

Feature–location effects in the Thatcher illusion

Experimental Psychology, University College London,
London, UK
Institute of Cognitive Neuroscience,
University College London, London, UK
Department of Psychology,
Justus Liebig Universität Giessen, Giessen, Germany



Benjamin de Haas

Experimental Psychology, University College London,
London, UK
Institute of Cognitive Neuroscience,
University College London, London, UK
School of Optometry & Vision Science,
University of Auckland, Auckland, New Zealand



D. Samuel Schwarzkopf

Face perception is impaired for inverted images, and a prominent example of this is the Thatcher illusion: “Thatcherized” (i.e., rotated) eyes and mouths make a face look grotesque, but only if the whole face is seen upright rather than inverted. Inversion effects are often interpreted as evidence for configural face processing. However, recent findings have led to the alternative proposal that the Thatcher illusion rests on orientation sensitivity for isolated facial regions. Here, we tested whether the Thatcher effect depends not only on the orientation of facial regions but also on their visual-field location. Using a match-to-sample task with isolated eye and mouth regions we found a significant Feature \times Location interaction. Observers were better at discriminating Thatcherized from normal eyes in the upper compared to the lower visual field, and vice versa for mouths. These results show that inversion effects can at least partly be driven by nonconfigural factors and that one of these factors is a match between facial features and their typical visual-field location. This echoes recent results showing feature–location effects in face individuation. We discuss the role of these findings for the hypothesis that spatial and feature tuning in the ventral stream are linked.

Young, 1986; Maurer, Le Grand, & Mondloch, 2002; Rhodes, 1988). One line of evidence interpreted to support this notion is face-inversion effects. Face perception is severely impoverished for images turned upside down (Schwaninger, Carbon, & Leder, 2003; Valentine, 1988; Yin, 1969), and this specifically applies to the recognition of configural aspects (Goffaux & Rossion, 2007; Leder & Bruce, 2000; Leder, Candrian, Huber, & Bruce, 2001; Schwaninger & Mast, 2005; but see Rakover & Teucher, 1997).

A prominent example is the Thatcher illusion (Thompson, 1980). When eyes and mouth within a face image are rotated by 180°, the resulting “Thatcherized” image looks grotesque. However, this manipulation is strikingly obvious only when the Thatcherized image is upright—observers perceive it as much subtler, and can miss it entirely, when the Thatcherized image itself is shown upside down. This dramatic difference (like other face-inversion effects) has been interpreted as the result of disrupted configural processing (Bartlett & Searcy, 1993; Maurer et al., 2002; Murray, Yong, & Rhodes, 2000; Rossion, 2009).

However, recent findings of a Thatcher illusion for isolated face parts challenge this interpretation (Psalta, Young, Thompson, & Andrews, 2014). This study showed eye or mouth regions (a horizontal strip across the face) in isolation and tested observers’ ability to discriminate Thatcherized from unaltered features. Participants performed near ceiling for upright images but at or below chance level when images were inverted (the Thatcher illusion). Consequently, Psalta et al. proposed that the Thatcher illusion is a result of

Introduction

Theories of face perception emphasize the importance of configural processing, referring to the arrangement and distance of face parts from each other as well as their integration into a whole (Bruce &

Citation: de Haas, B., & Schwarzkopf, D. S. (2018). Feature–location effects in the Thatcher illusion. *Journal of Vision*, 18(4):16, 1–12, <https://doi.org/10.1167/18.4.16>.

<https://doi.org/10.1167/18.4.16>

Received October 6, 2017; published April 23, 2018

ISSN 1534-7362 Copyright 2018 The Authors



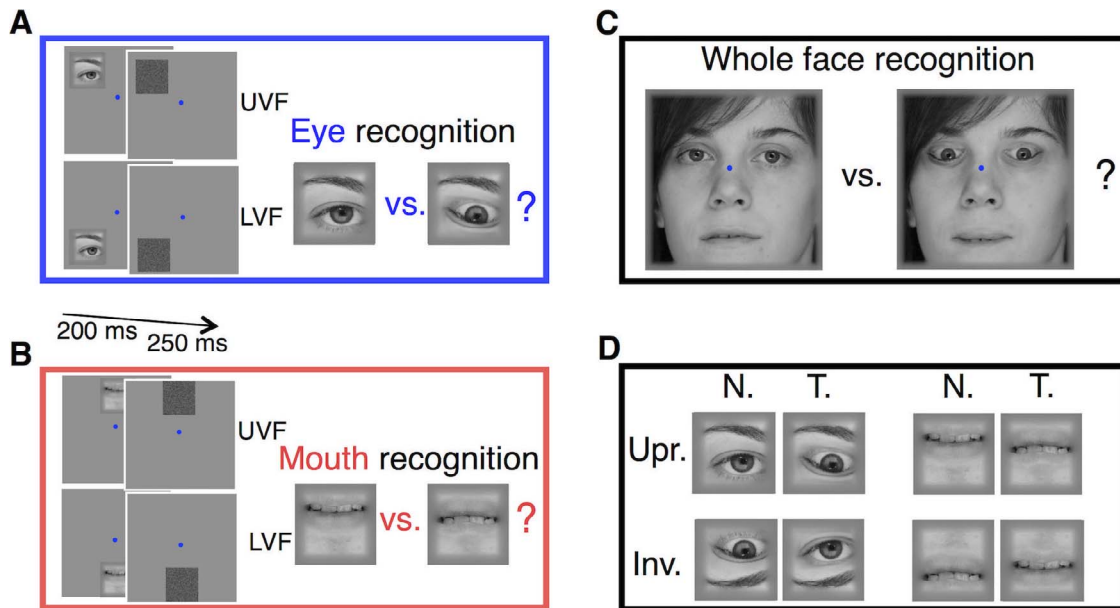


Figure 1. Stimuli and design. (A) Left-hand side: Participants saw a 200-ms flash of a left- or right-eye region shown at its typical visual-field location in an upright face (top) or vertically shifted to the corresponding location in the lower visual field (bottom). Images were overlaid with a dynamic noise mask that persisted for 250 ms after image offset. Right-hand side: The flashing target stimuli could be Thatcherized or not, and participants were asked to decide which version they saw in a match-to-sample task. Specifically, participants were presented with the Thatcherized and non-Thatcherized version of the corresponding image side by side after the offset of the target image and noise mask (on a gray background and without accompanying text, which is shown here only for illustration purposes). They could toggle a selection rectangle enclosing either candidate image (not shown) and confirm their choice with the keyboard. The selection rectangle briefly turned from blue to green or red for correct or incorrect choices, respectively. The choice period was self-paced, and participants were free to move their eyes until they confirmed their selection. (B) Mouth regions were presented on the vertical meridian, at either a typical lower or the corresponding upper visual-field position. (C) Flashing whole-face stimuli were centered on the fixation dot and otherwise followed the same design (not shown). (D) Example pairs of stimuli. In every trial of the match-to-sample task, participants judged whether they saw the normal (N) or Thatcherized (T) version of a feature. Images of the enclosing facial region could be upright (Upr.) or inverted (Inv.). Note that targets and candidates always had the same orientation.

orientation sensitivity for isolated facial regions rather than an example of disrupted configural processing. Note that Thatcherization and inversion refer to different kinds of image manipulations in this context. Thatcherization is the rotation of only the eyes or mouth (relative to the embedding context), while inversion refers to the inversion of the whole image, which shows a larger region of the face in which the respective feature is embedded. The images of facial regions used by Psalta et al. were full horizontal strips of the face, including eyes or mouth but also parts of the face outline. The images used in the present study were more restricted (see Figure 1D for examples), in order to keep the design as close as possible to that of de Haas et al. (2016), who reported behavioral and neural feature–location effects for facial-feature individuation (see later).

Another caveat regarding the configural interpretation of face-inversion effects comes from the observation that human observers show a stereotypical pattern of gaze behavior toward faces, which results in feature-

specific retinotopic biases. First fixations tend to land on the central upper nose region, just below the eyes (Hsiao & Cottrell, 2008), a landing point that is optimal for rapid face recognition (Peterson & Eckstein, 2012). Subsequent fixations typically remain restricted to inner face features, with a predominance of eye-directed fixations (van Belle, Ramon, Lefèvre, & Rossion, 2010) that is seen at least for static faces (cf. Vö, Smith, Mital, & Henderson, 2012). This pattern results in a strongly biased feature–location statistic for free viewing behavior, with eyes and mouths appearing mostly in the upper and lower visual field, respectively (de Haas et al., 2016).

A simultaneous reversal of these retinotopic biases for eyes and mouth is only possible for faces seen upside down. Moreover, eye-tracking studies investigating gaze behavior toward inverted faces indicate that the basic pattern of feature fixations remains similar to that for upright faces (Boutet, Lemieux, Goulet, & Collin, 2017; Williams & Henderson, 2007), although the predominance of the eye region appears

weakened for both first (Hills, Cooper, & Pake, 2013; Hills, Sullivan, & Pake, 2012) and subsequent (Barton, Radcliffe, Cherkasova, Edelman, & Intriligator, 2006; Xu & Tanaka, 2013) fixations. This similarity of gaze patterns implies a reversal of typical retinotopic feature locations in an inverted face: Eyes and mouth will typically appear in the lower and upper visual field, respectively.

A recent study on identity recognition of isolated features has shown that this input contingency is reflected in perceptual sensitivity (de Haas et al., 2016). This study used a match-to-sample task in which observers discriminated the identity of isolated eye or mouth regions from different faces. Recognition performance was significantly diminished by image inversion but also varied with visual-field position. Observers were better at individuating eye regions presented in the upper compared to the lower visual field, and vice versa for mouth regions.

This contingency has not been considered by most studies investigating the role of featural versus configural processing for face-inversion effects, including those that explicitly examined gaze behavior (e.g. Barton et al., 2006; Bombari, Mast, & Lobmaier, 2009; Boutet et al., 2017; Hills et al., 2012; van Belle, De Graef, Verfaillie, Rossion, & Lefèvre, 2010; Williams & Henderson, 2007; Xu & Tanaka, 2013). For instance, van Belle, De Graef, et al. (2010) conducted an elegant match-to-sample experiment in which candidate faces were presented in a gaze-contingent fashion. Foveal masking blocked the view of directly fixated features, rendering only parts of the face outside the fixated area visible. The reverse condition of a foveal window rendered *only* the fixated region visible, masking out the face context. Results showed an increased inversion effect for foveal masking and a decreased inversion effect for the foveal-window condition. The authors interpreted this as evidence that face inversion specifically disrupts the holistic processing of faces. Following their interpretation, the foveal-window condition probed configural processing, while foveal masking emulated holistic processing. However, these results would also be predicted by the hypothesis that feature–location interactions are crucial for the effect of face inversion. The typical contingency between facial features and retinotopic locations will be reversed for central fixations toward inverted faces in the foveal-mask condition but not the foveal window condition.

The perceptual effect of feature–location interactions has been hypothesized to reflect a match of spatial and feature tuning in cortical face areas (de Haas et al., 2016). This seems in line with the finding that eye and mouth representations in the occipital face area are more distinguishable from each other when they are presented at typical rather than reversed visual-field locations (as found when applying decoding techniques

to human neuroimaging data; de Haas et al., 2016). It is further in line with human neuroimaging results suggesting a maplike organization of facial-feature representations in the occipital face area (Henriksson, Mur, & Kriegeskorte, 2015; van den Hurk, Pegado, Martens, & Op de Beeck, 2015) and mirrors electrophysiological findings in macaques, where cells preferring images of another monkey’s contralateral eye region are also tuned to the contralateral upper quadrant (Issa & DiCarlo, 2012).

More generally, accumulating evidence casts doubt on the notion of location invariance in the ventral stream (Biederman & Cooper, 1991; Riesenhuber & Poggio, 1999). Neurons of the ventral stream (DiCarlo & Maunsell, 2003), face-adaptation effects (A. Afraz & Cavanagh, 2009; S.-R. Afraz & Cavanagh, 2008), and face-specific perceptual biases (A. Afraz, Pashkam, & Cavanagh, 2010) can show relatively narrow spatial selectivity. Furthermore, retinotopic organization along the ventral stream is correlated with feature selectivity (Kravitz, Saleem, Baker, Ungerleider, & Mishkin, 2013; Silson, Chan, Reynolds, Kravitz, & Baker, 2015) and an important precursor for the development of feature selectivity (Arcaro, Schade, Vincent, Ponce, & Livingstone, 2017). This functional importance of spatial tuning supports the notion that feature–location interactions could pose a general confound for face-inversion effects.

The visual-field displacement of features in an inverted face (see earlier) might also affect the ability to discriminate Thatcherized from unaltered features. If so, the Thatcher illusion might be better explained by sensitivity to retinotopic feature location and local orientation of isolated regions, rather than local orientation alone (cf. Psalta et al., 2014).

Here, we tested this hypothesis using a match-to-sample task. In each trial, observers saw a brief image of an isolated eye or mouth region and were asked to identify whether they saw the Thatcherized or unaltered version of the feature (Figure 1). We compared participants’ ability to discriminate Thatcherized from unaltered features for upright images presented at their respective typical visual-field positions to the same discrimination ability for inverted images at typical locations, upright images at reversed locations (eye and mouth regions in the lower and upper visual field, respectively), and inverted images at reversed locations (again, eye and mouth regions in the lower and upper visual field, respectively). This allowed us to quantify the Thatcher effect of local inversion, reversed retinotopic feature location, and the combination of both. In addition, we tested the (classical) Thatcher illusion for full faces using the same task.

Based on our main hypothesis of a feature–location interaction we predicted a *location*-induced Thatcher effect—that is, an advantage for discriminating the

Thatcherization of isolated eyes in the upper compared to the lower visual field, and vice versa for mouths. We further expected to replicate the findings of Psalta et al. (2014) of an advantage for discriminating the Thatcherization of upright compared to inverted images of isolated regions (that is, a Thatcher effect induced by local *inversion*), as well as the same effect for full faces (the classic Thatcher effect). Finally, we expected the *combination* of both manipulations (local inversion and atypical retinotopic feature location) to result in a stronger Thatcher effect than either manipulation on its own.

Materials and methods

Participants

Thirty-six healthy participants from the University College London participant pool took part in the experiment (ages: 21 to 55 years, M : 30 years, SD : 8 years; 25 women, 11 men; two left-handed, 24 right-handed). All participants had normal or corrected-to-normal vision. Written informed consent was obtained from each participant, all procedures adhered to the Declaration of Helsinki, and the University College London Research Ethics Committee approved the experiment.

Data from three additional participants were excluded because their performance across all conditions was $<60\%$ and three standard deviations below the group mean (robustly estimated using median absolute deviation; two participants) or because the eye-tracking recording failed (one participant). Control analyses confirmed that neither the direction nor the statistical significance of reported effects changed when data from these participants were included.

Stimuli

Face stimuli stem from the SiblingsDB set (Vieira, Bottino, Laurentini, & De Simone, 2014). Sixty frontal photographs of faces with neutral expression were cropped to a square region showing the inner face, stretching from chin to forehead. A Thatcherized version of each photograph was produced by inverting rectangular regions around either inner eye as well as around the lips (using the landmarks provided with the SiblingsDB set). Any resulting hard edges were blurred using the smudge tool in GIMP (<https://www.gimp.org/>; all other image manipulations using MATLAB [MathWorks, Natick, MA]). Isolated facial regions were extracted from both Thatcherized and original images by cropping a larger square region around

either eye (including the brow) and the mouth region (Figure 1D; cf. de Haas et al., 2016).

Whole-face images and isolated regions were presented at widths of 16.42° and 5° visual angle, respectively. The outer edge of each image was overlaid with a gray fringe that softened the edge between image and background and was 0.14° or 0.46° wide for region and whole-face images, respectively. All images were 8-bit grayscale, displayed on a gray background and with a dynamic-noise mask overlay (see later). Stimuli were shown on a liquid crystal display monitor (Samsung SyncMaster 2233RZ; Samsung, Seoul, South Korea) with a refresh rate of 120 Hz and a spatial resolution of $1,680 \times 1,050$ pixels.

Procedure

The design of the experiment closely followed the design of experiment 2 by de Haas et al. (2016). Participants rested their heads in a chin rest at a viewing distance of 44 cm. Each trial consisted of a 500-ms presentation of a blue fixation dot, followed by the target image flashing up for 200 ms, overlaid by a dynamic noise mask that lasted until 450 ms after image onset. Note that the central fixation dot persisted throughout target presentation, and participants were instructed that stable fixation during this period was crucial. Immediately after the offset of the noise mask, the fixation dot disappeared and two candidate images appeared on the screen, prompting the participant to indicate which of them flashed up earlier using a standard keyboard. The target was an isolated facial region or a whole face, and could be Thatcherized or not. Candidates consisted of the target and the corresponding (non-) Thatcherized version, shown side by side and centered vertically on the screen. Candidate images persisted on the screen until participants indicated their answer. Participants were free to move their eyes during this period. At the beginning of each choice period, a selection rectangle enclosed the left candidate image. Participants could toggle the position of the selection rectangle between candidate images using the left and right arrow keys and could confirm their choice using the space bar. Once participants confirmed their choice they received feedback via a 300-ms color change of the selection rectangle (from blue to either green or red, for correct and incorrect choices, respectively).

Target and candidate images were presented either upright or inverted, and isolated facial regions could appear at typical or reversed visual-field positions, yielding a 2×2 design with factors inversion and location (only inversion for whole-face images). Whenever the target was inverted, candidates were as well. Typical visual-field positions roughly corre-

sponded to the locations of the respective facial features (left eye, right eye, mouth) in the original image, assuming fixation slightly above the nose. Specifically, the mouth position was centered on the vertical meridian at 6.25° visual angle below fixation, and the left and right eye positions were centered 3.75° above fixation and shifted 5° to the left or right, respectively. In the reversed condition, mouth and eye regions were shown at corresponding locations in the upper and lower visual field, respectively (see Figure 1 for an illustration). Note that all center positions had an equal eccentricity of 6.25° .

Each of the three facial regions was shown upright or inverted (Figure 1D) and at its typical or reversed position in different trials. Together with upright and inverted whole-face images, this yielded a total of 14 trial types. Each of these trial types occurred 34 times in pseudorandom order, for a total of 476 trials per participant, split into 14 short blocks of 34 trials each. In each trial the exact stimulus position was determined as the center location for the corresponding trial type, plus a random horizontal and vertical offset of up to 0.7° visual angle to avoid adaptation or fatigue (drawn from Gaussian distributions with a standard deviation of 0.35° and centered on 0).

Fixation stability was monitored with an infrared eye tracker (Cambridge Research Systems, Rochester, UK) tracking the left eye at 200 Hz. Across participants, gaze direction could be tracked successfully for an average of $84.87\% \pm 2.45\%$ of trials (mean \pm standard error of the mean).

Analysis

All statistical analyses were performed in MATLAB R2016b (MathWorks) and JASP 0.8.1.2 (<https://jasp-stats.org/>). To test for an interaction between facial feature and visual-field position, the proportion of correct answers for upright isolated facial regions was averaged for each participant by feature (eye/mouth) and location (upper/lower visual field). The resulting values were compared across conditions using a repeated-measures general linear model and post hoc t tests. Additionally, we calculated the reduction in correct answers for each condition and facial region relative to its upright version shown at typical locations. In this way, we quantified the recognition cost of image inversion, atypical location, and the combination of both. We used t tests to test these Thatcher effects against zero and a repeated-measures general linear model with within-subject factors feature (eye/mouth) and manipulation (location/inversion/combination), and follow-up t tests to compare them against each other. Finally, we contrasted recognition performance for upright and inverted whole faces to

quantify the magnitude of the classical Thatcher effect for our design and sample. We used the proportion of correct responses as measure of recognition performance. Our experiment was not designed to yield meaningful reaction times, and involved self-paced responses that took a variable number of button presses (potential toggling of selection rectangle and confirmation; see earlier). Furthermore, participants were explicitly instructed about the self-paced nature of our experiment. Nevertheless, for completeness' sake we report response times in the Appendix.

Note that participants were instructed to fixate the central dot during stimulus presentation, and stimulus duration was deliberately limited to a duration below saccade latency to ensure fixation compliance (200 ms; see earlier). The main purpose of our experiment was to test an interaction between stimulus type and retinotopic location. Successful manipulation of the latter crucially depends on fixation compliance, to ensure a constant alignment of visual field and screen space. In addition to our fast stimulus presentation times, we explicitly tested fixation stability using eye-tracking data (see earlier). Specifically, we computed the vertical and horizontal median absolute deviation of gaze direction during target presentations. An index of gaze bias toward the stimulus was defined as the median difference of vertical gaze positions between trials in which stimuli were presented in the upper versus the lower visual field. Finally, to test whether lack of fixation compliance predicted the hypothesized effect, we tested a correlation between the individual discrimination advantage for typical feature locations and the individual magnitude of eye movements (defined as the median absolute deviation of gaze direction during stimulus presentation, averaged across the horizontal and vertical axes).

Data availability

All data and code to reproduce the results presented here can be found at <https://osf.io/cwsm2/>.

Results

We first tested our main hypothesis of a Thatcher effect induced by a mismatch of feature and retinotopic location. Specifically, we tested whether participants were better at detecting the Thatcherization of eyes in the upper compared to the lower visual field, and vice versa for mouths. A repeated-measures general linear model confirmed a significant Feature \times Location interaction, $F(1, 35) = 8.18$, $p = 0.007$ (Figure 2A). In line with our main hypothesis, participants were

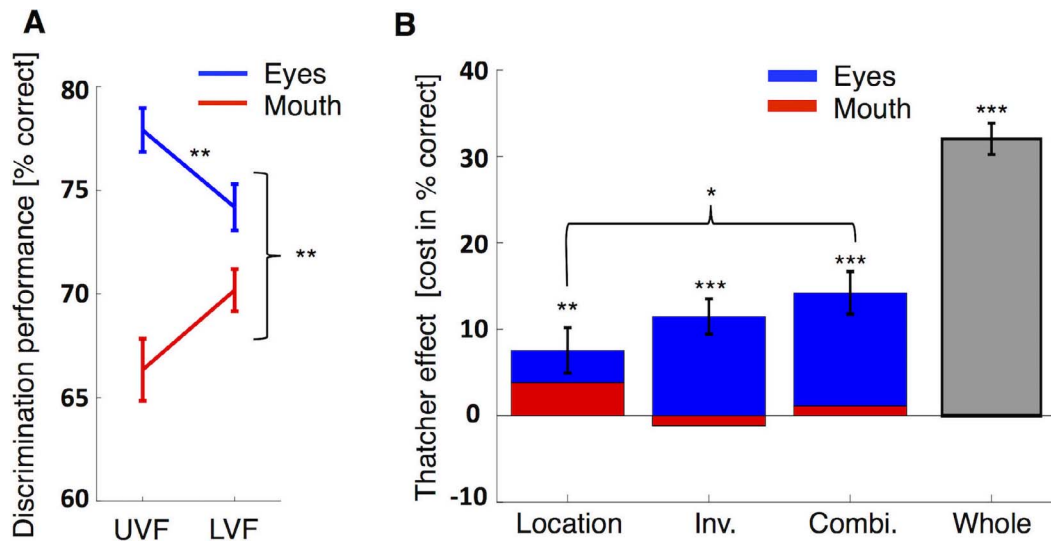


Figure 2. Behavioral results. (A) Performance for discriminating Thatcherized from unaltered eyes (blue) and mouths (red) in the upper and lower visual field. Data points and error bars indicate the mean ± 1 SEM across participants. There was a significant Feature \times Location interaction, with a significant advantage for eyes in the upper visual field and the opposite trend for mouths in the lower. (B) The Thatcher effect of different manipulations was quantified as a drop in discrimination accuracy relative to upright stimuli shown at typical visual-field locations. The first three bars show (from left to right) the sum of eye (blue, averaged across left and right eye) and mouth (red) effects for atypical locations, image inversion (Inv.), and the combination of both (Combi). All three manipulations induced a significant Thatcher effect, which was significantly larger for the combined compared to the location manipulation and approached about half that seen for the classic Thatcher effect for whole faces (rightmost bar). All bars and error bars indicate the mean ± 1 SEM across participants. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (Bonferroni corrected for multiple effect-size comparisons; see Results).

significantly better at discriminating Thatcherized from non-Thatcherized eyes in the upper visual field ($77.90\% \pm 1.40\%$ correct) compared to the lower ($74.18\% \pm 1.52\%$ correct), $t(35) = 2.80$, $p = 0.008$, and there was an opposite trend, $t(35) = 1.83$, $p = 0.08$, for mouths ($70.18\% \pm 1.40\%$ and $66.34\% \pm 1.74\%$ correct in the lower and upper visual field, respectively; percentages correct are given as mean \pm standard error of the mean throughout). An exploratory analysis of main effects showed that this also entailed a significant main effect of facial feature, $F(1, 35) = 31.34$, $p < 0.001$, with overall better performance for eyes than mouths, but no main effect of visual-field position, $F(1, 35) = 0.14$, $p = 0.96$.

Having established the predicted Thatcher effect of atypical visual-field location for facial features, we tested the effect of inversion on isolated facial regions (as reported by Psalta et al., 2014). We could replicate a significant Thatcher effect of inversion for eye regions, $t(35) = 8.21$, $p < 0.001$, but not for mouth regions, $t(35) = -0.91$, $p = 0.37$ (second bar in Figure 2B). Combining both types of manipulation also resulted in a significant Thatcher effect for eye regions, $t(35) = 8.24$, $p < 0.001$, but not for mouth regions, $t(35) = 0.69$, $p = 0.50$ (third bar in Figure 2B). Finally, we replicated the classic Thatcher effect, finding a significant advantage for discriminating the Thatcherization of full faces when

they were presented upright ($96.65\% \pm 0.60\%$ correct) compared to when they were inverted ($64.62\% \pm 1.96\%$ correct), $t(35) = 17.88$, $p < 0.001$ (rightmost bar in Figure 2B).

Testing Thatcher effects of location, inversion, and their combination for isolated facial regions in a single sample further allowed us to directly compare the magnitude of these effects in a repeated-measures fashion. To do this, we used the Thatcherization discrimination performance for upright facial regions presented at typical visual-field positions as a standard against which we compared the performance for each type of manipulation to quantify the respective Thatcher effects. Thus, the size of the location-induced Thatcher effect was defined as the accuracy advantage for typical versus reversed locations, summed across mouths and eyes (presented in upright facial regions). Its magnitude was $3.72\% \pm 1.33\%$ correct for eyes, $3.84\% \pm 2.10\%$ for mouths, and $7.56\% \pm 2.64\%$ collapsed across both features (leftmost bar in Figure 2B; effect sizes for eyes always defined as the average across left and right eye). Correspondingly, the size of the Thatcher effect induced by inversion was defined as the accuracy advantage for upright versus inverted isolated facial regions (presented at typical visual-field positions). Its magnitude was $12.62\% \pm 1.54\%$ for eyes, $-1.14\% \pm 1.26\%$ for mouths, and $11.48\% \pm 2.05\%$

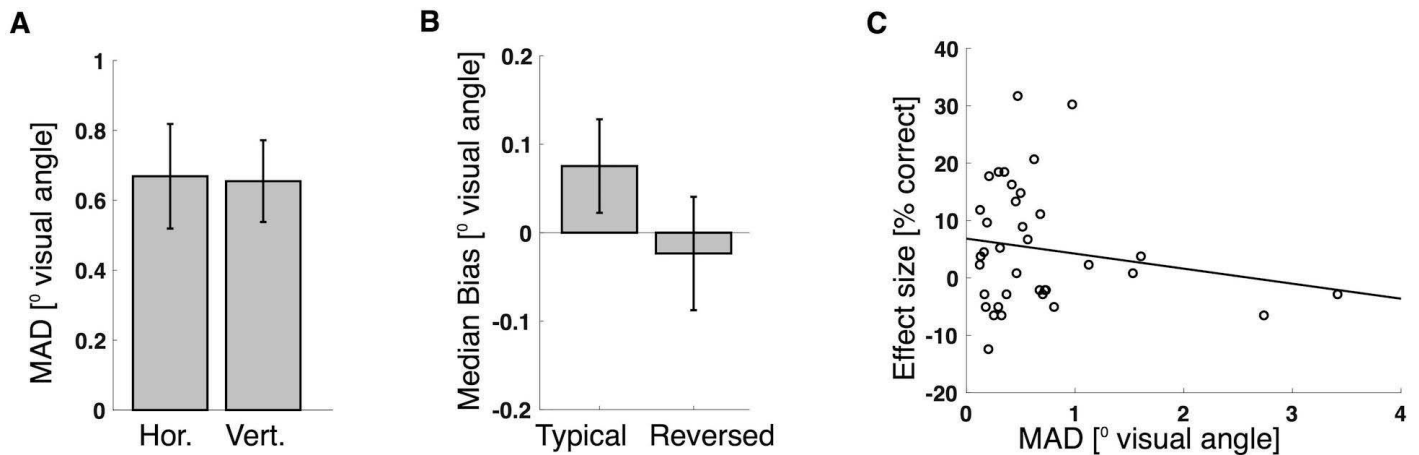


Figure 3. Fixation stability. (A) Median absolute deviation of horizontal (Hor.) and vertical (Vert.) gaze direction across stimulus presentations in degrees visual angle. (B) Median bias of vertical gaze direction toward stimulus locations for typical and reversed locations (in degrees visual angle). (C) Individual gaze stability, expressed as average median absolute deviation across the horizontal and vertical, plotted against the individual size of the hypothesized location effect. All bars and error bars indicate the mean ± 1 SEM across participants.

collapsed across both features (second bar in Figure 2B). Combining atypical visual-field positions and inversion resulted in an effect size of $13.07\% \pm 1.59\%$ for eyes, $1.14\% \pm 1.66\%$ for mouths, and $14.22\% \pm 2.46\%$ collapsed across features (third bar in Figure 2B). Finally, the magnitude of the classical Thatcher effect for full faces in our experiment was $32.03\% \pm 1.79\%$ (rightmost bar in Figure 2B).

We hypothesized that the size of the Thatcher effect for the combination of inversion and atypical location would be larger than that for either manipulation on its own. A repeated-measures general linear model on the size of Thatcher effects showed significant main effects of feature (i.e., eye or mouth), $F(1, 35) = 25.65$, $p < 0.001$; manipulation (atypical location/inversion/combination of both), $F(2, 70) = 4.38$, $p = 0.02$; and an interaction between these two factors, $F(2, 70) = 22.35$, $p < 0.001$. The preplanned comparison of manipulation effects showed that the Thatcher effect for combined inversion and atypical location was significantly larger than that for atypical location alone, $t(35) = 2.78$, $p = 0.03$ (combined across features; all Bonferroni corrected for three preplanned comparisons), but not than that for inversion alone, $t(35) = 1.31$, $p = 0.60$, and thus was only partly in line with our hypothesis. We also found no significant difference between the magnitude of inversion- and location-induced Thatcher effects, $t(35) = 1.72$, $p = 0.28$.

Post hoc t tests showed that Thatcher effects for eyes (collapsed across manipulations) were significantly larger than those for mouths, $t(35) = 5.06$, $p < 0.001$ (all Bonferroni corrected for seven post hoc comparisons), and that the Thatcher effect for eyes was larger for both inversion and the combined manipulations compared to the location manipulation alone, $t(35) >$

5.89 , $p < 0.001$. There was an opposite trend for the Thatcher effect for mouths, to be larger for the location manipulation compared to inversion, $t(35) = 2.52$, but this difference did not survive Bonferroni correction ($p = 0.12$). While the overall effect size of location- and inversion-induced Thatcher effects were not significantly different from each other (see earlier), these results point to the inversion effect being more heavily concentrated on eyes, while the location effect was more balanced between eyes and mouths.

The remaining post hoc t tests indicated no significant difference between the Thatcher effect for eyes induced by inversion compared to the combined condition, $t(35) = 0.35$, $p = 0.99$, and the same was true for mouths, $t(35) = 1.48$, $p = 0.99$. Finally, we observed no significant difference for mouths between the Thatcher effect induced by location compared to the combined condition, $t(35) = -1.65$, $p = 0.75$.

Analyses of fixation compliance showed that most observers kept very stable fixation as instructed, with a median absolute deviation of gaze position of less than 1° visual angle ($0.67^\circ \pm 0.15^\circ$ and $0.65^\circ \pm 0.12^\circ$ for the horizontal and vertical axis, respectively; Figure 3A). Furthermore, observers showed no significant bias of vertical gaze position toward stimulus location (Figure 3B) in either the typical condition, $t(35) = 1.43$, $p = 0.16$, or the reversed, $t(35) = -0.37$, $p = 0.72$, and there was no significant difference between these conditions, $t(35) = 0.71$, $p = 0.49$. There also was no significant correlation between the individual magnitude of the location-induced Thatcher effect and the variability of gaze position, $r = -0.19$, $p = 0.26$ (Figure 3C).

In summary, we found a strong feature–location interaction in Thatcherization discrimination. As predicted by our main hypothesis, discrimination perfor-

mance for eyes and mouths was better in the upper and lower visual field, respectively. We also replicated the inversion-induced Thatcher effect for isolated facial regions reported by Psalta et al. (2014), albeit only for eye regions. A general linear model comparing the size of Thatcher effects induced by atypical location and inversion showed no significant differences between them. Furthermore, the combination of these local manipulations resulted in a subadditive effect, which approached approximately half that seen for the classic Thatcher illusion in full faces. Post hoc tests further suggested that local inversion effects were almost exclusively driven by the eye region, while the effect of atypical visual-field location was more similar in magnitude across eyes and mouth. Analyses of fixation compliance confirmed that participants kept stable fixation during target presentations.

Discussion

Face-inversion effects like the Thatcher illusion are often interpreted to result from disrupted configural processing (Bartlett & Searcy, 1993; Maurer et al., 2002; Murray et al., 2000; Rossion, 2009). However, this hypothesis has recently been challenged by a Thatcher effect for inverted but isolated eye and mouth regions. This led to an alternative explanation of the Thatcher illusion as a result of orientation sensitivity for isolated facial regions (Psalta et al., 2014). Here, we tested whether another nonconfigural factor can induce Thatcher effects: retinotopic feature–location interactions. We presented isolated facial regions at retinotopic locations that matched those typical for upright or inverted faces and tested the ability to discriminate Thatcherized from unaltered features.

Participants saw brief flashes of isolated eye and mouth regions and indicated whether they were Thatcherized or not in a match-to-sample task. We found a significant Feature \times Location interaction, with better Thatcherization discrimination for eye regions in the upper visual field (Figure 2A, blue) and the opposite trend for mouth regions (Figure 2A, red). Compared to upright feature images presented at typical visual-field locations, atypical location, image inversion, and the combination of both all led to significant Thatcher effects for isolated eye regions (Figure 2B, blue bar parts). Atypical visual-field location induced a nonsignificant trend of a Thatcher effect for mouth regions, which was not seen for image inversion (Figure 2B, red bar parts). This Feature \times Manipulation interaction was significant; the Thatcher effect for mouth regions was bigger in the location compared to the inversion condition, while the Thatcher effect for eye regions was smaller in the

location compared to the inversion condition. Overall (i.e., collapsed across eye and mouth regions), there was no significant difference between the size of inversion- and location-induced effects, but the effect of the combined manipulations was significantly bigger than that of location alone.

Our results provide evidence against a purely configural account of the Thatcher illusion or an explanation based purely on the orientation of local facial regions. Thatcher effects can be found for isolated facial regions and can be induced by atypical retinotopic feature locations. Our findings suggest that high-fidelity perception of eyes and mouths is limited to upright facial regions presented at typical visual-field positions (or centrally; see Psalta et al., 2014). This is in line with recent findings of de Haas et al. (2016), who found that identity discrimination depends on typical feature locations and orientation. These nonconfigural effects may explain a considerable part of the face-inversion effect (de Haas et al., 2016; Yin, 1969), and our current results show that they generalize to the Thatcher illusion.

What explains feature–location interactions in face perception? Retinotopically specific biases (A. Afraz et al., 2010) and face-adaptation effects (A. Afraz & Cavanagh, 2009; S.-R. Afraz & Cavanagh, 2008) show that face processing is not location invariant. Furthermore, neurons in the macaque posterior lateral face patch show matching spatial and feature tuning to the contralateral eye and upper visual field (Issa & DiCarlo, 2012). This led de Haas et al. (2016) to propose a neural-tuning hypothesis for feature–location interactions. Visual-field coverage of eye- and mouth-processing neuronal populations would be biased toward the upper and lower visual field, respectively. This would reflect input regularities (cf. Arcaro et al., 2017; de Haas et al., 2016) and predict perceptual effects like the one reported here. It is further in line with human neuroimaging results suggesting a maplike organization of facial-feature representations in the occipital face area (Henriksson et al., 2015; van den Hurk et al., 2015) and that eye and mouth representations in the occipital face area are more distinguishable from each other when they are presented at typical rather than reversed visual-field locations (de Haas et al., 2016).

While we observed a significant Thatcherization-discrimination advantage for eye regions in the upper visual field, we saw only a trend toward an advantage for mouths in the lower. This difference was even more pronounced for inversion and the combined condition (Figure 2B). It is tempting to speculate about a connection between stronger effects for eye regions and the predominance of eye-preferring cells reported in macaque (Issa & DiCarlo, 2012). However, Psalta et al. (2014) found strong Thatcher effects for inverted

mouth regions, and de Haas et al. (2016) reported a significant location effect for individuating mouths. Thus, another explanation seems more likely. Our mouth-region stimuli showed little local context (Figure 1B) and were more restricted than those of Psalta et al. (2014), who used full horizontal face strips. The face outline might provide important contextual information for recognizing Thatcherization, and specifically for disambiguating it from inversion of the full feature image. This could also explain why participants were overall much better at Thatcherization discrimination for eye regions than mouth regions (Figure 2A). In our eye-region images the brow provided local context. Another factor may be that we used faces with neutral expressions and often closed lips. We used restricted mouth stimuli in order to match their size to that of our eye stimuli and keep our design as close as possible to that of de Haas et al. (2016). Nevertheless, future studies should probably use full horizontal face strips instead.

We also tested the magnitude of the Thatcher illusion for full faces in our match-to-sample design. This allows a first comparison of effect sizes between conditions. Nevertheless, it is difficult to estimate the relative contribution of local orientation, feature location, and (possibly) configural effects in the full-face condition. There were ceiling effects for upright full faces (approaching 97% correct), and it is unclear how the size of feature-specific effects translates to full-face stimuli. The effect sizes shown in Figure 2B average across the effects seen for the left and right eye regions and add to those seen for mouth regions. We do not mean to imply that this would indicate the expected contribution in the full-face situation, but rather aim to provide a convenient overview of effect sizes. Any model of full-face effects is further complicated by the fact that the combination of inversion and location effects is dramatically subadditive compared to either manipulation in isolation (cf. de Haas et al., 2016).

An additional factor to consider is that the size of our stimuli corresponded roughly to that of real faces at conversational distance. We chose this stimulus size to stay in line with salient real-life situations of scanning a face and with previously published experiments on feature–location interactions in face perception (de Haas et al., 2016). This stimulus size also afforded good control over the retinotopic stimulus locations we manipulated (controlling the exact location of stimuli presented closer to the fovea would be harder because small deviations from fixation would play a greater role). At the same time, it is important to note that many studies find a Thatcher effect for much smaller stimuli—including that of local orientation (Psalta et al., 2014). Future studies should test whether the Thatcher effect for whole

faces, as well as that for local inversion and retinal feature displacement, scales as a function of image size. If so, it would be of great interest whether these different effects scale in parallel or independently from each other.

Nevertheless, in our data and metric the size of nonconfigural, feature-based Thatcher effects is substantial. This shows that inversion effects like the Thatcher illusion were to be expected even in the absence of any configural contribution. We do not mean to deny the strong evidence for a prominent role of configural and holistic processing in face perception (Rossion, 2013; Tanaka & Farah, 1993). But inversion effects as such do not provide sufficient evidence for configural processing.

Keywords: Thatcher illusion, retinotopy, feature tuning, face perception

Acknowledgments

This work was supported by a research fellowship from the Deutsche Forschungsgemeinschaft (HA 7574/1-1; BdH), a Just’us fellowship (BdH), and European Research Council Starting Grant 310829 (DSS). The authors declare no competing financial interests.

Commercial relationships: none.

Corresponding author: Benjamin de Haas.

Email: benjamindehaas@gmail.com.

Address: Department of Psychology, Justus Liebig Universität Giessen, Giessen, Germany.

References

- Afraz, A., & Cavanagh, P. (2009). The gender-specific face aftereffect is based in retinotopic not spatio-topical coordinates across several natural image transformations. *Journal of Vision*, 9(10):10, 1–17, <https://doi.org/10.1167/9.10.10>. [PubMed] [Article]
- Afraz, A., Pashkam, M. V., & Cavanagh, P. (2010). Spatial heterogeneity in the perception of face and form attributes. *Current Biology*, 20(23), 2112–2116, <https://doi.org/10.1016/j.cub.2010.11.017>.
- Afraz, S.-R., & Cavanagh, P. (2008). Retinotopy of the face aftereffect. *Vision Research*, 48(1), 42–54, <https://doi.org/10.1016/j.visres.2007.10.028>.
- Arcaro, M. J., Schade, P. F., Vincent, J. L., Ponce, C. R., & Livingstone, M. S. (2017). Seeing faces is necessary for face-domain formation. *Nature Neu-*

- rosience*, 20(10), 1404–1412, <https://doi.org/10.1038/nn.4635>.
- Bartlett, J. C., & Searcy, J. (1993). Inversion and configuration of faces. *Cognitive Psychology*, 25(3), 281–316, <https://doi.org/10.1006/cogp.1993.1007>.
- Barton, J. J. S., Radcliffe, N., Cherkasova, M. V., Edelman, J., & Intriligator, J. M. (2006). Information processing during face recognition: The effects of familiarity, inversion, and morphing on scanning fixations. *Perception*, 35(8), 1089–1105.
- Biederman, I., & Cooper, E. E. (1991). Evidence for complete translational and reflectional invariance in visual object priming. *Perception*, 20(5), 585–593.
- Bombardi, D., Mast, F. W., & Lobmaier, J. S. (2009). Featural, configural, and holistic face-processing strategies evoke different scan patterns. *Perception*, 38(10), 1508–1521, <https://doi.org/10.1068/p6117>.
- Boutet, I., Lemieux, C. L., Goulet, M.-A., & Collin, C. A. (2017). Faces elicit different scanning patterns depending on task demands. *Attention, Perception, & Psychophysics*, 79(4), 1050–1063, <https://doi.org/10.3758/s13414-017-1284-y>.
- Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, 77(3), 305–327.
- de Haas, B., Schwarzkopf, D. S., Alvarez, I., Lawson, R. P., Henriksson, L., Kriegeskorte, N., & Rees, G. (2016). Perception and processing of faces in the human brain is tuned to typical feature locations. *Journal of Neuroscience*, 36(36), 9289–9302.
- DiCarlo, J. J., & Maunsell, J. H. R. (2003). Anterior inferotemporal neurons of monkeys engaged in object recognition can be highly sensitive to object retinal position. *Journal of Neurophysiology*, 89(6), 3264–3278, <https://doi.org/10.1152/jn.00358.2002>.
- Goffaux, V., & Rossion, B. (2007). Face inversion disproportionately impairs the perception of vertical but not horizontal relations between features. *Journal of Experimental Psychology: Human Perception and Performance*, 33(4), 995–1002, <https://doi.org/10.1037/0096-1523.33.4.995>.
- Henriksson, L., Mur, M., & Kriegeskorte, N. (2015). Faciotopy: A face-feature map with face-like topology in the human occipital face area. *Cortex*, 72, 156–167, <https://doi.org/10.1016/j.cortex.2015.06.030>.
- Hills, P. J., Cooper, R. E., & Pake, J. M. (2013). First fixations in face processing: The more diagnostic they are the smaller the face-inversion effect. *Acta Psychologica*, 142(2), 211–219, <https://doi.org/10.1016/j.actpsy.2012.11.013>.
- Hills, P. J., Sullivan, A. J., & Pake, J. M. (2012). Aberrant first fixations when looking at inverted faces in various poses: The result of the centre-of-gravity effect? *British Journal of Psychology*, 103(4), 520–538, <https://doi.org/10.1111/j.2044-8295.2011.02091.x>.
- Hsiao, J. H., & Cottrell, G. (2008). Two fixations suffice in face recognition. *Psychological Science*, 19(10), 998–1006, <https://doi.org/10.1111/j.1467-9280.2008.02191.x>.
- Issa, E. B., & DiCarlo, J. J. (2012). Precedence of the eye region in neural processing of faces. *The Journal of Neuroscience*, 32(47), 16666–16682, <https://doi.org/10.1523/JNEUROSCI.2391-12.2012>.
- Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The ventral visual pathway: An expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences*, 17(1), 26–49, <https://doi.org/10.1016/j.tics.2012.10.011>.
- Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 53(2), 513–536, <https://doi.org/10.1080/713755889>.
- Leder, H., Candrian, G., Huber, O., & Bruce, V. (2001). Configural features in the context of upright and inverted faces. *Perception*, 30(1), 73–83, <https://doi.org/10.1068/p2911>.
- Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, 6(6), 255–260, [https://doi.org/10.1016/S1364-6613\(02\)01903-4](https://doi.org/10.1016/S1364-6613(02)01903-4).
- Murray, J. E., Yong, E., & Rhodes, G. (2000). Revisiting the perception of upside-down faces. *Psychological Science*, 11, 492–496, <https://doi.org/10.1111/1467-9280.00294>.
- Peterson, M. F., & Eckstein, M. P. (2012). Looking just below the eyes is optimal across face recognition tasks. *Proceedings of the National Academy of Sciences, USA*, 109(48), E3314–E3323, <https://doi.org/10.1073/pnas.1214269109>.
- Psaltis, L., Young, A. W., Thompson, P., & Andrews, T. J. (2014). Orientation-sensitivity to facial features explains the Thatcher illusion. *Journal of Vision*, 14(12):9, 1–10, <https://doi.org/10.1167/14.12.9>. [PubMed] [Article]
- Rakover, S. S., & Teucher, B. (1997). Facial inversion effects: Parts and whole relationship. *Perception & Psychophysics*, 59(5), 752–761, <https://doi.org/10.3758/BF03206021>.
- Rhodes, G. (1988). Looking at faces: First-order and second-order features as determinants of facial appearance. *Perception*, 17(1), 43–63.

- Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2(11), 1019–1025, <https://doi.org/10.1038/14819>.
- Rossion, B. (2009). Distinguishing the cause and consequence of face inversion: The perceptual field hypothesis. *Acta Psychologica*, 132(3), 300–312, <https://doi.org/10.1016/j.actpsy.2009.08.002>.
- 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>.
- Schwaninger, A., Carbon, C. C., & Leder, H. (2003). Expert face processing: Specialization and constraints. In G. Schwarzer & H. Leder (Eds.), *Development of face processing* (pp. 81–97). Goettingen, Germany: Hogrefe.
- Schwaninger, A., & Mast, F. W. (2005). The face-inversion effect can be explained by the capacity limitations of an orientation normalization mechanism. *Japanese Psychological Research*, 47(3), 216–222, <https://doi.org/10.1111/j.1468-5884.2005.00290.x>.
- Silson, E. H., Chan, A. W.-Y., Reynolds, R. C., Kravitz, D. J., & Baker, C. I. (2015). A retinotopic basis for the division of high-level scene processing between lateral and ventral human occipitotemporal cortex. *The Journal of Neuroscience*, 35(34), 11921–11935, <https://doi.org/10.1523/JNEUROSCI.0137-15.2015>.
- Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology A*, 46(2), 225–245, <https://doi.org/10.1080/14640749308401045>.
- Thompson, P. (1980). Margaret Thatcher: A new illusion. *Perception*, 9(4), 483–484.
- Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, 79, 471–491.
- van Belle, G., De Graef, P., Verfaillie, K., Rossion, B., & Lefèvre, P. (2010). Face inversion impairs holistic perception: Evidence from gaze-contingent stimulation. *Journal of Vision*, 10(5):10, 1–13, <https://doi.org/10.1167/10.5.10>. [PubMed] [Article]
- van Belle, G., Ramon, M., Lefèvre, P., & Rossion, B. (2010). Fixation patterns during recognition of personally familiar and unfamiliar faces. *Frontiers in Psychology*, 1, 20, <https://doi.org/10.3389/fpsyg.2010.00020>.
- van den Hurk, J., Pegado, F., Martens, F., & Op de Beeck, H. P. (2015). The search for the face of the visual homunculus. *Trends in Cognitive Sciences*, 19(11), 638–641, <https://doi.org/10.1016/j.tics.2015.09.007>.
- Vieira, T. F., Bottino, A., Laurentini, A., & De Simone, M. (2014). Detecting siblings in image pairs. *The Visual Computer*, 30(12), 1333–1345, <https://doi.org/10.1007/s00371-013-0884-3>.
- Võ, M. L.-H., Smith, T. J., Mital, P. K., & Henderson, J. M. (2012). Do the eyes really have it? Dynamic allocation of attention when viewing moving faces. *Journal of Vision*, 12(13):3, 1–14, <https://doi.org/10.1167/12.13.3>. [PubMed] [Article]
- Williams, C. C., & Henderson, J. M. (2007). The face inversion effect is not a consequence of aberrant eye movements. *Memory & Cognition*, 35(8), 1977–1985, <https://doi.org/10.3758/BF03192930>.
- Xu, B., & Tanaka, J. W. (2013). Does face inversion qualitatively change face processing: An eye movement study using a face change detection task. *Journal of Vision*, 13(2):22, 1–16, <https://doi.org/10.1167/13.2.22>. [PubMed] [Article]
- Yin, R. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, 81(1), 141–145.

Appendix: Response latencies

Participants were instructed that the experiment was self-paced and they could take as much time as they wanted to respond. Additionally, responses involved a variable number of button presses to toggle a selection rectangle and confirm. Therefore, we deem the meaning of response latencies in our experiment to be very limited. Nevertheless, for completeness' sake here we report response latencies and corresponding analyses, similar to those for the proportion of correct responses in the main text. Note that data were pruned to trials with correct responses for analyses of response latencies (although control analysis showed no qualitative differences regardless of this step).

Considering upright facial regions, a repeated-measures general linear model showed a significant main effect of feature, with shorter response latencies for mouth regions compared to eye regions, $F(1, 35) = 8.83$, $p = 0.005$ (Figure A1A), but no significant main effect of location, $F(1, 35) = 0.41$, $p = 0.53$, nor a significant Feature \times Location interaction, $F(1, 35) = 0.36$, $p = 0.55$. Comparing different feature manipulations, we found no significant difference between manipulations, $F(2, 70) = 1.31$, $p = 0.27$, and no significant effect on response latencies for any of them: atypical feature locations, $t(35) = 0.60$, $p = 0.55$ (Figure A1B, leftmost bar); feature inversion, $t(35) = -0.96$, $p = 0.34$ (Figure A1B, second bar); combination, $t(35) = -0.92$, $p = 0.36$ (Figure A1B, third bar). Note that the

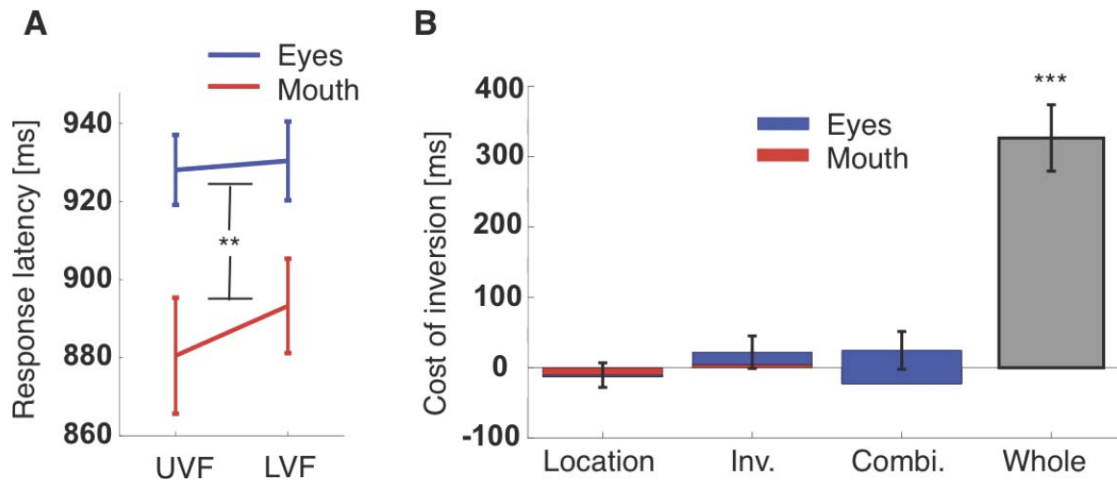


Figure A1. Response latencies. (A) Response latencies for eye regions (blue) and mouth regions (red) in the upper and lower visual field. Data points and error bars indicate the mean ± 1 SEM across participants. There was a significant main effect of feature, with shorter response latencies for mouth regions compared to eye regions. (B) Inversion cost in milliseconds for different manipulations, as compared to the response latency for upright facial regions at typical locations. There was no significant effect on response latencies for any of the local manipulations (from leftmost bar: atypical location, inversion, and combination of both), but there was for the inversion of full faces (rightmost bar). All bars and error bars indicate the mean ± 1 SEM across participants. $**p < 0.01$, $***p < 0.001$ (see Appendix).

absence of an effect for feature inversion is in contrast to the findings of Psalta et al. (2014), potentially underscoring the limited validity of response latencies in our design (see earlier). Nevertheless, we did observe

a significant modulation of response latencies for full faces, with faster responses for upright compared to inverted faces, $t(35) = 6.92$, $p < 0.001$ (Figure A1B, rightmost bar).