Contextual modulations at low and high levels of visual perception: functional (in)dependence and influence of local contrast

Mehmet Umut Canoluk[1], Marius Grandjean[1], Valerie Goffaux[1,2]

1. Research Institute for Psychological Science (IPSY), UCLouvain, Louvain-la-Neuve, Belgium,

2. Institute of Neuroscience (IoNS), UCLouvain, Louvain-la-Neuve, Belgium

KEYWORDS

Contextual modulation, local contrast dependence, contrast detection, face perception, holistic processing, individual differences

Abstract

Spatial context modulates perception at low- and high levels of visual processing, with a shared dependence on the strength of the local input. In the case of human face perception, contextual modulations are presumed to rely on high-level, specialized mechanisms, with little contribution of low-level contextual mechanisms. Low- and high-level contextual modulations are typically studied independently, further obscuring their possible functional link. The present study will investigate whether consistent individual differences exist in contextual modulations at these two distinct levels of visual processing using contrast detection and facial feature discrimination tasks while manipulating the contrast (i.e., strength) of a local input target (local grating and facial feature, respectively). The context in the contrast detection task refers to a larger surrounding grating while in the facial feature discrimination task it refers to the facial features surrounding the target. The detection/discrimination contrast thresholds will provide a metric to evaluate contextual modulation magnitudes. By investigating the shared variance between the contextual modulation magnitudes across tasks, this study will provide insight into the functional relationship between low- and high-level contextual modulations and re-address the specificity of the contextual mechanisms at stake when processing faces in the light of low-level contextual modulations.

Introduction

Contextual modulations are found at various stages of the visual processing hierarchy, from retinal ganglion cells to IT cortex (e.g., Brincat & Connor, 2004; Cadieu et al., 2007; Desimone, 1991; Dobbins et al., 1987; Hansen & Neumann, 2008; Hubel & Wiesel, 1965; Levick et al., 1972; McIlwain, 1964; Nauhaus et al., 2009; Nolt et al., 2004; Sceniak et al., 1999; Solomon, 2006; Versavel et al., 1990). They modulate the processing of diverse stimulus aspects, ranging from simple properties such as contrast (Angelucci et al., 2002; Levitt & Lund, 1997; Nolt et al., 2004) or orientation (O'Toole & Wenderoth, 1977; Schwartz et al., 2007) to more complex characteristics such as perceived shape (Young et al., 1987; Thompson & Wilson, 2012; Schloss et al., 2014; Fang et al., 2008; Schwarzkopf et al., 2011), facial expression (Calder et al., 2000; Palermo et al., 2011; Tanaka et al., 2012) and face identity (Sergent, 1984; Young et al., 1987; Tanaka & Farah, 1993).

The processing of the human face, among other visual objects, is particularly susceptible to contextual modulations. Facial features are more accurately discriminated or recognized when they are part of a complete face (context) compared to when they are presented in isolation or in an atypical context, such as a scrambled face (Bruce, 1988; Davidoff & Donnelly, 1990; Sergent, 1984; Tanaka & Farah, 1993), resulting in the so-called whole-part advantage. Another manifestation of face contextual modulations is observed with the composite face and congruency paradigms. In these paradigms, the observer's sensitivity to a local feature decreases if it is incongruent with the surrounding context, in terms of either similarity status (when performing a discrimination task) or identity (when performing a recognition task) (Hole, 1994; Farah et al., 1998; Anaki et al., 2011; Goffaux, 2009, 2012; Richler & Gauthier, 2014). Face contextual modulations are hypothesized to enable the fast and efficient processing of the whole stimulus for an accurate identification despite faces' high visual similarity (e.g., DeGutis et al., 2013; Diamond & Carey, 1986; Leder & Bruce, 2000; McKone et al., 2007; Richler et al., 2011).

Contextual modulations with human faces are disrupted by simple and reversible image manipulations such as inversion (Sergent, 1984; Young et al., 1987; Farah et al., 1995; Yovel & Kanwisher, 2005). Inversion causes weaker activation in face-specialized regions while still recruiting a significant portion of non face-specialized high-level visual cortex (Haxby et al., 1999; Yovel & Kanwisher, 2005; Taubert et al., 2015). This has been taken to suggest that the face contextual, so-called holistic, mechanisms are implemented at high-level stages of visual processing specifically tuned to the natural statistics of the human face (i.e., the regularities in the way faces appear to us in everyday life; e.g., Schiltz & Rossion, 2006; Mazard et al., 2006; Andrews et al., 2010; Schiltz et al., 2010; Goffaux et al., 2013; Liu-Shuang et al., 2015).

However, empirical evidence challenges a purely high-level account for face contextual modulations. The strength of contextual modulations for faces depends on image properties processed at primary stages of visual processing, such as spatial frequency (Boutet et al., 2003; Goffaux & Rossion, 2006; Goffaux, 2009) and orientation (Dakin & Watt, 2009; Goffaux & Dakin, 2010). Moreover, face contextual modulations share similarities with low level mechanisms in their dependence on local input strength. Specifically, recent evidence indicates that the impact of the (face) context in a local feature discrimination task is inversely related to the physical difference of the features at stake (Goffaux, 2012; Goffaux et al., 2013; Canoluk et al., 2023).

This inverse relationship between contextual influence and local input strength is shared with low-level contextual mechanisms. For example, neurons in the primate visual cortex are more likely to integrate information across the visual field when the contrast of the local input is weak (Levitt & Lund, 1997; Angelucci et al., 2002, 2017). Similarly, the extent of spatial summation is larger with weaker local contrast in retina, lateral geniculate nucleus (LGN; Nolt et al., 2004) and primary visual cortex (V1; Sceniak et al., 1999), and local field potential (LFP) waves, which reflect post-synaptic lateral connectivity, are stronger in magnitude and travel a longer distance with weak contrast in cat and monkey primary visual cortex (Nauhaus et al., 2009). Face contextual modulations' reliance on low-level properties, as well as this shared dependence on local input strength raises the rarely asked question of whether the inter-individual variability in face contextual mechanisms can be explained by low-level contextual mechanisms to some extent.

In a recent study, we aimed to answer this question by measuring the amount of shared variance in the magnitude of contextual modulations for low-level contrast detection and morphed facial feature discrimination (in upright and inverted faces) by exploiting their shared dependence on local input strength (Canoluk et al., 2023). We obtained contrast detection thresholds for the low-level task and point of subjective equality (PSE) for discrimination of two subsequently displayed, morphed facial features (eyes) in matched contextual conditions using a single Bayesian Hierarchical Model. We found no evidence for the covariance of the contextual modulation magnitude in upright facial feature discrimination and contrast detection, which indicates they are governed by separate mechanisms. On the other hand, there was substantial evidence that contextual modulation magnitude weakly covaried between contrast detection and inverted facial feature discrimination, suggesting a relationship between low-level and high-level, non face-specialized contextual modulations. Lastly, there was decisive evidence for a strong correlation between upright and inverted face contextual modulations which can be explained by the overlapping mechanisms that these two kinds of stimuli recruit, especially when the depicted identities are unfamiliar to the participant (Megreya & Burton, 2006; Burton, 2013).

It is important to consider the fact that in Canoluk et al. (2023) local input strength was defined using a different property of the stimulus across the two tasks (contrast versus feature shape difference defined through morphing). This difference in the local input properties of the two tasks may have reduced the dependence of the contextual mechanisms recruited by facial feature discrimination and contrast detection. Past findings indeed showed that contextual illusions employing a similar local input property (e.g., two different size illusions) are more likely to correlate (Grzeczkowski et al., 2017; see also Haberman et al., 2015 for a similar comparison in ensemble perception). Furthermore, in the facial feature discrimination task, the target (eye region) was displayed foveally, at ~1.3° eccentricity whereas the targets appeared at 5° of eccentricity in the contrast detection task. Past evidence indicates that contextual influences across fovea and periphery rely on non-overlapping mechanisms (Xing & Heeger, 2000). Therefore, we reasoned that the differences in the employed local input property and foveal position of the targets could be the reason that we could not observe a dependence between contrast detection and facial feature discrimination tasks.

In this study, we aim to re-address the existence of a functional link between low- and high-level contextual mechanisms by circumventing possible interferences caused by these differences across tasks in the previous study. First, we will manipulate contrast at the eye region for the facial feature discrimination task. By doing so, the two tasks will have the same local input property. By measuring the magnitude of face contextual modulations at varying degrees of contrast, we will investigate whether contextual modulations exhibit a dependence on local contrast during face processing like they do in basic visual tasks. In the contrast detection task, the targets will be displayed closer to fovea, aiming to tap into the foveal contextual mechanisms as in the facial feature discrimination task. The comparison of the discrimination/detection contrast thresholds across different context conditions will provide a metric of contextual modulation magnitude. A correlation of contextual modulation magnitudes between upright facial feature discrimination and contrast detection would demonstrate how much of the inter-subject variability in face contextual modulations can be explained by low-level contextual mechanisms (H1). An absence of correlation would bring further support to the specialization and independence of face contextual mechanisms. Second, we will investigate the correlation between inverted facial feature discrimination and contrast detection to see how the partial disruption of the face-specialized mechanisms caused by inversion influences the amount of shared variance with low-level mechanisms (H2). Finally, we will investigate the correlation between upright facial feature discrimination and inverted facial feature discrimination contextual modulations, in order to quantify the functional overlap of the contextual mechanisms involved for these two stimulus types (H3).

Methods

a. Participants

In this study, we will recruit participants from the young adult population, specifically targeting individuals between the ages of 18 and 35. Prior to the experiment, participants' handedness will be measured using Edinburgh Handedness Inventory (EHI; Oldfield, 1971), and their visual acuity will be tested with the Freiburg Visual Acuity Test (FrACT; Bach, 1996) to ensure they have normal or corrected-to-normal vision (LogMAR ≤ 0.0). Participants with a LogMAR value larger than 0 will be excluded and will be replaced with new participants. Additionally, facial recognition abilities will be tested using a computerized version of the Benton Facial Recognition Test (BFRT; Benton & Van Allen, 1968; Rossion & Michel, 2018), to ensure their face recognition performance is in the normal range (i.e., above the borderline score of 39/54; (Duchaine & Weidenfeld, 2003). The experimental protocol was approved by the ethical committee of the UCLouvain (approval number: 2016/13SEP/393). The participants who fail to reach the required performance level in the practice (80% accuracy) will be removed from the experiment and new participants will be added until the desirable sample size is reached. Additional exclusion criteria are described in the Analyses section.

To determine the necessary sample size, we performed an a priori Bayes Factor (BF) design analysis (Schönbrodt & Wagenmakers, 2018). For data collection, we decided to opt for a sequential design with a maximal sample size, which entails that data be gathered until enough evidence has been accumulated or the predefined maximum number of participants has been tested. Thus, we will continuously calculate the BF of the correlations between the contextual effects in the two tasks: facial feature discrimination (upright/inverted), and the contrast detection (for more details, see the Data Analysis section). Such a sequential design requires one to specify the BF_{10} boundaries, the maximal sample size (N_{max}), the sample size at which the BF_{10} is computed for the first time (N_{min}), and the number of participants to be added each time until N_{max} or one of the BF_{10} boundaries is reached ($N_{increment}$). We opted for the BF_{10} boundaries of 6 and 1/6, and N_{min} was set to 40 as in Schönbrodt & Wagenmakers (2018). $N_{increment}$ was set to 5 for practical purposes (i.e., scheduling participants). N_{max} was determined through simulations with the BFDA package (Schönbrodt & Wagenmakers, 2018) in R (R Core Team, 2020). We chose an informative kappa prior of 0.7 for BFDA simulations. A kappa prior of 0.7 creates a prior distribution for the correlation simulations in which the distribution resembles the one of our previous study, with a mean of roughly $r = 0.3$, 95% confidence intervals ranging from .05 to .49. This reflects the weak significant correlation we found in our previous study between low-and high-level non face-specialized contextual modulations (Canoluk et al., 2023). We conducted simulations using 10,000 hypothetical datasets to examine sequential testing with the specified kappa prior. Two scenarios were simulated: one assuming a correlation of 0.3 and the other assuming no effect (i.e. the null hypothesis, with an effect size of $r = 0$). Our goal was to estimate the maximum sample size necessary to reach one of the Bayes Factor (BF) boundaries in 80% of the cases. This analysis was performed for both scenarios to compare the required sample sizes. We selected the maximum value between these two values as N_{max}, resulting in 130. Next, we evaluated the rates of false-negative and false-positive evidence using this N_{min}/N_{max} combination. Both rates were < 0.05, therefore we opted for an N_{min} of 40 and an N_{max} of 130.

In summary, with this design, we will start calculating our BF_{10} values at every 5 subjects starting from 40 subjects (N_{min}). We will terminate the data collection as soon as a certain BF_{10} boundary (1/6 or 6) for all three correlations (Upright Face vs. Grating, Inverted Face vs. Grating, Upright Face vs. Inverted Face), or N_{max} (130 participants) is reached.

b. Stimuli

In the facial feature discrimination task, we will present pairs of whole faces or isolated eye regions. The original stimulus set consists of 32 full-frontal grayscale pictures (on a scale from 0 to 1; 600 x 700 pixels) of unfamiliar white neutral faces from each gender without glasses, facial hair, or makeup. Face models are young adult students and alumni (aged 18–25 years) of the UCLouvain (Belgium), who gave written consent for the use of their image (Laguesse et al., 2012).

The luminance and root-mean-square (RMS) contrast of all face images are set to a mean of .52 and .12, respectively (using MatLab, Natick, MA). Next, we ensure that the eye regions of faces are comparable in terms of inter-pupillary distance and physical similarity. We measured inter-pupillary distance in each face image, and paired the faces with comparable values (average pixel difference of 1.32 ± 1.31 pixels). After this initial pairing, pixelwise similarity within the eye region was calculated within each pair using *corr2* function in MatLab. The pixelwise similarity of the target eye region was .8 on average (SD = .014). For each gender, the eight pairs lying within two standard deviations around the mean pixelwise similarity distribution are paired; the remaining eight pairs are only used as "face contexts". An additional set of 12 female faces (six pairs of female faces) which had undergone the same procedure are selected for practice trials. We extract the target region (eyes and brows) of each selected pair using a single aperture within the same pair (so the subjects could not rely on the shape of the aperture between two images within a pair) and blend it onto two "face contexts" whose eye areas were *a priori* replaced with a smooth skin texture (using Adobe Photoshop CS., Berkeley, CA), therefore all presented stimuli are edited, to ensure similarity across experimental trials.

Each face subtends a visual angle of ~7° in height, and ~5.2° in width. This size corresponds to the size of a human face seen at an average social distance (a human face is approximately of 18cm in height and social distance is of 1.5m approximately, e.g. McKone, 2009). The target eye region subtends a visual angle of ~1.4° in height, and ~4.1° in width. In this arrangement, the pupils are roughly at ~1.3° lateral eccentricity from the image center. Stimuli will be shown in their canonical Upright (0° rotation) and Inverted (180° rotation) position, and the face image will be offset 1.18° vertically to ensure that the target eye region is always at the center of the screen regardless of the orientation.

We additionally created noise masks by iteratively phase-scrambling each face image in Fourier space (in 500 iterations; e.g. Jacobs et al., 2020; Petras et al., 2019, 2021; Canoluk et al., 2023; Roux-Sibilon et al., 2023). Iterative scrambling is used to prevent noise masks from being contaminated by uniform background pixels, and thus ensures that the spectral properties of the face and mask are similar.

The preparation steps and all the settings of the face stimulus set described up until this point are identical to what we used in Canoluk et al. (2023), hence were already tested and validated. The main novelty in the current design is that the manipulation of local input strength is better matched across tasks. To that end, the contrast levels of the target regions (i.e., eyes and brows) will be manipulated with the use of a rectangular, grey occluder, covering 1.8° in height and 4.15° in width (see Figure **2)** and an adaptive staircase, as detailed in the following section. We opted for an occluder instead of directly manipulating the eye region contrast because it provides a more naturalistic view of the face (similar to seeing faces through sunglasses) as opposed to faces with low-contrasted eyes. Occluder's edges will have a thickness of .1° and a pre-defined luminance of .35.

The stimuli and procedure of the contrast detection task were adapted from Mannion et al. (2017) to make the stimulus appearance more similar to the stimuli of facial discrimination task. Stimuli will consist of a circular target grating (diameter of 0.75° instead of the original 3°) that will appear in one of two target regions (positioned at 1.5° lateral eccentricity, similar to eye targets instead of the original four target regions at 5° eccentricity) and a circular context grating (diameter of 20°). The target will always be a horizontal grating whose contrast will be defined by a Psi Marginal staircase (as detailed further in the following section), while the context will be made up of a horizontally (parallel-) or vertically (orthogonal+ to the target orientation) oriented grating with a contrast of 25%. The contrast of both the context and target regions will be ramped smoothly to zero at the edges with the use of a raised cosine filter. Additionally, a small fixation dot (diameter of 0.25°) will be present at the center of the screen before and after the stimulus presentation.

c. Procedure

The experiment consists of two different sessions that last ~45 minutes for each task. The order of the sessions will be counterbalanced across participants. Participants will perform both tasks seated in a dark room (whose walls are painted black to ensure the lux/cd ^ m2 are lowest when the doors are closed) with their head position restrained by a forehead/chin rest, at a distance of 57 cm from the computer monitor. Prior to starting, all participants will be dark adapted for at least 30 seconds. Stimuli will be displayed on a ViewPixx monitor with a spatial resolution of 1920 x 1200 pixels, a temporal resolution of 120 Hz, a bit depth of up to 12 bits (achieved through M16 mode of ViewPixx), and a maximum luminance of 27 cd/m^2. Monitor luminance levels are linearized using a ColorCAL MKII colorimeter (Cambridge Research Systems, Kent, UK). Luminance levels, gamma correction and the temporal resolution of the monitor will be periodically controlled to ensure the same values are kept constant throughout the experiment. Both tasks will be run on PsychoPy (2020.2.4post1; Peirce, 2007) and Python 3.8.

The facial feature discrimination task consists of six conditions: 2 Orientation (Upright, Inverted) and 3 Context (Congruent+, Incongruent-, Isolated). The term "context" in this task includes all the elements of the face apart from the targeted eyes and brows, such as the nose, mouth, cheeks and face outline. The target regions will be presented either out of face context (i.e., Isolated), in a Congruent+ context, where the context and the target in a pair call for a similar response (e.g., “same” targets displayed with “same” face contexts, or “different” targets displayed with “different” face contexts), or in an Incongruent- context, where the context and target call for opposing responses (e.g., “same” targets displayed with “different” face contexts, or vice versa; **see Figure 1**).

For both tasks, the contextual conditions were assigned designations of + (facilitatory) and – (suppressive) indicators as they align with the anticipated effects (if any) of the context in relation to the Isolated condition in accordance with existing literature.

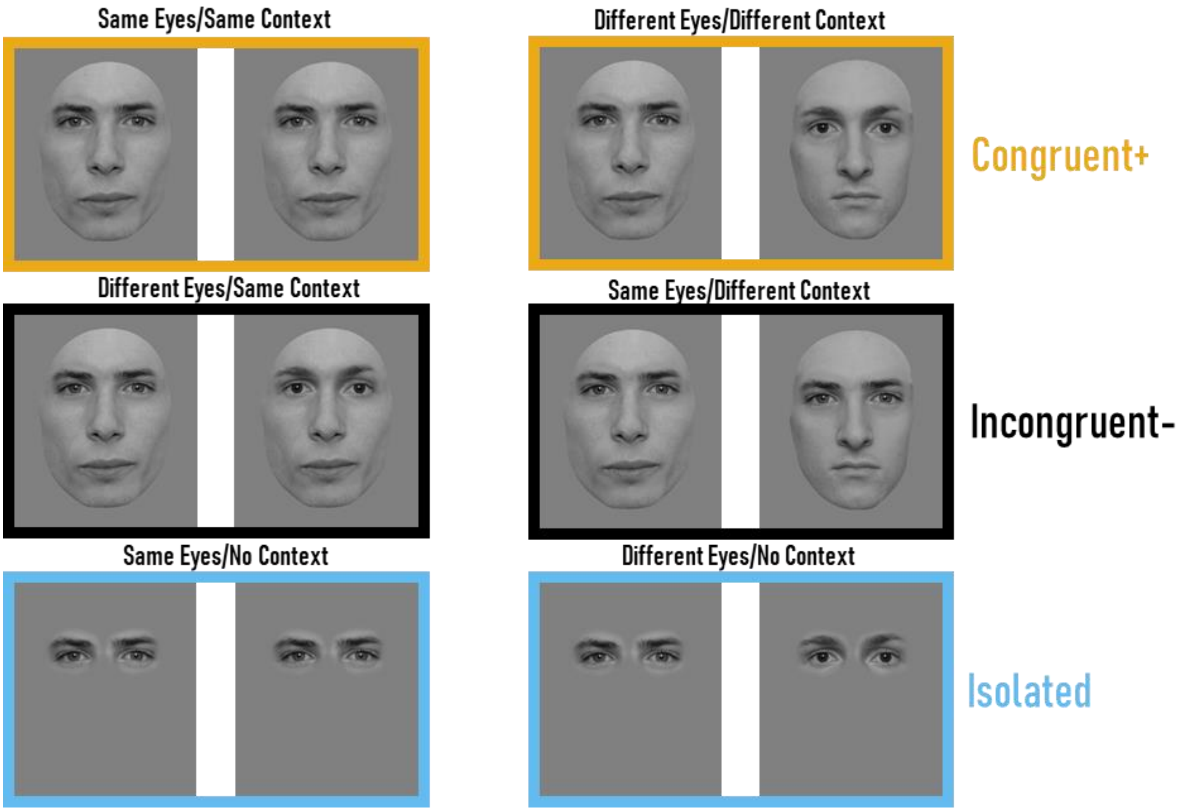

Figure 1. **Facial feature discrimination task contextual conditions.** The interaction of the context (face) and target (eyes) defines the conditions. In the congruent+ condition, the targets and the contexts of a pair are in accordance in terms of similarity (Same/Same or Different/Different) whereas there is a disagreement between them in the incongruent- condition (Same/Different or Different/Same). Target will be presented out of face context in the isolated condition. It can be seen here in Same Eyes/Different Context condition that context differences modulate the perception of the eyes, making them appear different when they are strictly identical (Young et al., 1987; Canoluk et al., 2023).

In each trial, the luminance of the occluder will be equalized to the average luminance of the region of the stimulus to be occluded. The occluder will be alpha-blended into the image. The amount of blending (i.e., transparency; e.g., 20% image, 80% occluder) will define the contrast (ranging from 2-40% contrast, or 98 - 60% occluder respectively) of the target region. Target contrast will be determined in each trial by two interleaved Psi Marginal adaptive staircases (Kontsevich & Tyler, 1999) for each condition. This interleaving procedure ensures that the algorithm does not settle into a narrow range of contrast levels within a condition.

The experimental session of the facial feature discrimination task will be divided into 16 blocks of 48 trials (768 trials in total), where the context conditions will be pseudo-randomized. Each block will comprise equal number of trials from each condition, therefore also equal number of trials that call for "Same" and "Different" responses in random order. Each trial will start with a fixation cross of 500 ms, followed by a blank screen of 500 ms. The first stimulus is then displayed for 500 ms, immediately followed by the phase-scrambled mask for 200 ms and the second stimulus will stay on the screen until the participant responds through keypress whether the eye regions of the first and second stimulus were same or different. The occluder will be overlaid on both the stimuli and the noise mask for the duration of a trial.

Prior to the experiment, participants will be required to do a practice run of 4 blocks (of 48 trials each) with 20% contrast in the eyes for the first two blocks and 10% contrast in the eyes for the second two blocks to familiarize with the task. They will be asked to re-do it until they reach 80% accuracy.

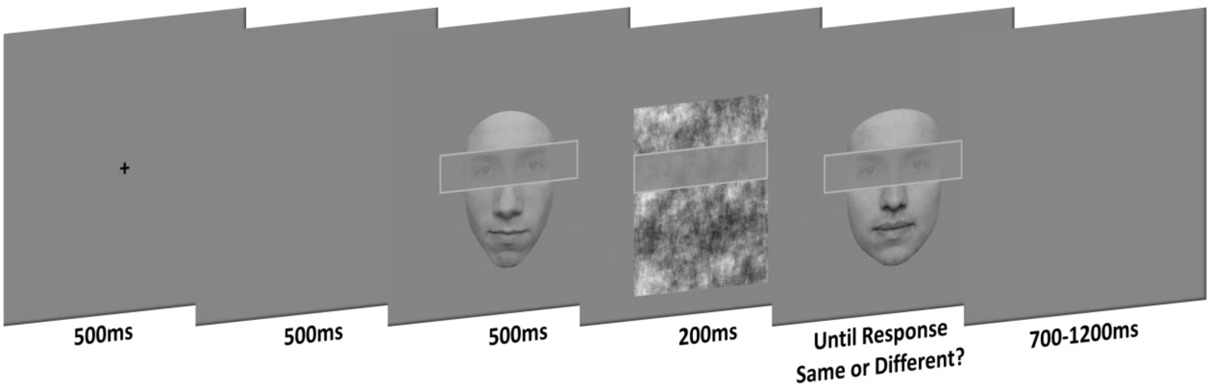

Figure 2. **Facial feature discrimination task trial sequence.** An example of an Incongruent- condition trial (different context, same eyes). Each trial starts with a fixation cross for 500 ms, followed by a blank screen. The first stimulus then appears for 500 ms, followed by the noise mask for 200 ms, and finally, by the second stimulus. The second stimulus stays on the screen until the participant indicates whether the target eye regions were "Same" or "Different" between the two stimuli by button press. This will be followed by a blank screen (randomly ranging between 700-1200 ms) before the next trial starts.

Similar to Mannion et al. (2017) and Canoluk et al. (2023), in the contrast detection task, the local target is a horizontal grating with the same spatial frequency (2 cycles per degree, and same phase as the context in the Parallel- condition) as the context, and with a variable, non-zero contrast will be presented in one of the two lateralized regions, within one of three context conditions: Isolated, Parallel- context, Orthogonal+ context (see **Figure 3**). The term "context" in this task refers to a surrounding, larger grating. Each context condition will be presented in separate blocks. Each block consists of 80 trials, with two staircases (as in the facial feature discrimination task) interleaved for 40 trials each. The contrast of the target grating will be determined by these staircases, which use a pool of 350 logarithmically spaced contrast values ranging from 0.001 to 1. There will be three block repetitions per condition, i.e. a total of 9 blocks (720 trials in total). Block order will be pseudo-randomized across participants, ensuring successive blocks belong to distinct context conditions. Each trial starts with a preparatory, black fixation dot of 500ms. Next, the local target (embedded in a context or presented in isolation, depending on the condition) is presented in one of the two possible locations for 100ms. After the stimulus presentation, the fixation dot turns to dark grey to signal the participant that they can give their response. We will ask the participant to report the side on which the target appeared (left or right). Before the experiment, participants will perform a short practice of 20 trials to familiarize with the task. No upper/lower limit is defined to pass the practice due to the simplicity of the task and instructions. Both tasks will emphasize accuracy over speed.

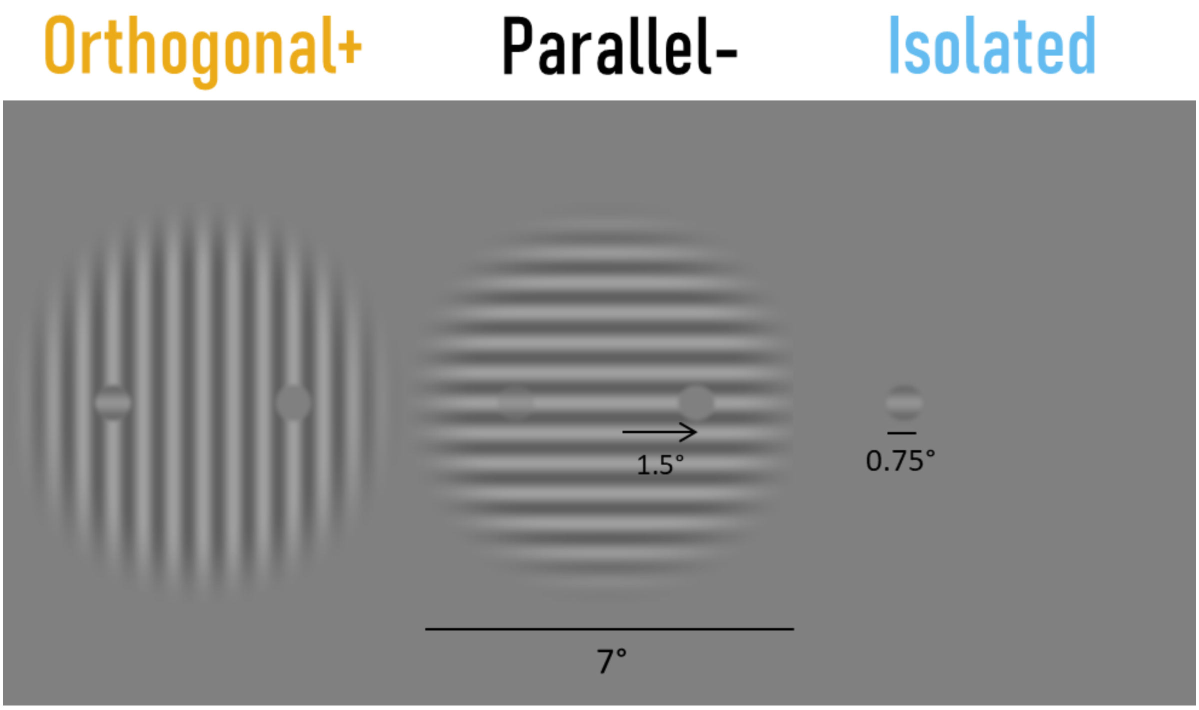

Figure 3. **Orthogonal+, Parallel- and Isolated context conditions of the contrast detection task, example of a left-lateralized local target.** It can be seen here that despite having the identical contrast value,

the target in the parallel- condition appears to have lower contrast because of the parallel context it is presented in.

d. Analyses

The present study aims at quantifying the variance of contextual modulation magnitude that is shared across tasks (i.e., processing levels). We will do so by comparing the detection/discrimination of each task thresholds in various context conditions. This approach largely replicates the one we successfully applied in Canoluk et al. (2023).

We will first log10 transform the contrast values from both tasks. Second, these values will be z-transformed for each task separately. After these transformations, contrast values will be within the range of -3 to 3 for both tasks. This normalization procedure on independent variables is a common practice when using complex hierarchical models, as it facilitates model convergence (Stan Development Team, 2023). Across tasks, we first investigate the function relating performance to local input strength (target eye region contrast or target grating contrast). Therefore, the z-scaled data from both tasks will be submitted to a Bayesian generalized linear mixed model (Yssaad-Fesselier & Knoblauch, 2006; Moscatelli et al., 2012) by means of *brms* package available in R (Bürkner, 2017, 2018). Using a mixed model makes it possible to include inter-individual variability in the statistical model where the estimates for the individuals are shrunk towards the general mean (due to the hierarchical nature of the model) based on trial-by-trial data, yielding richer (better informed) and more sensitive independent estimates and conferring more statistical power to the analysis compared to other analyses that rely on aggregated/averaged data (see Rouder & Haaf, 2019; Moors et al., 2020 for more details).

The dependent variable will be accuracy. The guess rate (i.e., lower asymptote of the psychometric function) is set at .5 for the contrast detection task (2-AFC, Knoblauch & Maloney, 2012) and the contrast value at .75 will be taken as “threshold”. Facial feature discrimination task being close to a yes-no task, so the guess rate will be set at 0 and the performance at .5 will be taken as “threshold” in this task (similar to Canoluk et al., 2023). Facial feature discrimination performance with Upright and Inverted faces will be treated separately. With this adjustment, the factor, “stimulus” has three different levels: “Upright Face”, “Inverted Face”, “Grating”. We will fit an inverse logit link function to model the sigmoidal pattern of the psychometric functions in each stimulus type (**Equation 1**). We put a normal distribution prior ($\mu = 0$, $\sigma = 5$) on all regression weights, and a lognormal distribution prior ($\mu = 0$, $\sigma = .1$) on the parameters capturing the random effects. All random effects will be included, yet will be assumed to be uncorrelated to each other (by means of || term in the formula) to increase sampling efficiency and ensure model convergence (Bürkner, 2018). Doing so, we allow a participant to deviate from the average, but we don't assume the size of these deviations to correlate across conditions (Kurz, 2019). The model will be fit with the following formula:

$$accuracy \sim .5 * \text{guess} + (1 - .5 * \text{guess}) * inv\_logit(\eta)$$

$$\eta \sim contrast * stimulus * context\ +\ (contrast * stimulus * context\ ||\ participant)$$

Equation 1. ***brms* formula to fit the sigmoidal functions.** The first formula implements the guess rate for both tasks (parameter "guess" is set to 1 for the contrast detection task, and 0 for the facial feature discrimination task, resulting in .5 and 0 guess rates respectively), to the inverse logit link function with the *η(eta)* hyperparameter, which is estimated by using the fixed (contrast, stimulus, context) and random (participant) effects described in the second formula.

Using these parameters and the formulae, we will run four Markov chain Monte Carlo (MCMC) with 6,000 iterations each (3,000 of each are considered warm-up), resulting in 12,000 iterations after warmup. Default values will be used for the *adapt_delta* and *max_treedepth* parameters, and they will be increased whenever it is deemed necessary for a better model fit. After fitting the data, we will run deviance tests as described in Wichmann and Hill (2001), to measure the goodness-of-fit of each individual psychometric function, in addition to the generic model diagnostics such as the R-hat value, effective sample size, and number of divergent transitions. The participants whose empirical deviance values of the psychometric functions are outside of the 95% confidence interval of the simulated deviance values will not be included in the analyses. Instead, they will be replaced with new participants until the maximum sample size is reached.

After fitting the data from both tasks, we will obtain the threshold value of the inverse logit link function for each stimulus and context condition, and for each participant using the formula, *threshold* = − *(intercept / slope)* as outlined in Knoblauch & Maloney (2012). This formula is used to transform intercept and slope values obtained from classical regression models to a threshold value. We will compare these threshold values across tasks and contexts using the model's threshold estimates, and the Highest Density Interval (HDI) of the differences in threshold comparisons. One approach in Bayesian statistics to interpret the magnitudes of these differences is to test whether the 95% HDI falls within the region of practical equivalence (ROPE). An HDI excluding the ROPE indicates that the difference is non-negligible, or states that the two sets of compared values are practically unlikely to be equivalent. An HDI including ROPE on the other hand suggests the values are likely equivalent and the difference between the two sets of values is practically zero (Kruschke, 2018). Specifying the ROPE limits is mostly arbitrary, but a general convention is to use ± 0.1 region around zero (Kruschke, 2018; Kelter, 2021) and we will follow this convention. In the facial feature discrimination task, we expect highest thresholds in the incongruent- condition than in the congruent+ condition (Outcome-Neutral Tests 1

(Upright Face) & 2 (Inverted Face); Richler et al., 2008; Goffaux, 2009; Meinhardt-Injac et al., 2014) and a smaller difference in thresholds between these conditions for inverted faces compared to upright faces (Outcome-Neutral Test 3), while in the contrast detection task, we expect the thresholds to be highest in the parallel- condition (Outcome-Neutral Test 4) as documented by the orientation-dependent contextual modulation effects (Mannion et al., 2017; Serrano-Pedraza et al., 2014). These differences will be used as confirmatory tools for replication of past findings and tests of soundness and feasibility of our methodology.

In order to quantify the contextual modulation magnitude of each participant, we will submit the individual threshold values from the three context conditions of each stimulus type to a two-step regression of regressions method (DeGutis et al., 2013; Canoluk et al., 2023). We chose to use this approach because the residuals of this two-step regression deliver an estimate of each individual contextual modulation that is not contaminated by the variance of the compared conditions, as opposed to simple subtraction method where the final outcome contains variance from both of the compared conditions. Each task comprises a control "Isolated" condition in which the target region (eyes and brows, or local grating) is presented on its own. In a first regression, we will regress participants' thresholds in the "Isolated" context condition from "+" context and "-" context thresholds, therefore removing all variance related to the control condition from the conditions where a context was present. In the second regression, we will regress the residuals of the "-" context condition from the residuals of the "+" context condition residuals. We confirmed in Canoluk et al. (2023) that this method returns sufficiently large intersubject variability, as well as good reliability.

Finally, we will conduct Bayesian correlation analyses on the final products of the two-step regressions from each stimulus type (Upright Face vs. Grating (H1), Inverted Face vs. Grating (H2), Upright Face vs. Inverted Face (H3)). We choose a Bayesian approach because of its ability to determine the strength of evidence presented by the data in supporting a particular hypothesis, hence, its potential to interpret a finding, even when it is null. We will compute Bayes Factors (BFs) of the correlation using BayesFactor package (Morey & Rouder, 2018) in R. For Bayesian correlation analyses, a uniform default prior (i.e., a stretched beta prior distribution with a width of 1, where any parameter value over the range of correlations between -1 and 1 are equally likely) can produce Bayes factors that can be biased towards null models (Vanpaemel, 2010; Quintana & Williams, 2018), so they are generally not recommended. Therefore, we will opt for a stretched beta prior distribution with a width of .33, centered at 0. Combining this prior distribution with the observed data forms the posterior distribution, and the Bayes factor is the ratio between the marginal likelihoods of the null model and the alternative model, given the data. A BF_{10} value indicates support for H_1 (e.g., correlation) over H_0 (e.g., no correlation) given the data (conversely, BF_{01} would indicate support for H_0 over H_1, and can be calculated with $1/BF_{10}$). We follow Jeffreys' guidelines (Jeffreys, 1998; Jarosz & Wiley, 2014) to interpret the BF values: a BF_{10} value of 1 would mean equal amount of evidence for both H_0 and H_1, and therefore inconclusive. A BF_{10} value between 1 and 3 is considered anecdotal evidence in favor of H_1 over H_0, whereas a BF_{10} between 3 and 10 indicates substantial evidence for H_1. $BF_{10} > 10$, $BF_{10} > 30$ and $BF_{10} > 100$ are regarded as strong, very strong and decisive evidence for H_1, respectively. We opted for evidence thresholds of $BF_{10} = 6$ (or 1/6) for concluding our analyses.

Table 1. Study Design Template

Question	Hypothesis	Sampling plan	Analysis Plan	Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis	Interpretation given different outcomes	Theory that could be shown wrong by the outcomes
To what extent are the low-level and high-level, face-specialized contextual mechanisms functionally independent?	Contextual modulation magnitudes in upright facial feature discrimination task and contrast detection task correlate. **(H1)**	BFDA - Sequential design with a maximal sample size	Bayesian correlation analysis of contextual modulation magnitudes between contrast detection and upright facial feature discrimination.	The BFDA simulations informed us that 80% of the samples stopped at (or before) a sample size of 130 (chosen as Nmax) because it reached either one of two boundaries (1/6 or 6) for both null or alternative hypotheses.	Evidence for correlation would support the functional dependence of the low-level and high-level face-specialized contextual mechanisms; Absence of a correlation would support their independence.	Full specificity of face contextual mechanisms.
Does inversion make the face contextual mechanisms more dependent on low-level contextual mechanisms?	Contextual modulation magnitudes in contrast detection task correlate more strongly with inverted than upright facial feature discrimination task. **(H2)**	Same as above	Bayesian correlation analysis of contextual modulation magnitudes between contrast detection and inverted facial feature discrimination.	Same as above	Evidence for a stronger correlation for inverted than upright facial feature discrimination task would support a stronger functional dependence to low-level contextual mechanisms; Similarly strong correlations would support the equal dependence of upright and inverted face contextual mechanisms on low-level contextual mechanisms.	Inversion disrupts the specificity of face contextual mechanisms.
To what extent are the high-level, face-specialized and high-level, non-face-specialized contextual mechanisms functionally independent?	Contextual modulation magnitudes in upright and inverted facial feature discrimination tasks correlate. **(H3)**	Same as above	Bayesian correlation analysis of contextual modulation magnitudes between upright facial feature discrimination and inverted facial feature discrimination.	Same as above	Evidence for correlation would support the functional dependence of the face-specialized and non-face-specialized contextual mechanisms. Absence of a correlation would support their independence.	Specificity of upright face processing.
Are upright faces processed contextually?	Discrimination thresholds in incongruent- trials are higher than the thresholds in congruent+ trials. **(Outcome Neutral test 1)**	Same as above	Assessing the overlap of the 95% HDI of threshold differences (congruent-incongruent) with zero.	Same as above	Evidence for a difference would demonstrate contextual processing in upright faces. Evidence for no difference would show part-based/local processing.	Upright faces are processed in a part-based manner.

Are inverted faces processed contextually?	Discrimination thresholds in incongruent- trials are higher than the thresholds in congruent+ trials. **(Outcome Neutral test 2)**	Same as above	Same as above	Same as above	Evidence for a difference would demonstrate contextual processing in inverted faces. Evidence for no difference would show part-based/local processing.	Inverted faces are processed in a part-based manner.
Are upright faces processed more contextually than inverted faces?	Smaller difference between incongruent- and congruent+ trials in inverted faces compared to upright faces. **(Outcome Neutral test 3)**	Same as above	Assessing the overlap of the 95% HDI of threshold differences (congruent-incongruent) between upright and inverted faces.	Same as above	Evidence for a bigger difference in upright faces would demonstrate stronger contextual effects for upright faces.	Upright faces are processed more contextually than inverted faces.
Does context orientation influence local target detection?	Detection thresholds in the parallel- trials are higher than the thresholds in orthogonal+ trials **(Outcome Neutral test 4)**	Same as above	Assessing the overlap of the 95% HDI of threshold differences ((parallel-)-(orthogonal+)) with zero.	Same as above	Evidence for a difference would demonstrate orientation-dependent contextual processing. Evidence for no difference would show that orientation of the surround has no effect on the local processing of gratings.	Context does not affect the processing of local contrast.

References

Anaki, D., Nica, E. I., & Moscovitch, M. (2011). Automatic Aspects in Face Perception: Evidence From Mandatory Processing of Distractor Facial Components. *Experimental Psychology*, *58*(1), 4–18. https://doi.org/10.1027/1618-3169/a000061

Andrews, T. J., Davies-Thompson, J., Kingstone, A., & Young, A. W. (2010). Internal and external features of the face are represented holistically in face-selective regions of visual cortex. *Journal of Neuroscience*, *30*(9), 3544–3552. https://doi.org/10.1523/JNEUROSCI.4863-09.2010

Angelucci, A., Bijanzadeh, M., Nurminen, L., Federer, F., Merlin, S., & Bressloff, P. C. (2017). Circuits and Mechanisms for Surround Modulation in Visual Cortex. *Annual Review of Neuroscience*, *40*(1), 425–451. https://doi.org/10.1146/annurev-neuro-072116-031418

Angelucci, A., Levitt, J. B., Walton, E. J. S., Hupé, J.-M., Bullier, J., & Lund, J. S. (2002). Circuits for local and global signal integration in primary visual cortex. *Journal of Neuroscience*, *22*(19), 8633–8646. https://doi.org/10.1523/JNEUROSCI.22-19-08633.2002

Bach, M. (1996). The Freiburg Visual Acuity Test—Automatic Measurement of Visual Acuity: *Optometry and Vision Science*, *73*(1), 49–53. https://doi.org/10.1097/00006324-199601000-00008

Benton, A. L., & Van Allen, M. W. (1968). Impairment in Facial Recognition in Patients with Cerebral Disease. *Cortex*, *4*(4), 344-IN1. https://doi.org/10.1016/S0010-9452(68)80018-8

Boutet, I., Collin, C., & Faubert, J. (2003). Configural face encoding and spatial frequency information. *Perception & Psychophysics*, *65*(7), 1078–1093. https://doi.org/10.3758/BF03194835

Brincat, S. L., & Connor, C. E. (2004). Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nature Neuroscience*, *7*(8), 880–886. https://doi.org/10.1038/nn1278

Bruce, V. (1988). *Recognising Faces* (1st ed.). Routledge. https://www.taylorfrancis.com/books/9781315471808

Bürkner, P.-C. (2017). **brms**: An *R* Package for Bayesian Multilevel Models Using *Stan*. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01

Bürkner, P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, *10*(1), 395. https://doi.org/10.32614/RJ-2018-017

Burton, A. M. (2013). Why has research in face recognition progressed so slowly? The importance of variability. *Quarterly Journal of Experimental Psychology*, *66*(8), 1467–1485. https://doi.org/10.1080/17470218.2013.800125

Cadieu, C., Kouh, M., Pasupathy, A., Connor, C. E., Riesenhuber, M., & Poggio, T. (2007). A Model of V4 Shape Selectivity and Invariance. *Journal of Neurophysiology*, *98*(3), 1733–1750. https://doi.org/10.1152/jn.01265.2006

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 527–551. https://doi.org/10.1037/0096-1523.26.2.527

Canoluk, M. U., Moors, P., & Goffaux, V. (2023). Contributions of low- and high-level contextual mechanisms to human face perception. *PLOS ONE*, *18*(5), e0285255. https://doi.org/10.1371/journal.pone.0285255

Dakin, S., & Watt, R. J. (2009). Biological “bar codes” in human faces. *Journal of Vision*, *9*(4), 1–10. https://doi.org/doi:10.1167/9.4.2

Davidoff, J., & Donnelly, N. (1990). Object superiority: A comparison of complete and part probes. *Acta Psychologica*, *73*(3), 225–243. https://doi.org/10.1016/0001-6918(90)90024-A

DeGutis, J., Wilmer, J., Mercado, R. J., & Cohan, S. (2013). Using regression to measure holistic face processing reveals a strong link with face recognition ability. *Cognition*, *126*(1), 87–100. https://doi.org/10.1016/j.cognition.2012.09.004

Desimone, R. (1991). Face-Selective Cells in the Temporal Cortex of Monkeys. *Journal of Cognitive Neuroscience*, *3*(1), 1–8. https://doi.org/10.1162/jocn.1991.3.1.1

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*(2), 107–117. https://doi.org/10.1037/0096-3445.115.2.107

Dobbins, A., Zucker, S. W., & Cynader, M. S. (1987). Endstopped neurons in the visual cortex as a substrate for calculating curvature. *Nature*, *329*(6138), 438–441. https://doi.org/10.1038/329438a0

Duchaine, B. C., & Weidenfeld, A. (2003). An evaluation of two commonly used tests of unfamiliar face recognition. *Neuropsychologia*, *41*(6), 713–720. https://doi.org/10.1016/S0028-3932(02)00222-1

Fang, F., Boyaci, H., Kersten, D., & Murray, S. O. (2008). Attention-Dependent Representation of a Size Illusion in Human V1. *Current Biology*, *18*(21), 1707–1712. https://doi.org/10.1016/j.cub.2008.09.025

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance*, *21*(3), 628–634. https://doi.org/10.1037/0096-1523.21.3.628

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka. (1998). What is “special” about face perception? *Psychological Review*, *105*(3), 482–498. https://doi.org/10.1037/0033-295X.105.3.482

Goffaux, V. (2009). Spatial interactions in upright and inverted faces: Re-exploration of spatial scale influence. *Vision Research*, *49*(7), 774–781. https://doi.org/10.1016/j.visres.2009.02.009

Goffaux, V. (2012). The discriminability of local cues determines the strength of holistic face processing. *Vision Research*, *64*, 17–22. https://doi.org/10.1016/j.visres.2012.04.022

Goffaux, V., & Dakin, S. (2010). Horizontal information drives the behavioral signatures of face processing. *Frontiers in Psychology*, *1*. https://doi.org/10.3389/fpsyg.2010.00143

Goffaux, V., & Rossion, B. (2006). Faces are “spatial”—Holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(4), 1023–1039. https://doi.org/10.1037/0096-1523.32.4.1023

Goffaux, V., Schiltz, C., Mur, M., & Goebel, R. (2013). Local Discriminability Determines the Strength of Holistic Processing for Faces in the Fusiform Face Area. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00604

Grzeczkowski, L., Clarke, A. M., Francis, G., Mast, F. W., & Herzog, M. H. (2017). About individual differences in vision. *Vision Research*, *141*, 282–292. https://doi.org/10.1016/j.visres.2016.10.006

Haberman, J., Brady, T. F., & Alvarez, G. A. (2015). Individual differences in ensemble perception reveal multiple, independent levels of ensemble representation. *Journal of Experimental Psychology: General*, *144*(2), 432–446. https://doi.org/10.1037/xge0000053

Hansen, T., & Neumann, H. (2008). A recurrent model of contour integration in primary visual cortex. *Journal of Vision*, *8*(8), 8–8. https://doi.org/10.1167/8.8.8

Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., & Martin, A. (1999). The Effect of Face Inversion on Activity in Human Neural Systems for Face and Object Perception. *Neuron*, *22*(1), 189–199. https://doi.org/10.1016/S0896-6273(00)80690-X

Hole, G. J. (1994). Configurational Factors in the Perception of Unfamiliar Faces. *Perception*, *23*(1), 65–74. https://doi.org/10.1068/p230065

Hubel, D. H., & Wiesel, T. N. (1965). Receptive Fields and Functioinal Architecture in Two Nonstriate Visual Areas (18 and 19) of the Cat. *Journal of Neurophysiology*, *28*(2), 229–289. https://doi.org/10.1152/jn.1965.28.2.229

Jacobs, C., Petras, K., Moors, P., & Goffaux, V. (2020). Contrast versus identity encoding in the face image follow distinct orientation selectivity profiles. *PLOS ONE*, *15*(3), e0229185. https://doi.org/10.1371/journal.pone.0229185

Jarosz, A. F., & Wiley, J. (2014). What Are the Odds? A Practical Guide to Computing and Reporting Bayes Factors. *The Journal of Problem Solving*, *7*(1). https://doi.org/10.7771/1932-6246.1167

Jeffreys, H. (1998). *Theory of probability* (3rd ed). Clarendon Press ; Oxford University Press.

Kelter, R. (2021). Bayesian Hodges-Lehmann tests for statistical equivalence in the two-sample setting: Power analysis, type I error rates and equivalence boundary selection in biomedical research. *BMC Medical Research Methodology*, *21*(1), 171. https://doi.org/10.1186/s12874-021-01341-7

Knoblauch, K., & Maloney, L. T. (2012). *Modeling psychophysical data in R*. Springer.

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, *39*(16), 2729–2737. https://doi.org/10.1016/S0042-6989(98)00285-5

Kruschke, J. K. (2018). Rejecting or Accepting Parameter Values in Bayesian Estimation. *Advances in Methods and Practices in Psychological Science*, *1*(2), 270–280. https://doi.org/10.1177/2515245918771304

Kurz, S. (2019). *Statistical Rethinking with brms, ggplot2, and the tidyverse*. (Retrieved from osf.io/97t6w).

Laguesse, R., Dormal, G., Biervoye, A., Kuefner, D., & Rossion, B. (2012). Extensive visual training in adulthood significantly reduces the face inversion effect. *Journal of Vision*, *12*(10), 14–14. https://doi.org/10.1167/12.10.14

Leder, H., & Bruce, V. (2000). When Inverted Faces are Recognized: The Role of Configural Information in Face Recognition. *The Quarterly Journal of Experimental Psychology Section A*, *53*(2), 513–536. https://doi.org/10.1080/713755889

Levick, W. R., Cleland, B. G., & Dubin, M. W. (1972). Lateral geniculate neurons of cat: Retinal inputs and physiology. *Investigative Ophthalmology*, *11*(5), 302–311.

Levitt, J. B., & Lund, J. S. (1997). Contrast dependence of contextual effects in primate visual cortex. *Nature*, *387*(6628), 73–76. https://doi.org/10.1038/387073a0

Liu-Shuang, J., Ales, J. M., Rossion, B., & Norcia, A. M. (2015). The effect of contrast polarity reversal on face detection: Evidence of perceptual asymmetry from sweep VEP. *Vision Research*, *108*, 8–19. https://doi.org/10.1016/j.visres.2015.01.001

Mannion, D. J., Donkin, C., & Whitford, T. J. (2017). No apparent influence of psychometrically-defined schizotypy on orientation-dependent contextual modulation of visual contrast detection. *PeerJ*, *5*, e2921. https://doi.org/10.7717/peerj.2921

Mazard, A., Schiltz, C., & Rossion, B. (2006). Recovery from adaptation to facial identity is larger for upright than inverted faces in the human occipito-temporal cortex. *Neuropsychologia*, *44*(6), 912–922. https://doi.org/10.1016/j.neuropsychologia.2005.08.015

McIlwain, J. T. (1964). RECEPTIVE FIELDS OF OPTIC TRACT AXONS AND LATERAL GENICULATE CELLS: PERIPHERAL EXTENT AND BARBITURATE SENSITIVITY. *Journal of Neurophysiology*, *27*(6), 1154–1173. https://doi.org/10.1152/jn.1964.27.6.1154

McKone, E. (2009). Holistic processing for faces operates over a wide range of sizes but is strongest at identification rather than conversational distances. *Vision Research*, *49*(2), 268–283. https://doi.org/10.1016/j.visres.2008.10.020

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences*, *11*(1), 8–15. https://doi.org/10.1016/j.tics.2006.11.002

Megreya, A. M., & Burton, A. M. (2006). Unfamiliar faces are not faces: Evidence from a matching task. *Memory & Cognition*, *34*(4), 865–876. https://doi.org/10.3758/BF03193433

Meinhardt-Injac, B., Persike, M., & Meinhardt, G. (2014). Holistic processing and reliance on global viewing strategies in older adults' face perception. *Acta Psychologica*, *151*, 155–163. https://doi.org/10.1016/j.actpsy.2014.06.001

Moors, P., Costa, T. L., & Wagemans, J. (2020). Configural superiority for varying contrast levels. *Attention, Perception, & Psychophysics*, *82*(3), 1355–1367. https://doi.org/10.3758/s13414-019-01917-y

Morey, R. D., & Rouder, J. N. (2018). *BayesFactor: Computation of Bayes Factors for Common Designs.* [Computer software]. https://CRAN.R-project.org/package=BayesFactor

Moscatelli, A., Mezzetti, M., & Lacquaniti, F. (2012). Modeling psychophysical data at the population-level: The generalized linear mixed model. *Journal of Vision*, *12*(11), 26–26. https://doi.org/10.1167/12.11.26

Nauhaus, I., Busse, L., Carandini, M., & Ringach, D. L. (2009). Stimulus contrast modulates functional connectivity in visual cortex. *Nature Neuroscience*, *12*(1), 70–76. https://doi.org/10.1038/nn.2232

Nolt, M. J., Kumbhani, R. D., & Palmer, L. A. (2004). Contrast-Dependent Spatial Summation in the Lateral Geniculate Nucleus and Retina of the Cat. *Journal of Neurophysiology*, *92*(3), 1708–1717. https://doi.org/10.1152/jn.00176.2004

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

O'Toole, B., & Wenderoth, P. (1977). The tilt illusion: Repulsion and attraction effects in the oblique meridian. *Vision Research*, *17*(3), 367–374. https://doi.org/10.1016/0042-6989(77)90025-6

Palermo, R., Willis, M. L., Rivolta, D., McKone, E., Wilson, C. E., & Calder, A. J. (2011). Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. *Neuropsychologia*, *49*(5), 1226–1235. https://doi.org/10.1016/j.neuropsychologia.2011.02.021

Petras, K., ten Oever, S., Dalal, S. S., & Goffaux, V. (2021). Information redundancy across spatial scales modulates early visual cortical processing. *NeuroImage*, *244*, 118613. https://doi.org/10.1016/j.neuroimage.2021.118613

Petras, K., ten Oever, S., Jacobs, C., & Goffaux, V. (2019). Coarse-to-fine information integration in human vision. *NeuroImage*, *186*, 103–112. https://doi.org/10.1016/j.neuroimage.2018.10.086

Quintana, D. S., & Williams, D. R. (2018). Bayesian alternatives for common null-hypothesis significance tests in psychiatry: A non-technical guide using JASP. *BMC Psychiatry*, *18*(1), 178. https://doi.org/10.1186/s12888-018-1761-4

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

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011). Holistic processing predicts face recognition. *Psychological Science*, *22*(4), 464–471. https://doi.org/10.1177/0956797611401753

Richler, J. J., & Gauthier, I. (2014). A meta-analysis and review of holistic face processing. *Psychological Bulletin*, *140*(5), 1281–1302. https://doi.org/10.1037/a0037004

Richler, J. J., Gauthier, I., Wenger, M. J., & Palmeri, T. J. (2008). Holistic processing of faces: Perceptual and decisional components. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(2), 328–342. https://doi.org/10.1037/0278-7393.34.2.328

Rossion, B., & Michel, C. (2018). Normative accuracy and response time data for the computerized Benton Facial Recognition Test (BFRT-c). *Behavior Research Methods*, *50*(6), 2442–2460. https://doi.org/10.3758/s13428-018-1023-x

Rouder, J. N., & Haaf, J. M. (2019). A psychometrics of individual differences in experimental tasks. *Psychonomic Bulletin & Review*, *26*(2), 452–467. https://doi.org/10.3758/s13423-018-1558-y

Roux-Sibilon, A., Peyrin, C., Greenwood, J. A., & Goffaux, V. (2023). Radial bias in face identification. *Proceedings of the Royal Society B: Biological Sciences*, *290*(2001), 20231118. https://doi.org/10.1098/rspb.2023.1118

Sceniak, M. P., Ringach, D. L., Hawken, M. J., & Shapley, R. (1999). Contrast's effect on spatial summation by macaque V1 neurons. *Nature Neuroscience*, *2*(8), 733–739. https://doi.org/10.1038/11197

Schiltz, C., Dricot, L., Goebel, R., & Rossion, B. (2010). Holistic perception of individual faces in the right middle fusiform gyrus as evidenced by the composite face illusion. *Journal of Vision*, *10*(2), 1–16. https://doi.org/10.1167/10.2.25

Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipito-temporal cortex. *NeuroImage*, *32*, 1385–1394. https://doi.org/10.1016/j.neuroimage.2006.05.037

Schloss, K. B., Fortenbaugh, F. C., & Palmer, S. E. (2014). The configural shape illusion. *Journal of Vision*, *14*(8), 23–23. https://doi.org/10.1167/14.8.23

Schönbrodt, F. D., & Wagenmakers, E.-J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, *25*(1), 128–142. https://doi.org/10.3758/s13423-017-1230-y

Schwartz, O., Hsu, A., & Dayan, P. (2007). Space and time in visual context. *Nature Reviews Neuroscience*, *8*, 522–535. https://doi.org/10.1038/nrn2155

Schwarzkopf, D. S., Song, C., & Rees, G. (2011). The surface area of human V1 predicts the subjective experience of object size. *Nature Neuroscience*, *14*(1), 28–30. https://doi.org/10.1038/nn.2706

Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology*, *75*(2), 221–242. https://doi.org/10.1111/j.2044-8295.1984.tb01895.x

Serrano-Pedraza, I., Romero-Ferreiro, V., Read, J. C. A., DiÃ©guez-Risco, T., Bagney, A., Caballero-GonzÃ¡lez, M., RodrÃguez-Torresano, J., & Rodriguez-Jimenez, R. (2014). Reduced visual surround suppression in schizophrenia shown by measuring contrast detection thresholds. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01431

Solomon, S. G. (2006). Suppressive Surrounds and Contrast Gain in Magnocellular-Pathway Retinal Ganglion Cells of Macaque. *Journal of Neuroscience*, *26*(34), 8715–8726. https://doi.org/10.1523/JNEUROSCI.0821-06.2006

Stan Development Team. (2023). *RStan: The R interface to Stan*. https://mc-stan.org/

Tanaka, & Farah, M. J. (1993). Parts and Wholes in Face Recognition. *The Quarterly Journal of Experimental Psychology Section A*, *46*(2), 225–245. https://doi.org/10.1080/14640749308401045

Tanaka, J. W., Kaiser, M. D., Butler, S., & Le Grand, R. (2012). Mixed emotions: Holistic and analytic perception of facial expressions. *Cognition & Emotion*, *26*(6), 961–977. https://doi.org/10.1080/02699931.2011.630933

Taubert, J., Van Belle, G., Vanduffel, W., Rossion, B., & Vogels, R. (2015). The effect of face inversion for neurons inside and outside fMRI-defined face-selective cortical regions. *Journal of Neurophysiology*, *113*(5), 1644–1655. https://doi.org/10.1152/jn.00700.2014

Thompson, P., & Wilson, J. (2012). Why Do Most Faces Look Thinner Upside Down? *I-Perception*, *3*(10), 765–774. https://doi.org/10.1068/i0554

Vanpaemel, W. (2010). Prior sensitivity in theory testing: An apologia for the Bayes factor. *Journal of Mathematical Psychology*, *54*(6), 491–498. https://doi.org/10.1016/j.jmp.2010.07.003

Versavel, M., Orban, G. A., & Lagae, L. (1990). Responses of visual cortical neurons to curved stimuli and chevrons. *Vision Research*, *30*(2), 235–248. https://doi.org/10.1016/0042-6989(90)90039-N

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, *63*(8), 1293–1313. https://doi.org/10.3758/BF03194544

Xing, J., & Heeger, D. J. (2000). Center-surround interactions in foveal and peripheral vision. *Vision Research*, *40*(22), 3065–3072. https://doi.org/10.1016/S0042-6989(00)00152-8

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configural information in face perception. *Perception*, *16*(6), 747–759. https://doi.org/10.1068/p160747

Yovel, G., & Kanwisher, N. (2005). The Neural Basis of the Behavioral Face-Inversion Effect. *Current Biology*, *15*(24), 2256–2262. https://doi.org/10.1016/j.cub.2005.10.072

Yssaad-Fesselier, R., & Knoblauch, K. (2006). Modeling psychometric functions in R. *Behavior Research Methods*, *38*(1), 28–41. https://doi.org/10.3758/BF03192747