

Natural contrast statistics facilitate human face categorization

Joan Liu-Shuang ^a, Yu-Fang Yang ^a, Bruno Rossion ^{a,b,c,d}, Valérie Goffaux ^{a,b,e}

^a Institute of Research in Psychology (IPSY), University of Louvain, 1348, Belgium

^b Institute of Neuroscience (IoNS), University of Louvain, 1348, Belgium

^c Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France

^d Université de Lorraine, CHRU-Nancy, Service de Neurologie, F-54000, France

^e Department of Cognitive Neuroscience, Maastricht University, Maastricht, 6229, the Netherlands

Abbreviated title: Natural contrast statistics facilitate human face categorization

Author Contributions:

JLS, Y-FY, BR, and VG designed research; JLS and Y-FY performed research; JLS, VG and Y-FY analyzed data; JLS, Y-FY, BR and VG wrote the paper

Manuscript accepted for publication in eNeuro (2022).

Corresponding Author:

Valérie Goffaux

valerie.goffaux@uclouvain.be

Psychological Sciences Research Institute (IPSY)

UC Louvain

Place Cardinal Mercier 10 box L3.05.01

1348 Louvain-la-Neuve, Belgium

Tel.: 0032 10473877

Funding sources

This work was supported by PDR Grant PDR T.0207.16 awarded by the Belgian Fund for Scientific Research (F.R.S.–FNRS) to BR, and by F.R.S.–F.N.R.S./Fonds Wetenschappelijk Onderzoek Excellence of Science grant HUMVISCAT 30991544 awarded to VG and BR. VG is an F.R.S.–FNRS Research Associate. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors report no conflict of interest.

Acknowledgement

We thank Dr Talia Retter for her help in analysing the data.

Keywords: face categorization, natural statistics, contrast polarity, Frequency tagging, EEG

Abstract

The ability to detect faces in the environment is of utmost ecological importance for human social adaptation. While face categorization is efficient, fast and robust to sensory degradation, it is massively impaired when the facial stimulus does not match the natural contrast statistics of this visual category, i.e., the typically experienced ordered alternation of relatively darker and lighter regions of the face. To clarify this phenomenon, we characterized the contribution of natural contrast statistics to face categorization. Specifically, 31 human adults viewed various natural images of non-face categories at a rate of 12 Hz, with highly variable images of faces occurring every eight stimuli (1.5 Hz). As in previous studies, neural responses at 1.5 Hz as measured with high-density EEG provided an objective neural index of face categorization. Here, when face images were shown in their naturally experienced contrast statistics, the 1.5 Hz face categorization response emerged over occipito-temporal electrodes at very low contrast (5.1%, or .009 root-mean-square, RMS, contrast), quickly reaching optimal amplitude at 22.6% of contrast (i.e., RMS contrast of .041). Despite contrast negation preserving an image's spectral and geometrical properties, negative contrast images required twice as much contrast to trigger a face categorization response, and three times as much to reach optimum. These observations characterize how the internally stored natural contrast statistics of the face category facilitate visual processing for the sake of fast and efficient face categorization.

Significance statement

Human faces share a universal property: the strict alternation of contrast, with the darker main features against the more uniform, lighter skin. The ability to categorise faces depends critically on the presence of these natural contrast statistics in the input stimulus. However, it is not yet known how natural contrast statistics facilitate the visual processing leading to categorisation. Using frequency tagging and high-density electroencephalography, we show that access to internally stored natural statistics reduces the amount of sensory input necessary for human face categorization.

Introduction

Among the many things that humans encounter in their visual environment, there is one category of utmost ecological relevance, and for which the human brain has developed an exceptionally high sensitivity: the faces of conspecifics. The newborn human preferentially orients their gaze towards faces and face-like patterns (Goren et al., 1975; Turati et al., 2002; Buiatti et al., 2019) and infants of a few months of age already show specialized (right-lateralized) neural responses to faces presented at a glance (de Heering and Rossion, 2015; Rekow et al., 2021 see also Tzourio-Mazoyer et al., 2002). At adulthood, humans automatically detect the presence of a face in the environment, and differentiate it from the other objects composing the scene (i.e., categorize it as a face) at presentation duration as short as 17ms (Keysers et al., 2001; Retter et al., 2020; see also Grill-Spector and Kanwisher, 2005; Fisch et al., 2009; Perrett et al., 2009; Mohsenzadeh et al., 2018). It only takes a few tens of milliseconds to initiate a saccade towards a face even if it is embedded in a complex background (Thorpe et al., 1996; Lewis and Edmonds, 2003; Rousselet et al., 2003; Hershler and Hochstein, 2005; Fletcher-Watson et al., 2008; Honey et al., 2008; Crouzet et al., 2010; Crouzet and Thorpe, 2011; Martin et al., 2018; Kauffmann et al., 2019; for a review see Fabre-Thorpe, 2011). The mechanisms underlying the automaticity and exceptional speed at which human adults categorize faces work efficiently even when access to sensory input is hampered due to brief exposure, masking, blur, noise, or clutter (e.g., Goffaux et al., 2011; Ales et al., 2012; Quek et al., 2018b). Yet, face categorization is dramatically impaired when the stimulus does not share the visual properties that have been experienced by the observer as being characteristic of the face category (Rosch and Mervis, 1975; Rosch et al., 1976), i.e., the regularities in how faces appear in everyday life (so-called natural statistics; e.g., Geisler, 2008; Clifford et al., 2015). One salient and universal regularity of the human face is the ordered alternation of contrast, with darker regions occupied by the main features (brows, eyes, nose and mouth) against the more uniform and relatively lighter skin surface (Figure 1; Watt, 1994; Dakin and Watt, 2009; Gilad et al., 2009). This relative distribution of contrast is due to the structural and reflectance properties of the face and the fact that lighting usually comes from above (Ramachandran, 1988; Liu et al., 1999). Such characteristic natural light/dark alternation is reversed with image negation. Although negation is a fully reversible manipulation of the image, which leaves the image geometric and spectral properties unchanged, it has profound consequences for the categorization of faces.

For example, contrast negation disrupts human infants' viewing preference for faces (Farroni et al., 2005; Otsuka et al., 2012). Later in adulthood, contrast negation slows down and impairs behavioral performance in face detection (e.g. Lewis and Edmonds, 2003; Liu-Shuang et al., 2015a) as well as face identity recognition tasks (i.e.; Galper, 1970; Galper and Hochberg, 1971; Bruce and Langton, 1994; Kemp et al., 1996; Liu et al., 1999; Lewis and Edmonds, 2003; Russell et al., 2006; Nederhouser et al., 2007; Garrido et al., 2008; Liu-Shuang et al., 2015a). This manipulation also reduces the human brain response to faces (George et al., 1999; White, 2001; Vuong et al., 2005; Nederhouser et al., 2007; Gilad et al., 2009; Nasr and Tootell, 2012; Rossion et al., 2012; Yue et al., 2013; Liu-Shuang et al., 2015a; see also Ohayon et al., 2012 for evidence in the primate brain). Altogether, these findings suggest that human efficiency at categorizing faces critically depends on whether their appearance obeys the naturally experienced statistics of the face category, i.e. whether it depicts the contrast pattern typically rendered by a top-lit face. Comparatively, contrast negation has a moderate impact on the categorization of non-face stimuli (e.g., Galper, 1970; Nederhouser et al., 2007; Yue et al., 2013), which has been taken to suggest that the unique susceptibility of human face categorization to contrast negation is mainly due to (1) face appearance obeying unique

universal ordinal contrast rules (Figure 1), and (2) their extensive learning by humans over lifespan.

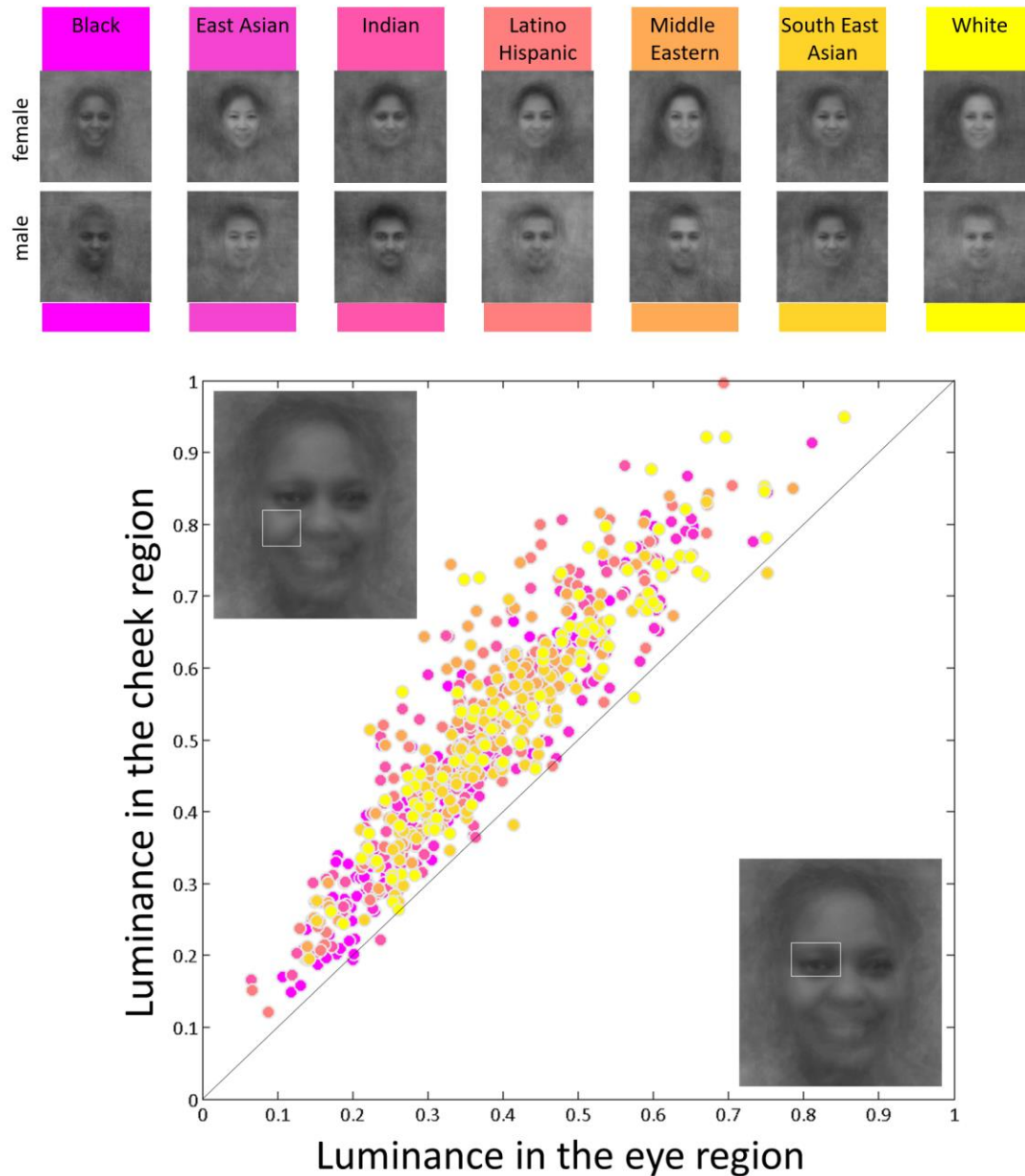


Figure 1. The human face obeys systematic and universal contrast statistics with the regions occupied by the main features being darker than the neighboring skin surface. To illustrate this point, we measured the luminance in the left eye and left cheek regions in a large set of ‘in-the-wild’ face images stemming from 7 distinct ethnic origins ($n=100$ images per ethnicity, half-female; age range: 30 to 39 year-old; Karkkainen and Joo, 2021). We used this image set for the luminance measurements exclusively (i.e., a different, validated, set of faces was used as experimental stimuli). We only used face images where these regions were free of hair, occlusion (by e.g., glasses, or a limb), and excessive make-up. Faces were of highly variable pose, expression, illumination, etc. The broad distribution of luminance values illustrate the large heterogeneity of skin tones measured. Top. Illustrative averages of the face images that were used for the analysis in each ethnic origin. Bottom. Scatterplot of the mean luminance of the eye and the cheek region for each measured faces ($n=700$ in total; see Gilad et al., 2009 for a similar analysis on a restricted set of 50 faces). White rectangles superimposed on the inset faces illustrate the eye and cheek regions as sampled in the present analysis. Except for a minority of outlier face images, one can see that all data points lie above the diagonal, which means that luminance in the eye region is systematically darker than around the cheek. This rule applies irrespective of face surface properties (absorbance, illumination, and specular reflection).

In studies addressing the influence of contrast polarity statistics on face categorization, the stimuli are typically a few homogeneous exemplars isolated from their natural context and selected from a restricted set of categories (e.g., Yue et al., 2013; Liu-Shuang et al., 2015a). However, in real life, categorization success depends on the dual ability to differentiate faces from their background (i.e., segmentation) and from diverse distracters that are present in the scene as well as to generalize across widely variable – and sometimes very dissimilar – category exemplars (see review by Rossion et al., 2018). Such naturalistic, i.e., variable and crowded, viewing conditions, place maximum demands on the sensitivity and tolerance of face categorization, and it is in such viewing conditions that the knowledge of natural statistics is expected to contribute the most to perception (e.g., Clifford et al., 2015; Karimi-Rouzbahani et al., 2021). In order to adequately define the influence of natural statistics on the genuine categorization of faces, it is therefore crucial to simulate the richness of real-world visual diet and, therefore, to use naturalistic, heterogeneous, and unsegmented face stimuli contrasted to a wide array of biological and man-made objects.

As noted above, while a well-known signature of human face recognition is its right hemispheric dominance, the causes/factors of this right-hemispheric lateralization remain a mystery (Behrmann and Plaut, 2020; Rossion and Lochy, 2022). When sensory input is reduced, e.g., by poor visual acuity during early development or with stimulus degradation, the right-hemispheric dominance is preserved, even enhanced (de Schonen and Mathivet, 1990; Quek et al., 2018b). Since contrast negation breaks the natural contrast statistics of the human face that are at the heart of the neural specialization of face recognition, their preservation may be necessary to observe a right-hemispheric dominance. If this reasoning is correct, the right-hemispheric dominance of human face categorization should disappear, or at least be significantly reduced, with contrast negation (cf. Liu-Shuang et al., 2015a).

Another fundamental issue is exactly how natural statistics support the successful categorization of faces. Past evidence has shown that contrast negation weakens neural and behavioral responses to faces but exactly how it impedes visual processing leading up to (genuine) face categorization is unclear.

In order to clarify these outstanding issues, the present study elicited genuine visual categorization processes by using a large and variable set of face and non-face (manmade objects, buildings, plants, animals) exemplars embedded in their original context. The depicted faces varied substantially in age, gender, expression, pose, size and viewpoint (Figure 2A-B). This wide variety of images places the visual system in a mode of operation closer to that activated in the natural conditions of face categorization, and necessarily engages its sensitivity and tolerance. We present these images at a fast periodic rate of 12 Hz (stimulus onset asynchrony, SOA, of 83.33 ms), which is optimal for categorization in these conditions (Retter et al., 2020), and record human brain responses using high-density scalp EEG. Since the human brain synchronizes its activity to the periodicity of the visual stimulus (Adrian and Matthews, 1934; Regan, 1966; see Norcia et al., 2015 for review), such a stimulus sequence elicits a common neural response to face and non-face stimuli at image frequency rate (12 Hz) and harmonics. Face stimuli appeared every 8th image (i.e., periodicity of 1.5 Hz; Figure 3A). If the periodically presented faces are successfully categorized, an additional category-selective neural response emerges at 1.5 Hz, i.e., at the same frequency as face occurrence.

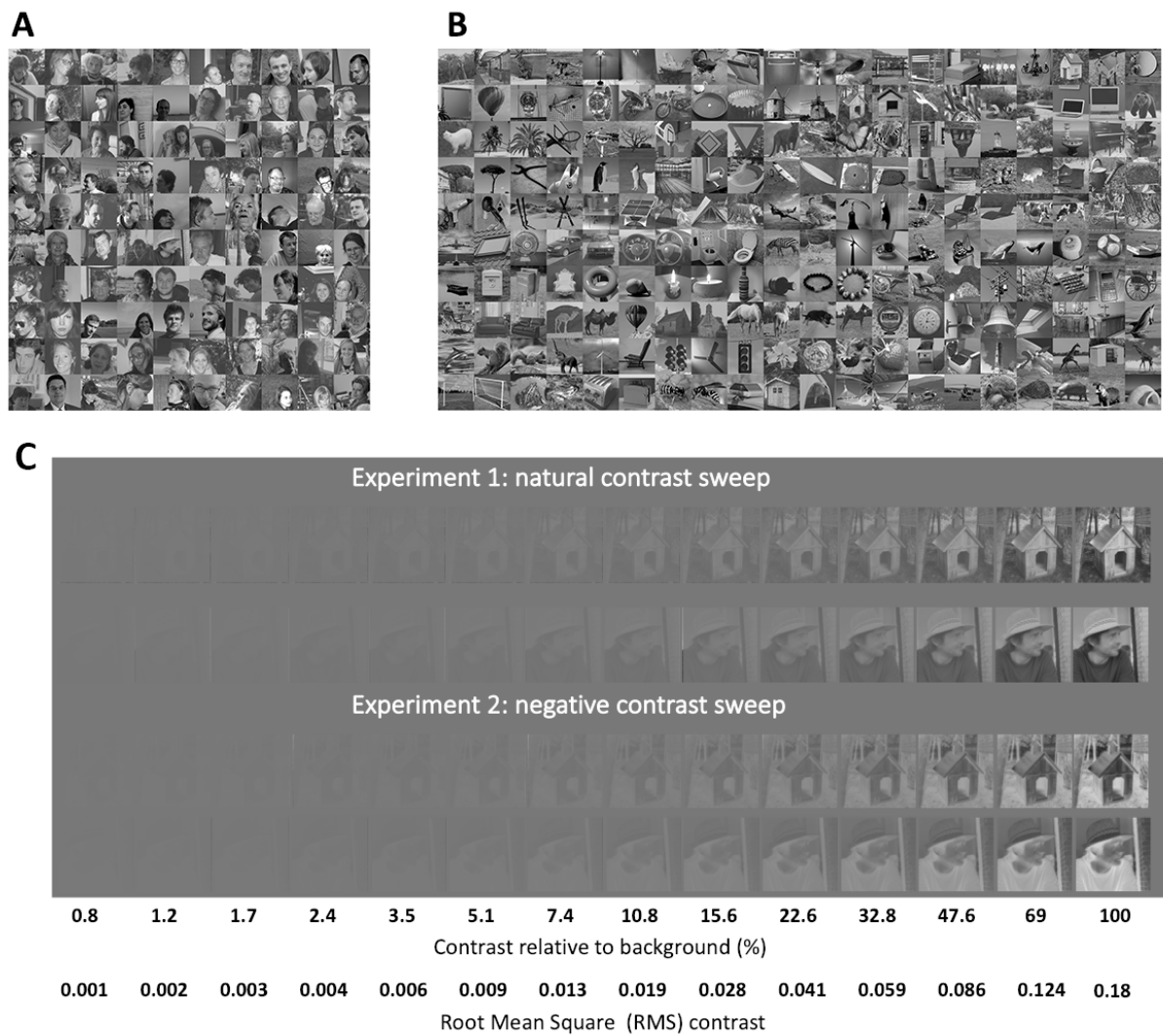


Figure 2. Experimental stimuli. **A**. Face images (N=100). **B**. Non-face images (N=200). **C**. Stimulus contrast was modulated via alpha blending (i.e. luminance values were calculated as a weighted average of the stimulus with the grey background). Example object and face image at the different contrast levels, in natural and negative contrast groups. Note that luminance and contrast were here adjusted for figure visibility and are not representative of the actual (gamma-corrected) values.

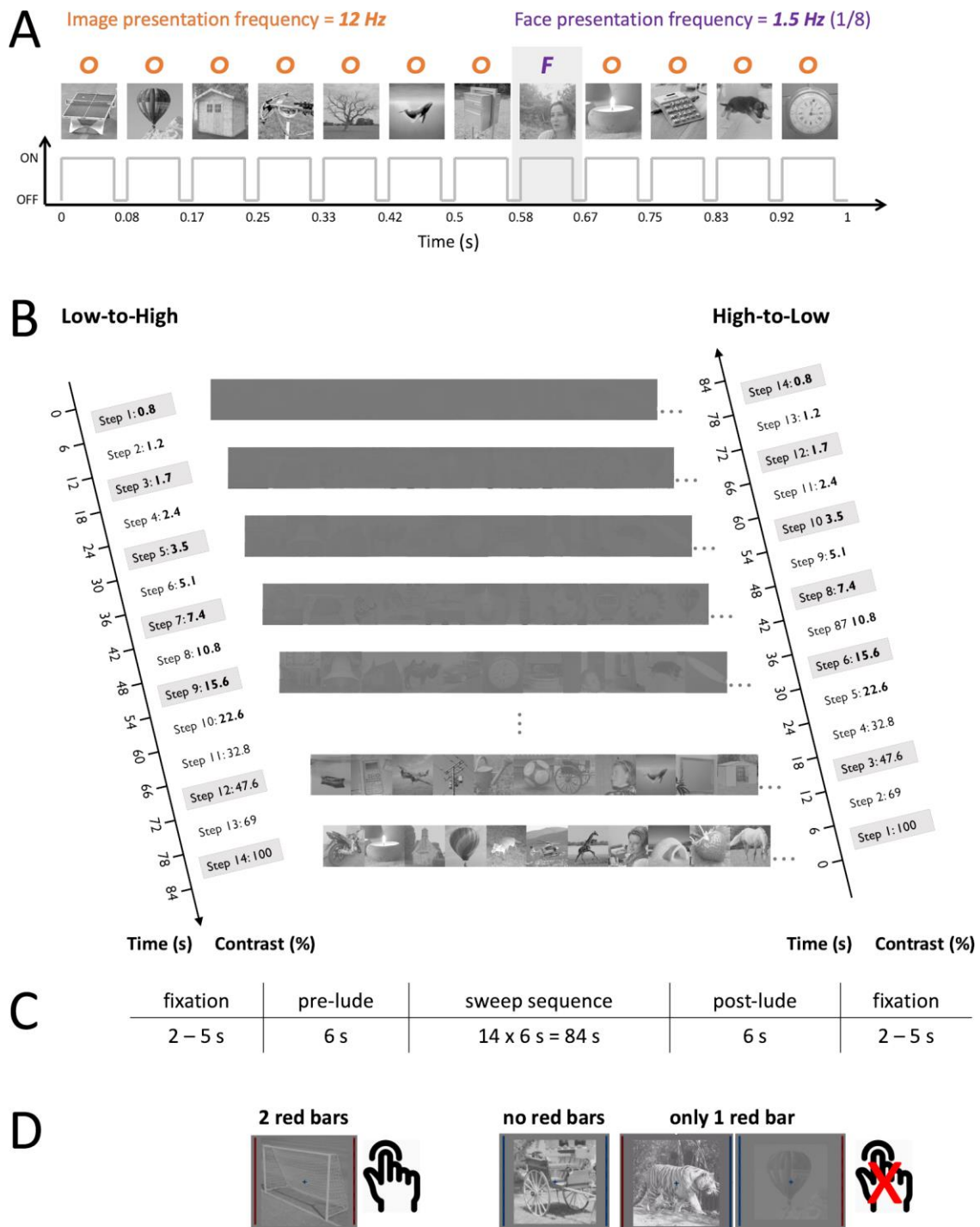


Figure 3. A. EEG Frequency-tagging paradigm. Images were presented in random order at a rate of 12 Hz (12 images/s, SOA of 83.33 ms), with faces embedded among objects at fixed intervals of 1/8 (face presentation frequency = $12/8 = 1.5$ Hz). B. The image contrast of the frequency-tagged sequences was parametrically varied over 14 x 6 s steps (84 s total), either gradually increasing (Low-to-High, *left*) or decreasing (High-to-Low, *right*). Sweeping in the two direction enabled to account for potential order effects on responses. C. Schematic timeline of a sweep trial. Following the appearance of a central fixation cross (lasting between 2 to 5 s), the sequence began with a pre-lude period (repetition of the first sweep step) before the sequence proper. At the end of the sequence, the last sweep step was repeated during the post-lude period, and the fixation remained on screen for another 2 to 5 s. D. Observers performed an orthogonal behavioural vigilance task during which they maintained central fixation while monitoring two flanking vertical bars. These bars independently changed colour from red to blue and observers had to respond only when both bars turned red simultaneously. Figure 3-1 illustrates the group-averaged accuracy and reaction times in each group and condition. Note that luminance and contrast were here adjusted for figure visibility and are not representative of the actual (gamma-corrected) values.

This neural response at 1.5 Hz is an implicit marker of genuine categorization, i.e., one that meets the dual requirements of face categorization: sensitivity to the face relative to the other categories present in each scene and sequence, as well as tolerance to the high variability of face appearance in the sequence (see Rossion et al., 2015; Retter and Rossion, 2016; Peykarjou et al., 2017; Quek and Rossion, 2017; Gao et al., 2018; Retter et al., 2018; Rossion et al., 2018; Gao et al., 2019; Or et al., 2019; Hauk et al., 2021). Its amplitude has recently been shown to scale systematically with the proportion of face images that are correctly categorized in a sequence (Retter et al., 2020). This implicit frequency-tagged marker of face categorization usually emerges over occipito-temporal scalp electrodes, mainly right-lateralized, and presumably reflect face-selective responses in the ventral occipito-temporal cortex (see Jonas et al., 2016 for evidence from human intracerebral recordings;; Gao et al., 2018 for functional magnetic resonance imaging evidence, and Hauk et al., 2021 with magnetoencephalography). A categorization response also emerges when non-face categories are frequency-tagged in fast stimulation sequences (e.g., houses and body parts Jonas et al., 2016; Hagen et al., 2020). However, it is of much weaker amplitude for non-face than for face categorization and associated with distinct patterns of lateralization and scalp topography. This paradigm has multiple advantages over other approaches. Not only does it lead to a more authentic categorisation, but it also measures the brain's response to categorisation implicitly and objectively, at time frequencies predetermined ('labelled') by the periodicity of the sequence (Retter and Rossion, 2016).

Here we used this well-validated approach to characterize how natural face contrast statistics affects the face categorization process. We compared the neural face categorization response as a function of sensory input, for image sequences displayed in a natural contrast polarity, to sequences with reversed contrast polarity (negative contrast). In each sequence, stimulus contrast ramped (i.e., was “swept”) up or down (Figure 3B). Contrast being the fundamental signal driving the visual system (e.g., Kaplan et al., 1987; Boynton et al., 1996), its sweeping enables manipulating the amount of sensory input available to categorize faces. With this approach, we defined for each individual participant the *minimal* amount of sensory input (i.e., contrast) required to elicit a genuine categorization response as well as the *optimal* level of sensory input, namely the level of contrast needed to see a full blown categorization response, i.e., reaching the amplitude as elicited by full contrast images.

We reasoned that the facilitation of categorization by natural statistics may take multiple forms. Humans may simply need less sensory input to start categorizing a face as a face. In this case, contrast negation should only increase the contrast minimum and leave the optimum level of contrast unchanged. It is however plausible that once the minimal amount of information has been reached and leads to a significant categorization response, the visual system may need less input to accrue the face categorization response until its optimum (e.g., Liu-Shuang et al., 2015a; Quek et al., 2018b; Retter et al., 2020). By deriving the distance separating minimum/optimum, we can define the dynamics of visual processing subtending face categorization. An increase in the contrast minimum for contrast-negated sequences will reduce such distance while an increase of optimum will enlarge the minimum/optimum distance. This approach allows us to characterize how the dynamics of visual processing subtending face categorization are modulated by natural face contrast statistics.

Methods

Participants

Thirty-one healthy adult participants were recruited online through the university's participation pool in exchange for monetary compensation (10€ per hour). Sixteen participants viewed stimulus sequences in their natural contrast polarity ('natural contrast' group). We excluded one individual whose EEG recording contained excessive artifact noise (deflections exceeding $\pm 100\mu\text{V}$ on multiple epochs). The final 'natural contrast' sample consisted of 15 participants (9 females, age = 23 ± 3 years). Another, so-called 'negative contrast' group of 15 participants (6 females, age = 23 ± 1.7 years) was tested with the contrast-negated stimulus sequences. We opted for a between-subject manipulation of contrast polarity to avoid any potential learning/order effects. All observers were right-handed as confirmed by an electronic version of the Edinburgh Handedness Inventory measurement Oldfield, 1971 and reported normal or corrected-to-normal vision. Participants gave written informed consent prior testing conformingly to the guidelines of the local Biomedical Ethical committee (B403201111965).

Stimuli

Stimuli were 256 x 256 pixel greyscale images of 100 white faces and a variety of 200 non-face items (e.g. animals, plants, man-made objects, buildings, etc.). All faces and objects were embedded in their natural backgrounds (i.e., unsegmented), and varied greatly in orientation, lighting, size, and overall appearance (Figure 2A-B). The same stimulus set was used in previous studies Quek and Rossion, 2017; Quek et al., 2018b. For the present study specifically, we equalised mean luminance and contrast of all the stimuli by first normalising the pixel luminance to a mean of zero and a standard deviation (i.e., RMS contrast) of one. Face and object images were contrast-negated by flipping the pixel intensities around the mean luminance value (Figure 2C). Contrast negation was used here to disrupt the natural statistics of the images thought to crucially support their visual processing while conserving their amplitude spectrum and contours (see below). All images were attributed a global luminance of .46 and a RMS contrast of .18. These values reflect the grand-averaged luminance and contrast values of the original images, adjusted to avoid luminance clipping. The resulting face and object images showed similar average pixel luminance distributions.

Finally, images were gamma-corrected according to the luminance profile of the BenQ XL2420T monitor (refresh rate of 120 Hz, screen resolution of 1920×1080 pixels). The mean luminance after gamma correction was 58.79cd/m^2 . Stimulation was carried out on Java SE version 8 in Windows 7. Stimuli were displayed at a viewing distance of 100cm and subtended a visual angle of 4.04 by 4.04° in a dimly lit room. During stimulation, image contrast was parametrically manipulated on a logarithmic scale via online alpha blending, which created a weighted sum of the luminance values of the image and a grey background while maintaining mean luminance. Fourteen contrast levels were presented, from .8% to 100% of the initial RMS contrast, with percentages referring to the weight of images relative to background. The respective RMS contrast values of the presented stimuli are listed on Figure 2. Logarithmic scaling was chosen in order to maximise sampling of the low contrast range where the human visual system is expected to be most sensitive.

Procedure

Prior to each EEG experiment, we measured the visual abilities of each participant using the Freiburg Visual Acuity and Contrast Test (FrACT; Landolt C optotype stimulus with 4AFC) at a viewing distance of 100 cm. Mean minimums in terms of Weber contrast were 1.02% for the

‘natural contrast’ group (range = .6-1.67%) and 1.05% in the ‘negative contrast’ group (range = .59-2.18%), and did not differ significantly between groups (independent samples t-test: $t(28) = -0.21$, $p = .84$).

In an experimental sequence, images were shown at a fast rate of 12 Hz (12 images/s, SOA = 83.33ms) through a squarewave luminance modulation with a 80% duty cycle (i.e. images were ON for 66.67ms and OFF for 16.666ms; Figure 3). Faces (F) were inserted among objects (O) at regular fixed intervals of 1/8 (or 12/8 = 1.5 Hz), resulting in the following stimulation pattern: OOOOOOFOOOOOOFOO... (Figure 3A as in e.g. Quek et al., 2018b). We opted for a 12 Hz periodicity due to consistent evidence that this rate elicits the largest EEG categorization responses as well near-ceiling behavioral performance (Retter et al., 2020). In order to minimise potential low-level edge and image adaptation effects, at each 12 Hz presentation cycle, stimulus size randomly varied between 90% and 110% (in 5% steps; i.e., 5 possible sizes). In agreement with previous studies (e.g., Quek et al., 2018b), we expected this periodic stimulation to elicit two critical neural responses: 1) a *general visual response* at the 12 Hz image presentation frequency (and harmonics), reflecting general visual processing common to faces and objects; and 2) a *face categorisation response* at the 1.5 Hz face presentation frequency (and harmonics). Considering that our study addresses the impact of natural statistics upon categorization efficiency, the subsequent analyses will mainly focus on the face categorization response.

The main experiment employed so-called *contrast sweep* sequences (Figure 3B), in which we varied image contrast over 14 sequential steps (6s per step), either in ascending (low to high) or descending order (high to low). Contrast was swept in both directions in order to compensate for potential order effects. Hence, images either gradually appeared or gradually disappeared against the grey background over the course of each 84s sweep sequence. At the face presentation rate of 1.5 Hz, each step contained 9 faces. A sequence was flanked by 6 extra seconds consisting respectively of repetitions of the first contrast step (e.g. .8% in the Low-to-High condition) and the last contrast step (e.g. 100% in the Low-to-High condition). As in previous studies (e.g., Quek et al., 2018b), these “pre-lude” and “post-lude” periods, aiming to minimise muscle artefacts and ocular movement elicited by the appearance and disappearance of flickering images, were not further analysed. We presented a total of 24 sequences (half in Low-to-High order and the other half in High-to-Low order), divided into four blocks of 6 sequences shown in pseudo-randomised order. With nine face images presented per sequence step, we reached a total of 3024 face image presentations in the main experiment (and a total of 21168 non-face images).

Tracking the amplitude of the 12 Hz general visual response and of the 1.5 Hz face categorisation response across contrast steps enabled us to define the minimum and optimum levels of contrast for each neural response (see below).

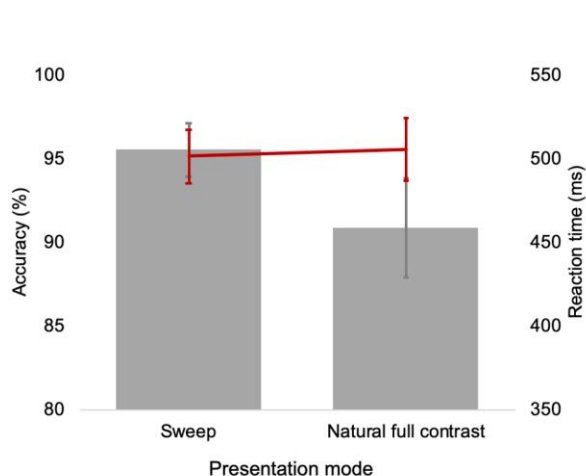
The main experiment was followed by one block of four 96-s (6 + 84 + 6 s as for contrast sweep sequences) long full-contrast sequences of natural contrast images, in which image contrast remained at 100% without modulation, except for the pre- and post-ludes. These latter sequences provided reference amplitudes for EEG responses recorded in the contrast sweep sequences. In the ‘negative contrast’ group of participants, we additionally presented four full-contrast sequences with negative contrast. Importantly, the response profile of each subject was evaluated against its own full positive contrast reference response, i.e. minima and optima were defined within subject.

Behavioral task & performance

A blue central fixation cross (.05 x .2° visual angle) was overlaid on top of the stimulation sequence, appearing 2 – 5 s before the onset of the sequence and remaining 2 – 5 s after its offset. Two vertical lateral bars (visual angle 0.08 x 4.3°) flanked the stimulus continuously at an eccentricity of 2.02°. In the beginning the trial, the both vertical bars were blue. At 12 (pseudo)random latencies within the 84 second sweep sequence, either one of the vertical bars or both of the vertical bars turned red for a duration of 300 ms. At all other times during the sequence, both bars were shown as blue. Each change lasted for 300 ms and there was a minimum interval of 800 ms between consecutive changes. The interval between the offset of the red bar(s) and the onset of the next red bar(s), that is, the time during which both bars were blue, was always at least 800 ms. Participants were instructed to maintain central fixation and to respond whenever both bars simultaneously changed colour by pressing on the spacebar of a keyboard. On average, there were approximately 6 targets per sequence (mean $\pm$ standard deviation: natural contrast group = 5.92 ± 1.74 ; negative contrast group = 5.98 ± 1.69). Responses were considered correct when falling within a time-window between 250 – 800 ms following colour change onset.

Previous studies have shown that the neural response to frequency-tagged face categorisation is robust whether participants perform this task, another vigilance task or no task at all (Quek et al., 2018a; Quek et al., 2018b). In the present study, a vigilance task was required to ensure that participants' attention to the stimulus was comparable across contrast levels and groups. Overall, behavioural performance in the full contrast and contrast sweep conditions was highly accurate across groups (Figure 3-1). Independent samples t-tests did not reveal any *Group* differences in accuracy, neither in the Full contrast ($t(28) = -0.052$, $p = .96$, 95% confidence interval of the group difference: $[-.078 .074]$), nor in the contrast sweep sequences ($t(28) = .4$, $p = .69$, 95% confidence interval of the group difference: $[-.036 .054]$). Mean correct response times were also similar across *Groups* (Full contrast sequence: $t(28) = -0.86$, $p = .4$, 95% confidence interval of the group difference: $[-.069 .028]$; Contrast sweep sequence: $t(28) = -.54$, $p = .6$, 95% confidence interval of the group difference: $[-.054 .031]$). Overall, these findings suggest that both groups were adequately and equally attentive to the stimulation sequences.

A. Natural Contrast Group



B. Negative Contrast Group

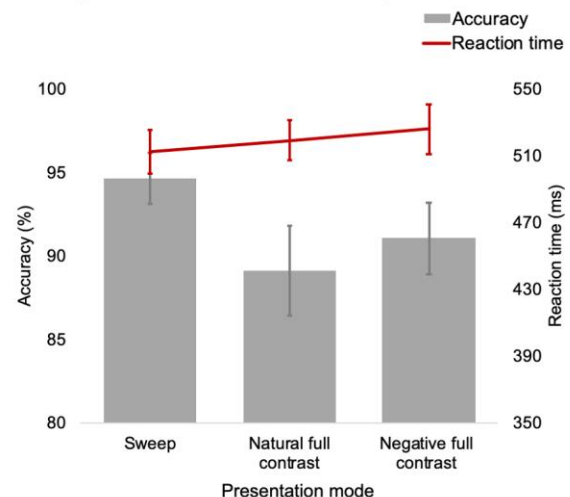


Figure 4-1. Group-averaged accuracy and correct reaction times in the vigilance task during the sweep and full-contrast sequences, separately. A. Behavioral performance in the natural contrast group. B. Behavioral performance in the negative contrast group. These histograms

show the means accuracy for each presentation mode. Secondary y-axis presents mean reaction time values (millisecond). Error bars represent 95% confidence intervals. Note that luminance and contrast were adjusted for figure visibility and are not representative of the actual (gamma-corrected) values.

EEG acquisition

Scalp EEG was recorded with 128 Ag-AgCl active-electrodes from the Biosemi ActiveTwo system (BioSemi B.V., Amsterdam, Netherlands) in a quiet and dimly-lit room. Default electrode labels (e.g., A1, B1, etc.) were renamed to approximate the 10/5 system (Oostenveld and Praamstra, 2001; Rossion et al., 2015). The vertical and horizontal electrooculogram (EOG) was monitored with four additional flat-type electrodes placed at the outer canthi of the eyes, and above and below the right eye. EEG data were sampled at 512 Hz and the magnitude of electrode offset was held below ± 25 mV.

Data analysis

Natural and negative contrast experiments followed the same analysis procedure outlined below.

Preprocessing

EEG data analysis was performed using *Letswave 5*, (www.letswave.org) running on MATLAB R2012b (MathWorks, MA, United States). We first de-trended and removed the DC component from the continuous EEG signal before applying a .1-120 Hz band-pass filter (4th order zero-phase Butterworth filter). Data were downsampled to 256 Hz for easier handling and segmented into sequences according to condition (High-to-Low contrast sweep, Low-to-High contrast sweep, and full-contrast), with two extra seconds before and after each sequence ($2 + 96 + 2$ s = 100 s). For each participant, we corrected artifacts related to eye blinks by applying independent component analysis (ICA) with a square mixing matrix on each participant and removing the single component corresponding to the blinks, identified based on the component waveform and scalp topography. Further artefact-ridden channels (i.e. sudden amplitude shifts exceeding ± 100 μ V over several epochs) were replaced by the average of the three neighbouring channels using linear interpolation. Next, the cleaned data were re-referenced to the average of the 128 scalp channels and averaged within conditions for each participant.

Frequency-domain analysis

In the sweep conditions, each preprocessed sequence was further cropped according to the 14 contrast steps into 6 s epochs, discarding the first and last 6 seconds corresponding to the pre-lude and post-lude. A fast Fourier transform (FFT) was applied to extract the frequency amplitude spectra for each step (i.e., normalized amplitude spectrum (μ V), with a frequency range of 0-128 Hz and frequency resolution of .167 Hz). For full-contrast sequences, we also cropped the four preprocessed sequences into 14 smaller epochs (total of 56 epochs) in order to obtain amplitude spectra with comparable frequency resolution as the sweep sequences. As there is no contrast modulation during these 14 epochs for full-contrast sequence, they were averaged together following the FFT.

To correct for variation in baseline noise level, the raw amplitudes were normalised relative to the background EEG activity by subtracting from the amplitude at the frequencies of interest (ie., 1.5 Hz and its harmonics for face categorization responses, and the stimulus presentation rate of 12 Hz and its harmonics; Figure 5) the mean amplitude of 8 neighbouring frequency bins (4 on either side, skipping the immediately adjacent bin) in each condition and

for each subject. To quantify general visual responses and face categorisation responses, we aggregated across their respective frequency harmonics (similarly to Quek et al., 2018b). The range of harmonics was defined a priori based on previous experiments using the current face categorisation frequency-tagging paradigm (e.g., Retter and Rossion, 2016; Quek et al., 2018b; Retter et al., 2020). Thus, we summed the baseline-corrected amplitudes of seven consecutive harmonics (i.e., 1.5, 3, 4.5, 6, 7.5, 9, & 10.5 Hz) for the face categorisation response, and four consecutive harmonics of 12 Hz up to 48 Hz for the general visual responses of each of the 128 channels. Going forward, the general visual responses and face categorisation responses will refer to these aggregated responses. We averaged the responses in the Low-to-High and High-to-Low contrast conditions according to the image contrast value. In other words, responses at step 1 in the Low-to-High condition were averaged with responses at step 14 in the High-to-Low condition, step 2 in Low-to-High with step 13 in High-to-Low, and so forth.

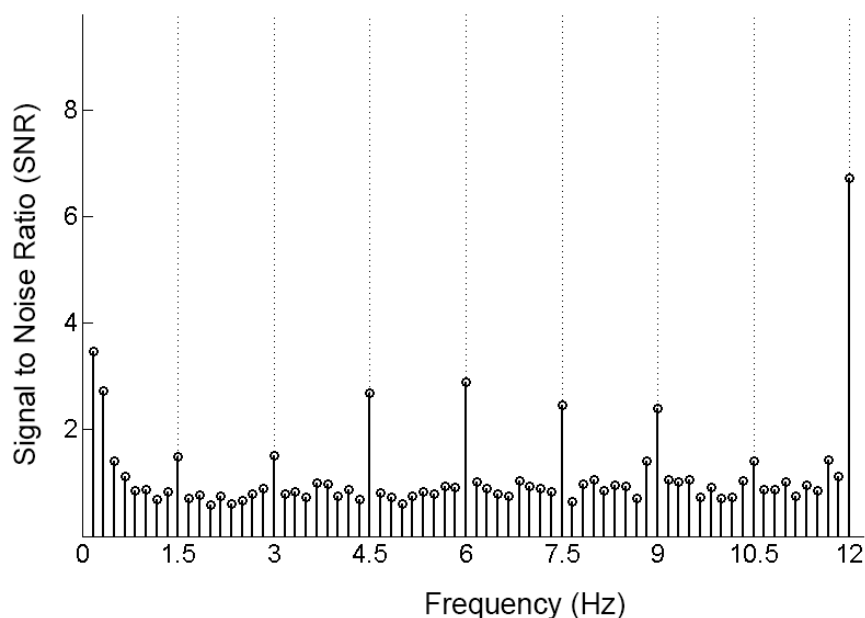


Figure 5. Frequency domain characterization of the neural response. Grand averaged SNR spectrum of the neural response to the contrast sweep at maximum visibility (i.e., step 14) at occipito-temporal channel (PO10). Note the robust response at the image presentation frequency (12 Hz and harmonics up to 48 Hz; not visible here). Responses to faces were tagged at 1.5 Hz and harmonics (1.5, 3, 4.5, 6, 7.5, 9, & 10.5 Hz).

Regions of interest

As in previous face categorisation frequency-tagging studies (e.g., Rossion et al., 2015), the scalp topography of both the general visual and face categorisation responses was comparable across observers; still there were moderate individual differences likely due to e.g., dipole generator orientation, cortical folding, skull thickness, and tissue conductivity (Luck, 2005; Nunez and Srinivasan, 2006; Woodman, 2010; Hauk et al., 2021). We therefore defined unique ROIs for each individual based on their EEG responses in the full-contrast condition. Thus, for each participant and for each type of response (general and categorization), we selected the four EEG channels showing the strongest response amplitude in the full-contrast sequences, and averaged amplitudes across these. ROIs selected this way replicated the previously reported dissociation between the general visual response (12 Hz) and the face categorisation response (1.5 Hz; Rossion et al., 2015; Retter and Rossion, 2016; Quek et al., 2018b). Overall, ROIs for the general visual response were located over medial occipital channels (the four channels most commonly selected for observers were Oiz-O2-POI2-Iz for

natural full-contrast and Oiz-Oz-PO11-O1 for negative full-contrast) whereas ROIs for the face categorisation response clustered over bilateral occipito-temporal channels (most commonly selected were channels P10-PO10-PO12-PO9 for natural full-contrast and PO12-P10-PO10-PO11 for negative full-contrast). There was little overlap between the ROI of the general visual response and the categorization response (5.83% overlap between channels in natural contrast group and 3.33% overlap in negative contrast group).

We investigated the hemispheric lateralisation of the face categorization response. To ascertain that the potential hemispheric differences result from functional lateralization, and not from differences in channel selection across hemispheres and individuals, we determined symmetrical right and left ROIs based on the group-averaged face categorisation responses in the full-contrast condition, i.e. the same channels were averaged for all participants for this analysis. The right occipito-temporal ROI (ROT) was composed of channels PO8, PO10, PO12, and P10, while the left occipito-temporal ROI (LOT) was composed of channels PO7, PO9, PO11, and P9.

Definition of minimal and optimal contrast

In each ROI and for each response, we defined the minimum and optimum levels of contrast for each neural response separately. The minimum contrast was the lowest contrast step to elicit a general visual or categorisation response amplitude significantly larger than 0 μ V. The optimal contrast was defined as the contrast step from which general visual or categorisation responses were not significantly different to responses to full-contrast stimuli. Note that full-contrast and individual contrast step amplitude values are based on different numbers of trials (n 56 and n 24, respectively), likely resulting in a difference of SNR. Yet, since this was the case in both the natural and negated samples, any potential SNR difference is expected to be similar across them, and unlikely to contaminate our finding of a relative group difference in the face categorization response dynamics. In Figure 5-3, we confirm that the amplitude of the face categorization response in full contrast sequences is robust to subsampling.

The contrast minimum and optimum were determined at the group-level using a bootstrapping procedure. First, for each type of response and in each experiment, we sampled 15 subjects with replacement from the original dataset. Next, we computed the 95% confidence interval of the mean of the bootstrapped sample at each contrast step of the sweep sequence. The contrast minimum was defined as the lowest contrast value where the response's 95% confidence interval exceeded zero. Similarly, we defined the optimum as the lowest contrast at which the 95% confidence interval included the mean amplitude of the full-contrast response in this bootstrapped sample. We repeated this entire process 10 000 times. Finally, we extracted the peak of the distributions of bootstrapped minimum and optimum values of the general visual or face categorisation responses.

Response amplitude comparisons

In addition to determining the minimum and optimum contrast, we examined whether the magnitude of the general and categorization responses differs across conditions. We expected the influence of contrast negation on the neural response amplitude to emerge at specific contrast levels, and show some stability across consecutive contrast levels, we ran permutation tests with cluster-based correction on the relevant datasets (see Maris and Oostenveld, 2007 for a detailed description of the procedure). However, since we were aware that the stringent cluster-based correction might favour the null hypothesis, we additionally report non-cluster-based statistics.

Briefly, we obtained the distribution of t-statistics under the null hypothesis of an absence of difference between conditions by permuting the condition labels of observer data

10,000 times. For each permutation iteration, we calculated the two-tailed t-statistic (for either pairwise or independent sample comparisons, depending the comparison) across all contrast steps, with an alpha of .05. We then saved the largest summed t-value among clusters of consecutively significant contrast steps. Finally, we computed the summed t-values within significant clusters found in the non-permuted data (alpha = .05) and calculated their respective p-values as their percentile relative to the permutation distribution of maximum cluster t-values.

Additional statistical analyses were carried out using mixed model or repeated measures ANOVAs. The Greenhouse-Geisser correction was applied whenever sphericity was violated and p-values were corrected via Bonferroni whenever applicable.

Results

In a fast periodic presentation paradigm, we measured the implicit neural signatures of the categorization of natural human face images (presented at a rate of 1.5 Hz) in a rapid 12 Hz stream of natural images of objects. We took advantage of periodic brain activity to extract the neural responses elicited by the image stream (i.e., the general visual response) and, most importantly, the periodic deviations from this response by the faces appearing at 1.5 Hz (i.e., the face categorization response; see Figure 5). We progressively modulated image contrast to estimate the amount of sensory input necessary for each type of neural responses to emerge (minimum) and reach its full-fledged amplitude (optimum) in the present design. Image sequences were presented either in their natural contrast polarity, or in a reversed, negative, contrast polarity. By comparing the minimum and optimum values of contrast for each neural response to emerge and reach their optimum, in natural and negative sequences, we quantified the influence of natural statistics on the dynamics of visual processing when categorizing faces in fast streams of naturalistic, in-the-wild, stimuli.

Face categorization response in full contrast viewing conditions

The face categorisation response reflects the brain's ability to consistently discriminate face images, despite their diversity, from the other object categories shown in the stimulus sequence (Rossion et al., 2015; Rossion et al., 2018). Across experiments, the responses to the full-contrast sequences of naturally-contrasted images were used as references to define the amplitude of the fully-fledged face categorization response, i.e., its amplitude at the maximum level of the tested contrast range. All observers showed periodic responses at the frequency of face presentation and its harmonics; all harmonic frequencies were aggregated to quantify the face-selective EEG response at the individual level (see Methods; Figure 5). In line with previous studies (e.g., Liu-Shuang et al., 2015b; Liu-Shuang et al., 2015a; Quek et al., 2018b; Retter et al., 2020; Or et al., 2021; Rekow et al., 2022), the group level categorisation response was most prominent over the bilateral occipito-temporal scalp electrodes (Figure 6). We defined occipito-temporal ROIs for each individual based on their response to full-contrast sequences.

Face categorization response (1.5Hz) Full contrast sequence

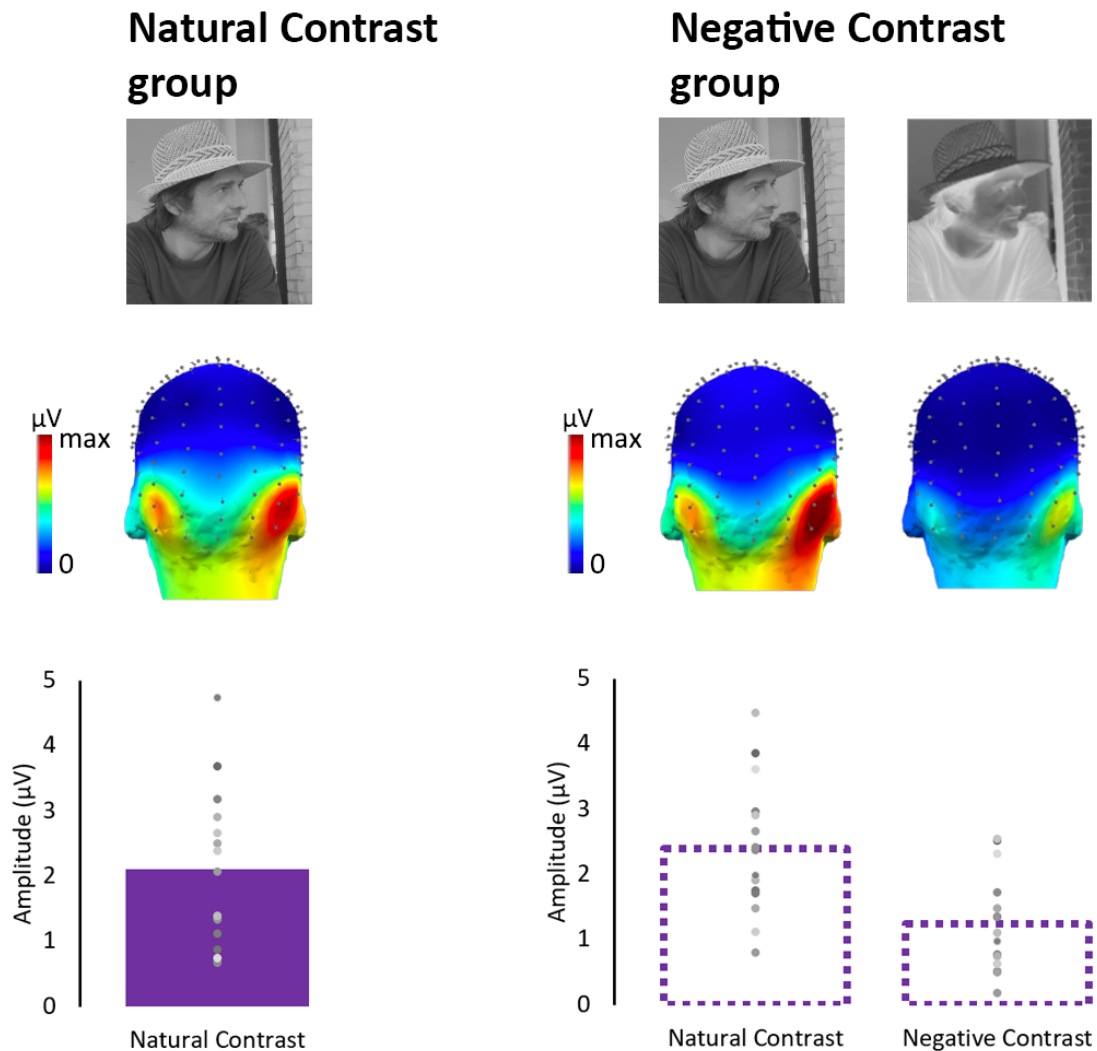


Figure 6. Face categorization response (1.5 Hz) in the full-contrast stimulation sequences, in natural and negative contrast groups. *top*, Group-level averaged scalp topographies. *bottom*, Mean baseline-corrected amplitudes averaged within individually defined ROIs (see Methods). Dots represent individual data points.

Figure 5-1. Individual ROIs in the Natural Contrast group. * = Participants which duplicate channels across general visual response ROIs and face categorisation ROIs.

	General visual response (12Hz)				Face categorisation response (1.5Hz)			
S02*	Iz	O2	Oiz	POI2	P10	PO10	PO12	POI2
S03	Oiz	Oz	POI1	POOz	I2	P10	PO11	PO12
S04	Iz	O2	Oiz	Oz	P10	PO10	PO12	PO7
S05	O1	Oz	POO5	POOz	P10	PO10	PO12	PO9
S06*	POO5	POO6	POOz	POz	O2	Oz	PO9	POO6
S07*	I1	Iz	Oiz	POI1	I1	P10	PO12	POI1
S08	Iz	Oiz	Oz	POOz	I2	P10	PO10	PO12
S09	I1	Iz	POI1	POI2	P10	PO10	PO12	PO9
S10*	O2	PO10	PO12	POI2	I2	P10	PO10	PO12
S11	O2	Oiz	POI1	POI2	P9	PO10	PO11	PO9
S12*	I2	O2	Oiz	POI2	P10	PO10	PO12	POI2
S13	O2	PO10	PO8	POI2	P9	PO11	PO7	PO9
S14	O2	Oiz	Oz	POO6	P10	PO10	PO8	POO6
S15	O1	Oz	POO5	POOz	P10	PO11	PO7	PO9
S16	I2	Iz	Oiz	POI2	PO11	PO7	PO9	POI1

Figure 5-2. Individual ROIs in the Negative Contrast group. * = Participants which duplicate channels across general visual response ROIs and face categorisation ROIs.

	General visual response (12Hz)				Face categorisation response (1.5Hz)			
S01	I1	O1	PO9	POI1	P10	PO10	PO11	PO12
S02	Iz	Oiz	POI1	POI2	P10	P9	PO11	PO12
S03	Iz	O2	Oiz	Oz	P10	PO10	PO12	PO8
S04*	O1	O2	Oiz	Oz	O2	PO10	PO12	PO8
S05	I1	Iz	Oiz	POI1	P10	P8	PO10	PO12
S06	I1	Iz	Oiz	Oz	P10	PO10	PO11	PO12
S07	O2	PO8	POI1	POI2	Iz	P9	PO11	PO9
S08	O2	Oz	POI2	POO6	P10	P9	PO11	PO9
S09	O2	Oiz	Oz	POI2	P10	P8	PO10	PO12
S10	O1	Oiz	Oz	POI1	P10	P8	PO10	PO12
S11	O1	Oiz	Oz	POOz	P9	PO11	PO12	PO9
S12	Oz	POO5	POOz	POz	P10	P8	PO10	PO12
S13*	I1	I2	POI1	POI2	I1	I2	PO12	POI2
S14	I1	O1	Oiz	POI1	P10	PO10	PO12	PO9
S15	O1	POI1	POO5	POO6	P10	P8	PO10	PO12

Figures 5-1 and 5-2 list the four channel ROIs for each individual in the Natural and Negative Contrast group, respectively. Asterisks indicate participants with duplicate channels across the general visual response and face categorisation ROIs.

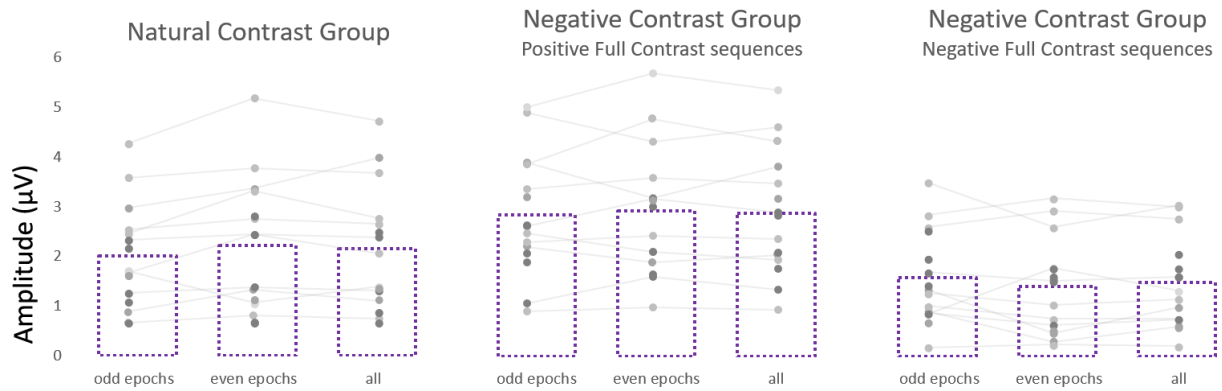


Figure 7-3. Halving the number of full-contrast trials does not influence the amplitude of the face categorization response. The mean amplitude of the face categorization response to full-contrast stimulation sequences was stable whether based on all trials or only (odd/even) half of them, in natural and negative contrast groups (in individually defined ROIs; see Methods). Dots represent individual response amplitude values. The number of full contrast and sweep trials ($n = 56$ and $n = 24$, respectively) differed equally across groups; any potential SNR difference arising due to this difference of trial number should thus be similar across groups, and not contribute to our main findings. Nevertheless, we subsampled trials from the full-contrast condition in order to investigate whether the amplitude of the face categorization response was influenced by subsampling. To do so, we split full contrast trials into odd and even 'steps' and tested whether face categorization response amplitude was similar across sampling methods (odd, even, all trials). As shown in the figure below, face categorization response in full contrast sequences is relatively stable in amplitude irrespective of the number of trials. In a mixed ANOVA with group (natural contrast, negated contrast) and sampling method (odd, even, all) as factors, we confirmed that the sampling method did not influence the amplitude of the face categorization response significantly ($F(1.37, 38.39) = 2.69, p = .1, \eta^2 = .09$). Furthermore, the interaction between sampling method and group was non-significant ($F(1.37, 38.39) = .83, p = .4, \eta^2 = .029$). While null results should always be taken with caution, they suggest that the face categorization response we measured in full contrast trials across groups was stable whether the full sample of trials was used or if we halve it to get closer to the n of sweep trials.

Although there was a slightly larger amplitude overall in the negative contrast group for the full natural contrast sequences (Figure 6), a one way ANOVA with *Group* (natural group, negative group) as between-subject factor did not reveal any difference across natural and negative contrast groups ($F(1,28) = .522, p = .476, \eta^2_p = .018$).

Furthermore, since the participants in the negative contrast experiment viewed a full-contrast sequence with both natural images and their contrast-negated versions, we were able to examine the impact of contrast negation on the full-contrast face categorization response. A one-way repeated-measures ANOVA with *Contrast polarity* (natural vs. negated) as within-subject factor revealed a significant main effect of *Contrast polarity* ($F(1,14) = 26.53, p < .0001, \eta^2_p = .65$) with contrast negation reducing the amplitude of the face categorisation response by 47% ($\pm 13\%$) on average (see Figure 6).

These findings confirm the drastic effect of contrast negation on face categorisation, quantifying it for the first time here with natural images. It further demonstrates that at least half of the amplitude of the face categorization response is not accounted for merely by the physical parameters of the stimulus but reflects the successful match between the incoming input and the stored representation of what a face generally looks like.

Face categorization response (1.5Hz)

Contrast sweep sequence – Group average

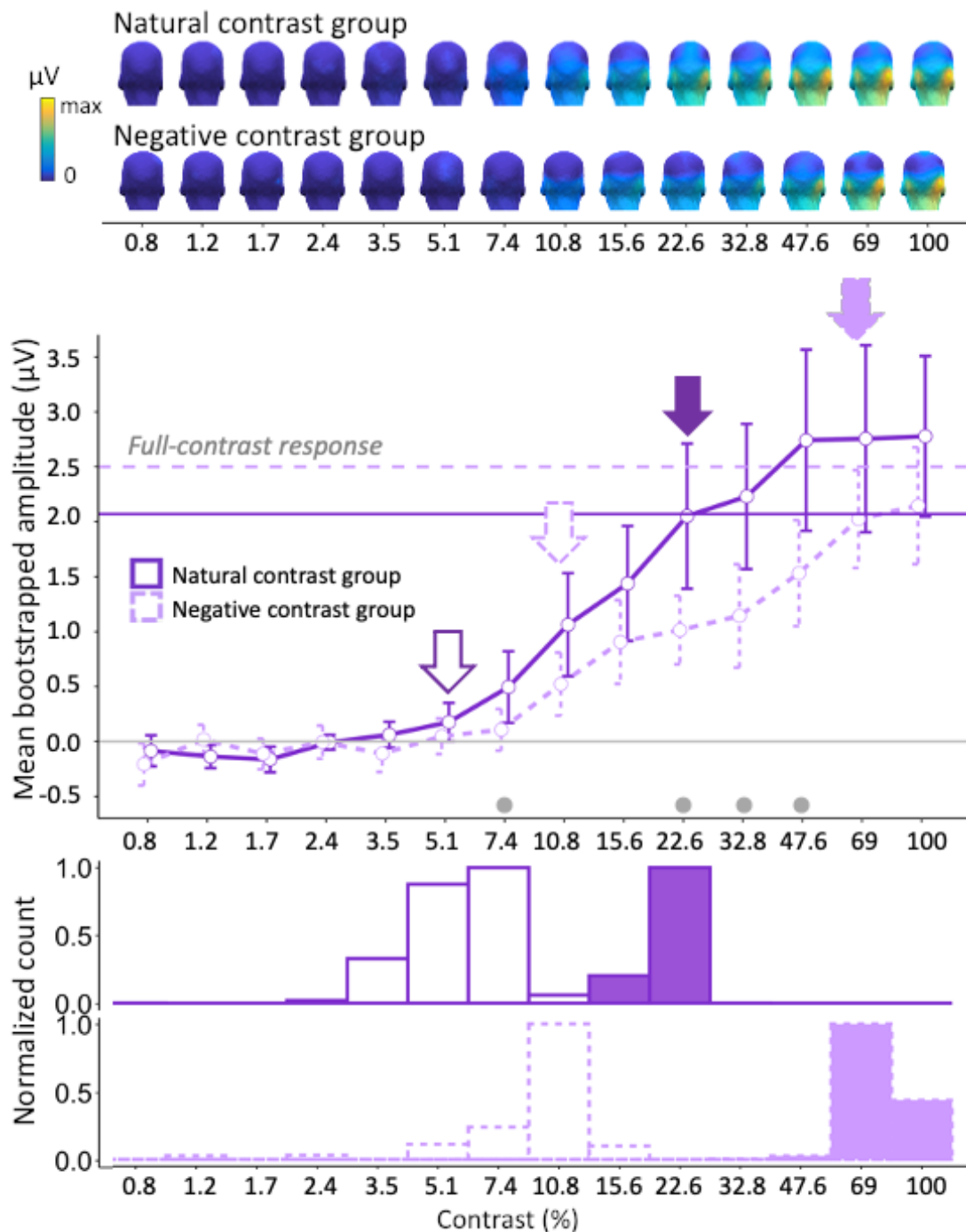


Figure 8. Group-level face categorization response in contrast sweep sequences, in natural and negative contrast groups. *Top.* Scalp topographies at different contrast steps. *Bottom.* Line plots depict the profile of the bootstrapped mean face categorization response (averaged within individually defined ROIs; see Methods) as a function of contrast (plain line for natural contrast and dotted line for negative contrast). Open and filled arrows indicate the minimum and optimum levels of contrast, respectively. Error bars represent mean bootstrapped 95% confidence intervals. Dots below the response curves represent significant amplitude differences between natural and negative contrast group responses at each contrast steps (grey dots for significance at $p < .05$ with no cluster-based correction; see Methods). Bar graphs represent the peak of the bootstrapped distributions (normalised within a range of 0 to 1) of the contrast minimum (open bars) and optimum (filled bars) in natural and negative contrast groups. Face categorisation based on negative contrast images elicited weaker response amplitudes and raised minimum and optimum thresholds relative to natural contrast images.

Contrast negation influences both the minimal and optimal contrast for human face categorisation

The profile of the 1.5 Hz face categorization response in the contrast sweep sequences is plotted separately for the natural and negative contrast polarity groups (Figure 8). In the natural contrast experiment, face categorisation responses hovered near zero in the lower end of contrast continuum and became significant already at 5.1% of contrast. This response increased with further image visibility until it no longer differed from full-contrast face categorisation responses at 22.6%. Thus, under ecological contrast polarity, the categorisation of faces required a minimum contrast of 5.1% to emerge and reached full-contrast levels at only 22.6% of contrast. In comparison, as illustrated in Figure 8, the face categorisation responses to contrast-negated images required not only a higher contrast to emerge (10.8%) but a much higher level of contrast (69%) to reach full-contrast response levels. These findings suggest that the disruption of natural statistics did not only increase the amount of minimal sensory input that is necessary to elicit a categorization response but also drastically protracts the dependence of the categorization response on contrast, suggesting a less efficient visual process. While there was some degree of inter-individual variability in the amplitude and steepness of EEG response as a function of contrast, the response profiles were largely consistent across participants (Figure 9).

Face categorization response (1.5Hz) Contrast sweep sequence - Individual data

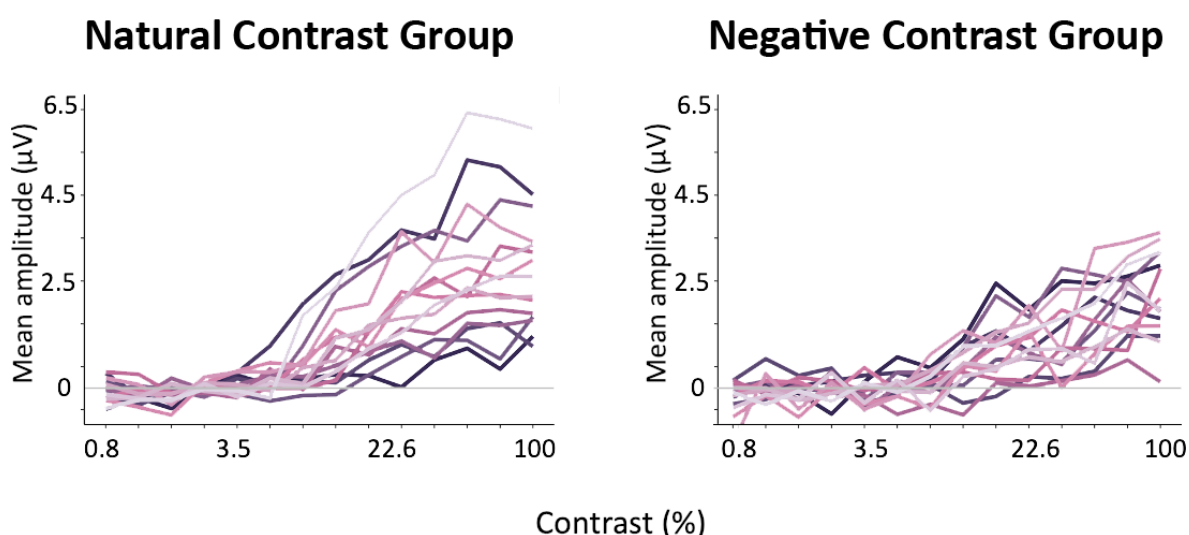


Figure 9. Line plots of the individual face categorization responses (1.5 Hz) in contrast sweep sequences, in natural and negative contrast groups.

The observation that the categorization of contrast-negated face images elicits neural responses with an amplitude that, at high contrast (69%; 100%) almost reaches the amplitude of the categorization response to natural images may seem surprising. Since contrast negation decreased the face categorization response by half in full-contrast sequences, one could have expected the categorization responses to contrast-negated in sweep sequences to never reach the categorization response amplitude for natural images in full contrast sequences. Yet, contrast negation reduced the amplitude of the face categorisation response in the contrast sweep sequences even at most supra-threshold contrast levels, showing again the detrimental

effect of contrast negation on categorization in sweep sequences (Figure 8; uncorrected p-value range = .007 – .047; these differences just failed to reach the cluster-wise significance minimum: cluster p-value range = .06 – .3).

Hemispheric lateralisation of the face categorisation response

To address the potential hemispheric differences in the modulation of the face categorization as a function of contrast, we computed the minimum and optimum contrast levels within left and right occipito-temporal ROIs in natural and negative contrast experiments, separately (Figure 10).

These analyses were conducted on symmetric ROIs predefined at the group level (see Methods) to ensure that hemispheric differences in face categorization responses reflect genuine functional lateralization, and not individual differences in channel selection. The resulting group-level ROIs contained channels with weaker amplitudes than the individual ROIs, since the latter were selected to contain the largest response amplitudes at the individual level.

First, we tested the hemispheric lateralization of the face categorization response to natural full-contrast sequences (Figure 10A). To do so, we conducted a mixed model ANOVA with *Hemisphere* as a within-subject factor and *Group* (natural group, negative group) as a between-subject factor. We found a significant main effect of *Hemisphere* ($F(1,28) = 7.14$, $p = .012$, $\eta^2p = .2$), with larger amplitudes over the right than the left hemisphere. Overall the face categorization response in natural full-contrast sequences was of similar magnitude across groups, corroborating the findings with the full-contrast responses in individual ROIs ($F(1,28) = .13$, $p = .72$, $\eta^2p = .005$). The interaction between *Hemisphere* and *Group* was not significant ($F(1,28) = .26$, $p = .61$, $\eta^2p = .009$).

We additionally tested the effect of *Contrast Polarity* as a within-subject factor in the full-contrast data of the negative contrast group. The 2x2 repeated measures ANOVA with *Contrast polarity* (natural vs. negative contrast experiment) and *Hemisphere* (LOT vs. ROT) as within-subject factors revealed a significant main effect of *Hemisphere* ($F(1,14) = 25.28$, $p < .0001$, $\eta^2p = .64$) and of *Contrast Polarity* ($F(1,14) = 5.8$, $p = .031$, $\eta^2p = .29$). The amplitude of the face categorization response was again larger in the right than the left hemisphere. As expected, full-contrast images presented in their natural contrast polarity resulted in the strongest face categorization response overall (Figure 10A). The absence of interaction between *Contrast polarity* and *Hemisphere* ($F(1,14) = 1.15$, $p = .3$, $\eta^2p = .08$) suggests a comparable influence of *Contrast Polarity* across hemispheres during the viewing of full contrast periodic sequences.

For contrast sweep sequences presented in a natural polarity, our bootstrapping procedure showed that face categorisation responses emerged at 7.4% contrast and saturated at 15.6% in both the left and right hemispheres (Figure 10B). The right hemisphere showed higher amplitudes at most contrast levels (from 10.8 to 32.8% of contrast: cluster p-value < .034). In other words, besides an overall right-lateralization of the face categorization responses, visual processing dynamics when categorizing ecological faces were comparable across hemispheres, at least with the resolution of the logarithmic contrast steps used in the current paradigm.

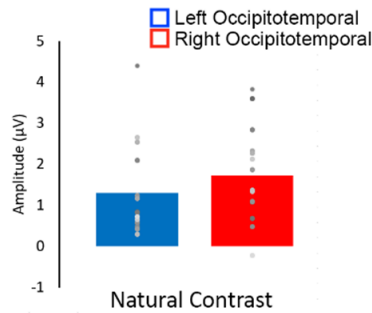
For contrast-negated image sequences, the face categorisation responses emerged later, at a contrast value of 10.8% and reached optimum later, at 69% in both the right and left occipito-temporal regions. Hemispheric differences in amplitude were only significant at the higher ends of the contrast continuum and at an uncorrected p value (uncorrected p-value range = .019-0.04; no clusters reached cluster-based significance minimum, cluster p-value =

.28). Thus, the right hemispheric lateralisation of the face categorization response is less reliable overall for negative than natural contrast images.

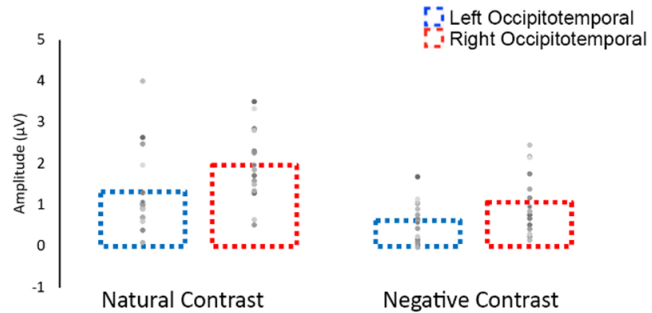
Lateralization of face categorization response (1.5Hz)

A. Full contrast sequence

Natural contrast group

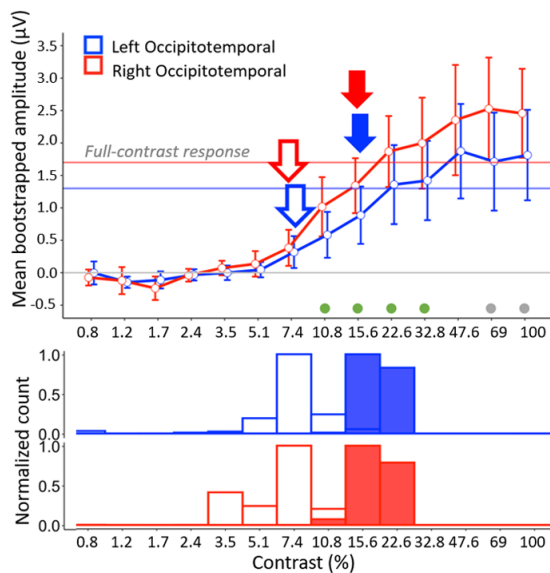


Negative contrast group



B. Contrast sweep sequence

Natural contrast group



Negative contrast group

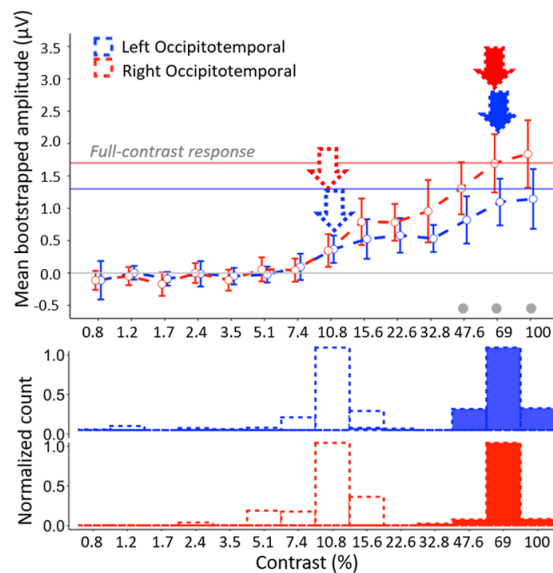


Figure 10. Group-level face categorisation responses (1.5 Hz) for left (LOT; in blue) and right occipito-temporal (ROT; in red) ROIs, defined at the group-level (see Methods). A. Full contrast sequences. Group-averaged amplitudes in LOT and ROT ROIs in the natural contrast group (left panel) and negative contrast group (right panel). Dots represent individual response amplitude values. B. Contrast sweep sequence. *Left Panel.* Natural contrast group. Bootstrapped sweep response profiles and corresponding minimum distributions for LOT and ROT group-level ROIs. The mean bootstrapped responses in the LOT and ROT are plotted in the top row (error bars represent mean bootstrapped 95% confidence intervals). Open and filled arrows indicate the minimum and optimum levels of contrast, respectively. The dots below the response curves represent significant hemispheric differences in amplitude (grey dots for significance at $p < .05$; green dots for $p < .05$ significance following cluster-based correction; see Methods). The rows below represent the bootstrapped distributions of the contrast minimum (open bars) and optimum points (filled bars) in the LOT (middle row) and the ROT (bottom row). Note that the histograms have been separately normalised within a range of 0 to 1. *Right Panel.* Negative contrast group. Bootstrapped sweep response profiles and corresponding minimum distributions for left (LOT) and right occipito-temporal (ROT) ROIs, displayed using the same conventions as in A.

In sum, we found no consistent impact of contrast negation on the right hemispheric lateralization of the face categorization response. However it is important to keep in mind that these findings stem from group-level ROIs, which only partially overlap with the maximally-responsive channels as defined at the individual level. Therefore, we cannot exclude the possibility that hemispheric differences would emerge if the maximally responsive channels would perfectly overlap across individuals. Besides addressing the lateralization of contrast negation effects upon face categorization, the present results confirm that the disruption of natural image statistics by means of contrast negation protracts the contrast dependence of the neural categorization response. This suggests that the impact of natural contrast statistics on visual processing dynamics replicates irrespective of the method used to select scalp electrodes (individual-level versus group-level) and attests the robustness of the findings.

To summarise, the disruption of image natural statistics through contrast negation not only decreased the neural signatures of face categorization at full contrast by almost half of their amplitude, the novel finding is that it influenced the underlying visual processing dynamics by protracting both the emergence (i.e., contrast minimum) and the optimum of the electrophysiological face categorization response to higher contrast levels. Contrast negation did not merely lead to higher contrast thresholds, it drastically expanded the dependence of the electrophysiological categorization response on contrast, and reduced its amplitude at most contrast steps.

Contrast negation does not affect the 12 Hz general visual response to full-contrast sequences

Besides the face categorization response, all observers showed a neural response at the 12 Hz frequency and harmonics, in response to the 12 Hz image stream (Figure 5). This neural signal reflects the general response of the visual system to the stimulus sequence, irrespective of the category membership of the presented images. As shown in Figure 11, the general visual response to full-contrast images was centred over medial occipital channels.

The medial occipital location of this response suggests that it reflects the low-level encoding of image basic properties such as amplitude spectrum and contours (Ales et al., 2012; Norcia et al., 2015; Hauk et al., 2021), which are comparable between positive and negative polarity images. There was no difference between groups for the 12 Hz neural response to the natural contrast polarity images ($F(1,28) = .38$, $p = .54$, $\eta^2_p = .014$). Similarly, in the negative contrast group, there was no effect of *Contrast Polarity* on the 12 Hz general visual response ($F(1,14) = .135$, $p = .72$, $\eta^2_p = .01$).

To further establish that contrast negation selectively affected face categorization while leaving generic visual response unchanged in full contrast viewing conditions, we ran a 2x2 repeated measures ANOVA with *Contrast polarity* (natural vs. negated) and *Response* (general visual vs. face categorisation) as within-subject factors. The *Contrast polarity* x *Response* interaction was significant ($F(1,14) = 30.54$, $p < .001$, $\eta^2_p = .69$) confirming that contrast negation selectively reduced the amplitude of the 1.5 Hz face categorization response ($P_{\text{bonf}} < .001$), leaving the general visual response to the 12 Hz sequence unaffected ($P_{\text{bonf}} = 1$). In interpreting this interaction, it is important to recall the difference in SNR between the response types (Figure 5-3), which reflects the generally higher number of events contributing to 12 Hz responses compared to 1.5 Hz responses. However, it is unlikely that the interaction reflects the difference in SNR as the effect of negation was only significantly observed for the 1.5 Hz categorisation, i.e. the response with the lowest SNR. Main effects were of no interest

in the present context; they are only reported here for the sake of completeness (main effect of *Response*: $F(1,14) = .68$, $p = .423$, $\eta^2_p = .046$; main effect of *Contrast polarity*: $F(1,14) = 21.58$, $p < .001$, $\eta^2_p = .61$).

These findings confirm the selective impact of contrast negation on the face categorisation response and thus validates this image manipulation as a means to disentangle the relative contributions of low- vs. high-level visual mechanisms to the presently observed contrast sensitivity profiles.

General visual response (12Hz) Full contrast sequence

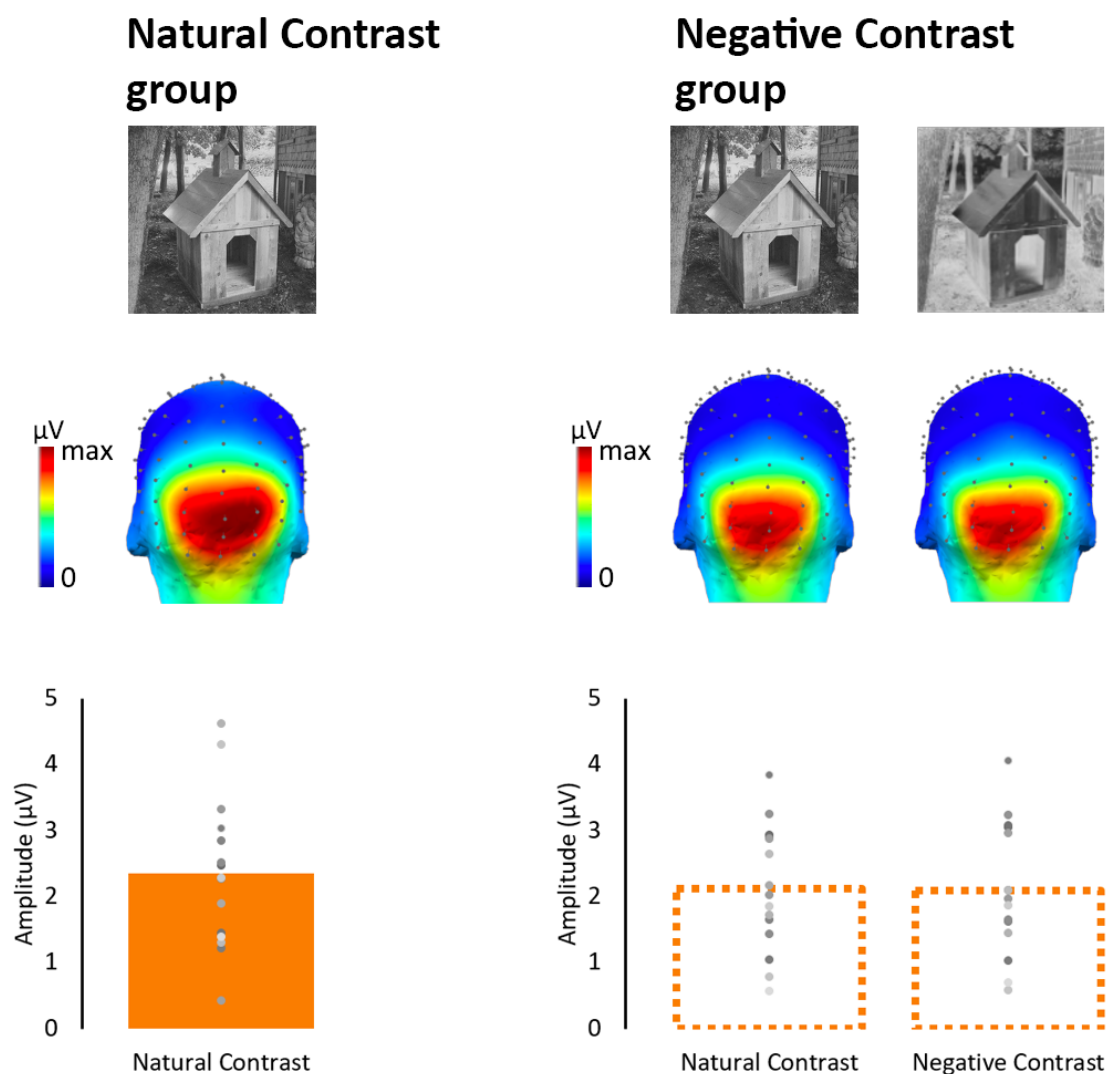


Figure 11. General visual response to the 12 Hz periodic rate in full-contrast stimulation sequences, in natural and negative contrast groups. *Top*. Group-level averaged scalp topographies. *Bottom*. Mean amplitudes averaged within individually defined ROIs (see Methods). Dots represent individual response amplitude values.

Contrast negation increases the minimum level of contrast necessary for the emergence of general visual responses

Next, we examined how the 12 Hz general visual response evolved as a function of contrast. The response profile of the contrast sweep sequences is plotted separately for the natural and negative contrast polarity groups (Figure 12). While there was some degree of inter-individual variability in the amplitude and steepness of EEG response as a function of contrast, the response profiles were largely consistent across participants (Figure 13).

General visual response (12Hz)

Contrast sweep sequence – Group average

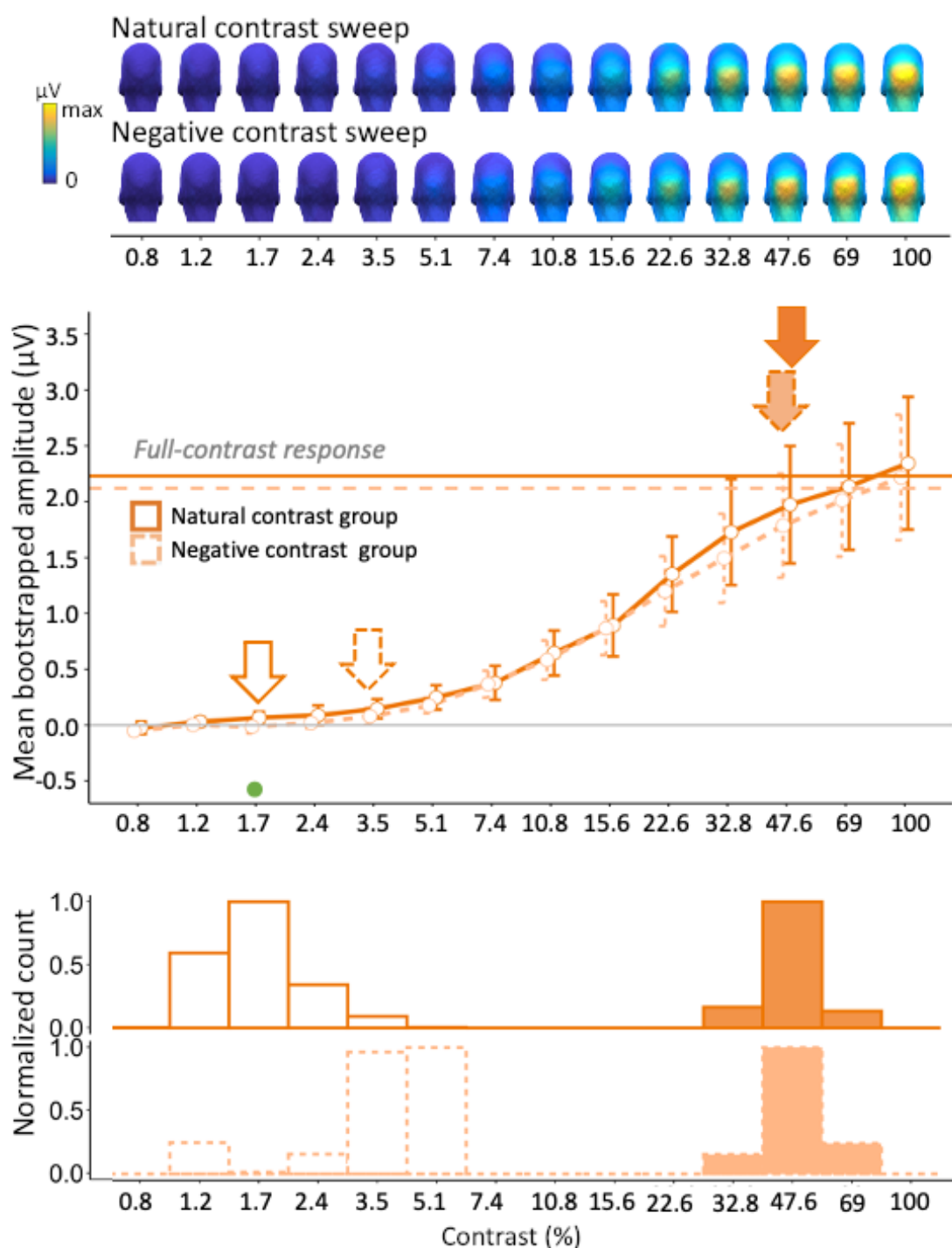


Figure 12. Group-level common visual response (12 Hz) in contrast sweep sequences, in natural and negative contrast groups. Scalp topographies at different contrast steps. Line plots depict the profile of the bootstrapped mean face categorization response as a function of contrast (averaged within individually defined ROIs; plain line for natural contrast and dotted line for negative contrast). Open and filled arrows indicate the minimum and optimum levels of contrast, respectively. Error bars represent mean bootstrapped 95% confidence intervals. Dots below the response curves represent significant amplitude differences between natural and negative contrast group responses at each contrast steps (grey dots for significance at $p < 0.05$; green dots for $p < 0.05$ significance following cluster-based correction; see Methods). Bar graphs represent the peak of the bootstrapped distributions of the contrast minimum (open bars) and optimum (filled bars) in natural and negative contrast groups. Note that the histograms have been separately normalised within a range of 0 to 1.

In natural image sequences, the group-level general response emerged significantly above zero at 1.7% contrast and then increased steadily until reaching the optimum level of full-contrast responses at 47.6% contrast (Figure 12). Hence, the general visual processes triggered by the 12 Hz stimulation sequence exhibited sensitivity to a wide range of contrast levels.

When contrast polarity was negated, the general visual response significantly emerged at a higher level of contrast (3.5% of contrast; Figure 12). Despite this disparity in response onset, the general visual response in the negative contrast experiment rose to optimum at the same contrast level (47.6%) for natural and negative images. In terms of response amplitude, only a single contrast step differed significantly between experiments (step 3 = 1.7% contrast, cluster p -value $< .046$).

General visual response (12Hz) Contrast sweep sequence - Individual data

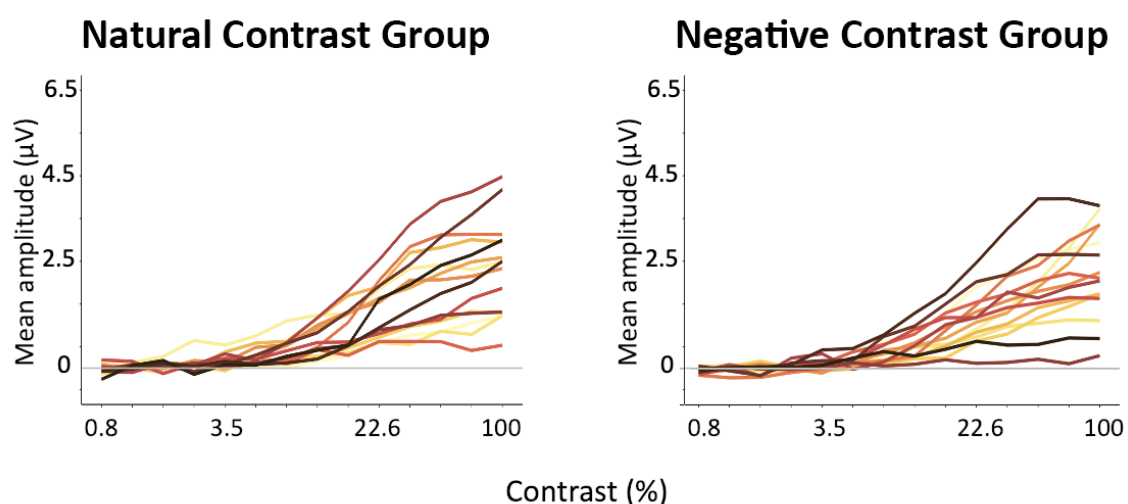


Figure 13. Line plots of the individual common visual response (12 Hz; averaged within individually defined ROIs) profiles in contrast sweep sequences, in natural and negative contrast groups.

Discussion

The present study reports a systematic investigation of the influence of natural image statistics on the visual processing dynamics occurring during human face categorization. Reversing the natural polarity of image contrast severely disrupts face categorization and reduces brain responses to faces compared to when their natural contrast polarity is preserved. Yet, *how* natural contrast statistics contribute to visual processing during naturalistic categorization has been so far elusive due to two main reasons. First, the effect of contrast negation on face categorization has typically been investigated using restricted sets of segmented images of faces and objects; these testing conditions are very different from everyday life, where faces and objects have to be categorised in a cluttered environment and despite a wide variety of appearances. A second limitation is that past studies did not investigate the nature of the facilitation triggered by natural statistics. It is unclear whether knowledge of natural statistics decrease the amount of sensory input that is necessary to see the face categorization emerge in the observer's brain (i.e., earlier onset), if it reduces the sensory input needed to generate a fully-fledged categorization response (i.e., earlier optimum), or both, making the categorization less dependent on contrast overall. We found that natural contrast statistics cause a drastic reduction in the amount of sensory input required for face categorization responses to emerge and also reach optimum, therefore severely reducing the minimum/optimum distance.

The medial occipital neural response elicited by the 12 Hz image stream and thought to originate primarily in low-level visual cortex (Hauk et al., 2021), increased monotonically with stimulus contrast in both natural and negated sequences. Contrast negation induced a moderate shift in the onset of the general visual response towards higher contrast values, but did not affect the level of contrast required to elicit an optimal general visual response or its overall magnitude. It is unclear whether this shift in the emergence of the general visual response is due to a differential activation of the center-surround receptive fields in the retina or the lateral geniculate nuclei (Meister and Berry, 1999), and/or whether it is due to contrast negation hindering general, low-level visual processing in the cortex (e.g. perception of brightness: Yang and Purves, 2004; Pickard-Jones et al., 2020; orientation; Girshick et al., 2011; apparent contrast: Bex and Makous, 2002; Haun and Peli, 2013). Hence, shading, i.e. the pattern of luminance across a surface, is informative of the structural properties of many things around us, not just faces. Combined with the prior assumption that light comes from above in the natural environment, the human visual system forges strong priors of how shape should be interpolated from shading in everyday life. For example, contrast negation impedes texture perception (Balas, 2012), and local shape computations based on texture cues (Fleming et al., 2004). Shading disruption through negation is therefore expected to disrupt the general visual processes reflected by the medial occipital 12 Hz response at least to some extent, especially when the stimuli are diverse, cluttered and complex scenes, i.e., challenging viewing conditions that maximally tap into prior experience (Hupe et al., 1998; Wyatte et al., 2014).

The impact of contrast negation was much more spectacular for human face categorization responses. For naturally-contrasted images, the face categorization response emerged at 5.1% of contrast (RMS contrast of .009) and quickly reached its optimal amplitude (22.6% or .041 RMS contrast). Considering the complexity and wide variety of face images in the sequence, this is an intriguingly low value. It indicates that when the sensory input matches ecological regularities, the human brain is able to categorize, i.e., consistently differentiate highly variable faces from other visual categories, even in seemingly suboptimal, low-contrast, viewing conditions.

In natural contrast viewing conditions, the rise in face categorization response was fast but gradual as it only reached its fully-fledged amplitude 4 steps later, at 22.6% of contrast.

This profile indicates the progressive but rapid direction of categorisation by natural statistics. Past evidence indicates that a fast rise in the face categorization response amplitude results from a proportionately rapid increase in the number of faces that are correctly categorized in the rapid image sequences (Retter et al., 2020). Whether such rise of the face categorization response reflects an increase in the sensitivity to face images, in the tolerance to face image variability, or both cannot be teased apart in the present study.

Our findings also imply that there is at least a restricted range where face categorization is contrast-dependent. This disagrees with previous evidence suggesting that the high-level cortical regions involved in the categorization of complex stimuli are (quasi-)invariant to basic properties of the stimulus such as e.g. contrast, position, size to enable efficient categorization (e.g. Ito et al., 1995; Andrews and Ewbank, 2004; Kovacs et al., 2008). Instead, our findings aligns with several studies that showed that high-level stages of visual processing are at least to some extent sensitive to contrast (e.g., Rolls and Baylis, 1986; Avidan et al., 2002; Murray and He, 2006; Yue et al., 2013), and with the more nuanced stance suggested by Nasanen et al. (2006) that human face identification is influenced by contrast only at low contrast levels ($\text{RMS} < .1$). Here, the RMS value at which the face categorization response reached its optimum level was .041, thus at a lower level than the one reported by Nasanen et al. (2006). This difference is most probably explained by methodological aspects. Nasanen and colleagues (2006) explicitly instructed their participants to recognize the identity of faces flashed in homogeneous sequences of segmented face images. The basic level of categorization, the high stimulus heterogeneity, as well the reliance on implicit neural measures likely explain the lower cutoff RMS value found here. Leaving aside these aspects, our finding that the face categorization response did not reach its optimum as soon as it emerged (minimum) is in line with the conclusions of these authors, namely that the underlying mechanisms are contrast-variant at the lower ends of the contrast continuum ($\text{RMS} < .041$), and resistant to contrast at higher levels of contrast.

In agreement with previous evidence, we found that the disruption of natural statistics due to contrast negation reduced by about a factor of 2 the overall amplitude of the face categorization response (Liu-Shuang et al., 2015a). More interestingly, contrast negation doubled the minimum level of contrast necessary for the human brain to categorize a face as a face. While face categorization responses emerged at 5.1% of contrast (RMS of .009) in natural contrast settings, a minimum of 10.8% (RMS of .019) of contrast was required in negated streams. Furthermore, contrast negation expanded visual processing, shifting the optimum of the categorization response to 69% (RMS of .124), which is three times the optimal contrast for face categorization in natural contrast settings (22.6%; RMS of .041).

Altogether this indicates that the access to internally stored knowledge about how faces appear in real-life, i.e., their natural statistics, does not merely enable stronger output responses, but actually empowers a faster and more efficient visual processing for the purpose of categorization. When these natural statistical priorities fail to guide visual processing, much more sensory input (contrast) is needed, which weakens and pushes the neural categorisation response to a considerably higher contrast. In contrast to its dramatic effect on the dynamics of the face categorization response, we found no consistent impact of contrast negation on right hemispheric specialization in face processing. This null finding must be interpreted with caution but suggests that the right-hemispheric specialization of face processing is not triggered by the contrast ordinal relationships of the human face.

In the fast periodic visual stimulation paradigm used here, each image appears very briefly and is masked both by the preceding and succeeding images (similar to rapid serial visual presentation; Potter and Levy, 1969; Keysers et al., 2001, 2005; Potter et al., 2014). This implies that most of the facilitation observed for the categorization of natural images happened during

the feedforward transmission of the visual signals (Lamme and Roelfsema, 2000; Keyser and Perrett, 2002; Bacon-Mace et al., 2005; Fahrenfort et al., 2007; although see Girard et al., 2001; Hupe et al., 2001 for monkey electrophysiological evidence of rapid feedback). Several authors have discussed the possibility of the influence of natural statistics priors in a purely feedforward architecture, with each ventral stream stage amplifying the features (and contingencies of those) that are relevant for categorization (Simoncelli and Olshausen, 2001; Martin et al., 2018 but see Kar et al., 2019). Clarifying the exact mechanisms of such amplification is out of the scope of the present work. Yet, the fact that natural statistics influenced the dynamic visual processing leading up to categorization but seemed to only moderately guide the general visual responses indicate that most of such amplification occurs when signals are read out by the high-level visual cortex responsible for categorization.

As human observers experience that members of a visual category share similar shape and function, access to such natural statistics is likely to facilitate visual processing dynamics for their categorization. Indeed, whether an input corresponds to stored natural statistics determines its categorizability. Since the current study was limited to categorizing human faces, our findings may not apply to other (visual) categories. However, we speculate that the facilitation of processing dynamics for other visual categories is unlikely to be as clear as for faces. Indeed, face categorization is one of the most critical visual functions for human social adaptation, and this likely exerted a strong evolutionary pressure for the development of fast and efficient categorization mechanisms. This may explain why the human species has developed an exceptional experience with and sensitivity to the universal properties of face stimuli as a strict and ordered alternation of contrast (Figure 1; e.g., Watt, 1994; Dakin and Watt, 2009; Gilad et al., 2009). Figure 1 illustrates the universality of this property and suggests that the current results can be generalised to all faces, regardless of skin tone. Such priors strongly drive the human visual system as illustrated by face pareidolia (e.g. Rekow et al., 2021). Non-face visual categories do not seem to lend such strict contrast rules, i.e., their processing is far less affected by contrast negation (i.e., dogs, Robbins and McKone, 2007; chairs, Itier et al., 2006), artificial objects (i.e., “Greebles”; Vuong et al., 2005), or abstract visual objects (i.e., “blobs”, Nederhouser et al., 2007; Yue et al., 2013). This implies that the influence of contrast negation observed here on the electrophysiological markers of face categorization are mostly driven by the face images. Nevertheless, non-face categories may share other visual regularities, stored in the human visual system as their natural statistics (see Long et al., 2018 for evidence that mid-level curvature features driving categorization into (in)animate categories). However, since the current study did not systematically compare the impact of natural contrast statistics for the categorization of face and non-face categories, whether such natural statistics facilitates visual processing as vigorously as for face stimuli requires further investigation.

Conclusion

How do natural statistics contribute to the visual processing dynamics leading up to human face categorization? Here we show that natural statistics not only boost the amplitude of face categorization responses in the human brain, but also modulate the underlying processing dynamics by halving the minimal amount of sensory input required to categorize faces. Once this minimum has been reached, the face categorization response samples further input, but reaches its optimum rapidly, i.e., at a three times lower level of contrast than for contrast-negative faces. These results have important implications for how internally stored natural statistics facilitate visual processing for rapid and efficient categorisation.

References

References.

- Adrian ED, Matthews BHC (1934) The Berger rhythm: potential changes from the occipital lobes in man. *Brain* 4:355–385.
- Ales JM, Farzin F, Rossion B, Norcia AM (2012) An objective method for measuring face detection thresholds using the sweep steady-state visual evoked response. *J Vis* 12.
- Andrews TJ, Ewbank MP (2004) Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *NeuroImage* 23:905-913.
- Avidan G, Harel M, Hendler T, Ben-Bashat D, Zohary E, Malach R (2002) Contrast sensitivity in human visual areas and its relationship to object recognition. *J Neurophysiol* 87:3102-3116.
- Bacon-Mace N, Mace MJ, Fabre-Thorpe M, Thorpe SJ (2005) The time course of visual processing: backward masking and natural scene categorisation. *Vision Res* 45:1459-1469.
- Balas B (2012) Contrast negation and texture synthesis differentially disrupt natural texture appearance. *Front Psychol* 3:515.
- Behrmann M, Plaut DC (2020) Hemispheric Organization for Visual Object Recognition: A Theoretical Account and Empirical Evidence. *Perception* 49:373-404.
- Bex PJ, Makous W (2002) Spatial frequency, phase, and the contrast of natural images. *Journal of the Optical Society of America A, Optics, image science, and vision* 19:1096-1106.
- Boynton GM, Engel SA, Glover GH, Heeger DJ (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16:4207-4221.
- Bruce V, Langton S (1994) The use of pigmentation and shading information in recognising the sex and identities of faces. *Perception* 23:803-822.
- Buiatti M, Di Giorgio E, Piazza M, Polloni C, Menna G, Taddei F, Baldo E, Vallortigara G (2019) Cortical route for facelike pattern processing in human newborns. *Proc Natl Acad Sci U S A* 116:4625-4630.
- Clifford CW, Mareschal I, Otsuka Y, Watson TL (2015) A Bayesian approach to person perception. *Consciousness and cognition* 36:406-413.
- Crouzet SM, Thorpe SJ (2011) Low-level cues and ultra-fast face detection. *Front Psychol* 2:342.
- Crouzet SM, Kirchner H, Thorpe SJ (2010) Fast saccades toward faces: face detection in just 100 ms. *J Vis* 10:16 11-17.
- Dakin SC, Watt RJ (2009) Biological "bar codes" in human faces. *Journal of Vision* 9:2 1-10.
- de Heering A, Rossion B (2015) Rapid categorization of natural face images in the infant right hemisphere. *Elife* 4.
- de Schonen S, Mathivet E (1990) Hemispheric asymmetry in a face discrimination task in infants. *Child Dev* 61:1192-1205.
- Fabre-Thorpe M (2011) The characteristics and limits of rapid visual categorization. *Front Psychol* 2:243.
- Fahrenfort JJ, Scholte HS, Lamme VA (2007) Masking disrupts reentrant processing in human visual cortex. *J Cogn Neurosci* 19:1488-1497.
- Farroni T, Johnson MH, Menon E, Zulian L, Faraguna D, Csibra G (2005) Newborns' preference for face-relevant stimuli: effects of contrast polarity. *Proc Natl Acad Sci U S A* 102:17245-17250.
- Fisch L, Privman E, Ramot M, Harel M, Nir Y, Kipervasser S, Andelman F, Neufeld MY, Kramer U, Fried I, Malach R (2009) Neural "ignition": enhanced activation linked to perceptual awareness in human ventral stream visual cortex. *Neuron* 64:562-574.

- Fleming RW, Torralba A, Adelson EH (2004) Specular reflections and the perception of shape. *J Vis* 4:798-820.
- Fletcher-Watson S, Findlay JM, Leekam SR, Benson V (2008) Rapid detection of person information in a naturalistic scene. *Perception* 37:571-583.
- Galper RE (1970) Recognition of faces in photographic negative. *Psychonomic Science* 19:207-208.
- Galper RE, Hochberg J (1971) Recognition memory for photographs of faces. *American Journal of Psychology* 84:351-354.
- Gao X, Gentile F, Rossion B (2018) Fast periodic stimulation (FPS): a highly effective approach in fMRI brain mapping. *Brain structure & function* 223:2433-2454.
- Gao X, Vuong QC, Rossion B (2019) The cortical face network of the prosopagnosic patient PS with fast periodic stimulation in fMRI. *Cortex* 119:528-542.
- Garrido L, Duchaine B, Nakayama K (2008) Face detection in normal and prosopagnosic individuals. *J Neuropsychol* 2:119-140.
- Geisler WS (2008) Visual perception and the statistical properties of natural scenes. *Annu Rev Psychol* 59:167-192.
- George N, Dolan RJ, Fink GR, Baylis GC, Russell C, Driver J (1999) Contrast polarity and face recognition in the human fusiform gyrus. *Nat Neurosci* 2:574-580.
- Gilad S, Meng M, Sinha P (2009) Role of ordinal contrast relationships in face encoding. *Proc Natl Acad Sci U S A* 106:5353-5358.
- Girard P, Hupe JM, Bullier J (2001) Feedforward and feedback connections between areas V1 and V2 of the monkey have similar rapid conduction velocities. *J Neurophysiol* 85:1328-1331.
- Girshick AR, Landy MS, Simoncelli EP (2011) Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nat Neurosci* 14:926-932.
- Goffaux V, Peters J, Haubrechts J, Schiltz C, Jansma B, Goebel R (2011) From coarse to fine? Spatial and temporal dynamics of cortical face processing. *Cereb Cortex* 21:467-476.
- Goren CC, Sarty M, Wu PY (1975) Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics* 56:544-549.
- Grill-Spector K, Kanwisher N (2005) Visual recognition: as soon as you know it is there, you know what it is. *Psychol Sci* 16:152-160.
- Hagen S, Jacques C, Maillard L, Colnat-Coulbois S, Rossion B, Jonas J (2020) Spatially Dissociated Intracerebral Maps for Face- and House-Selective Activity in the Human Ventral Occipito-Temporal Cortex. *Cereb Cortex* 30:4026-4043.
- Hauk O, Rice GE, Volfart A, Magnabosco F, Ralph MAL, Rossion B (2021) Face-selective responses in combined EEG/MEG recordings with fast periodic visual stimulation (FPVS). *NeuroImage* 242:118460.
- Haun AM, Peli E (2013) Perceived contrast in complex images. *J Vis* 13:3.
- Hershler O, Hochstein S (2005) At first sight: a high-level pop out effect for faces. *Vision Res* 45:1707-1724.
- Honey C, Kirchner H, VanRullen R (2008) Faces in the cloud: Fourier power spectrum biases ultrarapid face detection. *J Vis* 8:9 1-13.
- Hupe JM, James AC, Payne BR, Lomber SG, Girard P, Bullier J (1998) Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature* 394:784-787.
- Hupe JM, James AC, Girard P, Lomber SG, Payne BR, Bullier J (2001) Feedback connections act on the early part of the responses in monkey visual cortex. *J Neurophysiol* 85:134-145.
- Itier RJ, Latinus M, Taylor MJ (2006) Face, eye and object early processing: what is the face specificity? *NeuroImage* 29:667-676.

- Ito M, Tamura H, Fujita I, Tanaka K (1995) Size and position invariance of neuronal responses in monkey inferotemporal cortex. *J Neurophysiol* 73:218-226.
- Jonas J, Jacques C, Liu-Shuang J, Brissart H, Colnat-Coulbois S, Maillard L, Rossion B (2016) A face-selective ventral occipito-temporal map of the human brain with intracerebral potentials. *Proc Natl Acad Sci U S A* 113:E4088-4097.
- Kaplan E, Purpura K, Shapley RM (1987) Contrast affects the transmission of visual information through the mammalian lateral geniculate nucleus. *J Physiol* 391:267-288.
- Kar K, Kubilius J, Schmidt K, Issa EB, DiCarlo JJ (2019) Evidence that recurrent circuits are critical to the ventral stream's execution of core object recognition behavior. *Nat Neurosci* 22:974-983.
- Karimi-Rouzbahani H, Ramezani F, Woolgar A, Rich A, Ghodrati M (2021) Perceptual difficulty modulates the direction of information flow in familiar face recognition. *NeuroImage* 233:117896.
- Karkkainen K, Joo J (2021) FairFace: Face Attribute Dataset for Balanced Race, Gender, and Age for Bias Measurement and Mitigation. In: *IEEE/CVF Winter Conference on Applications of Computer Vision*, pp 1548-1558.
- Kauffmann L, Peyrin C, Chauvin A, Entzmann L, Breuil C, Guyader N (2019) Face perception influences the programming of eye movements. *Sci Rep* 9:560.
- Kemp R, Pike G, White P, Musselman A (1996) Perception and recognition of normal and negative faces: the role of shape from shading and pigmentation cues. *Perception* 25:37-52.
- Keysers C, Perrett DI (2002) Visual masking and RSVP reveal neural competition. *Trends Cogn Sci* 6:120-125.
- Keysers C, Xiao DK, Foldiak P, Perrett DI (2001) The speed of sight. *J Cogn Neurosci* 13:90-101.
- Keysers C, Xiao DK, Foldiak P, Perrett DI (2005) Out of sight but not out of mind: the neurophysiology of iconic memory in the superior temporal sulcus. *Cogn Neuropsychol* 22:316-332.
- Kovacs G, Cziraki C, Vidnyanszky Z, Schweinberger SR, Greenlee MW (2008) Position-specific and position-invariant face aftereffects reflect the adaptation of different cortical areas. *NeuroImage* 43:156-164.
- Lamme VA, Roelfsema PR (2000) The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci* 23:571-579.
- Lewis MB, Edmonds AJ (2003) Face detection: mapping human performance. *Perception* 32:903-920.
- Liu-Shuang J, Ales JM, Rossion B, Norcia AM (2015a) The effect of contrast polarity reversal on face detection: evidence of perceptual asymmetry from sweep VEP. *Vision Res* 108:8-19.
- Liu-Shuang J, Ales J, Rossion B, Norcia AM (2015b) Separable effects of inversion and contrast-reversal on face detection thresholds and response functions: a sweep VEP study. *J Vis* 15.
- Liu CH, Collin CA, Burton AM, Chaudhuri A (1999) Lighting direction affects recognition of untextured faces in photographic positive and negative. *Vision Res* 39:4003-4009.
- Long B, Yu CP, Konkle T (2018) Mid-level visual features underlie the high-level categorical organization of the ventral stream. *Proc Natl Acad Sci U S A* 115:E9015-E9024.
- Luck SJ (2005) *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: MIT Press.
- Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *Journal of neuroscience methods* 164:177-190.

- Martin JG, Davis CE, Riesenhuber M, Thorpe SJ (2018) Zapping 500 faces in less than 100 seconds: Evidence for extremely fast and sustained continuous visual search. *Sci Rep* 8:12482.
- Meister M, Berry MJ, 2nd (1999) The neural code of the retina. *Neuron* 22:435-450.
- Mohsenzadeh Y, Qin S, Cichy RM, Pantazis D (2018) Ultra-Rapid serial visual presentation reveals dynamics of feedforward and feedback processes in the ventral visual pathway. *Elife* 7.
- Murray SO, He S (2006) Contrast invariance in the human lateral occipital complex depends on attention. *Curr Biol* 16:606-611.
- Nasanen R, Ojanpaa H, Tanskanen T, Paallysaho J (2006) Estimation of temporal resolution of object identification in human vision. *Experimental brain research* 172:464-471.
- Nasr S, Tootell RB (2012) Role of fusiform and anterior temporal cortical areas in facial recognition. *NeuroImage* 63:1743-1753.
- Nederhouser M, Yue X, Mangini MC, Biederman I (2007) The deleterious effect of contrast reversal on recognition is unique to faces, not objects. *Vision Res* 47:2134-2142.
- Norcia AM, Appelbaum LG, Ales JM, Cottoreau BR, Rossion B (2015) The steady-state visual evoked potential in vision research: A review. *J Vis* 15:4.
- Nunez PL, Srinivasan R (2006) A theoretical basis for standing and traveling brain waves measured with human EEG with implications for an integrated consciousness. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology* 117:2424-2435.
- Ohayon S, Freiwald WA, Tsao DY (2012) What makes a cell face selective? The importance of contrast. *Neuron* 74:567-581.
- Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9:97-113.
- Oostenveld R, Praamstra P (2001) The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology* 112:713-719.
- Or CC, Retter TL, Rossion B (2019) The contribution of color information to rapid face categorization in natural scenes. *J Vis* 19:20.
- Or CC, Retter TL, Rossion B (2021) Does automatic human face categorization depend on head orientation? *Cortex* 141:94-111.
- Otsuka Y, Hill HC, Kanazawa S, Yamaguchi MK, Spehar B (2012) Perception of Mooney faces by young infants: the role of local feature visibility, contrast polarity, and motion. *J Exp Child Psychol* 111:164-179.
- Perrett DI, Xiao D, Barraclough NE, Keysers C, Oram MW (2009) Seeing the future: Natural image sequences produce "anticipatory" neuronal activity and bias perceptual report. *Q J Exp Psychol (Hove)* 62:2081-2104.
- Peykarjou S, Hoehl S, Pauen S, Rossion B (2017) Rapid Categorization of Human and Ape Faces in 9-Month-Old Infants Revealed by Fast Periodic Visual Stimulation. *Sci Rep* 7:12526.
- Pickard-Jones B, d'Avossa G, Sapir A (2020) 3D shape-from-shading relies on a light source prior that does not change with age. *Vision Res* 177:88-96.
- Potter MC, Levy EI (1969) Recognition memory for a rapid sequence of pictures. *J Exp Psychol* 81:10-15.
- Potter MC, Wyble B, Haggmann CE, McCourt ES (2014) Detecting meaning in RSVP at 13 ms per picture. *Atten Percept Psychophys* 76:270-279.
- Quek G, Nemdorov D, Rossion B, Liu-Shuang J (2018a) Selective Attention to Faces in a Rapid Visual Stream: Hemispheric Differences in Enhancement and Suppression of Category-selective Neural Activity. *J Cogn Neurosci* 30:393-410.

- Quek GL, Rossion B (2017) Category-selective human brain processes elicited in fast periodic visual stimulation streams are immune to temporal predictability. *Neuropsychologia* 104:182-200.
- Quek GL, Liu-Shuang J, Goffaux V, Rossion B (2018b) Ultra-coarse, single-glance human face detection in a dynamic visual stream. *Neuroimage* 176:465-476.
- Ramachandran VS (1988) Perceiving shape from shading. *Sci Am* 259:76-83.
- Regan D (1966) Some characteristics of average steady-state and transient responses evoked by modulated light. *Electroencephalography and Clinical Neurophysiology* 20:238-248.
- Rekow D, Baudouin JY, Brochard R, Rossion B, Leleu A (2022) Rapid neural categorization of facelike objects predicts the perceptual awareness of a face (face pareidolia). *Cognition* 222:105016.
- Rekow D, Baudouin JY, Poncet F, Damon F, Durand K, Schaal B, Rossion B, Leleu A (2021) Odor-driven face-like categorization in the human infant brain. *Proc Natl Acad Sci U S A* 118.
- Retter TL, Rossion B (2016) Uncovering the neural magnitude and spatio-temporal dynamics of natural image categorization in a fast visual stream. *Neuropsychologia* 91:9-28.
- Retter TL, Jiang F, Webster MA, Rossion B (2018) Dissociable effects of inter-stimulus interval and presentation duration on rapid face categorization. *Vision Res* 145:11-20.
- Retter TL, Jiang F, Webster MA, Rossion B (2020) All-or-none face categorization in the human brain. *NeuroImage* 213:116685.
- Robbins R, McKone E (2007) No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition* 103:34-79.
- Rolls ET, Baylis GC (1986) Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. *Experimental brain research* 65:38-48.
- Rosch E, Mervis CB (1975) Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology* 7:573-605.
- Rosch E, Mervis CB, Gray WD, Johnson DM, Boyes-Braem P (1976) Basic objects in natural categories. *Cognitive Psychology* 8:382-439.
- Rossion B, Lochy A (2022) Is human face recognition lateralized to the right hemisphere due to neural competition with left-lateralized visual word recognition? A critical review. *Brain structure & function* 227:599-629.
- Rossion B, Jacques C, Jonas J (2018) Mapping face categorization in the human ventral occipitotemporal cortex with direct neural intracranial recordings. *Ann N Y Acad Sci*.
- Rossion B, Torfs K, Jacques C, Liu-Shuang J (2015) Fast periodic presentation of natural images reveals a robust face-selective electrophysiological response in the human brain. *J Vis* 15:15 11 18.
- Rossion B, Prieto EA, Boremanse A, Kuefner D, Van Belle G (2012) A steady-state visual evoked potential approach to individual face perception: effect of inversion, contrast-reversal and temporal dynamics. *NeuroImage* 63:1585-1600.
- Rousselet GA, Mace MJ, Fabre-Thorpe M (2003) Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes. *J Vis* 3:440-455.
- Russell R, Sinha P, Biederman I, Niederhouser M (2006) Is pigmentation important for face recognition? Evidence from contrast negation. *Perception* 35:749-759.
- Simoncelli EP, Olshausen BA (2001) Natural image statistics and neural representation. *Annu Rev Neurosci* 24:1193-1216.
- Thorpe S, Fize D, Marlot C (1996) Speed of processing in the human visual system. *Nature* 381:520-522.
- Turati C, Simion F, Milani I, Umiltà C (2002) Newborns' preference for faces: what is crucial? *Dev Psychol* 38:875-882.

- Tzourio-Mazoyer N, De Schonen S, Crivello F, Reutter B, Aujard Y, Mazoyer B (2002) Neural correlates of woman face processing by 2-month-old infants. *NeuroImage* 15:454-461.
- Vuong QC, Peissig JJ, Harrison MC, Tarr MJ (2005) The role of surface pigmentation for recognition revealed by contrast reversal in faces and Greebles. *Vision Res* 45:1213-1223.
- Watt R (1994) A computational examination of image segmentation and the initial stages of human vision. *Perception* 23:383-398.
- White M (2001) Effect of photographic negation on matching the expressions and identities of faces. *Perception* 30:969-981.
- Woodman GF (2010) A brief introduction to the use of event-related potentials in studies of perception and attention. *Atten Percept Psychophys* 72:2031-2046.
- Wyatte D, Jilk DJ, O'Reilly RC (2014) Early recurrent feedback facilitates visual object recognition under challenging conditions. *Front Psychol* 5:674.
- Yang Z, Purves D (2004) The statistical structure of natural light patterns determines perceived light intensity. *Proc Natl Acad Sci U S A* 101:8745-8750.
- Yue X, Nasr S, Devaney KJ, Holt DJ, Tootell RB (2013) fMRI analysis of contrast polarity in face-selective cortex in humans and monkeys. *NeuroImage* 76:57-69.