Registered Report:

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

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

*Shared first authorship

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

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

Corresponding author :

Valerie 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

KEYWORDS

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

Abstract

What we see depends on context. From low to high levels of the visual system hierarchy, the amplitude of contextual influences is a function of local input strength. In the case of human face perception, the influence exerted by face context on local feature discrimination is determined by that feature discriminability. Face contextual modulations are presumed to rely on high-level, specialized contextual mechanisms, with little contribution of low-level mechanisms. The shared dependence of low- and high-level contextual modulations on local input strength however suggests they may be linked functionally. The present registered report addresses this gap by investigating whether consistent individual differences exist in contextual modulations across low-level contrast detection and high-level facial feature discrimination. Across tasks, we manipulated the contrast (i.e., strength) of a local input target (local grating and eye region, respectively) presented either in isolation or embedded in context. The context in the contrast detection task refers to a larger surrounding grating while it refers to the facial features surrounding the target eye region in the facial feature discrimination task. The comparison of contrast thresholds across conditions provided a metric to evaluate contextual modulation magnitudes. For both upright feature discrimination and contrast detection, context exerted a strong suppressive influence on local target processing; i.e., incongruent face context and parallel surround increased contrast threshold for feature discrimination and grating detection, respectively, compared to isolated viewing condition. As expected, face inversion drastically reduced the influence of context on local contrast threshold. We found no conclusive proof of a functional link between the magnitude of contextual modulation in local contrast detection and face feature matching; there was however extreme evidence for the correlation of contextual modulation magnitude in upright and inverted face feature matching. By investigating the shared variance between the contextual modulation magnitudes across visual processing levels, this study aimed at a system-wide understanding of the influence of spatial context on vision. Despite low- and high-level contextual modulations sharing similar operating principles, the present findings suggest they may involve independent mechanisms. We outline methodological approaches for further research development on this topic.

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, Zucker, & Cynader, 1987; Hansen & Neumann, 2008; Hubel & Wiesel, 1965; Levick et al., 1983; McIlwain, 1964; Nauhaus et al., 2009; Nolt, Kumbhani, & Palmer, 2004; Sceniak et al., 1999; Solomon, 2006; Versavel, Orban, & Lagae, 1990). They modulate the processing of diverse stimulus aspects, ranging from simple properties such as contrast (A. Angelucci et al., 2002; Levitt & Lund, 1997; Nolt, Kumbhani, & Palmer, 2004) or orientation (O'Toole & Wenderoth, 1977; Schwartz, Hsu, & Dayan, 2007) to more complex characteristics such as perceived shape (Fang et al., 2008; Schloss, Fortenbaugh, & Palmer, 2014; Schwarzkopf, Song, & Rees, 2011; Thompson & Wilson, 2012; Young, Hellawell, & Hay, 1987), facial expression (Calder et al., 2000; Palermo et al., 2011; Tanaka et al., 2012) and face identity (Sergent, 1984; Tanaka & Farah, 1993; Young, Hellawell, & Hay, 1987).

The processing of the human face, among other visual ~~objects~~categories, 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; Anaki, Nica, & Moscovitch, 2011; Farah et al., 1998; Goffaux, 2009, 2012; Hole, 1994; Richler & Gauthier, 2014). Face contextual modulations are hypothesized to enable the fast and efficient, so-called holistic, 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, Kanwisher, & Duchaine, 2007; Richler, Cheung, & Gauthier, 2011).

Contextual modulations with human faces are disrupted by simple and reversible image manipulations such as inversion (Farah, Tanaka, & Drain, 1995; Rossion, 2008; Sergent, 1984; Young, Hellawell, & Hay, 1987). I~~nversion~~ nversion largely preserves the physical properties of the face stimulus but drastically impacts high-level face processing: it ~~causes weaker~~ reduces activation in face-specialized regions ~~while still recruiting~~ and increases the contribution of a significant portion of non face-specialized high-level visual cortex (Haxby et al., 1999; Taubert et al., 2015; Yovel & Kanwisher, 2005). The vulnerability of face contextual modulations to inversion has therefore been taken to suggest that the face contextual 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, i.e., upright; e.g., Andrews et al., 2010; Goffaux et al., 2013; Liu-Shuang et al., 2015; Mazard, Schiltz, & Rossion, 2006; Schiltz et al., 2010; Schiltz & Rossion, 2006).

However, empirical evidence challenges a purely high-level account for face contextual modulations. First, tThe strength of contextual modulations for faces depends on image properties processed at primary stages of visual processing, such as spatial frequency (Boutet, Collin, & Faubert, 2003; Goffaux, 2009; Goffaux & Rossion, 2006) and orientation (Dakin & Watt, 2009; Goffaux & Dakin, 2010). Moreover, face contextual modulations share similarities with low- 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 (Canoluk, Moors, & Goffaux, 2023; Goffaux, 2012; Goffaux et al., 2013).

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 (Angelucci et al., 2017; Alessandra Angelucci et al., 2002; Levitt & Lund, 1997). Similarly, the extent of spatial summation is larger with weaker local contrast in retina, lateral geniculate nucleus (LGN; Nolt, Kumbhani, & Palmer, 2004) and primary visual cortex (V1; Sceniak et al., 1999), and l. 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 the 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, Moors, & Goffaux, 2023). We obtained contrast detection thresholds for the low-level task and point of subjective equality (PSE) for the 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-facenon-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 overlap in ping non-specialized

mechanisms that these two kinds of stimuli recruit, especially when the depicted identities are unfamiliar to the participant (Burton, 2013; Megreya & Burton, 2006).

Before reaching the conclusion that low-level and face-specialized contextual mechanisms are functionally independent, ~~I~~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~~ low-level task and high-level face task~~s~~ (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, Brady, & Alvarez, 2015 for a similar comparison for 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~~ manipulated contrast at the eye region for the facial feature discrimination task. By doing so, the low- and high-level ~~two~~ tasks ~~will have~~ involved the same local input property. By measuring the magnitude of face contextual modulations at varying degrees of contrast, we ~~will~~ investigated whether contextual modulations exhibit a dependence on local contrast during face processing like they do in basic visual tasks. ~~In~~ Moreover, in the contrast detection task, the targets ~~will be~~ were 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~~ provided 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~~ that the inter-subject variability in face contextual modulations ~~can be~~ is to some extent 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~~ investigated 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~~ investigated 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

Sixty-nine participants took part in the study (59 women, mean age 22.6 +/- 2.83 years, 10 left-handed), either volunteering for course credit or receiving financial compensation. Handedness was measured using the Edinburgh Handedness Inventory (EHI; Oldfield, 1971). Visual acuity was assessed with the Freiburg Visual Acuity Test (FrACT; Bach, 1996) before the experiment to ensure all participants had normal or corrected-to-normal vision. Four participants were excluded due to poor acuity (i.e., LogMAR value above 0; mean logMAR for the remaining participants: –0.10 ± 0.02). Additionally, facial recognition abilities were tested using a computerized version of the Benton Facial Recognition Test (BFRT; Benton & Van Allen, 1968; Rossion & Michel, 2018). Two participants performed below the borderline score of 39/54 and were excluded (Duchaine & Weidenfeld, 2003; ; mean score for remaining participants: 45.91 ± 3.64). The experimental protocol adhered to the Declaration of Helsinki and was approved by the ethical committee of the UCLouvain (approval number: 2016/13SEP/393). Three additional participants were removed due to technical issues. The remaining 60 participants did both tasks (53 women, mean age 22.5 ± 2.9, 10 left-handed).

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 calculated 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 (RCoreTeam, 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, Moors, & Goffaux, 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.

We deviated from the registered protocol and ended data collection at n = 60 owing to resource constraints, before the predetermined N_{max} or either evidential boundary (6, 1/6) was reached. Importantly, the Bayes factor remains a valid index of evidence irrespective of the stopping rule (Rouder, 2014).

b. Stimuli

In the facial feature discrimination task, we presented pairs of whole faces or isolated eye regions. The original stimulus set consisted 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 were 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 were set to a mean of .52 and .12, respectively (using MatLab, Natick, MA). Next, we ensured that the eye regions of faces were 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 were paired; the remaining eight pairs were only used as "face contexts". An additional set of 12 female faces (six pairs of female faces) which had undergone the same procedure were selected for practice trials. We extracted 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 blended 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 were edited, to ensure similarity across experimental trials.

Each face subtended 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 subtended a visual angle of ~1.4° in height, and ~4.1° in width. In this arrangement, the pupils were roughly at ~1.3° lateral eccentricity from the image center. Stimuli were shown in their canonical Upright (0° rotation) and Inverted (180° rotation) position, and the face image were offset 1.18° vertically to ensure that the target eye region was 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). Iterative scrambling was used to prevent noise masks from being contaminated by uniform background pixels and thus ensured 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 was that local input contrast was manipulated in both tasks. To that end, the contrast of the target regions (i.e., eyes and brows) was 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) compared to -non-occluded contrast variations between eye region and the rest of the face. Occluder's edges were 1°-thick and had a 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 consisted of a circular target horizontal grating (diameter of 0.75° instead of the original 3°) that appeared 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 used in Canoluk et al. (2023) embedded in a circular context grating (diameter of 20°). The contrast of the target local grating target ~~was~~ was defined by a Psi Marginal staircase (as detailed further in the following section), while the context was made up of a horizontally (~~parallel~~ context-) or vertically (~~orthogonal~~+context+ to the target orientation) oriented grating with a contrast of 25%. The contrast of both the context and target regions was ramped smoothly to zero at the edges with the use of a raised cosine filter. Additionally, a

small fixation dot (diameter of 0.25°) was presented at the center of the screen before and after the stimulus presentation.

c. Procedure

The experiment consisted of two different ~ 45 minute sessions ~~t~~ for each task (4 sessions in total). The order of the sessions was counterbalanced across participants. Participants performed 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 were dark adapted for at least 30 seconds. Stimuli were 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 were linearized using a ColorCAL MKII colorimeter (Cambridge Research Systems, Kent, UK). Luminance levels, gamma correction and the temporal resolution of the monitor were periodically controlled to ensure the same values are kept constant throughout the experiment. Both tasks were run on PsychoPy (2020.2.4post1; Peirce, 2007) and Python 3.8.

The facial feature discrimination task consisted of six conditions: 2 Orientation (Upright, Inverted) and 3 Context (~~Congruent+~~Context+, ~~Incongruent-~~Context-, 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 were presented either out of face context (i.e., Isolated), in a ~~Congruent+~~Context+ context, where the context and the target in a pair call for a similar response (i.e., “same” targets displayed with “same” face contexts, or “different” targets displayed with “different” face contexts), or in an ~~Incongruent-~~Context- context, where the context and target call for opposing responses (i.e., “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.

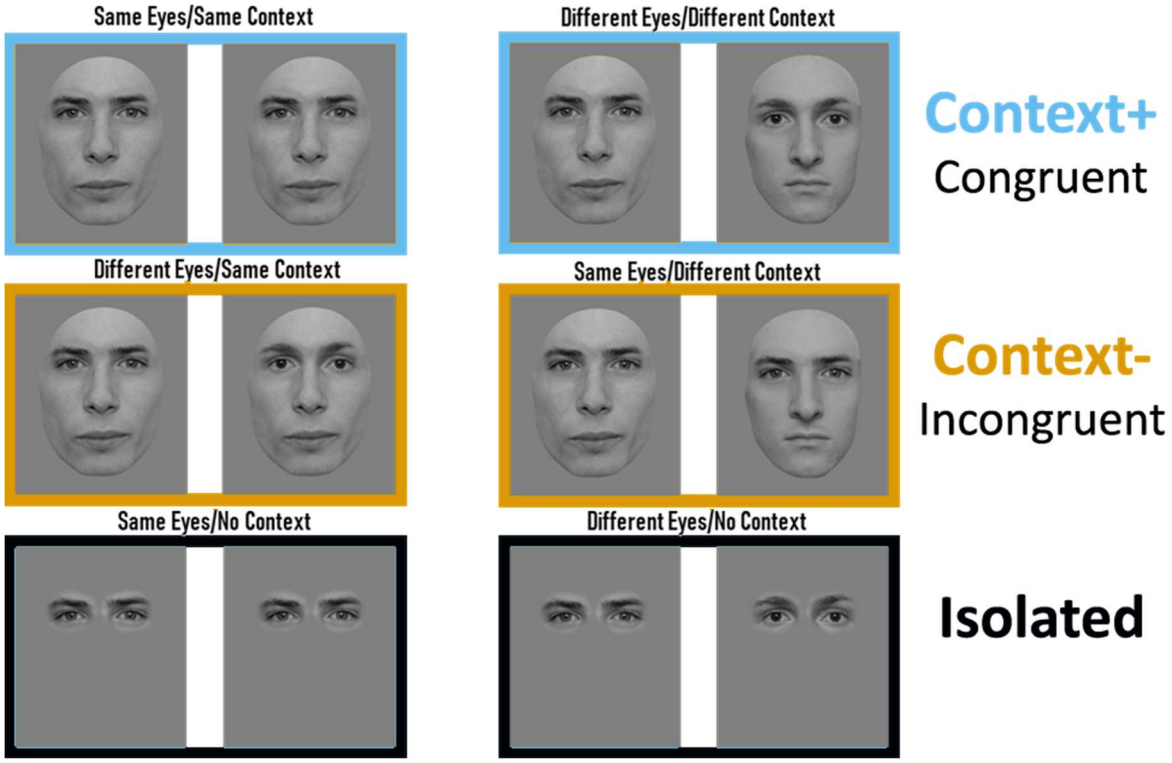


Figure 1. **Facial feature discrimination task contextual conditions.** The interaction of the context (face) and target (eyes) defines the conditions. In the ~~congruent+~~context+ condition, the targets and the contexts of a pair are in accordance in terms of similarity (Same/Same or Different/Different) whereas there is a disagreement between them in the ~~incongruent-~~context- condition (Same/Different or Different/Same). Target ~~will be~~ were presented out of face context in ~~the isola~~the ~~ted~~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 (e.g., Canoluk, Moors, & Goffaux, 2023; Young, Hellawell, & Hay, 1987).

In each trial, the luminance of the occluder was equalized to the average luminance of the region of the stimulus to be occluded. The occluder was alpha-blended into the image. The amount of blending (i.e., transparency; e.g., 20% image, 80% occluder) defined the contrast (ranging from 2-40% contrast, or 98 - 60% occluder respectively) of the target region. Target contrast was determined in each trial by two interleaved Psi Marginal adaptive staircases (Kontsevich & Tyler, 1999) for each condition. This interleaving procedure aimed at ~~avoiding that~~avoiding the algorithm settles into a narrow range of contrast levels within a condition.

The experimental session of the facial feature discrimination task was divided into 16 blocks of 48 trials (768 trials in total), where the context conditions were pseudo-randomized. Each block comprised 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 started with a fixation cross of 500 ms, followed

by a blank screen of 500 ms. The first stimulus was then displayed for 500 ms, immediately followed by the phase-scrambled mask for 200 ms and the second stimulus~~.~~ stayed on the screen until the participant responded through keypress whether the eye regions of the first and second stimulus were same or different. The occluder was overlaid on both the stimuli and the noise mask for the duration of a trial.

Prior to the experiment, participants were 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 themselves with the task. They were asked to re-do it until they reached 80% accuracy.

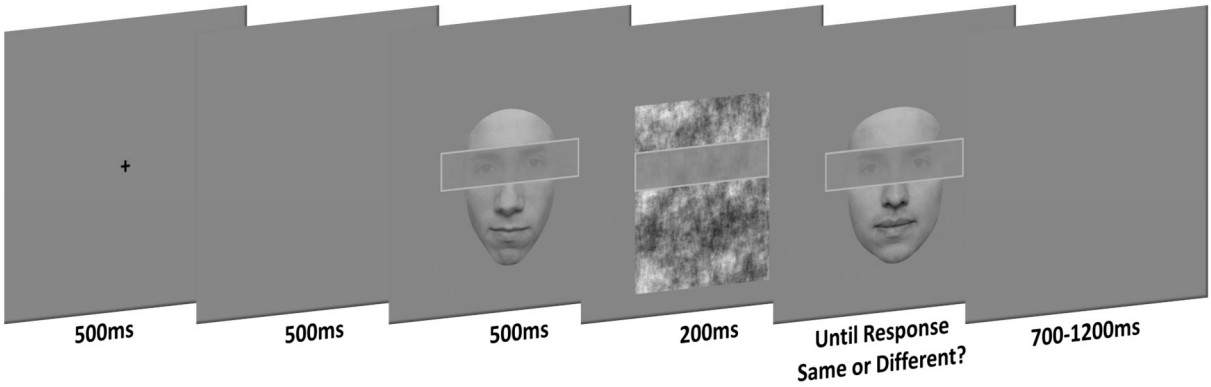


Figure 2. **Facial feature discrimination task trial sequence.** An example of an ~~Incongruent-~~Context- condition trial (different context, same eyes). Each trial started with a fixation cross for 500 ms, followed by a blank screen. The first stimulus then appeared for 500 ms, followed by the noise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

Similar to Mannion et al. (2017) and Canoluk et al. (2023), in the contrast detection task, the local target was a horizontal grating with the same spatial frequency (2 cycles per degree, and same phase as the context in the ~~Parallel-~~Context- condition) as the context, and with a variable, non-zero contrast was presented in one of the two lateralized regions, within one of three context conditions: Isolated, ~~Parallel-~~Context- context, ~~Orthogonal+~~Context+ context (see **Figure 3**). The term "context" in this task refers to a surrounding, larger grating. Each context condition was presented in separate blocks. Each block consisted of 80 trials, with two staircases (as in the facial feature discrimination task) interleaved for 40 trials each. The contrast of the target grating was determined by these staircases, which used a pool of 350 logarithmically spaced contrast values ranging from 0.001 to 1. There were three block

repetitions per condition, i.e. a total of 9 blocks (720 trials in total). Block order was pseudo-randomized across participants, ensuring successive blocks belong to distinct context conditions. Each trial started with a preparatory, black fixation dot of 500ms. Next, the local target (embedded in a context or presented in isolation, depending on the condition) was presented in one of the two possible locations for 100ms. After the stimulus presentation, the fixation dot turned to dark grey to signal the participant that they can give their response. We asked the participant to report the side on which the target appeared (left or right). Before the experiment, participants performed a short practice of 20 trials to familiarize with the task. No upper/lower limit was defined to pass the practice due to the simplicity of the task and instructions. Both tasks emphasized accuracy over speed.

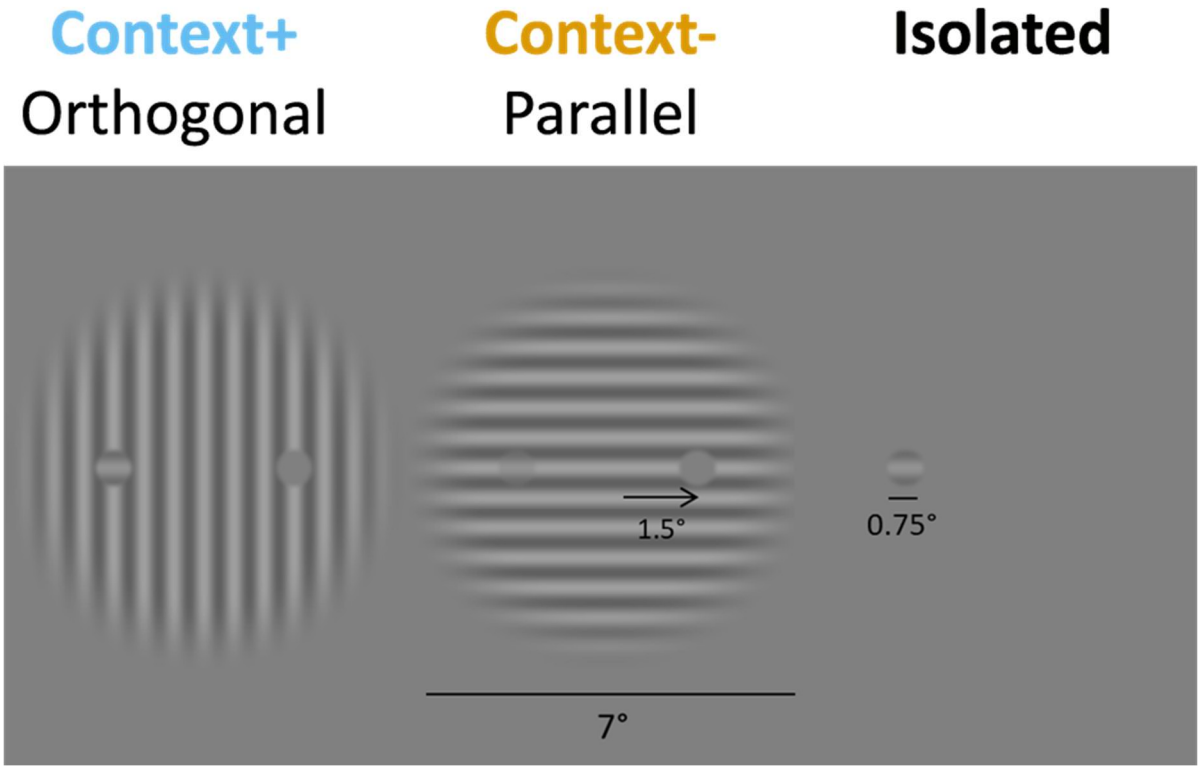


Figure 3. **Contrast detection task.** ~~Orthogonal+~~Context+, ~~Parallel-~~Context- and Isolated context conditions ~~**of the contrast detection task, example of**~~ for **a left-lateralized local target.** It can be seen here that despite having the identical contrast value, the target in the ~~parallel-~~context- condition appears to have lower contrast because of the parallel context it is presented in.

d. Analyses

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

Contrast values ~~We first~~ were log10 transformed ~~the contrast values from both tasks. Second, these values were~~ and z-transformed for each task separately. After these transformations, contrast values were 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). Because the staircase procedure led to uneven sampling of contrast across participants and because performance of five participants fluctuated around a single contrast level, we had to deviate from the preregistered analyses and split the normalized contrast values into five quantile-based bins. Consequently, accuracy could no longer be represented as a binary outcome (0 or 1) at the single-trial level as we initially planned in the registered analyses~~-~~. Instead, each bin reflected a proportion of correct responses over multiple trials, a value between 0 and 1. To appropriately model this, we used a binomial likelihood, treating the observed number of correct responses as the number of "successes" out of a known number of trials per bin. This allowed us to estimate the underlying probability of a correct response (i.e., success) for each combination of contrast bin, task, and condition. The guess rate (i.e., lower asymptote of the psychometric function) was set at .5 for the contrast detection task (2-AFC; Knoblauch & Maloney, 2012) and the contrast value at .75 was 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) rather than bounded at chance. 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”.

Across tasks, we first investigated the function relating performance to local input strength (contrast of eye region or local grating). The binned z-scaled data from both tasks was submitted to a Bayesian generalized linear mixed model (Bayesian GLMM; Moscatelli, Mezzetti, & Lacquaniti, 2012; Yssaad-Fesselier & Knoblauch, 2006) 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/~~group-averaged data (see Moors, Costa, & Wagemans, 2020; Rouder & Haaf, 2019). The factor, “stimulus” has three different levels: “Upright Face”, “Inverted Face”, “Grating”. We- fitted 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 ($\mu = 0$, $\sigma = 2.5$) on all regression weights, and a lognormal distribution prior ($\mu = 0$, $\sigma = .2$) on the parameters capturing the random effects. All random effects were included, yet were 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 allowed a participant to deviate from the average, but we did not assume the size of these deviations to correlate across conditions (Kurz, 2019). The model was fitted with the following formula:

$$probability\ of\ success \sim .5 * \text{guess} + (1 - .5 * \text{guess}) * inv_logit(\eta)$$
$$\eta \sim contrast * stimulus * context\ +\ (contrast * stimulus * context\ ||\ participant)$$

Equation 1. ***brms*** **formula to fit the sigmoidal functions.** The probability of success is estimated from proportion of correct responses~~the number of correct responses out of the total number of trials,~~ within each contrast bin. 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 ran 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 were used for the *adapt_delta* and *max_treedepth* parameters. After fitting the data, we checked generic model diagnostics (i.e., R-hat value, effective sample size, and number of divergent transitions). We initially planned to run deviance tests as described in Wichmann & Hill, 2001), to measure the goodness-of-fit for each psychometric function. However, the staircase issues described above made this procedure inadequate for the facial feature discrimination task. In some staircases, participants alternated between correct and incorrect responses rather than improving systematically as contrast varied, so accuracy changed little across contrast levels and the fitted psychometric function was nearly flat (slope close to zero). Because flat data are themselves well described by a flat curve, these functions could yield a good model fit even though contrast did not predict accuracy as the staircase was designed to ensure, which in turn produced extremely high or undefined thresholds. To prevent extreme values from skewing our group-level analyses, we used the Interquartile Range method (IQR, Hoaglin, Iglewicz, & Tukey, 1986) to identify and exclude participants whose calculated thresholds were statistical outliers from the group distribution.

After fitting the data from both tasks, we obtained 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). At this value the inverse-logit link reaches its midpoint; combined with the task-specific guess rates, the threshold therefore corresponds to .75 proportion correct in the contrast detection task and .50 in the facial feature discrimination task.

We compared 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 lying entirely outside the ROPE indicates that the difference is credibly non-negligible. Conversely, an HDI completely contained within the ROPE indicates that the difference is practically negligible (equivalent to zero). Partial overlap between the HDI and the ROPE indicates an inconclusive result (Kruschke, 2018). Specifying the ROPE limits is mostly arbitrary, but a general convention is to use 0 ± 10% of the posterior standard deviation region (Kelter, 2021; Kruschke, 2018) and we followed this convention. In the facial feature discrimination task, we expected highest thresholds in the ~~incongruent~~ context- condition than in the ~~congruent+~~context+ condition (Outcome-Neutral Tests 1 (Upright Face) & 2 (Inverted Face); Goffaux, 2009; Meinhardt, Meinhardt-Injac, & Persike, 2014; Richler et al., 2008) 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 expected the thresholds to be highest in the ~~parallel-~~context- condition (Outcome-Neutral Test 4) as documented by the orientation-dependent contextual modulation effects (Mannion, Donkin, & Whitford, 2017; Serrano-Pedraza et al., 2014). These differences were 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 submitted the individual threshold values from the three context conditions of each stimulus type to a two-step regression of regressions method (Canoluk, Moors, & Goffaux, 2023; DeGutis et al., 2013). 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~~ regressed condition~~s~~, as opposed to ~~simple~~ subtraction method ~~where the final outcome contains variance from both of the compared conditions~~. Each task comprised a control "Isolated" condition in which the target region (eyes and brows, or local grating) was presented on its own. In a first regression, we regressed participants' thresholds in the "Isolated" context condition from "+" context and "-" context thresholds, therefore removing ~~all~~ the variance related to the control condition from the conditions where a context was present. In the second regression, we regressed 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 conducted 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 chose a Bayesian approach because Bayes factors quantify the relative evidence provided by the data for one hypothesis over another, allowing evidence in favor of both the alternative and null hypotheses to be quantified.

We computed 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) is not recommended because it can yield Bayes factors that can be biased towards null models (Quintana & Williams, 2018; Vanpaemel, 2010), so they are generally not recommended. Therefore, we opted for a stretched beta prior distribution with a width of .33, centered at 0. The Bayes factor BF_{10} is the ratio of the marginal likelihood of the data under the alternative model to that under the null model. 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 followed Jeffreys' guidelines (Jeffreys, 1998; Jarosz & Wiley, 2014) to interpret the BF values: a BF_{10} of 1 indicates no evidence H_1 over H_0, values between 1 and 3 indicate anecdotal evidence in favor of H_1 over H_0, whereas values between 3 and 10 indicate substantial evidence. Values between 10 and 30, 30 and 100, and greater than 100 indicate strong, very strong, and extreme evidence for H_1, respectively.

e. Results

Contrast threshold comparisons

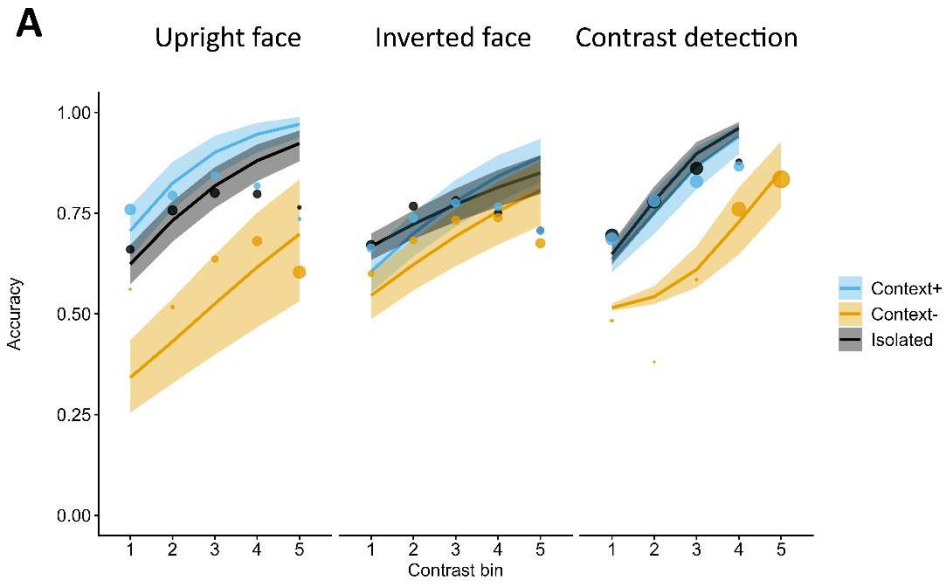
In the contrast detection task, our results replicated the orientation-dependent surround-suppression (Figure 4; **Outcome Neutral Test 4**; Mannion et al., 2017; Serrano-Pedraza, Romero-Ferreiro, Read, Diéguez-Risco, et al., 2014). Using a ROPE of ± 0.19, thresholds were credibly higher in the context-condition compared to the context+ condition (median $= 2.13$, 95% HDI $[1.95, 2.30]$; HDI entirely outside the ROPE). Thresholds were also credibly higher in the context- condition relative to the isolated condition (median $= 2.36$, 95% HDI $[2.21, 2.50]$; HDI entirely outside the ROPE). By contrast, the difference between the context+ and isolated conditions was inconclusive (median $= 0.24$, 95% HDI $[0.06, 0.41]$; HDI overlapping the ROPE). These findings confirm the suppressive influence of contextual modulation on local contrast detection.

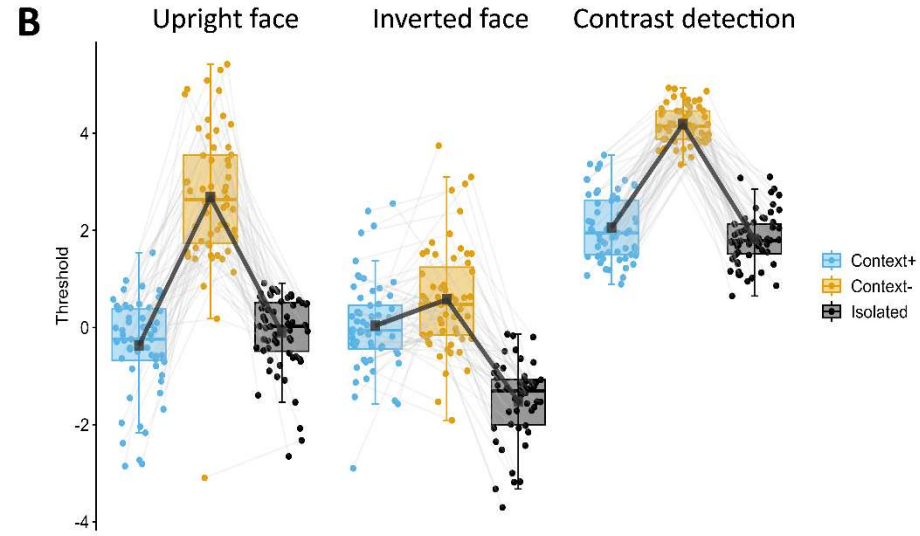
Upright face feature discrimination depicted a similar profile of contrast threshold modulations across context conditions. Using a ROPE of ± 0.23, thresholds were credibly higher (i.e., more contrast was required to discriminate local target features) in the context- condition than in the context+ condition (median $= 3.04$, 95% HDI $[2.66, 3.42]$; **Outcome-Neutral Test 1**; HDI entirely outside the ROPE; Figure 4). Thresholds in the context- condition were also credibly higher than in the isolated condition (median $= 2.75$, 95% HDI $[2.43, 3.10]$; HDI entirely outside the ROPE), whereas the difference between

the context+ and isolated conditions remained inconclusive (median $= -0.28$, 95% HDI $[-0.51, -0.05]$; HDI overlapping the ROPE; see Figure 4C). This again indicated a primarily suppressive influence of context upon local target processing in line with past reports that contextual incongruence impairs feature discrimination (Goffaux, 2009; Meinhardt-Injac, Persike, & Meinhardt, 2014; Richler et al., 2008).

For inverted faces, contextual modulations were also present (**Outcome-Neutral Test 2**), but their overall magnitude was smaller than for upright faces. Using a ROPE of ± 0.11, we observed a credible difference between the context- and context+ conditions, with higher thresholds for context- (median $= 0.56$, 95% HDI $[0.22, 0.87]$; HDI entirely outside the ROPE). Both contextual conditions showed credibly higher thresholds than the isolated condition (context- vs. isolated: median $= 2.12$, 95% HDI $[1.85, 2.38]$; context+ vs. isolated: median $= 1.56$, 95% HDI $[1.30, 1.82]$, with both HDI entirely outside the ROPE).

To assess whether upright faces are processed more contextually than inverted faces (**Outcome Neutral Test 3),** we compared the magnitude of the contextual effects across the two tasks. Using a ROPE of ± 0.18, the contextual effect was credibly larger for upright faces (median $= 3.04$, 95% HDI $[2.66, 3.42]$) than for inverted faces (median $= 0.56$, 95% HDI $[0.22, 0.87]$), as indicated by the posterior distribution of the difference between these two effects (median difference $= 2.48$, 95% HDI $[2.00, 2.95]$, HDI laying entirely outside the ROPE). These results confirm that contextual processing is substantially more pronounced for upright faces.


A
Upright face
Inverted face
Contrast detection
Accuracy
1.00
0.75
0.50
0.25
0.00
1 2 3 4 5
Contrast bin
Context+
Context-
Isolated


B
Upright face
Inverted face
Contrast detection
Threshold
4
2
0
-2
-4
Context+
Context-
Isolated

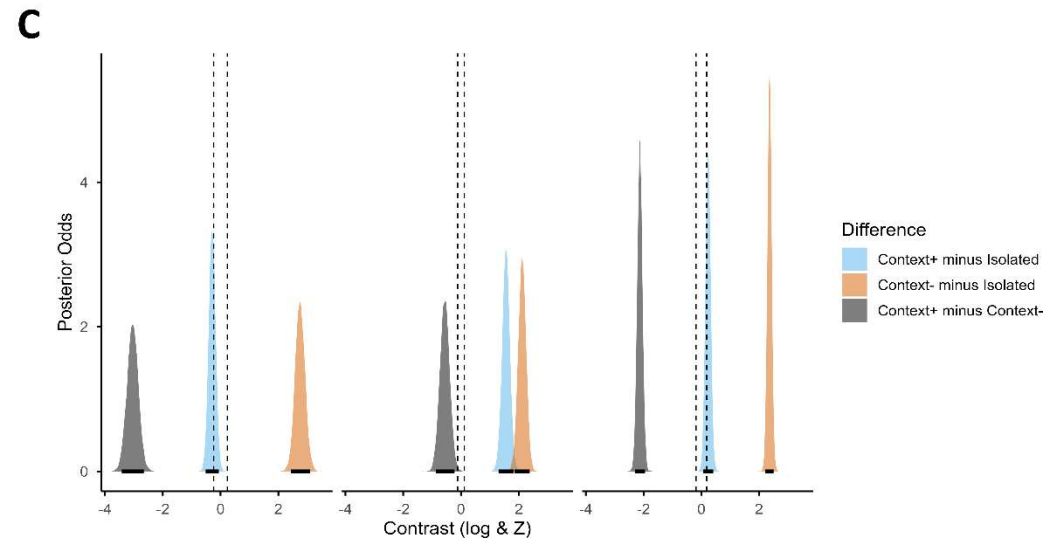
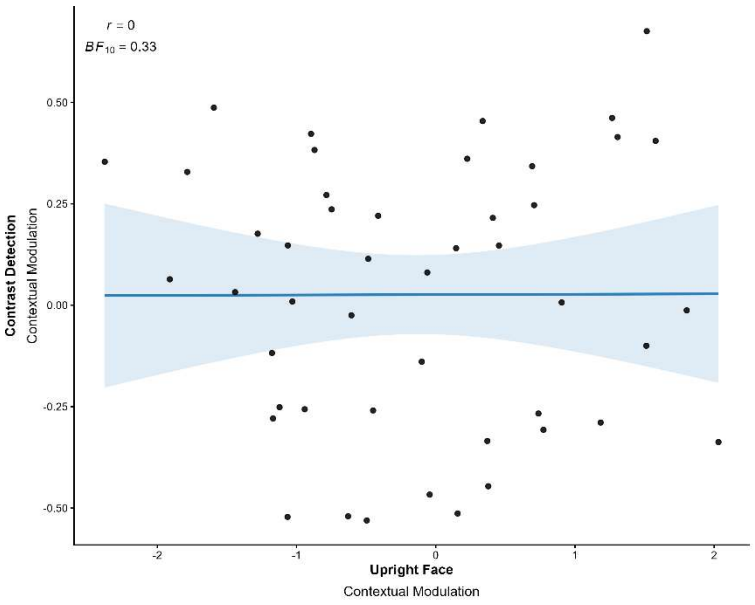

C
Posterior Odds
4
2
0
-4 -2 0 2
Contrast (log & Z)
Difference
Context+ minus Isolated
Context- minus Isolated
Context+ minus Context-

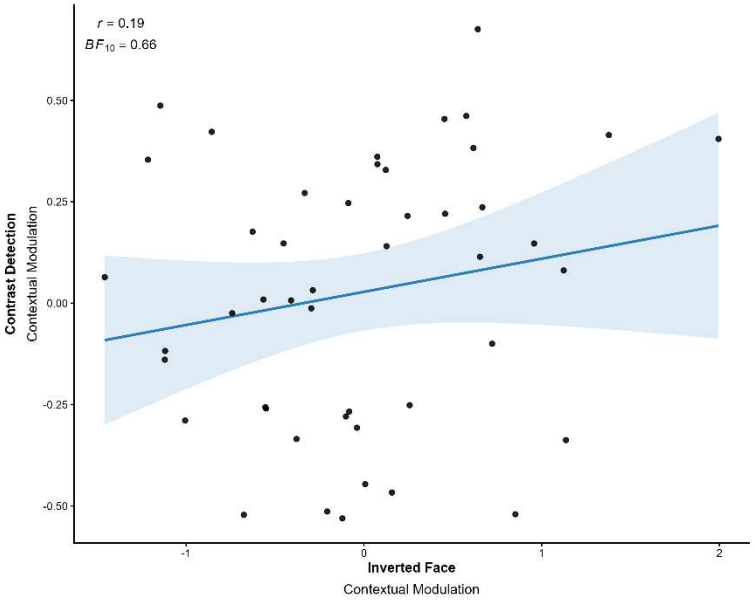
Figure 4. A. Group average psychometric function. Dots represent binned group-level data and curves represent the model fits, with 95% HDI shaded ribbons. B. Normalized threshold estimates across conditions in the facial features discrimination (left) and contrast detection (right) tasks. Each dot represents a participant's z-scored threshold for a given condition and task. Box plots indicate the interquartile range (IQR; box), median (horizontal line), and whiskers extending to 1.5 × IQR. Faint gray lines link observations from the same participant across conditions. Black points and lines represent condition-wise means. C. Posterior distributions of pairwise condition differences for each task. Colored density plots show the posterior estimates of the threshold differences. Dashed lines indicate the Region of Practical Equivalence (ROPE), defined as ±10% of the posterior standard deviation. Thin black bars denote the 95% Highest Density Interval (HDI). The HDIs for the context+ vs. isolated difference (Upright Face task and Contrast Detection task) overlap with the ROPE, rendering the difference inconclusive. All other comparisons fall outside the ROPE and are thus interpreted as credible differences.

Correlation analyses

To investigate a potential functional relationship between low-level and high-level (non-)specialized contextual mechanisms, we estimated the magnitude of contextual modulation by applying a two-step regression-of-regression approach to participants' z-scored thresholds. The residual values obtained from this procedure were then used to compute correlations across tasks (see Methods). These magnitude correlations provided a quantitative assessment of the relationship between contextual modulations across tasks (see Figure 5).

We found moderate evidence for the absence of a correlation of contextual modulation magnitudes between contrast detection and upright face discrimination ($r = .003$, $BF_{01} = 3.03$, $BF_{10} = .33$; short of the preregistered boundary of $BF_{01} = 6$), which suggests the functional independence of low-level surround suppression and high-level face-specialized contextual processing (H1). The correlation between contrast detection and inverted face discrimination was numerically larger but evidentially weak, with only anecdotal evidence for the null ($r = .19$, $BF_{01} = 1.51$, $BF_{10} = .66$). Although the difference between these dependent correlations went in the predicted direction, the comparison was inconclusive (95% HDI of the difference [-0.10, 0.47], overlapping the ROPE of ±0.015), so it remains unresolved whether detection couples more strongly with inverted than with upright face contextual modulation as found previously (Canoluk, Moors, & Goffaux, 2023; H2). Upright and inverted face discrimination magnitudes were strongly correlated, with extreme evidence for a positive relationship ($r = .54$, $BF_{10} > 100$), supporting a shared, high-level contextual process (H3).


r = 0
BF10 = 0.33
Contrast Detection
Contextual Modulation
0.50
0.25
0.00
-0.25
-0.50
-2
-1
0
1
2
Upright Face
Contextual Modulation


r = 0.19
BF10 = 0.66
Contrast Detection
Contextual Modulation
0.50
0.25
0.00
-0.25
-0.50
-1
0
1
2
Inverted Face
Contextual Modulation

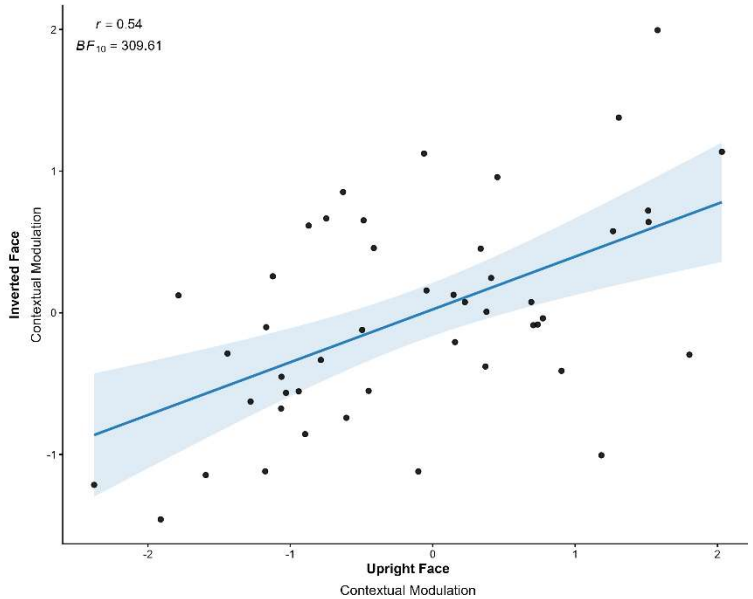

r = 0.54
BF10 = 309.61
Inverted Face
Contextual Modulation
2
1
0
-1
-2
-1
0
1
2
Upright Face
Contextual Modulation

Figure 5. **Correlation of contextual modulation magnitudes across tasks.** Contextual modulation magnitude values are the product of regression of regressions method for each task. Black line indicates the regression line and shaded area covers the standard error bounds.

Discussion

Modulation by context is a fundamental computational principle of visual processing (Mareschal, Sceniak, & Shapley, 2001; Nurminen & Angelucci, 2014) thought to reduce spatial redundancies of sensory input and enable sparse, efficient and feature-selective neural code at all levels of the visual system's hierarchy (Angelucci et al., 2017; Field, 1987; Vinje & Gallant, 2002). Research on contextual modulation focuses on specific processing levels such as the low-level coding of basic visual features like orientation and contrast using artificial narrowband stimuli (but see Coen-Cagli, Kohn, & Schwartz, 2015; Kiniklioglu, Pitcher, & Kaiser, 2026; Vinje & Gallant, 2002) or the high processing levels involved in face perception. Such level-specific approach prevents a system-wide understanding of contextual modulations (Issa, Cadieu, & DiCarlo, 2018; Krause & Pack, 2014; Vanni & Rosenstrom, 2011). The fact that low- and high-level contextual modulations exhibit a similar dependence on basic visual properties and on the local intensity of the input signals indicates the potential existence of a common neural factor (Jastrzebowska et al., 2026) and therefore the heuristic potential of a more systemic approach. A first step in this direction is to investigate the functional relationship of contextual modulations across visual processing levels. To do so, we compared contextual modulations at low- and high-level stages of processing in the same group of participants using a comparable metric, namely the level of contrast necessary to reach threshold performance to detect or discriminate local targets presented either in isolation or in context. We reasoned that using a similar metric would optimize the chance to observe a functional link between low- and high-level contextual modulations (see Canoluk, Moors, & Goffaux, 2023 for a similar approach). In the low-level task (adapted from Canoluk, Moors, & Goffaux, 2023; Mannion, Donkin, & Whitford, 2017), participants detected a grating presented in either a parallel (context-), an orthogonal context (context+), or isolation. In the high-level tasks, participants matched pairs of upright or inverted eye regions presented either in isolation, congruent (context+) or in incongruent (context-) upright or inverted face contexts. In both tasks, local target contrast was adaptively stair-cased.

In the low-level task, the local grating was less well detected when embedded in a parallel context than when presented in an orthogonal context or in isolation. There was no conclusive evidence for the influence of the orthogonal context on contrast detection. These findings are in line with past evidence of a mainly suppressive influence of parallel surround on grating detection (Canoluk, Moors, & Goffaux,

2023; Mannion, Donkin, & Whitford, 2017). In upright faces, incongruent context increased local contrast thresholds for facial feature discrimination compared to congruent and isolated condition. While it is standard practice to quantify low-level contextual modulations in terms of contrast threshold, we have no knowledge of past use of this metric to quantify contextual modulations in face perception. Whether congruent context facilitated facial feature discrimination relative to the isolated condition was inconclusive. The mainly suppressive influence of (incongruent) face context on local feature perception has been extensively documented in past studies (Goffaux, 2012; Rossion, 2013; Suzuki & Cavanagh, 1998; Young, Hellawell, & Hay, 1987), which showed that incongruent context shifts the level of shape dissimilarity necessary to discriminate local facial features (Canoluk, Moors, & Goffaux, 2023; Goffaux, 2012). The results presented here suggest that contextual suppression extends to a more fundamental visual property of the face stimulus, namely local contrast.

In line with past empirical evidence that inversion shrinks the spatial window of face processing (Poltoratski et al., 2021; Rossion, 2008; Van Belle et al., 2010), contextual influence in inverted faces drastically decreased compared with that observed for upright faces. However, context influence was not eliminated by inversion: contrast thresholds were slightly higher when local facial features were presented in context than when they were presented in isolation. Contextual suppression was stronger for incongruent than congruent context as observed for upright faces. However, both context types exerted a suppressive influence on local facial feature discrimination. This departs from our past findings that context moderately but significantly improved (i.e. lowered) shape dissimilarity threshold for feature discrimination in inverted faces (Canoluk, Moors, & Goffaux, 2023). It is unclear whether this polarity flip of contextual influence across studies for inverted faces might be due to the use of local contrast compared to shape dissimilarity.

The central objective of our study was to examine the potential functional link between contextual mechanisms at different levels of the visual processing hierarchy and of high-level processing specialization. In line with past evidence, we found extreme evidence for the functional dependence of high-level contextual mechanisms between specialized and non-specialized (upright and inverted, respectively) face processing (Canoluk, Moors, & Goffaux, 2023). The mechanisms shared between upright and inverted face processing accounted for ~30% of the variance in contextual modulation magnitude. The amount of variance in contextual modulation that remains after removing this shared variance is further support to the functional specificity of the contextual mechanisms engaged in the processing of upright faces.

To address functional relationship of contextual modulations across the levels of the visual hierarchy, we systematically quantified the shared variance in contextual modulation magnitude between the low-level contrast detection and the high-level feature discrimination. Correlation results did not support the hypothesis of a functional link between low-level with high-level contextual modulations, neither for specialized (upright, H1) nor for non-specialized (inverted, H2) face processing. For upright faces (H1), the data were about three times more likely under the absence than under the presence of a correlation (BF_{01} = 3.03), which is moderate evidence for functional independence, although it falls short of our preregistered evidential threshold (BF_{01} > 6). We regard this as an indication that high-level mechanisms specialized in face processing take precedence and detach face perception from low-level contextual representations (Canoluk, Moors, & Goffaux, 2023; Suzuki & Cavanagh, 1998). Indeed, our sample, despite being below the maximum limits established by our design analysis, was relatively large relative to psychophysics research standards and of reasonable size for correlational analyses in this field of research (e.g. Canoluk, Moors, & Goffaux, 2023; Jastrzebowska et al., 2026). Moreover, the flat distribution of low-level against high-level contextual modulations for upright faces (Figure 5) is consistent with this interpretation, although a larger sample would be needed to reach our preregistered evidential threshold.

Yet, it seems reasonable to defend that V1 encoding determines the perception of complex natural stimuli. V1 anatomy and connectivity have been shown to determine the perceptual experience of object size and orientation (Moutsiana et al., 2016; Song et al., 2015; Song et al., 2013 but see Jastrzebowska et al., 2026). Moreover, the shared dependence on local input strength indicates similar working principles for contextual modulations at low- and high-level stages of visual processing. However, our findings indicate that contextual modulations in upright faces recruit mechanisms that are functionally similar to but independent from primary contextual mechanisms (Kiniklioglu & Kaiser, 2026; Kiniklioglu, Pitcher, & Kaiser, 2026). The fact that the low-level task used here induced near surround contextual modulation may have hindered the emergence of a link with high-level face contextual modulations. Indeed, while near (<3° distance) surround suppression results from local lateral interactions within V1, far surround modulations (>3° distance) are driven by top-down influences from higher-level brain regions (Bair, Cavanaugh, & Movshon, 2003; Nurminen & Angelucci, 2014; Pokorny, Weldon, & Olman, 2024). Using far-surround contextual modulations may involve system-wide neural communications and thus be a more promising avenue to address the link in contextual mechanisms across visual processing levels. Another aspect that may have hindered the observation of the functional link between low- and high-level contextual modulations is the reliance on purely behavioral measures. Behavioral responses reflect a long chain of processing, ranging from early sensory encoding to the later stages of decision-making, and may lack the sensitivity and

specificity required to highlight variations that are selectively shared by contextual mechanisms. Indeed, within or between visual domains, contextual modulations do not or only weakly correlate (Cretenoud, Francis, & Herzog, 2020; Jastrzebowska et al., 2026). The different types of decisions involved in low- and high-level tasks (detection as opposed to discrimination) may have further obscured the functional link.

The absence of a reliable correlation between low- and high-level contextual effects may also result from the use of gratings in the low-level task. Contextual modulations are influenced by the high-order properties of natural input and simple narrow-band grating stimuli have been shown to be too limited to inform about how complex natural images are contextually processed (Balla et al., 2025; Carandini et al., 1997; Coen-Cagli, Kohn, & Schwartz, 2015; Kiniklioglu & Kaiser, 2026; Pokorny, Weldon, & Olman, 2024; Ziemba et al., 2018). To address the functional link of contextual modulations across computational levels and, more broadly, better understand how context influences natural vision, it is more promising to use complex natural images and track contextual influences directly on the neural responses at different levels of the visual system hierarchy (for example Goffaux et al., 2016; Kiniklioglu, Pitcher, & Kaiser, 2026; Schuurmans et al., 2023; see Jastrzebowska et al., 2026).

To conclude, we investigated whether the behavioral signatures of contextual processing attributed to low and high levels of the visual system hierarchy are functionally linked at the individual participant level. Despite their sharing of common functional principles, we found no conclusive evidence for a link between low- and high-level contextual modulations. A more promising avenue to progress toward a much-needed system-wide understanding of the influence of spatial context in everyday life vision requires directly investigating contextual influences on the neural representation of natural images across the levels of the visual system hierarchy.

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 modulation magnitudes in contrast detection task correlate more strongly with inverted than upright	Same as above	Bayesian correlation analysis of contextual modulation magnitudes	Same as above	Evidence for a stronger correlation for inverted than upright facial feature discrimination task would support a stronger functional	Inversion disrupts the specificity of face contextual

contextual mechanisms?	facial feature discrimination task. **(H2)**		between contrast detection and inverted facial feature discrimination.		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.	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~~ context-trials are higher than	Same as above	Assessing the overlap of the 95% HDI of threshold differences	Same as above	Evidence for a difference would demonstrate contextual processing in upright faces. Evidence for no difference would	Upright faces are processed in a part-

	the thresholds in ~~congruent+~~context+ trials. **(Outcome Neutral test 1)**		(congruent-incongruent) with ~~zero~~ the rope~~.~~		show part-based/local processing.	based manner.
Are inverted faces processed contextually?	Discrimination thresholds in ~~incongruent-~~context- trials are higher than the thresholds in ~~congruent+~~context+ 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-~~context- and ~~congruent+~~context+ 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 incogruent~~parallel-~~ trials are higher than	Same as above	Assessing the overlap of the 95% HDI of threshold	Same as above	Evidence for a difference would demonstrate orientation-dependent contextual	Context does not affect the

	the thresholds in congruent~~orthogonal~~+ trials **(Outcome Neutral test 4)**		differences ((incongruent~~parall el~~-)-(congruent~~orthogo nal~~+)) with zero.		processing. Evidence for no difference would show that orientation of the surround has no effect on the local processing of gratings.	processing of local contrast.

Funding.

This work was supported by the Fonds de la Recherche Scientifique – FNRS under the funding “Projet de Recherche” (FNRS T.0102.19), and “Excellence of Science” (HUMVISCAT-30991544) grants. M.U.C. and M.G. were PhD Students funded by the F.R.S-F.N.R.S.. V.G. is an F.R.S-F.N.R.S. Senior Research Associate.

References

Anaki, D., Nica, E. I., & Moscovitch, M. (2011). Automatic aspects in face perception: evidence from mandatory processing of distractor facial components [Research Support, Non-U.S. Gov't]. *Experimental psychology*, *58*(1), 4-18. https://doi.org/10.1027/1618-3169/a000061

Andrews, T. J., Clarke, A., Pell, P., & Hartley, T. (2010). Selectivity for low-level features of objects in the human ventral stream. *Neuroimage*, *49*(1), 703-711. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19716424

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*, 425-451. https://doi.org/10.1146/annurev-neuro-072116-031418

Angelucci, A., Levitt, J. B., Walton, E. J., Hupe, J. M., Bullier, J., & Lund, J. S. (2002). Circuits for local and global signal integration in primary visual cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *22*(19), 8633-8646. http://www.ncbi.nlm.nih.gov/pubmed/12351737

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

Bair, W., Cavanaugh, J. R., & Movshon, J. A. (2003). Time course and time-distance relationships for surround suppression in macaque V1 neurons. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *23*(20), 7690-7701. https://doi.org/10.1523/JNEUROSCI.23-20-07690.2003

Balla, E., Nabbefeld, G., Wiesbrock, C., Linde, J., Graff, S., Musall, S., & Kampa, B. M. (2025). Broadband visual stimuli improve neuronal representation and sensory perception. *Nat Commun*, *16*(1), 2957. https://doi.org/10.1038/s41467-025-58003-1

Benton, A. L., & Van Allen, M. W. (1968). Impairment in Facial Recognition in Patients with Cerebral Disease. *Cortex; a journal devoted to the study of the nervous system and behavior*, *4*(4), 344-IN341. https://doi.org/10.1016/S0010-9452(68)80018-8

Boutet, I., Collin, C., & Faubert, J. (2003). Configural face encoding and spatial frequency information. *Percept Psychophys*, *65*(7), 1078-1093. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=14674634

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* (1 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. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10811161
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
Carandini, M., Barlow, H. B., O'Keefe, L. P., Poirson, A. B., & Movshon, J. A. (1997). Adaptation to contingencies in macaque primary visual cortex [Research Support, Non-U.S. Gov't
Research Support, U.S. Gov't, P.H.S.]. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *352*(1358), 1149-1154. https://doi.org/10.1098/rstb.1997.0098
Coen-Cagli, R., Kohn, A., & Schwartz, O. (2015). Flexible gating of contextual influences in natural vision. *Nature neuroscience*, *18*(11), 1648-1655. https://doi.org/10.1038/nn.4128
Cretenoud, A. F., Francis, G., & Herzog, M. H. (2020). When illusions merge. *J Vis*, *20*(8), 12. https://doi.org/10.1167/jov.20.8.12
Dakin, S. C., & Watt, R. J. (2009). Biological "bar codes" in human faces. *Journal of Vision*, *9*(4), 2 1-10. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19757911
Davidoff, J., & Donnelly, N. (1990). Object superiority: a comparison of complete and part probes. *Acta Psychol (Amst)*, *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. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=2940312
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. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7790837
Farah, M. J., Wilson, K. D., Drain, M., & Tanaka. (1998). What is "special" about face perception? *Psychol Rev*, *105*(3), 482-498. https://doi.org/10.1037/0033-295X.105.3.482
Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *J Opt Soc Am A*, *4*(12), 2379-2394. https://doi.org/10.1364/josaa.4.002379
Goffaux, V. (2009). Spatial interactions in upright and inverted faces: re-exploration of spatial scale influence. *Vision research*, *49*(7), 774-781.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19236892
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 behavioural signatures of face processing. *Frontiers in psychology*. https://doi.org/10.3389/fpsyg.2010.00143
Goffaux, V., Duecker, F., Hausfeld, L., Schiltz, C., & Goebel, R. (2016). Horizontal tuning for faces originates in high-level Fusiform Face Area. *Neuropsychologia*, *81*, 1-11. https://doi.org/10.1016/j.neuropsychologia.2015.12.004
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.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16846295
Goffaux, V., Schiltz, C., Mur, M., & Goebel, R. (2013). Local discriminability determines the strength of holistic processing for faces in the Fusiform Face Area [Original Research]. *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. http://www.ncbi.nlm.nih.gov/pubmed/10027301
Hoaglin, D. C., Iglewicz, B., & Tukey, J. W. (1986). Performance of Some Resistant Rules for Outlier Labeling. *Journal of the American Statistical Association*. https://www.tandfonline.com/doi/abs/10.1080/01621459.1986.10478363
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
Issa, E. B., Cadieu, C. F., & DiCarlo, J. J. (2018). Neural dynamics at successive stages of the ventral visual stream are consistent with hierarchical error signals. *Elife*, *7*. https://doi.org/10.7554/eLife.42870
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
Jastrzebowska, M. A., Ozkirli, A., Cretenoud, A. F., Draganski, B., & Herzog, M. H. (2026). Is there a neural common factor for visual illusions? *J Vis*, *26*(10), 13. https://doi.org/10.1167/jov.26.10.13
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
Kiniklioglu, M., & Kaiser, D. (2026). Characterizing surround suppression with dynamic natural scenes. *J Vis*, *26*(3), 9. https://doi.org/10.1167/jov.26.3.9
Kiniklioglu, M., Pitcher, D., & Kaiser, D. (2026). Distinct mechanisms of contextual modulation for dynamic natural scenes in motion- and scene-selective cortex. *Journal of neurophysiology*, *136*(1), 61-76. https://doi.org/10.1152/jn.00063.2026
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
Krause, M. R., & Pack, C. C. (2014). Contextual modulation and stimulus selectivity in extrastriate cortex. *Vision research*, *104*, 36-46. https://doi.org/10.1016/j.visres.2014.10.006
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 ed.).
Laguesse, R., Dormal, G., Biervoye, A., Kuefner, D., & Rossion, B. (2012). Extensive visual training in adulthood significantly reduces the face inversion effect. *J Vis*, *12*(10), 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. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10881616
Levick, W. R., Thibos, L. N., Cohn, T. E., Catanzaro, D., & Barlow, H. B. (1983). Performance of cat retinal ganglion cells at low light levels. *J Gen Physiol*, *82*(3), 405-426. https://doi.org/10.1085/jgp.82.3.405
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
Mareschal, I., Sceniak, M. P., & Shapley, R. M. (2001). Contextual influences on orientation discrimination: binding local and global cues. *Vision research*, *41*(15), 1915-1930. https://doi.org/10.1016/s0042-6989(01)00082-7
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. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16229867
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. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19022276
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. *Mem Cognit*, *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
Meinhardt, G., Meinhardt-Injac, B., & Persike, M. (2014). The complete design in the composite face paradigm: role of response bias, target certainty, and feedback. *Frontiers in human neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00885

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*. In 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. *J Vis*, *12*(11). https://doi.org/10.1167/12.11.26
Moutsiana, C., de Haas, B., Papageorgiou, A., van Dijk, J. A., Balraj, A., Greenwood, J. A., & Schwarzkopf, D. S. (2016). Cortical idiosyncrasies predict the perception of object size. *Nat Commun*, *7*, 12110. https://doi.org/10.1038/ncomms12110
Nauhaus, I., Busse, L., Carandini, M., & Ringach, D. L. (2009). Stimulus contrast modulates functional connectivity in visual cortex. *Nature neuroscience*, *12*(1), 70-76. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19029885
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
Nurminen, L., & Angelucci, A. (2014). Multiple components of surround modulation in primary visual cortex: multiple neural circuits with multiple functions? *Vision research*, *104*, 47-56. https://doi.org/10.1016/j.visres.2014.08.018
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
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
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
Peirce, J. W. (2007). PsychoPy--Psychophysics software in Python. *J Neurosci Methods*, *162*(1-2), 8-13. https://doi.org/10.1016/j.jneumeth.2006.11.017
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
Pokorny, V. J., Weldon, K. B., & Olman, C. A. (2024). Surround Suppression of Broadband Images. *bioRxiv*. https://doi.org/10.1101/2024.05.15.594329
Poltoratski, S., Kay, K., Finzi, D., & Grill-Spector, K. (2021). Holistic face recognition is an emergent phenomenon of spatial processing in face-selective regions. *Nat Commun*, *12*(1), 4745. https://doi.org/10.1038/s41467-021-24806-1
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
RCoreTeam. (2020). *R: A language and environment for statistical computing. (Version 4.3.2) [R].* In R Foundation for Statistical Computing.
Richler, J., Gauthier, I., Wenger, M., & Palmeri, T. (2008). Holistic processing of faces: perceptual and decisional components. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *34*, 328-342.
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. *Psychol Bull*, *140*(5), 1281-1302. https://doi.org/10.1037/a0037004
Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica*, *128*, 274-289.

Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition*, *21*(2), 139-253. https://doi.org/10.1080/13506285.2013.772929
Rossion, B., & Michel, C. (2018). Normative accuracy and response time data for the computerized Benton Facial Recognition Test (BFRT-c). *Behav Res Methods*, *50*(6), 2442-2460. https://doi.org/10.3758/s13428-018-1023-x
Rouder, J. N. (2014). Optional stopping: no problem for Bayesians. *Psychon Bull Rev*, *21*(2), 301-308. https://doi.org/10.3758/s13423-014-0595-4
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
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. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10412063
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. *J Vis*, *10*(2), 25 21-16. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=20462326
Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipito-temporal cortex. *Neuroimage*, *32*(3), 1385-1394. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16870475
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
Schuurmans, J. P., Bennett, M. A., Petras, K., & Goffaux, V. (2023). Backward masking reveals coarse-to-fine dynamics in human V1. *Neuroimage*, *274*, 120139. https://doi.org/10.1016/j.neuroimage.2023.120139
Schwartz, O., Hsu, A., & Dayan, P. (2007). Space and time in visual context. *Nature Reviews Neuroscience*, *8*(7), 522-535. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17585305
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*, 221-242.
Serrano-Pedraza, I., Romero-Ferreiro, V., Read, J. C. A., Diéguez-Risco, T., Bagney, A., Caballero-González, M., Rodríguez-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
Song, C., Schwarzkopf, D. S., Kanai, R., & Rees, G. (2015). Neural population tuning links visual cortical anatomy to human visual perception. *Neuron*, *85*(3), 641-656. https://doi.org/10.1016/j.neuron.2014.12.041
Song, C., Schwarzkopf, D. S., Lutti, A., Li, B., Kanai, R., & Rees, G. (2013). Effective connectivity within human primary visual cortex predicts interindividual diversity in illusory perception. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *33*(48), 18781-18791. https://doi.org/10.1523/JNEUROSCI.4201-12.2013

Suzuki, S., & Cavanagh, P. (1998). A shape-contrast effect for briefly presented stimuli. *Journal of experimental psychology. Human perception and performance*, *24*(5), 1315-1341. https://doi.org/10.1037//0096-1523.24.5.1315
Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, *46*(2), 225-245. http://www.sciencedirect.com/science/article/B6WVB-45CV0TH-4MX/1/bcc9a51e9650d1f79be49ba2798fee66
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
Van Belle, G., De Graef, P., Verfaillie, K., Rossion, B., & Lefevre, P. (2010). Face inversion impairs holistic perception: evidence from gaze-contingent stimulation. *J Vis*, *10*(5), 10. https://doi.org/10.1167/10.5.10
Vanni, S., & Rosenstrom, T. (2011). Local non-linear interactions in the visual cortex may reflect global decorrelation. *J Comput Neurosci*, *30*(1), 109-124. https://doi.org/10.1007/s10827-010-0239-2
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
Vinje, W. E., & Gallant, J. L. (2002). Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *22*(7), 2904-2915. https://doi.org/10.1523/JNEUROSCI.22-07-02904.2002
Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept Psychophys*, *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). Configurational 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 : CB*, *15*(24), 2256-2262. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16360687
Yssaad-Fesselier, R., & Knoblauch, K. (2006). Modeling psychometric functions in R. *Behavior Research Methods*, *38*(1), 28-41. https://doi.org/10.3758/BF03192747
Ziemba, C. M., Freeman, J., Simoncelli, E. P., & Movshon, J. A. (2018). Contextual modulation of sensitivity to naturalistic image structure in macaque V2. *Journal of neurophysiology*, *120*(2), 409-420. https://doi.org/10.1152/jn.00900.2017