2021

Received: 28 January 2021
Revised: 25 May 2021
Accepted: 10 June 2021

This Early Release article has been peer-reviewed and accepted, but has not been through the composition and copyediting processes. The final version may differ slightly in style or formatting and will contain links to any extended data.

Alerts: Sign up at www.eneuro.org/alerts to receive customized email alerts when the fully formatted version of this article is published.
Modulations of depth responses in the human brain by object context: Does biological relevance matter?

**Abbreviated Title:** Depth responses in biological contexts

Idy W. Y. Chou¹, Hiroshi Ban²,³, Dorita H. F. Chang¹,⁴

¹Department of Psychology, The University of Hong Kong, Hong Kong
²Center for Information and Neural Networks (CiNet), National Institute of Information and Communications Technology, Japan
³Graduate School of Frontier Biosciences, Osaka University, Japan
⁴State Key Laboratory of Brain and Cognitive Sciences, The University of Hong Kong, Hong Kong

**Author contributions:**

IC – Data curation; Formal analysis; Writing – original draft; Writing – review & editing.
HB – Data curation; Formal analysis; Writing – original draft; Writing – review & editing; Funding acquisition.
DC – Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Writing - original draft; Writing - review & editing.

**Correspondence should be addressed to:**

Dorita H. F. Chang
Department of Psychology
Jockey Club Tower
The University of Hong Kong
Pokfulam, Hong Kong
Tel: (852) 3971-2383
E-mail: changd@hku.hk

**Counts:**

Number of Figures: 4
Number of Tables: 1
Number of Multimedia: 0
Number of words for Abstract: 219
Number of words for Significance Statement: 89
Number of words for Introduction: 748
Number of words for Discussion: 2157

**Acknowledgments:**

This research was supported by Early Career Scheme, General Research Fund (27612119 and 17612920, Research Grants Council, Hong Kong), and the Foreign Researcher
Conflicts of Interest:

The authors declare no conflicts of interest.

Funding sources:

Early Career Scheme and General Research Fund Grants (27612119 and 17612920, Research Grants Council, Hong Kong)

The Foreign Researcher Invitation Program (National Institute of Information and Communications Technology, Japan)

JSPS KAKENHI (21H00968), and ERATO (JPMJER1801) to HB.
Abstract

Depth sensitivity has been shown to be modulated by object context (plausibility). It is possible that it is behavioural relevance rather than object plausibility per se which drives this effect. Here, we manipulated the biological relevance of objects (face or a non-face) and tested whether object relevance affects behavioural sensitivity and neural responses to depth-position. In a first experiment, we presented human observers with disparity-defined faces and non-faces, and observers were asked to judge the depth position of the target under signal-noise and clear (fine) task conditions. In the second experiment, we concurrently measured behavioural and fMRI responses to depth. We found that behavioural performance varied across stimulus conditions such that they were significantly worse for the upright face than the inverted face and the random shape in the SNR task, but worse for the random shape than the upright face in the feature task. Pattern analysis of fMRI responses revealed that activity of FFA was distinctly different during depth judgments of the upright face versus the other two stimuli, with its responses (and to a stronger extent, those of V3) appearing functionally-relevant to behavioural performance. We speculate that FFA is not only involved in object analysis, but exerts considerable influence on stereoscopic mechanisms as early as in V3 based on a broader appreciation of the stimulus’ behavioural relevance.
Significance Statement

We asked how disparity sensitivity is modulated by object (biological) relevance using behavioural and neuroimaging paradigms. We show that behavioural sensitivity to depth-position changes in biological (face) vs non-biological (random surface) contexts, and that these changes are task-dependent. Imaging results highlight a potentially key role of the fusiform region for governing the modulation of stereo encoding by object relevance. These findings highlight powerful interactions between object recognition mechanisms and stereoeencoding, such that the utility of disparity information may be up/down weighed depending on the biological relevance of the object.
1 Introduction

Binocular disparity, referring to the difference in the left- and right-eye’s image that is a consequence of the two eyes’ separation on the head, is crucial to our everyday ability to recognize and interact with objects (Burke, 2005; Dal Mutto et al., 2011). Neurophysiological work have indicated a wide network of regions in the primate brain, starting from V1, and extending dorsally to V3 and the MT complex, as well as ventrally to V4 and IT, that respond selectively to binocular disparity (Hubel & Wiesel, 1970; Poggio & Fischer, 1977; Poggio & Talbot, 1981; Poggio et al., 1988; Roy et al., 1992; Cumming & Parker, 1997, 1999; DeAngelis et al., 1998; Uka et al., 2000; Hinkle & Conor, 2001). Many of these same areas are responsible for more complex representations rendered by stereoscopic cues. For instance, the caudal intraparietal (CIP) region along dorsal cortex, and V4 and along ventral cortex, respond to 3D slant (P. Janssen et al., 1999, 2000; Hegdé & Van Essen, 2005; Sakata et al., 2005; Durand et al., 2007).

More recent studies suggest that the particular region engaged during the perception of depth from disparity depends on the task employed. In the primate, the MT complex appears to respond during judgments of disparity signals from noise (DeAngelis et al., 1998; Uka & DeAngelis, 2003, 2004, 2006). On the other hand, V4 and IT seem to be central to the discrimination of fine, but clear depth differences (Uka et al., 2005; Shiozaki et al., 2012). This apparent segregation of the roles of dorsal and ventral regions in different depth discrimination tasks has been observed in humans as well (Chang et al., 2014).

Task-based modulation of stereo responses observed in the cortex suggest that the disparity-processing circuitry is sensitive to context. Recent work has shown that in addition to task-requirements, disparity mechanisms can be sensitive to an entirely different form of context – object identity/class. In particular, depth sensitivity is surprisingly modulated by object plausibility (Wong et al., 2019). Specifically, during depth position-judgment tasks, performance was worse for physically plausible versus implausible objects. This effect is particularly striking as both classes of stimuli were
equivalently complex, and perhaps more importantly, the plausibility of the object was
task-irrelevant! Wong et al. (2019) demonstrated further that the physical plausibility
of the object appeared to modulate disparity-related multivariate response patterns in
intermediate and higher dorsal (V3A, V3B, V7) and ventral (LO) regions. These findings
highlight the importance of object-level information in processing stereoscopic signals.

Here, we sought to better understand the nature of the dependence of depth
responses on object context by asking whether the biological relevance of objects can
affect behavioural sensitivity and neural responses to stereoscopic depth. We
manipulated the relevance of objects by presenting stereoscopic face stimuli (at upright
and inverted orientations) and non-face surfaces matched in depth-power (i.e.
amplitude). Indeed, object (biological) relevance has been shown to play a role in the
perception of many visual attributes. For instance, the perceived colour of food items is
biased toward the familiar colour of the object stored in long-term memory (Hansen et
al., 2006; Olkkonen et al., 2008). Moreover, the race category of a face can affect the
perceived lightness of the face (Levin & Banaji, 2006; Chang et al., 2018). The human
visual system’s apparent sensitivity to faces is of course now well known to be served by
dedicated regions for processing these stimuli, including the fusiform face area (FFA;
Kanwisher et al., 1997; McCarthy et al., 1997), the superior temporal sulcus (Puce et al.,
1995, 1996) and the occipital face area (Gauthier et al., 2000; Pitcher et al., 2007).
Interestingly, the illusory percepts of faces are sufficient to elicit activity in these same
regions (Liu et al., 2014).

The apparent importance of faces to our visual system suggests that it is an
appropriate stimulus context to be used to explore the effect of biological relevance on
disparity processing. In two experiments, we investigated how behavioural sensitivity
(Experiment 1) and fMRI responses (Experiment 2) to depth are modulated by biological
relevance. We created random-dot stereograms (RDS) depicting faces (upright and
inverted) and depth-matched random shapes. We elected to include an upside-down
variant of the face stimulus as it is well established that faces are more poorly perceived
when they are presented inverted (Yin, 1969). Turning the face upside-down thus,
presents an interesting stimulus condition that allows us to preserve the structural
integrity of an upright face, but is yet reduced in terms of its biological relevance.

2 Materials and Methods

2.1 Experiment 1 (Behaviour)

In Experiment 1, we measured depth thresholds in two tasks: a signal-in-noise (SNR)
task of judging the depth position (near/far) of a central target relative to the
surrounding plane when the signal-to-noise ratio (SNR) varied between 0 and 100%, and
a feature task that required discriminating which of two consecutively presented clear
objects was nearer. We deemed it important to index performance according to both
tasks as they have been shown to engage distinct regions in cortex (Chang et al., 2014;
Patten & Welchman, 2015).

2.1.1 Participants

A total of sixty observers were recruited for this first experiment: Thirty
participants (aged 19 – 34 years; mean 24.9 years; 10 males), were tested on the SNR
task and a separate group of thirty participants (aged 18 – 34 years; mean 22.0 years; 12
males) participated in the feature task. All participants were screened for stereo deficits
(Titmus test), had normal or corrected-to-normal vision as screened with the Snellen
linear acuity chart, and provided written informed consent in line with ethical approval
by the institutional Human Research Ethics Committee (HREC).

2.1.2 Apparatus

Stimuli were generated using MATLAB (MathWorks) with extensions from the
Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and presented through a mirror
stereoscope, in which two eyes viewed the left and right halves of a 24-inch monitor
(1920 x 1080 resolution; 60 Hz refresh rate) through four front-surface mirrors. Viewing
distance was 65 cm, and a chinrest was used to limit head movements and maintain viewing distance.

2.1.3 Stimuli

Three kinds of stimuli, including the upright face, the inverted face, and the depth-power-matched random shape, were rendered as random-dot stereograms (Figure 1a). 3D laser-scanned head models were firstly obtained from the Face Database of the Max Planck Institute for Biological Cybernetics (Troje & Bülthoff, 1996; Blanz & Vetter, 1999). All heads were in frontal view with hair removed. Four face identities, including two male and two female head models, were randomly selected and rendered as depth maps such that intensity coded for depth. The size and shape of all faces were standardized by applying an oval-shape cropping mask that eliminated the ears and the neck. The inverted faces were generated by rotating the upright faces by 180° in the picture plane. The depth-power-matched random shapes were created by taking the depth map of each upright face, applying a Fourier transform, and randomizing its phases while preserving the amplitudes. The resulting images thus carry no structural coherence but maintain identical depth information as the upright and inverted variants. A total of four random shapes were created, each corresponding to one face identity.

The depth maps were finally rendered as random-dot stereograms (RDSs), where the amount of pixel-wise disparity was indexed by intensity. Each RDS had a maximum disparity of ±5.3 arcmin and a size of 8.68 deg × 8.68 deg, within which the target subtended 6.45 deg vertically and 4.55 deg horizontally. Each target had an average disparity of ±4.27 arcmin. Dots were randomly assigned as black or white, with a size of 0.05 deg and a density of 20 dots/deg². The RDS was presented against a black background and surrounded by a grid of grey and white squares (each 0.5 deg in size) that served to provide an unambiguous background reference and stabilize vergence position.

2.1.4 Tasks
**SNR task**

On each trial, participants were required to judge whether a central target was ‘near’ or ‘far’ as compared to the surrounding region (the ‘surround’). Disparity of the surround was fixed at zero (i.e. at the plane of fixation). For the target, disparity was fixed at ±5.3 arcmin at the maximum point across the target’s surface. Task difficulty was manipulated by varying the signal-to-noise (SNR) ratio, altering the proportion of dots that defined the target (‘signal’) relative to noise dots that were assigned a random disparity within the range of ±5.3 arcmin (Figure 1b). The initial test value was 80% SNR. On each trial subsequently, the SNR of the stimulus was adjusted according to the QUEST staircase procedure, estimating thresholds at the 82% correct-level (Watson & Pelli, 1983).

**Feature task**

On each trial, participants were required to judge which of two consecutively presented objects was ‘nearer’. One of the objects was designated the ‘reference’ and had a fixed disparity of 7.8 arcmin. Task difficulty was manipulated by varying the disparity difference between the target and the reference within a range of 150 arcsec (Figure 1b). Both the target and reference had crossed disparity and were presented at 100% SNR. The initial test value of the target was set at 120 arcsec (relative to the reference). As for the SNR task, on each trial, disparity difference was adjusted according to the QUEST staircase procedure yielding threshold estimates at the 82% accuracy level (Watson & Pelli, 1983).

Prior to each task, participants were provided with oral instructions, viewed examples of the stimuli, and completed twelve practice trials with audio feedback in order to get familiarized with the tasks. Each task consisted of six test runs [3 conditions (upright face, inverted face, random-surface) × 2 stimulus identities] presented in random order. A test run contained one type of object with two interleaved staircases of stimulus gender (i.e. the gender of the head models from which the stimuli were derived). Each staircase had 64 trials that comprised of 4 practice trials and 60 test trials.
The test value of the first test trial was determined by the threshold obtained in the practice trials.

In both tasks, each trial began with a nonius-type fixation box (0.25 deg in size) that consisted of dichoptically presented nonius lines (each eye perceives only two of the four lines), which lasted for 500 ms. In the SNR task, each stimulus was presented for 300 ms. In the feature task, each stimulus was presented for 300 ms with an inter-stimulus interval of 500 ms. In both tasks, stimulus presentation was followed by a response period that ended after a response was made or 3,000 ms, whichever came first. Trials without a response after the maximum response period elapsed were deemed incorrect. Participants were required to respond using the arrow keys on a computer keyboard.

2.2 Experiment 2 (fMRI)

In Experiment 2, we measured fMRI responses when subjects were performing the SNR task, and examined both univariate and multivariate responses to depth under the different stimulus conditions. In this second experiment, we selected to test the SNR task as it elicited the more robust behavioural effects both here, and in previous work (Wong et al., 2019).

2.2.1 Participants

Twenty-two participants (aged 21 – 34; mean 23.2 years; 17 males) who did not participate in the behaviour experiments were recruited for the fMRI experiment. All participants had normal or corrected-to-normal vision and provided written informed consent in line with ethical approval by the institutional Human Research Ethics Committee (HREC).

2.2.2 Apparatus
Left and right images were back-projected using two projectors (WUX4000, Canon), each equipped with a polarizing filter (resolution: 1280 × 1024 pixels; 60 Hz refreshing rate) and placed at 96 cm from the back of the fMRI bore. Participants viewed the stimuli with the corresponding polarized glasses through a 45-deg tilted mirror mounted in front of the head.

2.2.3 Stimuli

Stimuli were the same as those used in Experiment 1 except for the following differences. First, for each condition (upright face/inverted face/random shape), only two stimulus exemplars were used, one per stimulus gender, resulting in a total of six unique images used in this experiment. This change was made to minimize the noise of fMRI signals across stimulus conditions that could result with large stimulus variance. Second, the target subtended 5.23 deg vertically and 3.70 deg horizontally to accommodate a smaller field of view in the bore setup. Third, the maximum disparity on the target surface was 5.4 arcmin.

2.2.4 fMRI acquisition

Imaging data were acquired at the Center for Information and Neural Networks (CiNet), National Institute of Information and Communications Technology (NICT), Osaka, Japan, using a 3-T Siemens Trio MR scanner with a half of the 32-channel, phase-array (whole) head coil that covered the occipital lobe. Head movements were limited by a foam padding inside the coil. For both localizer and experimental runs, blood oxygenation level–dependent (BOLD) signals were measured with a multiband echo-planar imaging (EPI) sequence [Voxel size = 2 × 2 × 2 mm³, Echo time (TE) = 30 ms, Repetition time (TR) = 2,000 ms, field of view = 192 × 192, Flip angle = 75°, 78 axial slices (slightly oblique along the AC–PC line), multiband factor = 3; 205 volumes where the first 5 were discarded to eliminate the effects of start-up transients] provided by the University of Minnesota (under a C2P contract). Additionally, high-resolution T1-weighted images were collected for each participant (Voxel size = 1 mm³, TE = 2.48 ms, TR = 1900 ms, field of view = 256 × 256, Flip angle = 9°, 208 slices).
2.2.5 ROI definition

For each participant, we defined regions-of-interest (ROIs) V1, V2, V3, V4, V3A, and V7 using standard phase-encoded retinotopic mapping procedures that mapped polar angle with a checkerboard wedge stimulus that was rotating periodically (clockwise or anti-clockwise; Sereno et al., 1995). Area V3B/KO (kinetic occipital) was defined using the same map, as the region falling anterior to V3A and inferior to V7 (Dupont et al., 1997). Using a single-run functional localizer, we identified hMT+ (human motion complex) as a set of contiguous voxels that showed significantly stronger responses (p < .01) to an array of moving dots that receded or expanded coherently than to an array of static dots (Huk et al., 2002). Note that due to time constraints, five participants did not complete the hMT+ localizer scan. For these participants, we defined hMT+ as fixed 5-mm spherical ROIs centered on Talairach coordinates of [−51, −72, 0] and [51, −69, 3] in the left and right hemispheres, respectively (Orban et al., 2003). LO (lateral occipital) was identified as the region in the lateral occipito-temporal cortex that was significantly more responsive (p < .01) to intact than scrambled images of objects and shapes (Kourtzi & Kanwisher, 2001). Bilateral FFAs (fusiform face area) were defined as 5-mm spherical ROIs for all participants, with Talairach coordinates of [−37, −42, −16] in the left hemisphere and [39, −40, −16] in the right hemisphere (Grill-Spector et al., 2004). Note that FFA was not localized in individual subjects as we saw no reason to extract responses from solely face-stimulus-specific voxels in this region.

2.2.6 Design and procedures

Prior to scanning, each participant firstly completed three runs of the SNR task inside the MRI bore (with the scanner idle). Thresholds estimated from these in-bore behaviour-only runs were used to determine individually-tailored stimulus test values for subsequent fMRI acquisition runs. This was done in order to standardize task difficulty across conditions and participants, thus reducing the noise of fMRI signals in our data. From these behaviour-only runs, we obtained subject-specific test values for the fMRI runs as follows: For each staircase (stimulus condition), we took the average of
the test values in last 30 trials and defined a range of ±1 SD from this mean value.

Subsequently, on each fMRI trial (of the corresponding condition), the SNR value was
randomly sampled within this range. For any instances where a single staircase could
not yield reliable estimates [due to: (i) a lack of asymptotic convergence, and/or (ii) 95%
confidence intervals reaching floor] stimuli were sampled at 80% SNR in the

Subsequently, on each fMRI trial (of the corresponding condition), the SNR value was
randomly sampled within this range. For any instances where a single staircase could
not yield reliable estimates [due to: (i) a lack of asymptotic convergence, and/or (ii) 95%
confidence intervals reaching floor] stimuli were sampled at 80% SNR in the
corresponding fMRI trials. This implicated a total of 19 (of 132 total staircases) run in
bore.

fMRI runs were arranged in a block design. Each run comprised seven block
types, corresponding to six stimulus conditions [3 (type) × 2 (gender)] and fixation. Each
block lasted for 16 s. A stimulus block comprised 8 trials with equal number of ‘near’
and ‘far’ trials. On each trial, a stimulus was presented for 500 ms, and was followed by
a fixed-duration response period of 1,500 ms. A background grid of randomly grey and
white squares was present throughout the entire run.

A fixation block contained a nonius-type fixation (see above for details). In each
run, each stimulus block type was repeated twice, with condition order randomised, and
interleaved with fixation blocks. Thus, there were a total of 12 stimulus blocks and 13
fixation blocks (each run began and ended with a fixation block). Each run consisted of
96 trials and lasted 6.4 min. The entire scanning session lasted around 90 min.

2.2.6 fMRI data analysis

Data were processed using BrainVoyager QX (Goebel et al., 2006; Formisano et
al., 2006). For each participant, T1-weighted anatomical scans were transformed into
Talairach space (Talairach & Tournoux, 1988). Inflated and flattened cortical surfaces
were reconstructed for both hemispheres. Functional data were pre-processed with
slice time correction, three-dimensional rigid-body motion correction, linear trend
removal and high-pass filtering (frequency space filter; 3 cycles per run) without spatial
smoothing.

We performed two analyses that examined both univariate [general linear model
(GLM)] responses and pattern-level responses using multivoxel pattern analysis (MVPA).
The GLM consisted of three regressors representing the three stimulus conditions (upright face, inverted face and random shape) and six regressors of motion parameters (three translation parameters and three rotation parameters) and a constant term. Stimulus-condition regressors were defined by modelling stimulation periods separately for each stimulus condition by a boxcar model convolved with a (two-gamma) haemodynamic response function (HRF; Glover, 1999). The averaged time course signal across all voxels of each ROI was then modelled as a linear combination of the regressors. The regressor coefficients of different stimulus conditions and ROIs were then contrasted.

MVPA classification analysis was done with a linear support vector machine implemented in libSVM tool (Chang & Lin, 2011) with a Recursive Feature Elimination (RFE) technique (De Martino et al., 2008). The RFE is a multivariate feature selection algorithm that eliminates uninformative voxels. The time course signals of all voxels were converted into z-scores. Also, to account for the typical delay of the haemodynamic response, signals were shifted by 4 seconds (= 2 TRs; Serences, 2004). For each ROI, the SVM was trained to classify the patterned responses between the upright face and the random shape, between the inverted face and the random shape, and between the upright face and the inverted face.

A leave-one-run-out cross-validation procedure was used for the MVPA. Specifically, in each cross-validation, functional data of one run were used as test dataset and the rest of the data were used as training dataset for computing SVM weights. In each RFE step, the SVM classifier was trained with 80% of the training dataset, which were resampled 20 times without partition of blocks. Therefore, each voxel was assigned with 20 SVM weights. All voxels were ranked according to the average weights in that particular RFE step, and the 5 voxels that had the lowest average weights were eliminated. The remaining data were used to decode the test pattern, and the prediction accuracy for that particular voxel pattern (in the particular RFE step) was obtained. This procedure was repeated until the voxel count fell below 250, after which it would proceed to the next cross-validation. As such, the final pattern
size for all ROIs here was 250 voxels. Note that the analysis was in fact repeated at all
possible fine voxel counts between 50 – 800 voxels (stepping in 50 voxel increments),
but we ultimately decided to report results from the 250 pattern size as this
corresponded to the smallest pattern size at which accuracies reached asymptotic levels.
For each ROI and classification, the final prediction accuracy was tested against
permutation baseline level (0.527), as determined by running 1,000 SVMs with shuffled
labels. In addition to an ROI-based MVPA, we conducted a searchlight analysis in which
the pattern discriminability between stimulus conditions was tested by moving a 6-mm
(radius) spherical ROI across the cortex.

We explored the relationship between MVPA classification accuracies and
behavioural thresholds. For each set of comparisons [(i) upright face vs. random shape,
(ii) inverted face vs. random shape, and (iii) upright face vs. inverted face], we extracted
its SVM accuracy and calculated a behavioural index, corresponding to the difference in
behavioural thresholds between the two conditions in the SNR task. We then computed
the (Pearson’s) correlation coefficient (r) between the two metrics for each ROI. We
also performed a further analysis concatenating the inverted face and random shape
conditions, forming a ‘non-upright-face’ condition. The corresponding SVM was run, and
accuracies subsequently correlated with a behavioural index similarly computed by
averaging thresholds from the two non-upright-face conditions.

3 Results

3.1 Behaviour

Behavioural performance for the SNR and feature tasks, indexed in terms of
signal-noise (SNR task) or disparity difference thresholds (feature task) are presented in
Figure 2. Sample staircases from one subject are shown in Figure 2c. Thresholds for each
task were analysed using a one-way repeated-measures ANOVA comparing thresholds
across stimulus conditions (Statistical Table 1). The analysis for the SNR task (Figure 2a)
indicated a significant main effect of stimulus condition, \( F(2, 58) = 5.65, p = .006, \eta^2_p = .163 \). Post-hoc Bonferroni-corrected t-tests revealed that thresholds were higher (poorer) for the upright face than the inverted face, \( t(29) = 3.00, p = .018, d = .547 \) and for the upright face than the random shape, \( t(29) = 2.89, p = .021, d = .528 \). However, there was no significant difference between the thresholds for the inverted face and that for the random shape \( t(29) = .337, p > .999, d = .067 \).

The analysis for feature task thresholds (Figure 2b) revealed a significant main effect of stimulus condition, \( F(2, 58) = 4.26, p = .019, \eta^2_p = .128 \). Post-hoc Bonferroni-corrected t-tests indicated that thresholds were significantly lower (better) for the upright face than the random shape \( t(29) = 2.79, p = .027, d = .510 \), but there was no significant difference between the thresholds for the inverted face and the upright face, \( t(29) = .431, p > .999, d = .079 \), and that between the inverted face and the random shape, \( t(29) = 2.04, p = .153, d = .372 \).

3.2 fMRI

3.2.1 In-bore Behaviour (SNR task)

We verified firstly that behavioural effects for the task chosen for the MRI could be replicated by the new group of participants whilst inside the bore. A one-way repeated-measures ANOVA indicated a significant effect of stimulus condition, \( F(2,36) = 3.45, p = .043, \eta^2_p = .161 \). Post-hoc t-tests indicated that, consistent with findings from Experiment 1, thresholds for the upright face were significantly higher than those for the random shape: \( t(18) = 2.33, p = .032, d = .448 \), but there were no significant differences between thresholds for the upright face and the inverted face \( t(18) = 1.74, p = .098, d = .244 \), and between the inverted face and the random shape \( t(18) = 1.05, p = .309, d = .169 \).
Turning to the threshold-sampled stimuli of the functional runs, we then assessed whether the difference in disparity distribution between stimuli at the ‘near’ and ‘far’ stimulus positions (relative to the fixation plane) might have an influence on response accuracies by means of a 2(depth-position: near/far) × 3(condition) ANOVA. The analyses indicated no significant main effect of depth-position, $F(1,21) = .053$, $p = .820$, $\eta^2_p = .003$, nor interactions involving depth-position [mean accuracy (SD) for Near\_uprface=.81(.13); Far\_uprface=.79(.18); Near\_invface=.80(.17); Far\_invface=.81(.15); Near\_ran=.78(.13); Far\_ran=.82(.13)].

3.2.2 General Linear Model

We firstly examined univariate fMRI signals across stimulus conditions. For each hemisphere, we extracted the GLM beta weights (percent signal change) corresponding to the three conditions from the ten ROIs (Figure 3). These were then entered into a 2(hemisphere)×3(type)×10(ROI) repeated-measures ANOVA (Statistical Table 1) that revealed significant main effects of hemisphere [$F(1,21) = 5.28$, $p = .032$, $\eta^2_p = .201$] and ROI [$F(9,189) = 7.60$, $p < .001$, $\eta^2_p = .266$]. Beta weights did not differ across conditions [$F(2,42) = 1.87$, $p = .167$, $\eta^2_p = .082$], and there were no significant interactions: [hemisphere × condition: $F(2,42) = .609$, $p = .549$, $\eta^2_p = .028$; hemisphere × ROI: $F(9,189)$ = .656, $p = .748$, $\eta^2_p = .030$; ROI × condition: $F(18,378)$ = .610, $p = .892$, $\eta^2_p = .028$; hemisphere × condition × ROI: $F(18,378) = .1.37$, $p = .141$, $\eta^2_p = .036$]. A comparison of the mean beta weights indicated that signals were generally stronger in the left hemisphere ($M = .145$, $SE = .083$) than the right hemisphere ($M = .073$, $SE = .079$).

Note that while the GLM results indicated stronger responses in the left hemisphere as compared to the right hemisphere, amplitudes differences were homogenous across ROIs and stimulus conditions. For this reason, and, as subsequent multivariate analyses remove the univariate amplitude, we elected to concatenate data from the two hemispheres for subsequent analyses.

[Figure 3 about here]
3.2.3 Multivoxel Pattern Analysis

Multivariate responses to different stimulus conditions were examined by contrasting the multivoxel fMRI signals between the upright face and the random shape, and those for the inverted face and the random shape (see Methods). MVPA classification accuracies are presented in Figure 4a. Classification accuracies were compared to baseline (.53) while holding familywise error rate at .05 (qFDR < .05) (Statistical Table 1).

SVM accuracies for the classification between the upright face and the random shape were significantly above permutation-baseline in V2, V3, V7, V4, LO and FFA, whilst those between the inverted face and the random shape were above-baseline in V1, V2, V3A, V7, V4 and LO. For the comparison between the upright face and inverted face, SVM accuracies were above-baseline in V3A, V3B, V4, and FFA. In other words, brain regions that had above-baseline classification performance in the contrasts involving the random-shape (which may reflect simply the classification of differences in lower-order features) included V2, V7, V4 and LO. Regions that showed above-baseline discriminability for the contrast between the upright face and the random shape only were V3 and FFA, whilst those above-baseline for the contrast between the inverted face and the random shape only were V1 and V3A. Only V3B uniquely distinguished between responses of the upright vs inverted faces. Finally, only FFA exhibited above-baseline discrimination accuracies that were common and unique to both SVMs involving the upright face stimulus (i.e., above-baseline for both the upright face vs random, and upright face vs inverted face comparisons, but not the inverted vs random comparison).

Lastly, the searchlight analysis reflected that our choice of ROIs well covered the regions showing above-baseline pattern discriminability. Sample searchlight maps for the pattern discriminability (t-value) between the upright face and random shape conditions are presented in Extended Data Figure 4-1. Note that these searchlight maps...
were not further used to emphasize our main findings together with the other results to avoid circularity problems (Kriegeskorte et al., 2009).

3.2.4 Brain-behaviour relationships

Finally, to clarify the relevance of these brain areas for the effects observed behaviourally, we further examined the relationship between behavioural thresholds and MVPA accuracies within each ROI (Figure 4) in two sets of analyses. First, for each possible comparison of the three stimulus conditions [(i) upright face vs. random shape, (ii) inverted face vs. random shape, (iii) upright face vs. inverted face], we derived a behavioural index computed as the difference of SNR thresholds between the two conditions. We then calculated the (Pearson’s) correlation coefficient ($r$) between this behavioural index and the SVM accuracies from the classification of the corresponding stimulus conditions for each ROI across all participants. We found that for the ‘upright face vs. random shape’ comparison, SVM accuracies in V4 correlated positively to the behavioural index ($r = .513$, $p = .025$); for the ‘inverted face vs. random shape’ comparison, SVM accuracies in V3A correlated negatively to behavioural index ($r = -.465$, $p = .045$); and finally, for the ‘upright face vs. inverted face’ comparison, SVM accuracies in V3 correlated positively to the behavioural index ($r = .707$, $p < .001$). Amongst these results, only the last involving V3 withstood Bonferroni correction for multiple comparisons.

Second, we performed an additional MVPA-behaviour correlational analysis with data of the inverted face and random shape conditions combined to form a ‘non-upright-face’ condition. Indeed, previous work has suggested that upright faces represent a class of objects that is distinct from non-face objects, as well as inverted faces (Yin, 1969; Valentine, 1988; Farah et al., 1995; Kanwisher et al., 1997; Kanwisher & Yovel, 2006). Behavioural indices (upright face minus non-upright-face conditions) were then correlated with results from a separate MVPA that compared classification accuracies for discriminating between the upright face and non-upright-face (i.e.,
inverted face + random shape) conditions. We found a significant correlation between
behaviour and SVM accuracies in FFA only ($r = .481$, $p = .041$; Extended Figure 4-2),
though the correlation did not survive further statistical correction.

4 Discussion

We tested for differences in behavioural sensitivity and fMRI responses to
stereoscopic depth, in face and non-face contexts. In Experiment 1, we compared
behavioural depth discrimination performances under three conditions with varying
biological relevance: upright faces, inverted faces, and phase-randomised surfaces. We
showed that the behavioural sensitivity to depth differed across conditions, but
curiously, also across tasks. In the SNR task, thresholds were higher for the upright face
than the inverted face and the random shape; that is, SNR depth sensitivity was poorer
in a biologically relevant context. In contrast, in the feature task, depth thresholds were
higher for the random shape than the upright face; that is, fine thresholds were instead,
better for biologically relevant stimuli. In Experiment 2, we selected the SNR task - as it
elicited more robust differences in responses across stimulus contexts, both in
Experiment 1 of the present study and in previous work (Wong et al., 2019)—and
examined fMRI responses to the stimuli concurrently with behaviour. Analysis of the
univariate responses showed rather homogenous activation across stimulus conditions.
However, results from the multivoxel pattern analysis (MVPA) were more revealing.

Early and intermediate retinotopic areas (V1, V2, V3, V3A, V3B, V4) and extrastriate
dorsal (V7) and ventral (LO, FFA) regions showed discriminable response patterns across
contexts. Notably, of these regions, responses FFA in particular appear to be classifying
information beyond simple lower-order features (i.e., are uniquely and commonly
discriminative for contrasts that involved the upright face). Moreover, multivariate
responses of V3, V4 and FFA are predictive (albeit weakly) of behavioural responses. We
consider the behavioural and imaging findings in turn.
4.1 Modulation of behavioural sensitivity to depth by object relevance

We firstly consider our behavioural findings. Our data indicate that depth sensitivity differs between face and non-face contexts and, importantly, this difference is task-dependent. Our results showed that a face context, in particular in the upright orientation, facilitates feature (disparity) discrimination but worsens the segmentation of disparity signal from noise. The reversed trends observed between the tasks, at first glance, may seem paradoxical. However, as it has been proposed that performance in tasks similar to those employed here can be thought to involve two different mechanisms, namely noise filtering and feature readout (Dosher & Lu, 2005), we propose that the differences between the SNR task and feature task can be explained within this framework. The level of external noise is high in the SNR task but absent in the feature task, implying that noise filtering is the dominating visual process during the SNR task but not the feature task. The two tasks also require different levels of precision: the SNR task requires a lower level of precision (near/far depth discrimination) whereas the feature task requires a higher level of precision (discrimination of fine depth difference). This implies a heavy engagement in feature readout for the feature task, but only a modest amount of feature extraction could suffice depth judgments in the SNR task.

Why then, is noise filtering better in non-face objects, and feature readout better in face contexts? Firstly, that the detection of a target from noise is better for novel, unfamiliar objects has been well documented in the literature. For example, when participants are presented with an array of four objects and then asked to recall the location of an item in the array, accuracy is significantly higher when the test item is novel than when it is familiar (Johnston et al., 1990). This ‘novel popout’ effect, referring to the fact that novel objects are easier to detect when noise is present, may relate to the higher saliency of novel (vs. familiar) objects that captures attention and hence facilitates target detection. In our SNR task, the ‘novel popout’ effect can be manifest as an improvement in the quality of signal extraction for the random shape (compared to
the upright face), thus reducing thresholds. Furthermore, it has been shown that the 
detection of faces appears to be especially susceptible to noise compared to other non-
face daily objects (Chen et al., 2015). In particular, the presence of noise impairs feature 
discrimination to a larger extent for faces than for cars. In the context of this previous 
work then, our data for the SNR task appear to be in alignment.

Secondly, we found that feature readout is better for the upright face than for 
the random shape. In everyday functioning, the readout of fine features is crucial to face 
individuation and the recognition of facial expressions. Indeed, that a higher level of 
precision is required for the perception of faces than other objects has been shown 
empirically. Participants inclined to name faces at the level of its identity (i.e. the name 
of the face), but dogs at the level of its species or subspecies (e.g. ‘golden retriever’ or 
simply ‘dog’), indicating that the processing of faces, compared to other objects, is more 
specific (Tanaka, 2001). Other work has shown that human participants are strikingly 
sensitive to fine differences in faces (Lehky, 2000). It is thus very likely that fine feature 
readout is more relevant to everyday face perception tasks. By contrast, precise feature 
discrimination is rarely required for non-face objects. Here, it is important to note that 
any behavioural enhancements of feature readout in face contexts shall not necessarily 
manifest in the results of the SNR task as only a limited amount of feature readout was 
needed in this task.

Finally, we observed a clear difference of thresholds between the upright and 
inverted faces in the SNR task. This is in line with the well-documented face inversion 
effect, in which the inversion of face affects perception, both in terms of recognition 
performance (Yin, 1969; Valentine, 1988; Farah et al., 1995) and neural responses 
(Kanwisher et al., 1998) of face-selective regions. Yet, behavioural thresholds did not 
differ substantially between the upright face and the inverted face in the feature task. 
Interestingly, these findings seem to suggest that the face inversion effect is more 
prominent in high noise, compared to the discrimination of fine feature differences in 
low visual noise.
Importantly, the opposite behavioural trends observed in our SNR and feature tasks distinguishes these findings from those observed by Wong et al. (2019), who showed that the physical plausibility of objects affects depth sensitivity similarly for SNR and feature discrimination tasks. Particularly, depth sensitivity was poorer for physically plausible than implausible objects. At first glance, our data are at odds with the findings of Wong et al. (2019). However, we view that the findings generated by both studies are in fact complementary to each other. The data from Wong et al. (2019) suggest a reweighing of stereo processes based on the physical plausibility of the object. They suggested that during the perception of familiar (physically plausible) objects, stereo signals are down-weighed while object-level processing is up-weighed, thus optimizing the interpretation and interaction with these familiar, naturally relevant objects. While Wong et al. (2019) showed a general modulation of disparity processes by the object’s context (physical plausibility), we argue that our data, on the other hand, suggest a more specific contextual effect on stereo processing, that depends on the biological relevance of the object. Here, all three kinds of stimulus used in our study, namely the upright face, the inverted face and the random shape, are physically plausible. If the stereo re-weighing mechanism as suggested by Wong et al. (2019) only cares about whether the stimulus is physically plausible or not, one should expect no sensitivity differences to be observed across our stimuli. On the other hand, if the re-weighing mechanism is based more broadly on familiarity, then, one should expect a down-weighing of stereo sensitivity for face, especially the upright face, compared to the inverted face and the random shapes. Our data in the SNR task are in line with this prediction. However, there is something more to unravel for the feature task, for which we observed a somewhat reversed behavioural trend. Here, we suggest that the everyday face perception tasks, such as face identification, recognition of facial expressions, etc., require the processing of precise details in the stimulus. Therefore, during the perception of faces, it would be likely that a pre-existing circuitry that favours detailed visual processing would be triggered, which in turns facilitates the readout of features. Critically, such facilitation in feature readout should be robust and able to...
counteract or even out-weigh the stereo down-weighing effect that results from the more general process as suggested by Wong et al. (2019).

4.2 fMRI responses during depth judgments under varying object contexts

We next turn to our fMRI data. The GLM revealed rather homogenous univariate activation across stimulus conditions during depth judgments. Intriguingly, there was an overall higher amplitude of responses in the left versus right hemisphere. We speculate that this asymmetry has nothing to do with stereoscopic vision per se, as indeed, existing work has indicated an absence of hemispheric effects (Akay & Celebi, 2009) in depth perception. We fathom instead that our results might have been caused by the strong imbalance of genders in our recruited participants, with many more males than females in our fMRI component. To this end, there is evidence of globally stronger hemispheric asymmetry in males than in females (Hirstein et al., 2019).

Our MVPA results were more revealing. We reasoned that we may be able to detect functionally-relevant responses to face-object modulations of stereo sensitivity, by teasing out any regions that can uniquely and commonly discriminate response patterns for the two stimulus contrasts involving the upright face stimulus (i.e., upright face versus random shape; upright face versus inverted face), but not for the comparison involving the inverted face and random shape. This allows us to avoid, as far as possible, picking up regions that might be responding to lower-order stimulus differences. We found that of our ROIs, only FFA showed unique above-baseline SVM accuracy for those classifications that involve the upright face, suggesting a potentially key role for this region for explaining the strong modulations of stereo-sensitivity by biological relevance that was observed behaviourally.

V3B, which was uniquely discriminative of upright face vs inverted face responses here, has been previously implicated in the discrimination of correlated and anticorrelated RDS (Preston et al., 2008). Both V3 (Anzai et al., 2011; Backus et al., 2001) and the inferotemporal cortex (Janssen et al., 2000, 2003) more broadly, have also been well-demonstrated to be stereo-sensitive. Of these two regions, V3 showed here
particular relevance to behaviour (i.e., the upright vs inverted face, brain-behaviour correlation). Here, larger discriminability of the two stimuli exhibited by V3 translated to larger discrepancies in the behavioural sensitivities to the two stimuli. While modulations in early visual cortex, long believed to care only about local stimulus features appear perplexing, this findings appears well fit previous work showing modulation of responses in V1 to ambiguous stereograms by changing perceived depth via luminance changes of the surround (Rideaux & Welchman, 2019).

Responses of retinotopic V1, V2, V3A, V7 found here likely reflect decoding of lower order structural differences between the objects during depth judgments as these regions showed above-baseline SVM accuracy in both classifications involving the random shape (upright face vs. random shape and inverted face vs. random shape). Further, above-baseline classification accuracies in the ventral regions (V4, LO) found in both comparisons likely reflect the decoding of the object’s identity and/or lower-order stereo information: V4 has been shown to respond to absolute depth (Shiozaki et al., 2012) and structured objects, regardless of the exact object identity (Desimone et al., 1985; Merigan & Pham, 1998). LO is well-demonstrated to respond more strongly to intact than scrambled objects (Malach et al., 1995), and shows similar responses to familiar and novel shapes (Margalit et al., 2016). These findings are in line with our results where pattern discriminability was similar across classifications, even though the upright and inverted faces differ in the level of familiarity.

We showed that depth sensitivity is affected by the biological relevance of the objects, defined here in terms of whether the target object depicted a face or non-face, and that such modulations are task-dependent. Using fMRI, we observed modulations of multivariate fMRI responses during depth judgments, particularly in FFA, by object type. We speculate that FFA is not only involved in object analysis as discussed above, but exerts considerable influence on stereoscopic mechanisms as early as in V3 such that the utility of disparity information is up/down regulated depending on the relevance of the object (i.e., whether it is a face or non-face). The role of FFA in face perception is well documented (Kanwisher et al., 1997, 1998) but its up- and down-
stream interactions with more rudimentary processes (here, stereo-encoding, in particular carried by V3) is well worth continued exploration, perhaps through using more temporally-resolved techniques. Lastly, we fathom that our findings in the FFA likely reflect our choice of biological manipulations here (i.e., faces or non-faces). In this vein, it would be further interesting to explore object-stereo interactions under different indices that vary biological importance.

References


Figure Legends

Figure 1. (a) Sample depth maps and random-dot stereograms (RDS). (Top) The depth map of the upright face was presented in an oval region that cropped the ears and the neck. (Center) An inverted face was created by rotating the upright face by 180º in the picture plane. (Bottom) A depth-power matched random shape was generated by randomizing the phases of the Fourier transform of the upright face while preserving the amplitudes. Anaglyphic examples of the RDS in ‘near’ and ‘far’ conditions, at 100% SNR, are also shown for illustration purposes only. (b) In the SNR task (top), signal-to-noise ratio (SNR) varied between 0% and 100%. In the feature task (bottom), two stimuli were presented consecutively on each trial. Disparity difference varied within 150 arcsec.

Figure 2. (a,b) Depth discrimination thresholds in the SNR and feature tasks. In the SNR task, thresholds were higher (i.e. poorer) for the upright face than the inverted face and the random shape. In the feature task, thresholds for the upright face were significantly lower (i.e. better) than those for the random shape, whereas thresholds for the inverted face did not differ significantly from those of the other two stimulus conditions. (c) Sample staircases from a randomly-sampled observer, showing threshold estimates across trials within one experimental run, for the upright face condition. (d) Behavioural thresholds obtained in-bore (SNR task). Thresholds were higher for the upright face than the random shape. Individual subject thresholds are overlaid in a-c. Error bars represent ±1 SEM.

Figure 3. GLM beta weights (% signal change) of the three stimulus conditions in the (a) left and (b) right hemispheres. Univariate signals were globally higher in the left hemisphere than the right hemisphere. Signals were generally higher in V1, V2 and LO. Error bars represent ±1 SEM (N = 22).

Figure 4. (a) SVM accuracies for all ROIs in the three exhaustive classifications comparing all stimulus conditions. Classification accuracies (Bonferroni-corrected) that are significantly above-baseline (.527; dashed line) are indicated by asterisks. Error bars represent ±1 SEM (N = 22). (b-d) Brain-behaviour correlations in V4 for upright-face vs. random-shape (b); V3A for inverted-face vs. random-shape (c); and V3 for upright-face vs. inverted-face (d). The horizontal axis represents a behavioural index computed as the threshold difference between conditions. Additional results from the whole-brain classification (searchlight) and from an additional brain-behaviour correlation following the concatenation of the inverted and non-face conditions (FFA) are presented in Extended Figure 4-1 and 4-2, respectively.
Extended Data Legend

**Figure 4-1.** Sample searchlight maps for SVM discriminations between the upright face and random shape conditions. Results are superimposed onto the representative flattened surface maps of a single participant with ROIs delineated. Gyri are coloured in light grey and sulci in dark grey.

**Figure 4-2.** Brain-behaviour correlation in FFA for the upright-face versus non-face (i.e., concatenating the inverted face and random shape conditions) comparison. The horizontal axis represents a behavioural index computed as the threshold difference between the upright face condition and the mean of the two non-face conditions.
**Figure 1**

### Depth maps vs. Random-dot stereograms

<table>
<thead>
<tr>
<th>Depth Maps</th>
<th>Random-dot stereograms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upright face</td>
<td>![Example Image]</td>
</tr>
<tr>
<td>Inverted face</td>
<td>![Example Image]</td>
</tr>
<tr>
<td>Random shape</td>
<td>![Example Image]</td>
</tr>
</tbody>
</table>

### Experimental Setup

- **SNR task**
  - Signal + noise
  - Plane of fusion (5 arcmin)
  - Stimuli 1: 5.3 arcmin
  - Stimuli 2: 5.3 arcmin

- **Feature task**
  - Signal + noise
  - Plane of fusion (5 arcmin)
  - Stimuli 1: 7.8 arcmin
  - Stimuli 2: 10.3 arcmin

Note: Stimulus 2 is reverse.
Figure 2
Figure 3
Figure 4
Extended Data, Figure 4-1

Extended Data, Figure 4-2
<table>
<thead>
<tr>
<th>Analysis</th>
<th>Figure</th>
<th>Data structure</th>
<th>Type of test</th>
<th>Variables</th>
<th>Observed power</th>
</tr>
</thead>
</table>
| Experiment 1 – SNR task  | 2a     | Assumed normal | One-way repeated-measures ANOVA     | Independent: Stimulus condition  
Dependent: Depth threshold | .843    |
| Experiment 1 – Feature task | 2b   | Assumed normal | One-way repeated-measures ANOVA     | Independent: Stimulus condition  
Dependent: Depth threshold | .722    |
| Experiment 2 – Behaviour | 2d     | Assumed normal | One-way repeated-measures ANOVA     | Independent: Stimulus condition  
Dependent: Depth threshold | .610    |
| Experiment 2 – GLM       | 3      | Assumed normal | Three-way repeated-measures ANOVA   | Independent: Hemisphere  
Stimulus condition  
ROI | .580    |
| Experiment 2 – MVPA      | 4      | Assumed normal | One-sample t-test                  | SVM accuracy against baseline  
Upright face vs. Random shape Classification  
V1  
V2  
V3  
V3A  
V3B  
V7  
hMT+  
v4  
LO  
FFA | .690  
> .999  
.972  
.744  
.422  
.919  
.703  
.880  
.905  
.762 |
<table>
<thead>
<tr>
<th>Analysis</th>
<th>Figure</th>
<th>Data structure</th>
<th>Type of test</th>
<th>Variables</th>
<th>Observed power</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V3</td>
<td>.639</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V3A</td>
<td>.999</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V3B</td>
<td>.479</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V7</td>
<td>.998</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hMT+</td>
<td>.175</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V4</td>
<td>.999</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>L0</td>
<td>.764</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FFA</td>
<td>.396</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>Upright face vs. Inverted face Classification</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V1</td>
<td>.437</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V2</td>
<td>.418</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V3</td>
<td>.064</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V3A</td>
<td>.999</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V3B</td>
<td>.875</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V7</td>
<td>.504</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hMT+</td>
<td>.088</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V4</td>
<td>.903</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>L0</td>
<td>.540</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FFA</td>
<td>.838</td>
</tr>
<tr>
<td>Experiment 2 – MVPA-behaviour correlation</td>
<td>4</td>
<td>Assumed normal</td>
<td>Pearson correlation</td>
<td>Variable 1: Threshold difference</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V1</td>
<td>.050</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V2</td>
<td>.378</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V3</td>
<td>.288</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V3A</td>
<td>.071</td>
</tr>
<tr>
<td>Analysis</td>
<td>Figure</td>
<td>Data structure</td>
<td>Type of test</td>
<td>Variables</td>
<td>Observed power</td>
</tr>
<tr>
<td>---------------------------</td>
<td>--------</td>
<td>----------------</td>
<td>--------------</td>
<td>-----------</td>
<td>----------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V3B</td>
<td>.055</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V7</td>
<td>.074</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hMT+</td>
<td>.051</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V4</td>
<td>.690</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LO</td>
<td>.258</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FFA</td>
<td>.075</td>
</tr>
<tr>
<td>Inverted face vs. Random shape</td>
<td></td>
<td></td>
<td></td>
<td>V3</td>
<td>.312</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V2</td>
<td>.087</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V3</td>
<td>.087</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V3A</td>
<td>.579</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V3B</td>
<td>.117</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V7</td>
<td>.342</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hMT+</td>
<td>.053</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V4</td>
<td>.053</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LO</td>
<td>.075</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FFA</td>
<td>.072</td>
</tr>
<tr>
<td>Upright face vs. Inverted face</td>
<td></td>
<td></td>
<td></td>
<td>V3</td>
<td>.533</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V2</td>
<td>.449</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V3</td>
<td>.984</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V3A</td>
<td>.058</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V3B</td>
<td>.353</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V7</td>
<td>.250</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hMT+</td>
<td>.202</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V4</td>
<td>.488</td>
</tr>
<tr>
<td>Analysis</td>
<td>Figure</td>
<td>Data structure</td>
<td>Type of test</td>
<td>Variables</td>
<td>Observed power</td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
<td>----------------</td>
<td>--------------</td>
<td>-----------</td>
<td>----------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LO</td>
<td>.193</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FFA</td>
<td>.544</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Upright face vs. Non-upright-face</td>
<td>V1</td>
<td>.057</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V2</td>
<td>.129</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V3</td>
<td>.170</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V3A</td>
<td>.060</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V3B</td>
<td>.110</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V7</td>
<td>.065</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hMT+</td>
<td>.061</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V4</td>
<td>.121</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LO</td>
<td>.084</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FFA</td>
<td>.562</td>
</tr>
</tbody>
</table>