Vision Research 227 (2025) 108535

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Perceived intrinsic 3D shape of faces is robust to changes in lighting direction, image rotation and polarity inversion

Jordi M. Asher [a,b,*], Paul B. Hibbard [a,b], Abigail L.M. Webb [b,c]

[a] University of Stirling, Department of Psychology, Stirling, FK9 4LA, Scotland, UK
[b] University of Essex, Department of Psychology, Wivenhoe Park, Colchester, CO4 3SQ, Essex, UK
[c] University of Suffolk, Institute of Health and Wellbeing, Ipswich, IP4 1QJ, Suffolk, UK

ARTICLE INFO

Keywords:
Pictorial relief
Affine transformation
3D surface shape
Face recognition
Shape from shading

ABSTRACT

Face recognition from 2D images is influenced by various factors, including lighting conditions, viewing direction, rotation, and polarity inversion. It has been proposed that these techniques affect face recognition by distorting shape from shading. This study investigates the perception of 3D face shape in 2D images using a gauge figure task. Two experiments were conducted where participants adjusted a gauge figure across multiple locations within a 3D image to assess its surface structure. We manipulated face orientation, lighting direction, and polarity inversion (exp 2). While these manipulations resulted in variations from the true surface structure, they could be explained by an affine transformation. This suggests that the perception of the intrinsic 3D shape of faces is stable across these image manipulation techniques. The effects of viewing conditions on face recognition may thus be better interpreted through their influence on the perception of material properties such as pigmentation, or on information closer to the level of the retinal image itself.

1. Introduction

Face perception is fundamental to human interaction. Faces allow us to recognise others, appreciate their emotions and navigate complex social cues. While a seemingly effortless human skill, recognition is a challenging computational problem, that needs to be stable across changes in pose angle, viewing direction and lighting conditions, amongst other factors. A representation of the 3D surface shape of a face may be important in achieving this stability (Johnston, Hill, & Carman, 1992). When considered as a physical structure, a face can be defined by its three-dimensional surface shape, biological characteristics such as pigmentation, and its optical material properties, such as its reflectance and glossiness (Bruce & Langton, 1994; Henderson, Holzleitner, Talamas, & Perrett, 2016). These factors combine with the lighting and viewing conditions to create the image of a viewed face. It is a challenging computational task to recover 3D shape from 2D images.

Face recognition from 2D images is affected by many factors including lighting and viewing direction, but also rotation and polarity inversion of images (Bruce & Langton, 1994; Favelle, Hill, & Claes, 2017; Johnston et al., 1992; Kemp, Pike, White, & Musselman, 1996; Liu, Collin, & Chaudhuri, 2000; O'Toole, Vetter, & Blanz, 1999; Palmer, Goddard, & Clifford, 2022; Peterson, Susilo, Clifford, & Palmer, 2023; Russell, Biederman, Nederhouser, & Sinha, 2007; Russell & Sinha, 2007; Sinha, Balas, Ostrovsky, & Russell, 2006). These effects have been attributed to their influence on the perception of both the 3D surface structure and the material properties of the face.

Image manipulation strategies

Lighting that illuminates a face from above the head casts shadows and highlights that align with our common experience and expectations of how faces are typically seen and lit (Berbaum, Bever, & Chung, 1983; Gibson, 1950; Mamassian & Goutcher, 2001; Ramachandran, 1988). In contrast, when light comes from below the chin, it creates an unusual pattern of shadows and highlights. Lighting faces, or any shape, from below disrupts our ability to extract shape from shading, and the pattern of shading and shadows across the face (Palmer et al., 2022). These changes in turn impair the recognition of faces.

A second method to investigate recognition is the **inversion (rotation)** effect whereby recognition is poorer for upside down rather than upright faces. This shows that facial recognition relies on a degree of holistic information, for example the spatial configuration of key facial features and their collective organisation. Favelle et al. (2017) compared the effects of viewing direction, lighting direction, and image rotation using scanned faces rendered with uniform reflectance. Image rotation had the largest effect on recognition, while lighting direction had less of an effect than the orientation of the face. The authors

* Corresponding author at: University of Stirling, Department of Psychology, Stirling, FK9 4LA, Scotland, UK.
E-mail address: jordi.asher@stir.ac.uk (J.M. Asher).

https://doi.org/10.1016/j.visres.2024.108535
Received 29 July 2024; Received in revised form 10 December 2024; Accepted 10 December 2024
Available online 24 December 2024
0042-6989/

suggested that changes in lighting alters the depth and contours of the face and therefore the effects can best be explained in terms of their effects on perceived 3D shape. Johnston et al. (1992) assessed the effects of rotation on recognition for photographs of faces lit from above, in front, or below. When inverting the orientation of faces, top-lit photographs were affected by rotation, but bottom-lit photographs were not. These results were interpreted in terms of the role of shape-from-shading. It was proposed that lighting from below disrupts the formation of a representation of 3D shape, due to the default assumption that scenes are lit from above (Berbaum et al., 1983; Gibson, 1950; Mamassian & Goutcher, 2001; Ramachandran, 1988).

A third method is the use of **photographic negatives.** Inverting the luminance of an image reverses light and dark areas from the original scene and is known to substantially impair recognition (Johnston et al., 1992; Kemp et al., 1996; Sandford & Rego, 2019). Inverting the luminance polarity creates a complicated and unfamiliar change in the image that cannot be attributed to a simple change in the lighting direction. This plausibly explains why the perception of shape and structure is substantially disrupted in luminance inverted images. In contrast, while luminance inversion impaired recognition, the inversion of colour hue did not (Kemp et al., 1996). The authors of this study concluded that, since only luminance provides shape information, polarity inversion impedes face recognition through its effect on 3D shape representation.

Taken together, these experiments suggest that errors in recognition that result from these manipulations are the result of errors in the representation of 3D shape. These errors in 3D shape are however inferred from performance on face recognition tasks, but have not been tested directly.

1.1. 3D shape perception

The studies summarised above emphasise the inherent difficulty in estimating the 3D shape and material properties of a face from a single 2D image. Yet, despite the ambiguity of this information, we experience a stable, unambiguous percept of the 3D shape. Pictorial relief describes this stable perception of 3D shape in 2D images. Visual cues within the image such as shading, perspective, and texture gradients provide cues for the brain to perceptually interpret depth and distance within the image (Koenderink, 1998; Koenderink & Van Doorn, 1995). Multiple techniques have been developed for measuring pictorial relief (Koenderink, 1998; Koenderink, Van Doorn, Kappers, & Todd, 2001) and how it is affected by changes in viewing conditions such as lighting and viewing direction (Todd, Koenderink, Van Doorn, & Kappers, 1996). In one of the most commonly used tasks, the 'Tissot Indicatrix' (or the 'gauge figure task'), observers rotate the three-dimensional orientation of an elliptical gauge figure so that is appears to lie on a surface (Koenderink, Van Doorn, & Kappers, 1992; Mingolla & Todd, 1986). The orientation of the gauge figure may be described in terms of its slant and tilt. Slant is the angle of rotation out of the reference plane. Tilt is the orientation of the surface normal projected into the reference plane. Slant and tilt together define a unique 3D surface orientation. The joint slant-tilt vector defines a point on the surface of a unit sphere (Fig. 1). Local measurements of slant and tilt of the gauge figure can be used to recover the global relief structure of the surface (Koenderink et al., 1992). The gauge figure task therefore provides a direct way to assess whether the influence of lighting, viewing direction, image rotation and polarity inversion on face recognition are the result of their effect of the perception of 3D shape, as has been proposed (Johnston et al., 1992; Kemp et al., 1996).

1.2. Incidental versus intrinsic differences in pictorial relief

Pictorial relief is ambiguous. For a given image of a 3D surface, there is an infinite number of other shapes that, combined with appropriate changes to the lighting conditions or viewing direction, would give rise to an identical image (Belhumeur, Kriegman, & Yuille, 1999). This ambiguity is restricted to a generalised bas-relief transformation, that allows for a stretching and shearing of the surface in the depth dimension. This transformation consists of four components (see Fig. 2): (a) **Translation:** Shifting all points by a fixed vector; this error is a shift in distance but does not change the shape of the object; (b & c) **Horizontal and Vertical Depth shear:** Shifting each point in depth by an amount that is proportional to its horizontal or vertical location; these shears can be interpreted as a rotation of the observer's location around a vertical or horizontal axis, respectively. (d) **Stretching in depth:** This elongates or flattens the object in the depth direction.

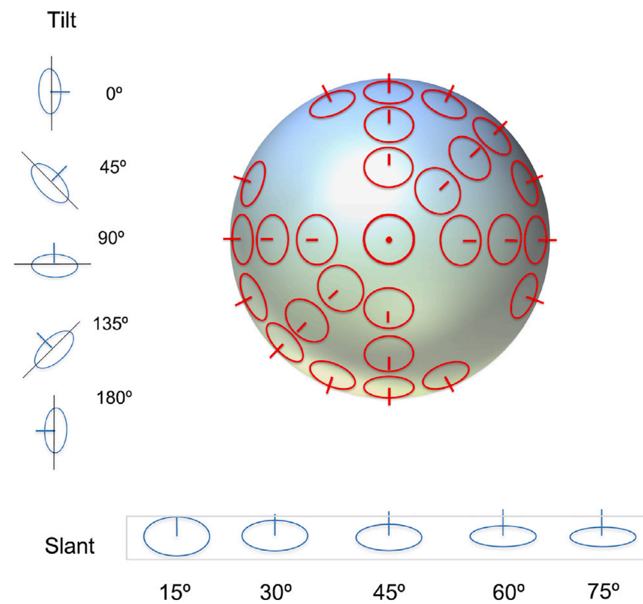

Fig. 1. Using a gauge figure to obtain slant and tilt.
Source: Adapted from Burge, McCann, and Geisler (2016) under license CC BY-NC-ND 4.0.

Belhumeur et al. (1999) show how multiple affine transformations of a face, accompanied by corresponding changes in scene lighting, would give rise to the same retinal image (their figure 2). Koenderink and Van Doorn (2003) show that the same affine transformation can also result from a change of viewing direction relative to the face. This means that our ability to extract information about the three-dimensional shape of a face from a single image is subject to a fundamental ambiguity that is equivalent to a change in lighting, viewing direction, and/or pigmentation (Georghiades, Belhumeur, & Kriegman, 2001).

When we assess the difference between the surface shape in pictorial space and the true depth from which the image was created, it is thus helpful to distinguish between the component that can be accounted for by an affine transformation, and the residual difference that remains once this affine transformation has been taken into account. The affine component captures the ambiguity in recovering 3D surface shape from a 2D image, and may be interpreted in terms of changes in lighting direction or the position of the observer relative to the object. The (non-affine) differences that remain once the incidental (affine) differences have been factored out reflect changes in the perception of the intrinsic surface shape. In the case of face recognition, we would thus expect any image manipulations that degrade recognition through their effects on perceived surface shape to do so via intrinsic rather than incidental changes. That is, only effects on pictorial relief that cannot be accounted for by an affine transformation would be expected to impede face recognition. The affine transformation, while altering metric depth, preserves key shape properties for recognition, such as parallelism, ratios of distances, planarity, and the locations

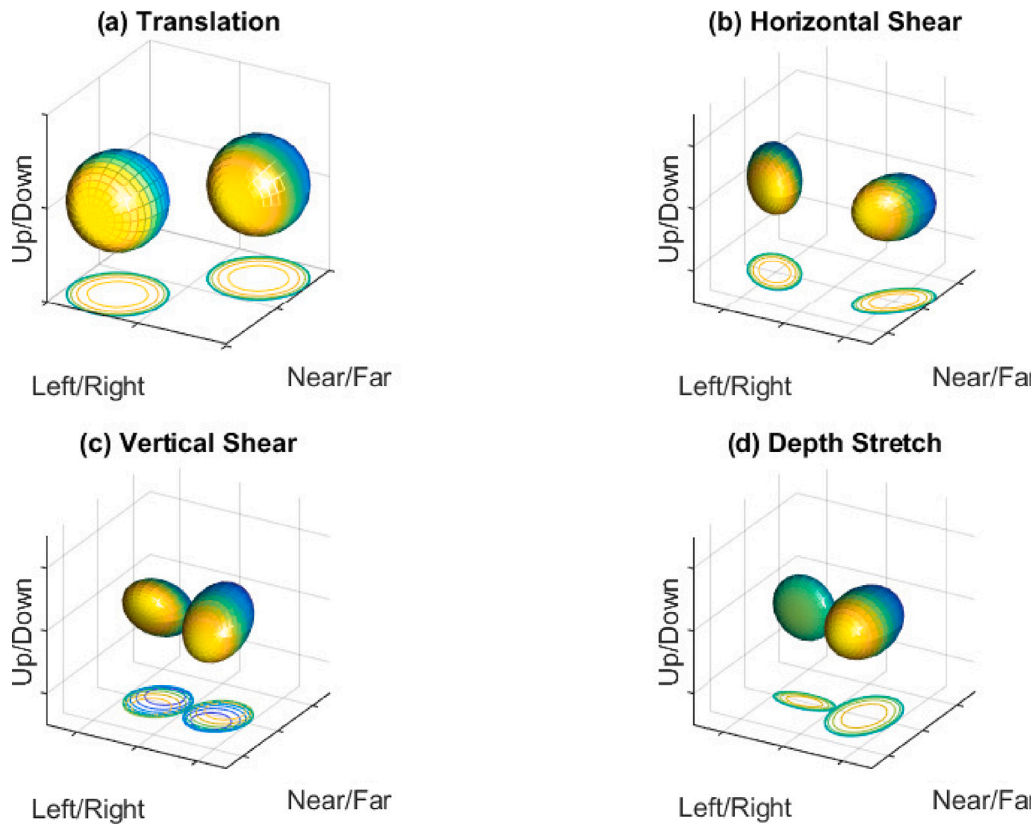

Fig. 2. The four components of the affine transformation used to distinguish between incidental (affine) and intrinsic (non-affine) differences between 3D surfaces. **(a) Translation**, in which each point is shifted in depth by an equal amount **(b) Horizontal Shear**, in which the change in depth of each point depends on it horizontal position **(c) Vertical Shear**, in which the change of depth depends on vertical position and **(d) Depth stretching or flattening**, in which the depth differences between points are magnified or minified.

of concavities and convexities (Todd & Petrov, 2022). Consistent with these invariances, previous research has shown that the recognition of faces from photographs is very robust to stretching and shearing transformations in the image plane (Hole, George, Eaves, & Rasek, 2002; Sandford & Rego, 2019; Sandford, Sarker, & Bernier, 2018).

1.3. Current study

In the current study, we used these well-established techniques for the measurement of pictorial relief to assess the effects of lighting, image rotation and polarity inversion on the perceived surface shape of faces. These techniques allow us to measure the effects of viewing conditions on the perception of 3D shape directly, rather than inferring this from their influence on face recognition.

In the first experiment, we assessed the effects of image rotation, and lighting from above or below, on pictorial relief for a stimulus with a uniform reflectance.

2. Experiment one: Image rotation and inversion of lighting direction

Method

2.1. Participants

10 participants completed the study, (7 female) with a median age of 25.5 (Mean = 29.6, SD = 9.94). This study was approved by the University of Essex Ethics Sub-committee 3 (ETH2122-0448). All participants provided written, informed consent.

2.2. Apparatus

Stimuli were presented on a 21" iMac with a display resolution of 4096 × 2304 pixels and a 60 Hz refresh rate. Stimuli were generated and presented with Matlab 2020b, using Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) on a 2.7 GHz IMac running MacOS Big Sur 11.7. Stimuli were viewed from a distance of 400 mm, and one pixel subtended 0.94 arc min.

2.3. Stimuli

The stimulus was created from a scan of a real male face. This was the demonstration scan from 3DScanStore (2020), obtained in 2019–2020, and was used with permission. A copy of this model can be found here (Asher, Hibbard, & Webb, 2024). The Wavefront .obj file three-dimensional mesh of this face was loaded into MATLAB and rendered using OpenGL via the Psychophysics toolbox extensions (Hibbard, Goutcher, Hornsey, Hunter, & Scarfe, 2023). The face was positioned at a distance of 20 cm from the camera, rendered with a uniform matte lambertian mid-grey texture and presented against a grey background. OpenGL lighting with a (0.3, 0.3, 0.3) magnitude ambient component and a (0.7, 0.7, 0.7) diffuse component was used. The diffuse component provides the contribution that depends on the orientation of the surface relative to the directional light source, which in this case was a spotlight located 100 cm at an angle of 45° above or below the centre of the scene and directed at the centre of the screen. The true distance for each pixel was calculated by ray tracing the three-dimensional coordinates of each vertex to the image plane. Stimuli were lit from above or below, and presented either upright or rotated through 180°. In defining the lighting direction, we always define this

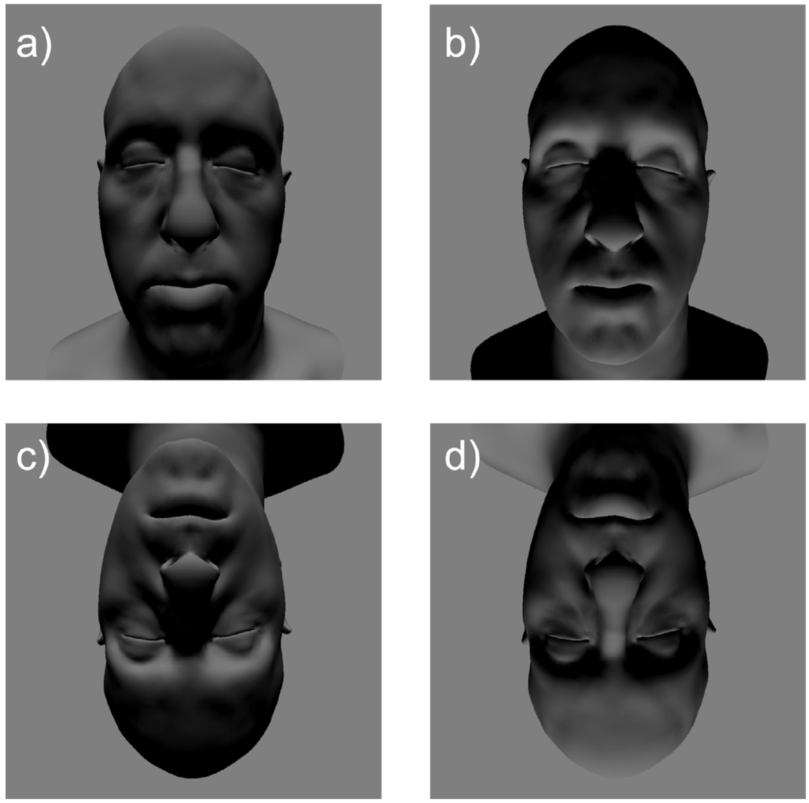

Fig. 3. The face stimuli (3DScanStore, 2020; Asher et al., 2024) used in Experiment 1: Upright (a) lit from above, (b) lit from below. Rotated (c) lit from above (d) lit from below.

relative to the image. Thus, for a face that is rotated through 180° the lighting comes from the direction towards the chin in the lit from above condition, and the direction towards the forehead in the lit from below condition. They were presented as a 512×512 pixel image in the centre of the screen (see Fig. 3).

2.4. Design

The purpose of the first experiment was to use local measures of perceived 3D orientation to create global depth relief maps that could be compared across conditions and against the true depth. These were used to assess (1) how closely each relief map matched the true depth (2) the extent to which deviations from true depth could be explained by the bas-relief ambiguity in terms of an affine transformation, rather than a more fundamental change in the depth structure (3) the nature of the affine transformation between true depth and perceived surface relief and (4) how each of these metrics was affected by the orientation of the stimulus and the lighting direction (lit from above versus below).

2.5. Procedure

On each trial, the face stimulus was presented in the centre of the screen. The gauge figure was superimposed on the stimulus in red. Participants indicated the apparent 3D orientation of the surface using a gauge figure that consisted of an ellipse and a straight line. The gauge figure was presented in red; the thickness of the line was 1 pixel. The length of the line when it was parallel to the screen was 10 pixels, and the diameter of the ellipse, when circular, was 20 pixels. The participant's task was to orient the gauge figure so that the ellipse appeared aligned with the orientation of the surface, and the line coincident with the surface normal (Hibbard, Hornsey, & Asher, 2023). A triangular mesh grid, with an edge length of 60 pixels, was created within an elliptical region that covered the central region of the face, without extending beyond the chin or ears. The number of triangular faces in the mesh was 371. The barycentres (centres of mass) of the triangles were used as the sampling locations for the gauge figure. On each trial, a barycentre was chosen at random as the location for the gauge figure. Participants used a mouse to vary the slant and tilt of the gauge figure, taking as long as they required. When they were happy that the gauge figure appeared aligned with the surface, they clicked the mouse. This recorded their response and moved the gauge figure to a new location for the next trial (see Fig. 4 for example). Each block of trials consisted of a single gauge figure setting for each barycentre.

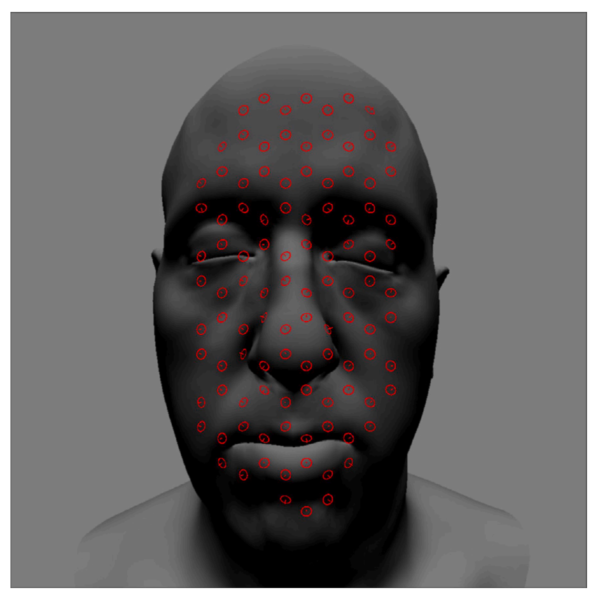

Fig. 4. The face stimuli with sample gauge placement.

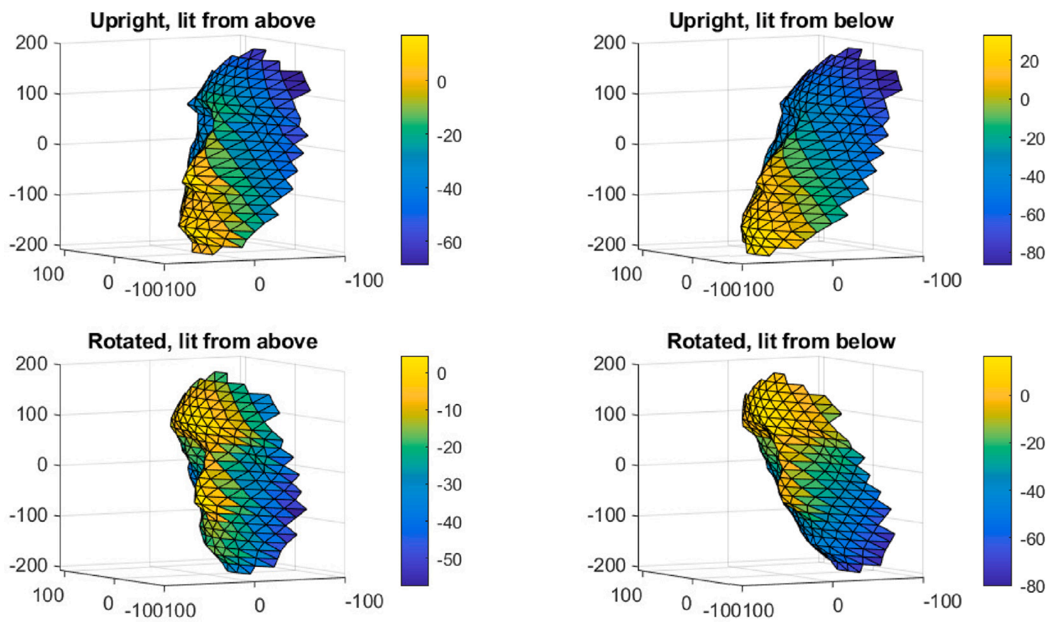

Fig. 5. Mean depth relief from participants' settings prior to affine transformation. All units are in pixels. A comparison between these results and the mesh derived from the true depth (6) shows a close agreement between participant's responses and the underlying surface shape, up to an affine shearing that varies between viewing conditions.

Prior to the trial blocks, the task was explained and demonstrated to the participant by the experimenter. The participant then completed a few practice trials of their own until they were happy that they understood the task. For each face, all trials were presented in a single block. Participants were instructed to take as many, self-paced breaks between trials as they needed, and also to take breaks between blocks.

2.6. Results

For each trial, the slant and tilt of the gauge figure were recorded and used to generate the depth relief map providing the best fit with the orientation settings made (Koenderink et al., 1992; Nefs, 2008). This map consists of a set of connected triangular faces, each defined by three vertices. The depth relief was determined by finding the vertex depth coordinates that provided the best fit with the gauge figure orientation settings. Fig. 5 shows the relief maps for the four viewing conditions (all combinations of upright and rotated faces, with lighting from above or below). In all cases, the 3D surface structure of the face is well-captured by participants' settings, and clearly visible in the relief maps.

Relief maps based on psychophysical data were compared against the true depth, created as follows. For each stimulus, the depth value of each pixel (its distance from the observer in the depth dimension) is known. These values were used to provide an estimate of the slant and tilt of the surface in the region of each barycentre. A linear regression was performed on the depth values within a square 20×20 pixel region surrounding the barycentre, with the horizontal and vertical location of each pixel relative to the barycentre as predictors. An intercept term was included in the regression. This regression provides an estimate of the depth gradient in the horizontal and vertical directions (g_x and g_y respectively), as well as the distance at its centre. The estimated horizontal and vertical gradients were used to calculate slant (σ) and tilt (τ) values (Stevens, 1983):

$$\sigma = \tan^{-1}\sqrt{g_x^2 + g_y^2} \tag{1}$$

$$\tau = \tan^{-1}\left(\frac{g_y}{g_x}\right) \tag{2}$$

We then used the same regression algorithm as was used for the psychophysical data to create a true depth relief map from these values. This provides us with the surface that we would expect if participants' slant and tilt settings matched the true local depth orientation. This true depth relief map is shown in Fig. 6.

We calculated a number of statistics to assess the relationship between the true depth and psychophysical relief maps, for each participant in each condition. The first was a simple correlation, to assess how well the psychophysical depth coordinates (Z_P) reflected the true depth coordinates (Z_G).

The correlations were analysed using a linear mixed effects model, with viewing condition as a categorical predictor, and random intercepts across participants (Table 1, Fig. 7a). In the upright, lit from above condition, there was a strong positive correlation (0.68). This was reduced by lighting from below, and was lowest when the image was both lit from below and rotated.

This same deviation from true depth was also seen in the Root Mean Square (RMS) differences between the pictorial relief and true depth (Table 1, Fig. 7b), which were lowest in the upright, lit from above condition, and increased with lighting from below for both upright and rotated faces.

To determine the degree to which this mismatch could be accounted for by the bas-relief ambiguity (Belhumeur et al., 1999; Koenderink et al., 2001) we calculated the affine transformation that provided the best fit between the psychophysical and true depth relief maps:

$$Z_p = b_0 + b_X X + b_Y Y + b_Z Z_G \tag{3}$$

Data were centred on the point closest to the median X and Y values prior to performing the regression. The RMS error following the application of this transformation was also calculated. This determines the extent to which the mismatch between psychophysical and true depth maps could be accounted for by the bas-relief ambiguity, rather than a more complex change in depth structure, and how this varied across conditions. The regression parameters also provide a description of the affine relationship between the true depth and psychophysical relief maps. b_0 is a simple shift in distance, which does not alter the shape of the relief map. b_x and b_y define depth shears, and b_z is a stretching of the relief map in the depth direction (see Fig. 2a).

The best fitting affine transformation was then applied to align the true depth and psychophysical relief maps. This substantially reduced the RMS difference between the two, such that the difference following

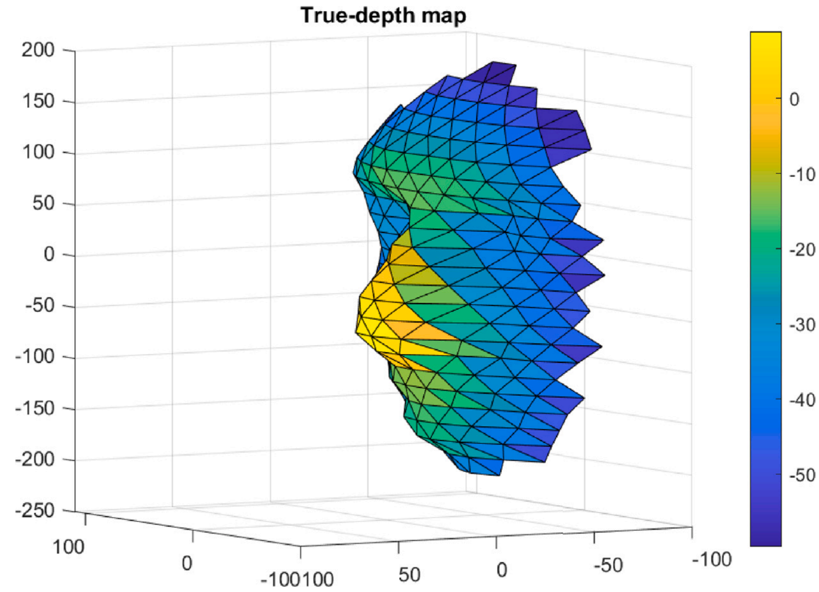

Fig. 6. True depth relief map for the face stimulus. All units are in pixels.

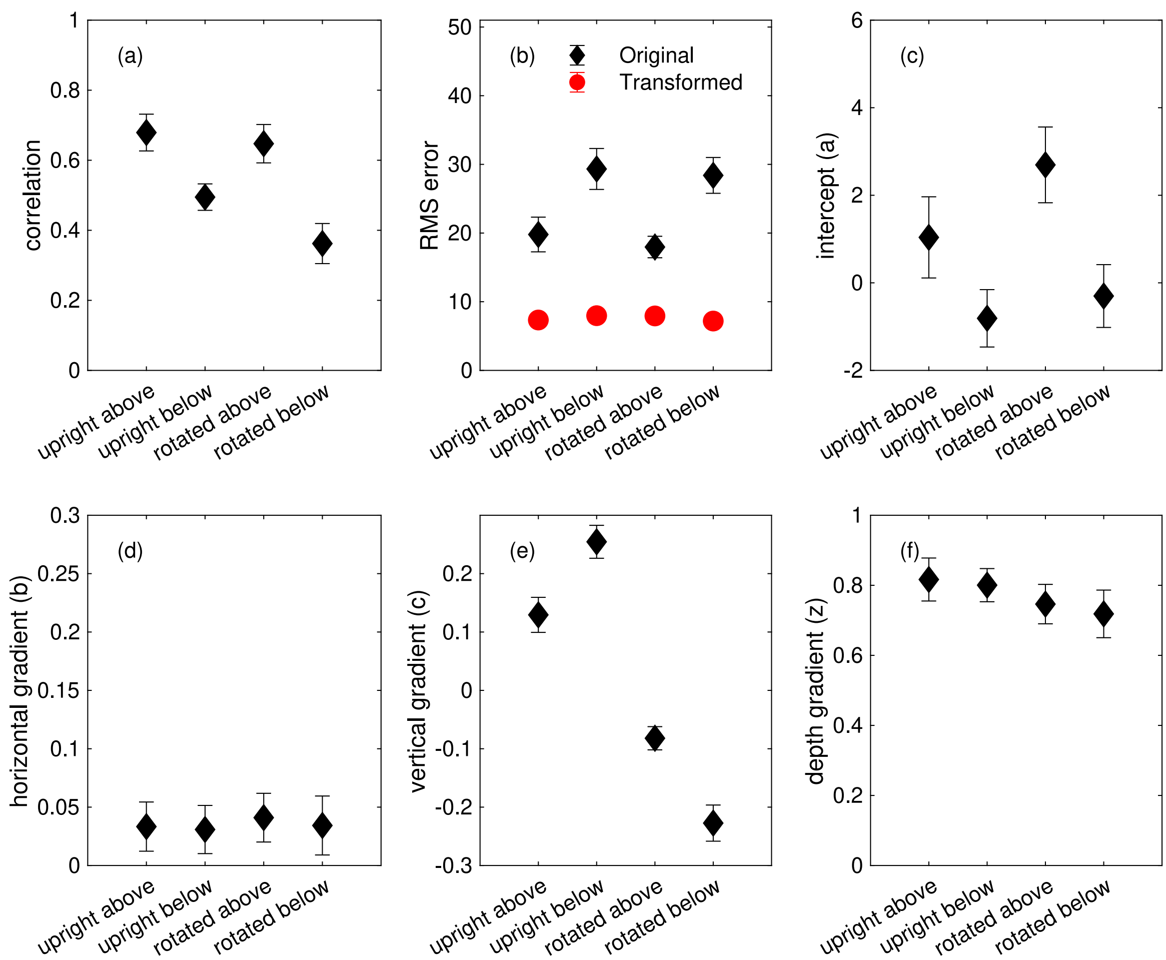

Fig. 7. Plots showing the **(a)** correlation between psychophysical depth and the true depth coordinates **(b)** original RMS error and the transformed error (after affine transformations). The four affine transformations are shown for the **(c)** Intercept, **(d)** Horizontal shear, **(e)** Vertical shear **(f)** Depth Stretch. The most salient effect of viewing condition is a rotation about a horizontal axis when the face is lit from below rather than above (vertical shear). The positive values when upright, and negative values when rotated, both represent a rotation of the face to look upwards in physical space, as seen in figure 5. Confidence intervals for the model fits are provided in Tables 1–2 Error bars are $\pm SEM$.

transformation was reduced to 37% of the original value. This shows that the majority of the deviation from the true depth relief maps could be explained by the bas-relief ambiguity, rather than an intrinsic change in the apparent 3D surface structure of the face. Moreover, there

Table 1
Experiment 1 (orientation & lighting direction) correlation, original RMS and transformed RMS.

(a) Condition	Correlation	SE	t	df	p	Lower	Upper
intercept[a]	0.679	0.049	14.00	36	< **0.001**	0.581	0.777
Upright below	−0.184	0.0425	−4.33	36	< **0.001**	−0.270	−0.098
Rotated above	−0.032	0.0425	−0.747	36	0.46	−0.118	0.054
Rotated below	−0.317	0.0425	−7.45	36	< **0.001**	−0.401	−0.231
Condition	Original RMS	SE	t	df	p	Lower	Upper
intercept[a]	19.79	2.35	8.42	36	< **0.001**	15.03	24.56
Upright below	9.542	1.98	4.82	36	< **0.001**	5.53	13.56
Rotated above	−1.83	1.98	−0.922	36	0.362	−5.84	2.19
Rotated below	8.60	1.98	4.34	36	< **0.001**	4.59	12.62
Condition	Transformed RMS	SE	t	df	p	Lower	Upper
intercept[a]	7.34	0.48	15.15	36	< **0.001**	6.35	8.32
Upright below	0.63	0.47	1.35	36	0.184	−0.31	1.57
Rotated above	0.59	0.47	1.27	36	0.211	−0.35	1.54
Rotated below	−0.15	0.47	−0.32	36	0.748	−1.09	0.79

[a] The intercept is the Upright lit from Above condition. See Fig. 7a & b.

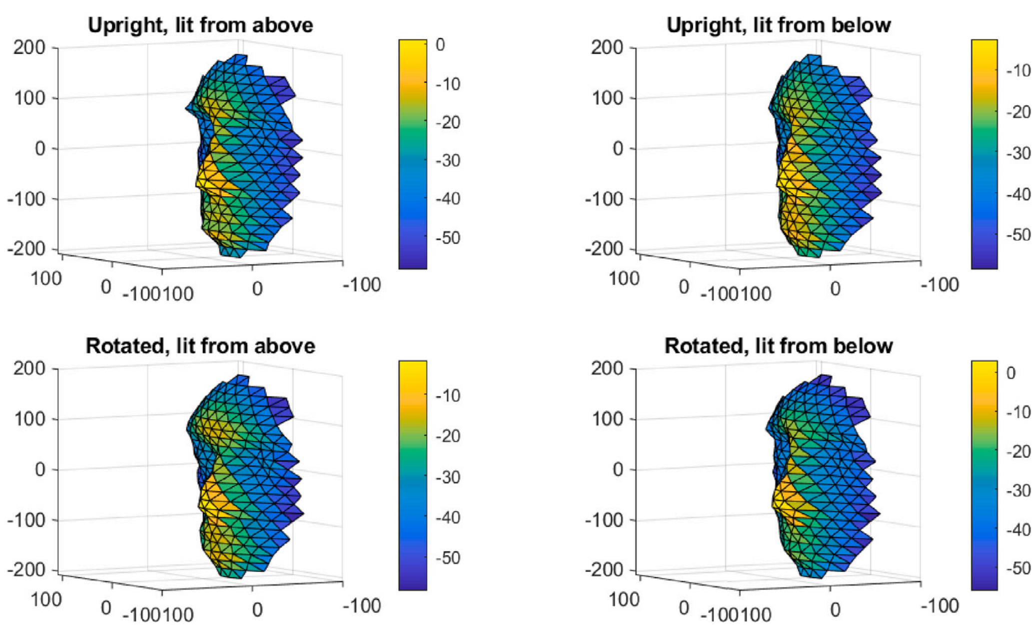

Fig. 8. Mean depth relief from participants' settings, following affine transformation to align with the true depth. All units are in pixels.

was no significant variation in the RMS difference across conditions following the transformation. The differences between relief maps that were found can therefore be characterised as an affine transformation of the relief map between conditions (Fig. 7b, Table 1).

Fig. 7(c) b_0 shows the error in the intercept, which is a shift in distance that does not alter the shape of the relief map (also see Table 2).

We also assessed the nature of the affine transformation that aligned the psychophysical relief maps with the true depth. The horizontal gradient of depth b_x is a rotation of the image surface around a vertical axis relative to the observer (a 'shaking of the head'). No significant horizontal gradient, or variation across conditions, was found (see Fig. 7d, Table 2 b).

The vertical gradient of depth b_y is a rotation of the image surface about a horizontal axis relative to the observer (a 'nodding of the head'). A positive value was found in the upright conditions and a negative value in the rotated conditions (see Fig. 7e, Table 2).

Finally, the b_Z gradient indicates how apparent depth varies with true depth. With perfect alignment, this parameter would have a value of 1. On average, this value was 0.77, indicating that the psychophysical relief maps were flattened relative to the true depth. Across the four viewing conditions, b_Z was reduced by around 9% in the two rotated conditions, indicating a slightly flatter depth relief (see Fig. 8).

Summary & Interim discussion

Experiment 1 explored effects of stimulus inversion and lighting direction on a gauge figure surface orientation task using a matte grey 3D face image. The effect of image rotation can be summarised as a small rotation of the face around a horizontal axis, combined with a flattening of the depth relief. No effect of lighting direction was observed.

Depth relief maps in all conditions were highly correlated with the true depth, indicating observers' ability to accurately perceive the three-dimensional surface structure of the face. Deviations from true depth were lowest in the upright, lit-from-above condition, and highest when lit from below, consistent with the notion that lighting an object from below disrupts the ability to discern shape from shading distributed across a surface (Gibson, 1950; Palmer et al., 2022). The differences between conditions were fully accounted for by an affine transformation, which consisted of a slight rotation of the image plane about a horizontal axis, and a squashing in the depth dimension, when the image was rotated.

Table 2
Experiment 1 (orientation & lighting direction) affine transformations.

(a) Translation							
Condition	b_0	SE	t	df	p	Lower	Upper
intercept[a]	1.0382	0.758	1.37	36	0.18	−0.50	2.58
Upright below	−1.8487	0.997	−1.85	36	0.072	−3.87	0.17
Rotated above	1.6562	0.997	1.66	36	0.11	−0.37	3.68
Rotated below	−1.3397	0.997	−1.34	36	0.19	−3.36	0.68
(b) Horizontal Shear							
Condition	b_X	SE	t	df	p	Lower	Upper
intercept[a]	0.033	0.021	1.59	36	0.120	−0.009	0.076
Upright below	−0.0025	0.011	−0.237	36	0.814	−0.024	0.019
Rotated above	0.0077	0.011	0.724	36	0.474	−0.138	0.029
Rotated below	0.001	0.011	0.090	36	0.929	−0.021	0.022
(c) Vertical Shear							
Condition	b_Y	SE	t	df	p	Lower	Upper
intercept[a]	0.129	0.0261	4.938	36	< **0.001**	0.0762	0.182
Upright below	0.125	0.037	3.38	36	**0.0018**	0.050	0.200
Rotated above	−0.211	0.037	−5.71	36	< **0.001**	−0.287	−0.136
Rotated below	−0.357	0.037	−9.63	36	< **0.001**	−0.432	−0.282
(d) Depth Shear							
Condition	b_Z	SE	t	df	p	Lower	Upper
intercept[a]	0.817	0.056	14.66	36	< **0.001**	0.704	0.930
Upright below	−0.016	0.043	−0.371	36	0.713	−0.104	0.072
Rotated above	−0.070	0.043	−1.617	36	0.115	−0.159	0.018
Rotated below	−0.098	0.043	−2.26	36	**0.030**	−0.186	−0.010

[a] The intercept is for the Upright lit from Above condition.
Where r represents the b_Z gradient, how apparent depth varies with true depth (Stretch). Also see Fig. 7(c–f).

As such, these differences do not reflect a change in the intrinsic surface structure, since they are consistent with the bas-relief ambiguity in shape from shading, and may be characterised either as a change in the viewing direction (Koenderink et al., 2001), lighting direction or surface pigmentation (Belhumeur et al., 1999). These results suggest the effects of image rotation and lighting direction that have been found in face recognition are not necessarily a result of changes in the intrinsic 3D surface structure (Johnston et al., 1992).

The stimuli in the first experiment were created using a 3D model of a scanned face, and the surface was a uniform, matte grey material. This provided information about three-dimensional shape-from-shading only. However, the perception and recognition of faces depends not only on shape information, but also variations in colour and reflectance across the surface (also referred to as pigmentation cues) (Bruce & Langton, 1994; Favelle et al., 2017; Liu, Collin, Burton, & Chaudhuri, 1999; O'Toole et al., 1999; Russell, Sinha, Biederman, & Nederhouser, 2006). We replicated this experiment with full-colour stimuli, to assess whether similar effects are evident when both sources of information are available.

3. Experiment two: Image rotation, lighting direction and polarity inversion

3.1. Introduction

In the first experiment we used a matte greyscale face, this removed any surface reflectance and texture. For the second experiment, we extended this to include faces that also contained spatial variations in colour and reflectance, with lighting from the front and the side, and to assess the effects of polarity inversion.

Along with replicating the findings in the first experiment, these experiments also provide a direct test of the hypothesis that changes in lighting direction affect face recognition through the disruption of the perception of 3D surface shape.

The second experiment had three goals. The first was to provide a replication of experiment one, for colour stimuli that included surface material (pigmentation) information.

Our ability to detect and recognise a face depends on its 3D surface shape and variation in pigmentation, and the two-dimensional image features that these create (Bruce & Langton, 1994; Favelle et al., 2017; Liu et al., 1999; O'Toole et al., 1999; Palmer et al., 2022; Russell et al., 2006). In the first experiment, by using a uniform matte grey reflectance pattern, we isolated 3D surface shape available from shape-from-shading and edge contours (Nefs, 2008). As in previous studies using this technique, excluded other cues to 3D shape, such as colour and surface texture. The second experiment therefore used faces with non-uniform reflectance to assess the effects of lighting and orientation when multiple pictorial depth cues were available.

The second experiment also explored the effect of lighting direction in more detail. Lighting from below can have pronounced effects on shape-from-shading because it conflicts with our more typical experience of lighting from above (Mamassian & Goutcher, 2001; Palmer et al., 2022). Cues to shape from shading are influenced by less unusual directions. For example, it is predicted that depth relief will be increased by lighting from the side, and decreased by lighting from in front, due to the effect of increasing or decreasing cast shadows (Hunter, Biver, Fuqua, & Reid, 2021; Todd et al., 1996).

Finally, we also assessed the effect of inverting contrast polarity on perceived three-dimensional shape. Polarity inversion strongly disrupts face recognition. Since there is no simple interpretation of the effects of inversion in terms of lighting direction or surface shape, this condition quantified the potential distortion of surface shape caused by polarity inversion.

3.2. Method

The method used in the current study is similar to those outlined in the previous study

3.3. Participants

There were 15 participants (8 female), who were randomly assigned a male (7) or female (8) stimulus. All participants were over the age of 18. The study was approved by the University of Essex Ethics Sub-committee 3 (ETH2122-0448) and all participants provided written, informed consent.

3.4. Apparatus

The apparatus were the same as in Experiment one.

3.5. Stimuli

Facegen was used to randomly generate one male and one female face with a neutral expression, with age and ethnicity selected randomly. This was saved as a Wavefront .obj file and loaded to Blender. RGB images were rendered with a camera with a focal length of 36 mm at a location 90 cm directly in front of the face, and a single light source at a distance of 60 cm above, below, to the right of or directly in front of the face, depending on the condition. Stimuli were rendered with a matte lambertian reflectance, with the RGB texture created by Facegen. True depth values were extracted using a Z pass render, and saved as an OpenEXR file which provided true distance information.

3.6. Design

The first set of stimuli used the same design as experiment one. The light source was positioned either below or above the face. (see Figs. 9a & b). Furthermore, stimuli were presented either upright or

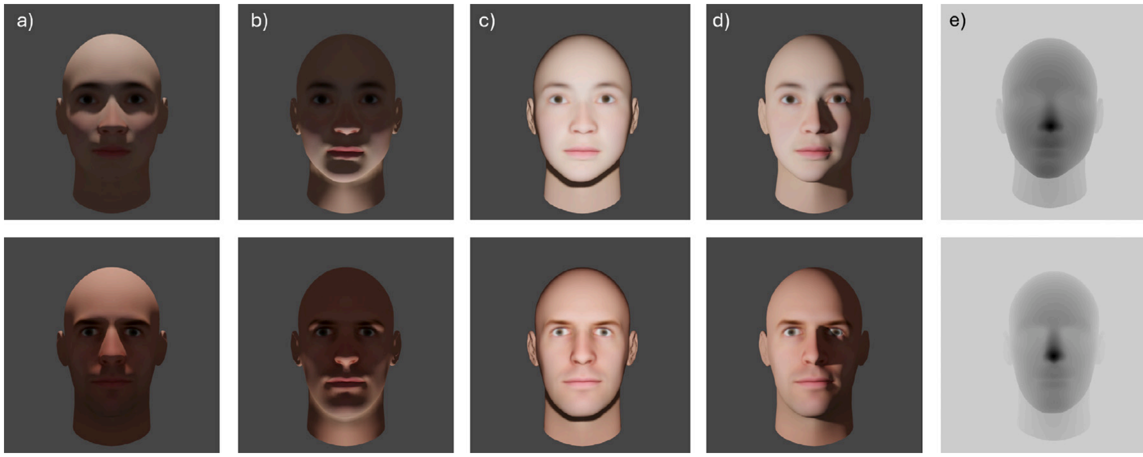

Fig. 9. The face stimuli used in Experiment 2 where top row is the female face and the bottom row the male face. (a) lit from above, (b) lit from below (c) lit from the front, (d) lit from the right. (e) Luminance polarity reversed image of the greyscale upright, lit-from-above stimulus.

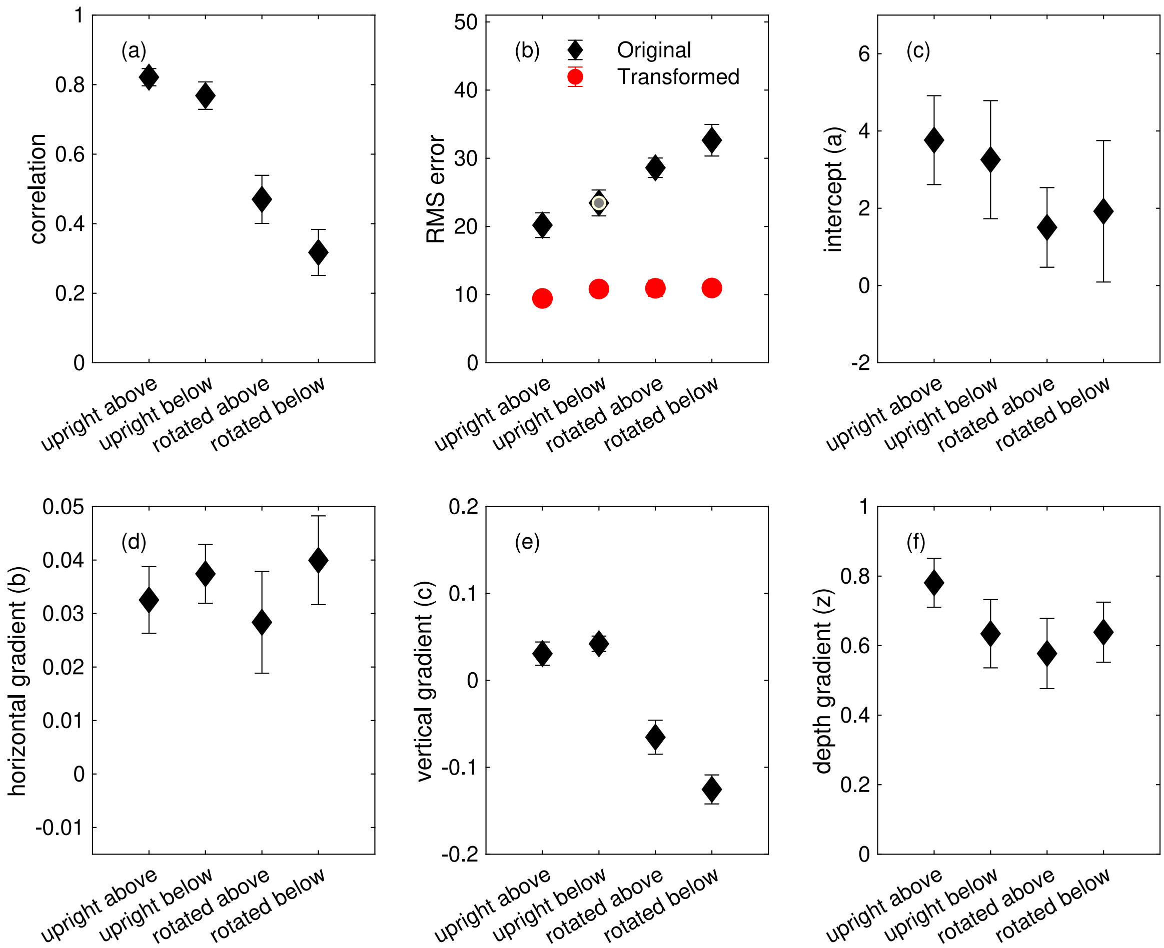

Fig. 10. **Experiment 2: Image Rotation (a)** correlation between psychophysical depth and the true depth coordinates; **(b)** RMS Error for Original (diamonds) and Transformed (circles) data; The four affine transformations are shown for the **(c)** Intercept, **(d)** Horizontal shear, **(e)** Vertical shear **(f)** Depth Stretch. Error bars are $\pm SEM$.

rotated through 180°. This resulted in the same four conditions as were compared in experiment one (upright lit from above; upright lit from below; rotated lit from above; rotated lit from below).

The second set of stimuli introduced two new lighting direction positions: in front of, or to the right of the face (see Figs. 9c & d). This was used to compare the effects of lighting across three directions: from above, in front, or the side.

A final stimulus was created by inverting the luminance polarity of a greyscale version of the upright, lit-from-above stimuli (see Fig. 9e). Greyscale images were used to avoid the possible influence of colour hue or saturation (Kemp et al., 1996). This manipulation allowed us to compare the pictorial relief for a polarity inverted face with both the true depth, and a (colour) stimulus with correct luminance polarity.

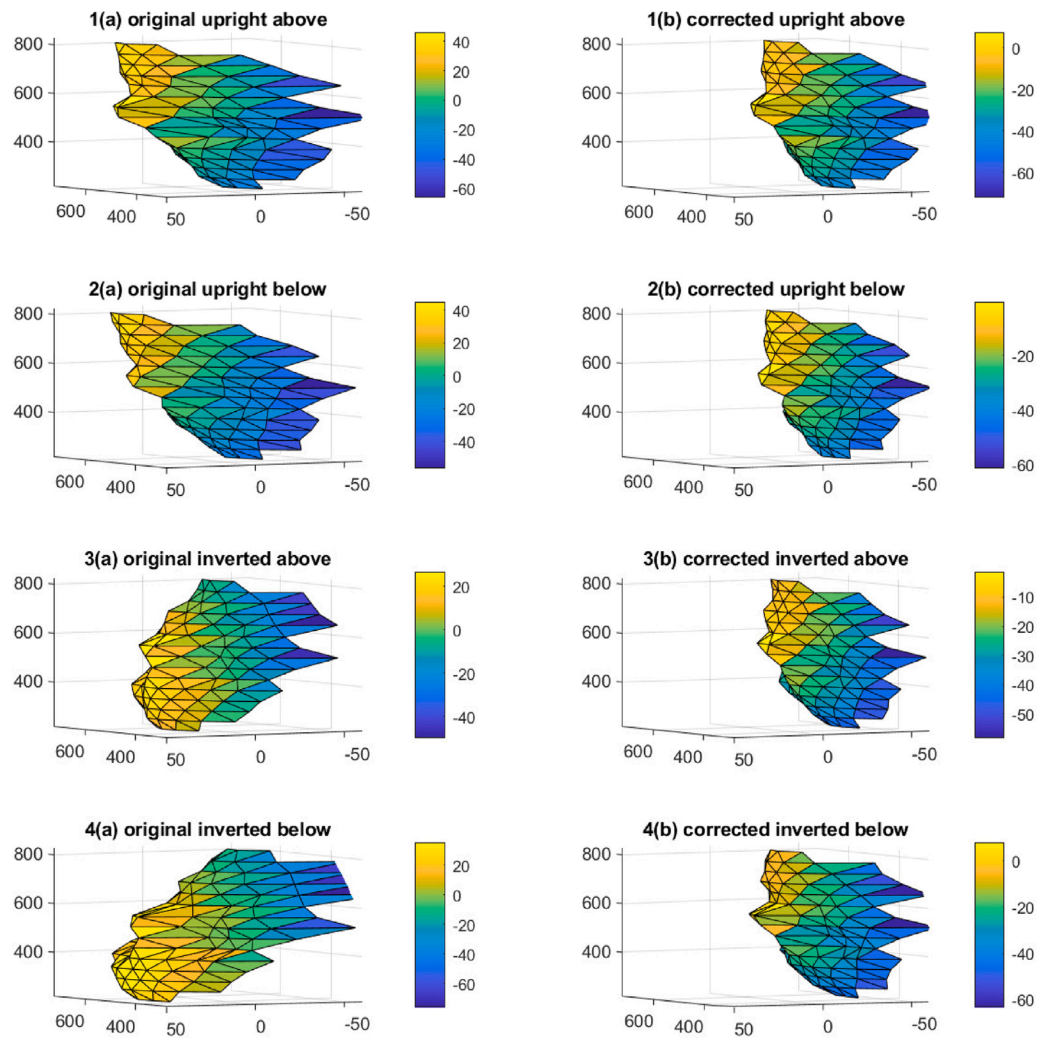

Fig. 11. Image Rotation: Mean depth relief from participants' settings for the generated female face. Pre (original) and Post (corrected) affine transformation. All units are in pixels.

3.7. Procedure

The procedure was the same as used in the first experiment, with the stimuli for the seven conditions presented in separate blocks of trials. Each condition consisted of 188 trials, which was the number of triangular faces in the mesh.

3.8. Results

3.8.1. Image rotation

Average pictorial relief maps across the four conditions are shown in Fig. 10 and the corresponding regression reports can be found in Table 3. In the upright, lit-from-above condition depth values correlated highly with the true depth (mean r = 0.82) (a). This correlation was reduced in the two rotated conditions to a mean of r = 0.44. This was reflected in the RMS errors, which were significantly greater in the rotated conditions. This error was reduced on average to 62% of the original error following affine transformation, and the transformed RMS errors did not vary across conditions (b).

As reported in Table 4 there was a small positive horizontal depth shearing (b_x) of the pictorial relief in comparison with the true depth, equivalent to a rotation of less that 2°. This was not affected by either image rotation or lighting direction. For the upright faces, there was a small positive vertical depth (b_y) shearing relative to true depth. This shearing was in the opposite direction for rotated stimuli.

Depth relief (b_z) was compressed by a factor of 0.79 relative to the true depth, and was reduced by each of three image manipulations (also in Table 4).

Summary

Overall, the results for image rotation replicate those in the first experiment. Pictorial relief was closely matched with the true depth, and those differences that were found across viewing conditions could be accounted for by an affine transformation (see the b-column in Figs. 11 (female) and 12 (male)). This amounted to a small rotation of the face around a horizontal axis, combined with a flattening of the depth relief, when the face was rotated through 180°.

3.8.2. Lighting direction

A strong correlation between depth relief values and true depth was found (see Fig. 13a, & Table 5). This was on average 0.80, and did not differ between lighting from above, in front or to the side. RMS errors relative to the true depth were slightly higher (8%) in the lit from the front condition (see Fig. 13b). RMS error reduced to 38% of the original value on average following affine alignment with the true depth, and did not vary between conditions. No difference in the two directions of depth shear were evident across conditions. However, the depth relief was flattened by 20% in the lit from in front condition (see Fig. 13f and Table 6). Pictorial relief maps across the three conditions can be seen here Figs. 14 (female) and 15 (male).

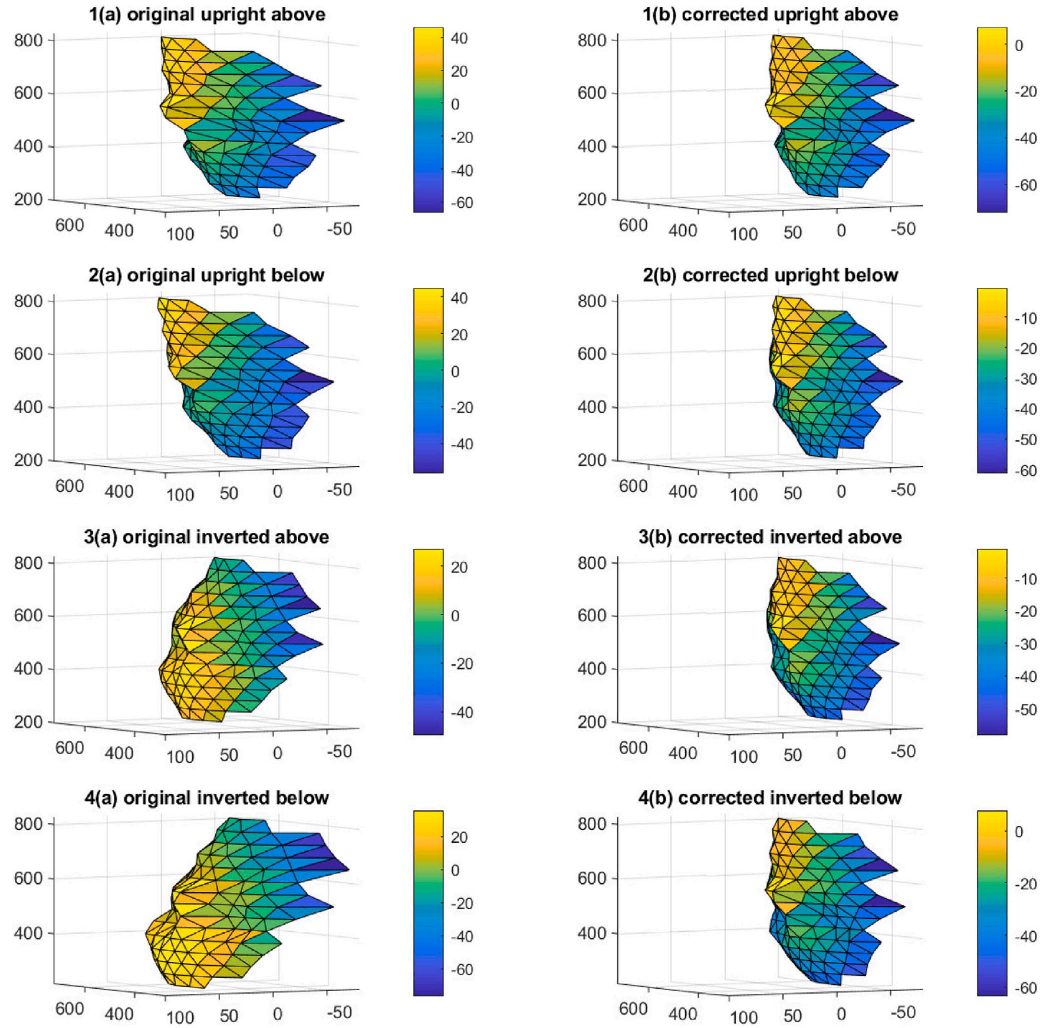

Fig. 12. Image Rotation: Mean depth relief from participants' settings for the generated male face. Pre (original) and Post (corrected) affine transformation. All units are in pixels.

Table 3
Experiment 2 (image rotation) correlation, original RMS and transformed RMS.

(a) Condition	Correlation	SE	t	df	p	Lower	Upper
intercept[a]	0.821	0.051	15.98	56	< **0.001**	0.718	0.924
Upright below	−0.053	0.058	−0.909	56	0.367	−0.170	0.064
Rotated above	−0.351	0.058	−6.03	56	< **0.001**	−0.468	−0.235
Rotated below	−0.504	0.058	−8.64	56	< **0.001**	−0.621	−0.387
(b) Condition	Original RMS	SE	t	df	p	Lower	Upper
intercept[a]	20.18	1.83	11.05	56	< **0.001**	16.52	23.84
Upright below	3.26	1.80	1.81	56	0.075	−0.35	6.86
Rotated above	8.42	1.80	4.68	56	< **0.001**	4.82	12.03
Rotated below	12.46	1.80	6.93	56	< **0.001**	8.86	16.06
(c) Condition	Transformed RMS	SE	t	df	p	Lower	Upper
intercept[a]	9.44	0.87	10.84	56	< **0.001**	7.70	11.19
Upright below	1.38	0.85	1.62	56	0.110	−0.32	3.08
Rotated above	1.48	0.85	1.75	56	0.086	−0.22	3.19
Rotated below	1.53	0.85	1.80	56	0.077	−0.17	3.23

[a] The intercept is the Upright lit from Above condition. See Fig. 10(a & b).

3.8.3. *Polarity inversion*

Psychophysical depth relief maps are shown in Figs. 17 and 18 for the female and male face respectively. These show that gauge settings were consistent with the 3D surface structure of the faces, despite the inversion of luminance polarity. There was a strong positive correlation between psychophysical and true depth (mean 0.742; $t(14) = 10.03$, $p < 0.0001$). RMS error relative to true depth was not significantly greater for the polarity reversed stimuli (see Table 7). The affine regression against true depth produced a small (4°) rotation out of the and were flattened by a factor of 0.42 relative to the true depth (see Table 8).

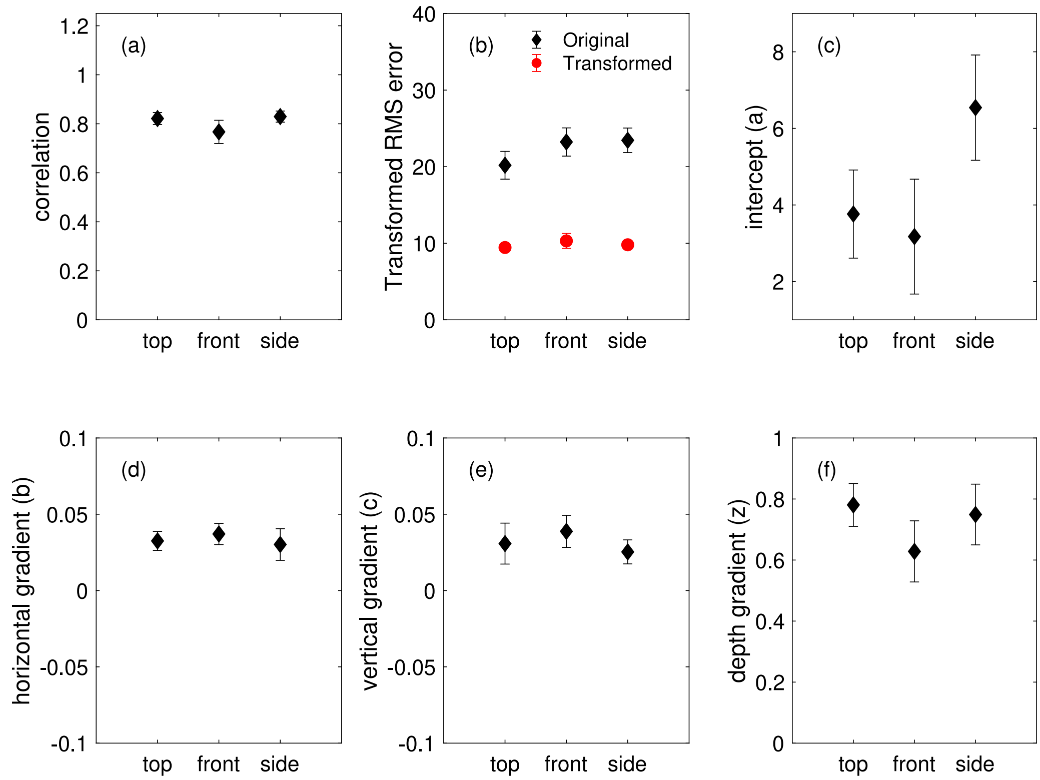

Fig. 13. Experiment 2: Lighting Direction (a) correlation between psychophysical depth and the true depth coordinates; **(b)** RMS Error for Original (diamonds) and Transformed (circles) data; The four affine transformations are shown for the **(c)** Intercept, **(d)** Horizontal shear, **(e)** Vertical shear **(f)** Depth Stretch. Error bars are $\pm SEM$.

Table 4
Experiment 2 (image rotation) affine transformations.

(a) Translation							
Condition	b_0	SE	t	df	p	Lower	Upper
intercept[a]	3.938	2.27	1.73	56	0.089	−0.61376	8.4892
Upright below	−0.506	1.63	−0.31	56	0.757	−3.77	2.76
Rotated above	−2.259	1.63	−1.39	56	0.171	−5.52	1.00
Rotated below	−1.843	1.63	−1.13	56	0.263	−5.11	1.42
(b) Horizontal Shear							
Condition	b_X	SE	t	df	p	Lower	Upper
intercept[a]	0.033	0.0078	4.19	56	< **0.001**	0.017	0.048
Upright below	−0.0050	0.0078	0.629	56	0.532	−0.011	0.020
Rotated above	−0.0042	0.0078	−0.537	56	0.593	−0.020	0.011
Rotated below	0.0074	0.0078	0.955	56	0.344	−0.008	0.0023
(c) Vertical Shear							
Condition	b_Y	SE	t	df	p	Lower	Upper
intercept[a]	0.031	0.015	2.10	56	**0.040**	0.0014	0.060
Upright below	0.011	0.021	0.549	56	0.585	−0.301	0.053
Rotated above	−0.096	0.021	−4.64	56	< **0.001**	−0.138	−0.055
Rotated below	−0.156	0.021	−7.54	36	< **0.001**	−0.198	−0.115
(d) Depth Shear							
Condition	b_Z	SE	t	df	p	Lower	Upper
intercept[a]	0.786	0.112	6.99	56	< **0.001**	0.561	1.01
Upright below	−0.147	0.067	−2.19	56	**0.033**	−0.280	−0.013
Rotated above	−0.203	0.067	−3.04	56	**0.036**	−0.337	−0.067
Rotated below	−0.142	0.067	−2.13	56	**0.038**	−0.276	−0.008

[a] The intercept is for the Upright lit from Above condition. Also see Fig. 10(c–f).

Although stimuli varied from experiment one in both being polarity inverted, and greyscale rather than full colour, we also compared results between natural and polarity inverted stimuli in the upright, lit from above condition. Correlation and RMS errors in comparison with true depth were not affected by polarity inversion, and there were no differences in the affine transformations between psychophysical and true pictorial relief.

Table 5
Experiment 2 (lighting direction) correlation, original RMS and transformed RMS.

Condition	r	SE	t	df	p	Lower	Upper
intercept[a]	0.821	0.032	25.37	42	< **0.001**	0.756	0.887
Front	−0.054	0.035	−1.55	42	0.129	−0.125	0.016
Side	0.008	0.035	0.228	42	0.821	−0.629	0.0790
Condition	r	SE	t	df	p	Lower	Upper
intercept[a]	20.184	1.70	11.89	42	< **0.001**	16.76	23.61
Front	3.04	1.49	2.04	42	**0.048**	0.028	6.05
Side	3.26	1.49	2.18	42	**0.035**	0.246	6.27
Condition	r	SE	t	df	p	Lower	Upper
intercept[a]	9.44	0.67	14.00	42	< **0.001**	8.08	10.80
Front	0.86	0.67	1.29	42	0.205	−0.49	2.21
Side	0.36	0.67	0.54	42	0.596	−0.99	1.71

[a] The intercept is lit from above (top) condition.

4. General discussion

In this study, we investigated the effect of lighting, image rotation, polarity inversion and surface reflectance on the perceived 3D surface shape of faces. Shape was assessed using a gauge figure task, which provided a measure of surface orientation (from the slant and tilt of the placed gauge figure) for a defined point within a 3D image.

In the first experiment, we assessed the effects of image rotation and lighting from above or below on pictorial relief for a stimulus with uniform reflectance. Depth relief maps from user settings were highly correlated with the true depth in all conditions, indicating an accurate ability to perceive the three-dimensional surface structure of the face. Deviations from true depth were lowest in the upright, lit-from-above condition, and highest when lit from below. The differences between

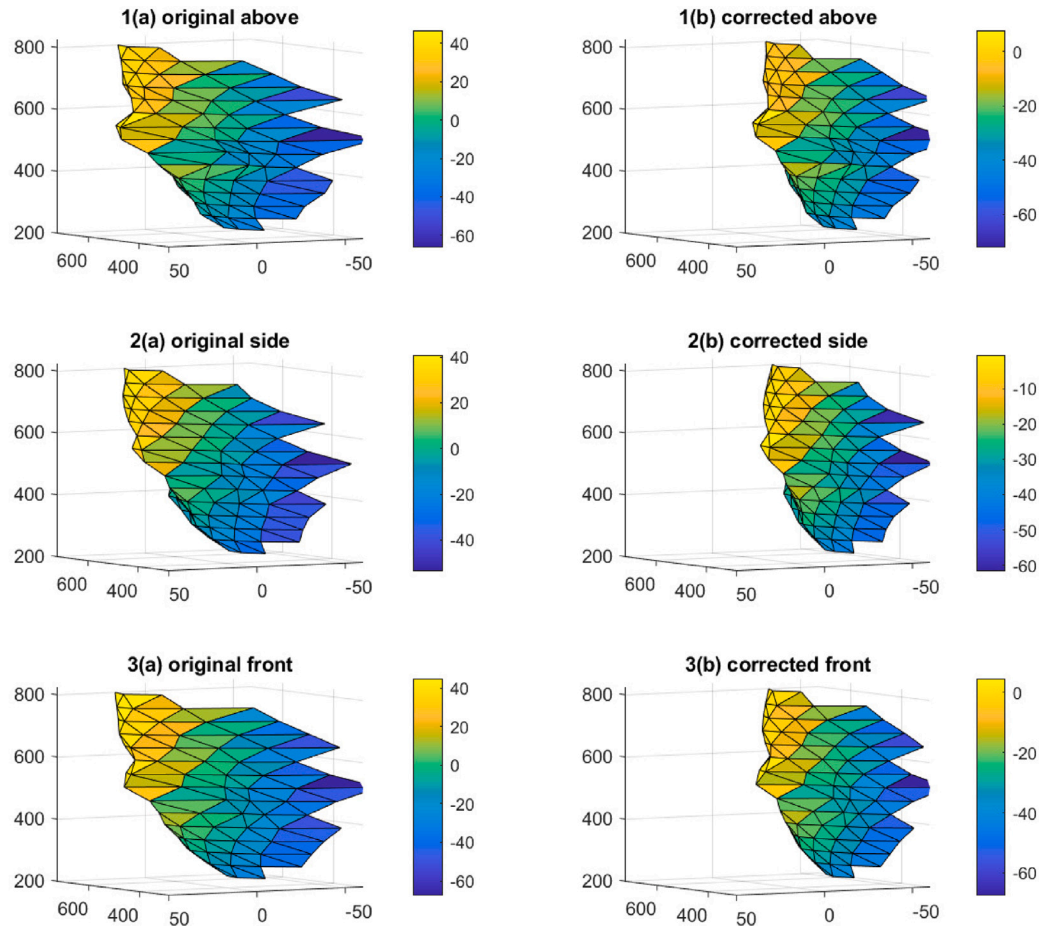

Fig. 14. Lighting Direction: Mean depth relief from participants' settings for the generated female face. Pre (original) and Post (corrected) affine transformation. All units are in pixels.

Table 6
Experiment 2 (lighting direction) affine transformations.

(b) Translation							
Condition	b_0	SE	t	df	p	Lower	Upper
intercept[a]	3.9524	2.358	1.676	42	0.10	−0.81	8.71
Front	−0.5892	1.393	−0.423	42	0.67	−3.40	2.22
Side	2.7790	1.393	1.995	42	**0.05**	−0.03	5.59
(b) Horizontal Shear							
Condition	b_X	SE	t	df	p	Lower	Upper
intercept[a]	0.033	0.009	3.74	42	< **0.001**	0.015	0.051
Front	0.0046	0.011	0.422	42	0.675	−0.172	0.0263
Side	−0.002	0.011	−0.22	42	0.829	−0.024	0.019
(c) Vertical Shear							
Condition	b_Y	SE	t	df	p	Lower	Upper
intercept[a]	0.031	0.010	2.94	42	**0.005**	0.010	0.052
Front	0.008	0.008	1.06	42	0.30	−0.007	0.023
Side	−0.005	0.008	−0.72	42	0.48	−0.021	0.010
(d) Depth Shear							
Condition	b_Z	SE	t	df	p	Lower	Upper
intercept[a]	0.791	0.14	5.51	42	**0.005**	0.50	1.08
Front	−0.152	0.074	−2.05	42	**0.047**	−0.30	−0.002
Side	−0.032	0.074	−0.424	42	0.67	−0.18	0.12

[a] The intercept is lit from above (top) condition. See Fig. 13(c–f).

conditions were fully accounted for by an affine transformation, which consisted of a slight rotation of the image plane about a horizontal axis, and a squashing in the depth dimension, when the image was rotated. As such, these differences do not reflect a change in the intrinsic surface structure, since they are consistent with the bas-relief ambiguity in shape from shading, and may be characterised either as a change in the viewing direction (Koenderink et al., 2001), lighting direction or surface pigmentation (Belhumeur et al., 1999; Palmer et al., 2022). These results suggest the effects of image rotation and lighting direction that have been found in face recognition are not necessarily a result of changes in the intrinsic 3D surface structure (Johnston et al., 1992).

In the second experiment, we extended our investigation to include faces with spatial variations in reflectance, the lighting conditions to include front and side lit, and finally an additional polarity-inverted condition. The effects of rotation and inversion replicated those of experiment one. Deviations from true depth varied across viewing conditions, but in all cases could be accounted for by an affine transformation. In both experiments, this took the form of a vertical depth shear and depth compression when the images were rotated. When faces were lit from in front, depth relief was compressed in comparison with lighting from above, as predicted (Todd et al., 1996). However, the predicted expansion of pictorial relief with lighting from the side was not found. Polarity inversion had no effect on pictorial relief. These findings align with the perspective that the impact of lighting and polarity inversion on face recognition is linked to their influence on apparent pigmentation, rather than three-dimensional surface shape.

To assess the statistical power of our methods to find differences in apparent 3D shape, simulation-based power analysis was performed (Kumle, Võ, & Draschkow, 2021). We simulated 1000 repetitions of the experiment, using the model estimated standard deviations of the fixed and random effects, for sample sizes of 10 and 15 participants . These showed that, in the first experiment (n = 10), changes of correlation

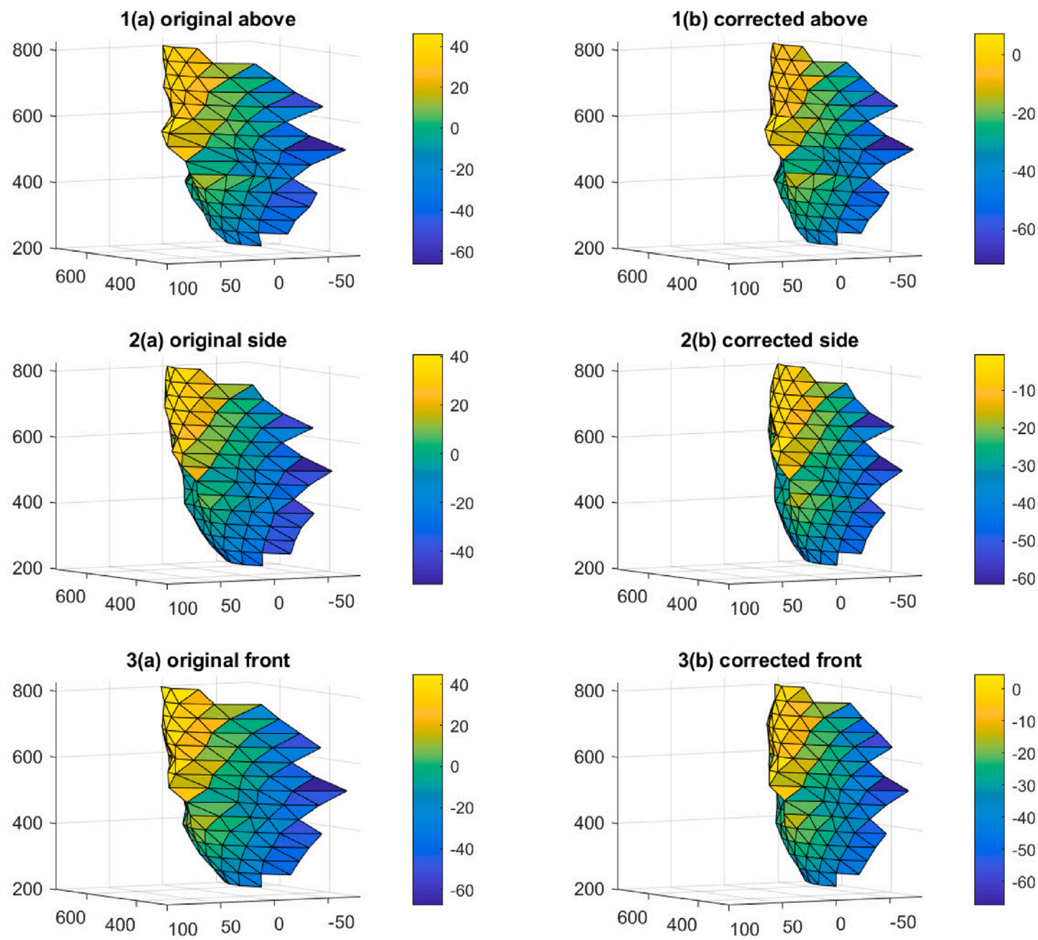

Fig. 15. Lighting Direction: Mean depth relief from participants' settings for the generated male face. Pre (original) and Post (corrected) affine transformation. All units are in pixels.

Table 7
Experiment 2 (inverted polarity) correlation, original RMS and transformed RMS.

(a) Condition	Correlation	SE	t	df	p	Lower	Upper
intercept[a]	0.821	0.058	14.12	28	< **0.001**	0.702	0.941
Reversed	−0.086	0.061	−1.412	28	0.168	−0.2.13	0.039
(b) Condition	Original RMS	SE	t	df	p	Lower	Upper
intercept[a]	20.184	2.278	8.56	28	< **0.001**	15.52	24.85
Reversed	2.762	2.260	1.22	28	0.232	−1.869	7.392
Condition	Transformed RMS	SE	t	df	p	Lower	Upper
intercept[a]	9.442	0.764	12.364	28	< **0.001**	7.88	11.01
Reversed	0.997	0.748	1.33	28	0.193	−0.536	2.53

[a] The intercept is the Upright lit from Above condition. See Fig. 16(a & b).

of 0.26, original RMS error of 11 pixels, and an affine-transformed RMS error of 2.5 pixels, could all be detected with 80% power. For the second experiment (n = 15), changes of correlation of 0.26, original RMS error of 10 pixels, and an affine-transformed RMS error 2.25 pixels, could all be detected with the same power.

While the gauge figure task is a well established tool for measuring surface orientation, there is no single recommendation for an optimal size for the gauge figure. The size of the gauge required will depend on the size and texture of the object and the detail of information required. For this experiment each triangle had an edge of 60 pixels, used as the sampling location for each gauge. Using a smaller or larger gauge would likely result is slightly differing results on a local or global level of the 3D structure. Further research could explore manipulating the scale of the gauge figure to capture varying shapes and textures, in an attempt to define optimal gauge to sampling ratios.

The full 3D shape of a face can be described by specifying the location of each point on its surface in Euclidean space. It can also be described at different levels of specificity, that preserve some, but not all, spatial relationships on the surface. The affine ambiguity that exists in 2D images, for example, allows for the surface to be stretched or sheared, but preserves important shape properties such as locations of maxima and minima of curvature (Phillips, Todd, Koenderink, & Kappers, 2003).

Our results suggest that the effects of lighting direction and image rotation affect 3D shape at the level of metric depth, however, the intrinsic properties that are necessary for recognition are invariant at the level of affine or projective geometry (Todd & Petrov, 2022), and are therefore not affected by these perceptual changes. The effects of these manipulations on recognition may thus be better interpreted through their influence on the perception of material properties such as

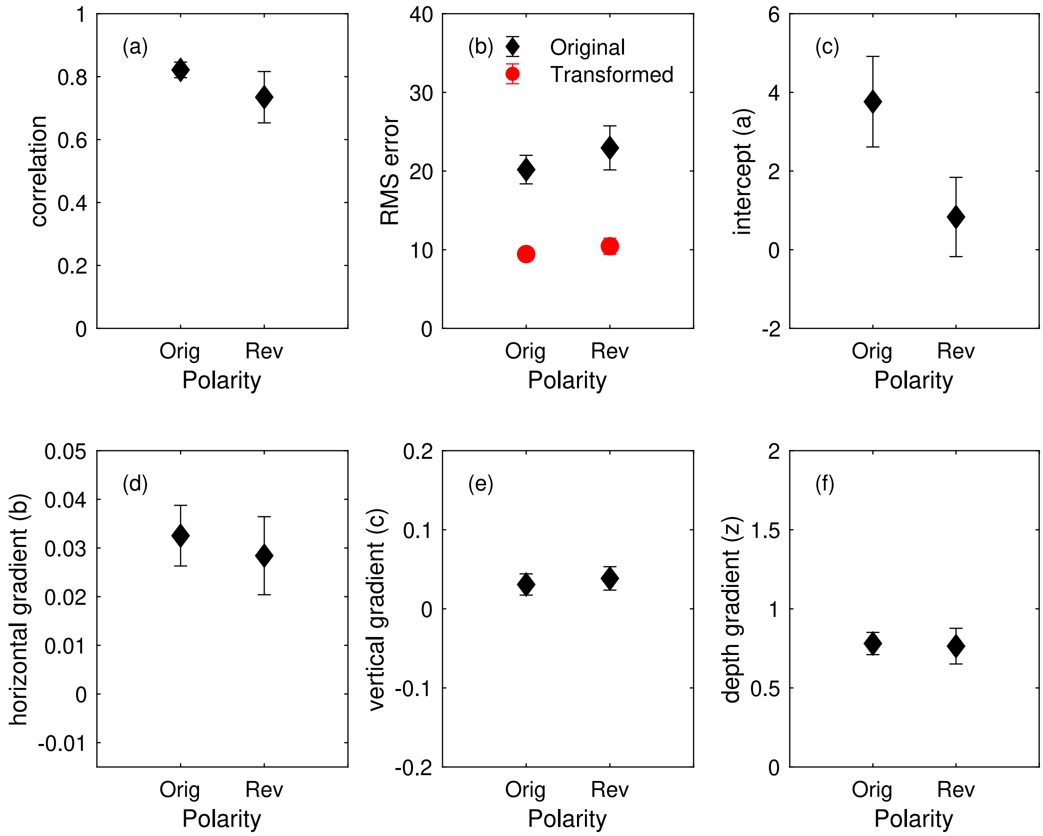

Fig. 16. **Experiment 2: Polarity Inversion (a)** correlation between psychophysical depth and the true depth coordinates; **(b)** RMS Error for Original (diamonds) and Transformed (circles) data; The four affine transformations are shown for the **(c)** Intercept, **(d)** Horizontal shear, **(e)** Vertical shear **(f)** Depth Stretch. Error bars are $\pm SEM$.

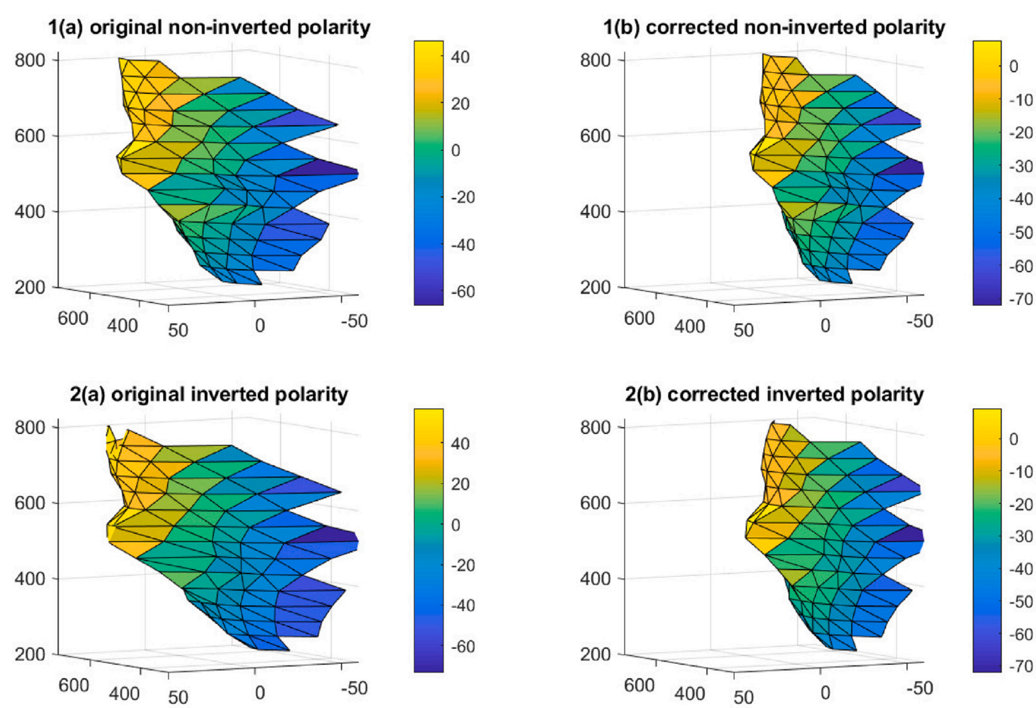

Fig. 17. Inverted Polarity: Mean depth relief from participants' settings for the generated female face. Pre (original) and Post (corrected) affine transformation. All units are in pixels.

pigmentation, or in recognition mechanisms operating on information closer to the level of the retinal image, such as the pattern of shading and shadows (Palmer et al., 2022). Under this account, it is the pattern of image features, rather than the perceived three-dimensional shape that is affected by changing the lighting direction. Note that the methods used in the current study allow us to quantify the 3D surface

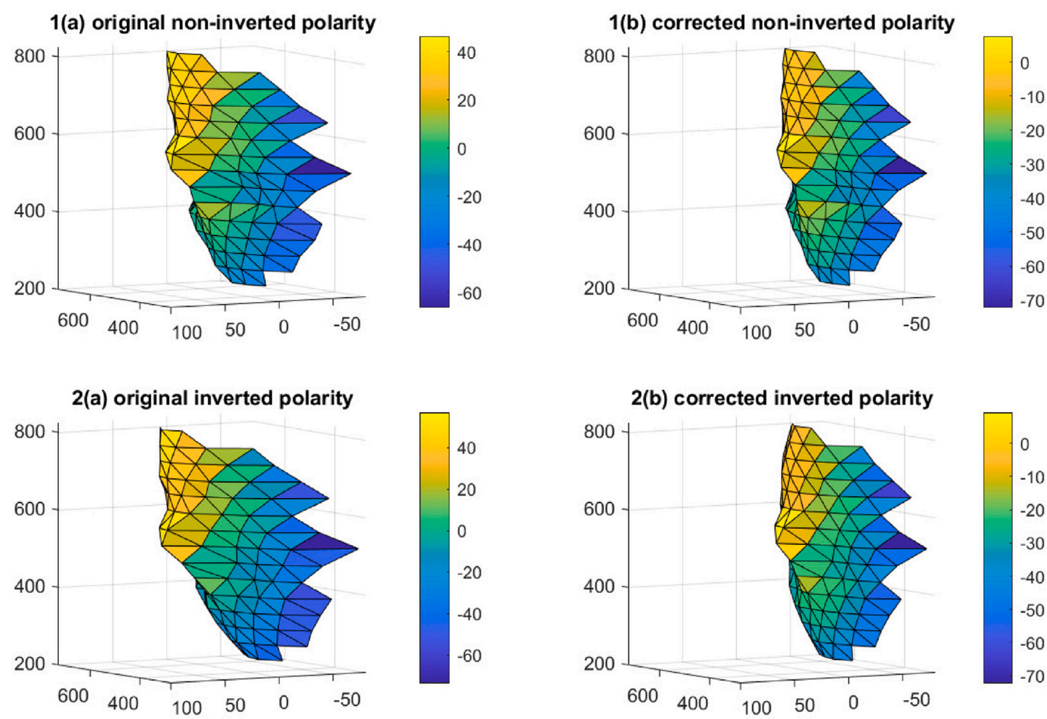

Fig. 18. Inverted Polarity: Mean depth relief from participants' settings for the generated male face. Pre (original) and Post (corrected) affine transformation. All units are in pixels.

Table 8
Experiment 2 (inverted polarity) affine transformations.

(a) Translation							
Condition	b_0	SE	t	df	p	Lower	Upper
intercept[a]	3.893	1.740	2.24	28	**0.033**	0.328	7.458
Reversed	−2.932	1.261	−2.325	28	**0.0276**	−5.515	−0.349
(b) Horizontal Shear							
Condition	b_X	SE	t	df	p	Lower	Upper
intercept[a]	0.032	0.007	4.59	28	< **0.001**	0.0176	0.0461
Reversed	−0.004	0.006	−0.72	28	0.4775	−0.0157	0.0076
(c) Vertical Shear							
Condition	b_Y	SE	t	df	p	Lower	Upper
intercept[a]	0.032	0.017	1.81	28	0.081	−0.0042	0.0667
Reversed	0.008	0.016	0.450	28	0.635	−0.0251	0.0406
(b) Depth Shear							
Condition	b_Z	SE	t	df	p	Lower	Upper
intercept[a]	0.795	0.174	4.55	28	< **0.001**	0.437	1.152
Reversed	−0.0178	0.051	−0.327	28	0.746	−0.121	0.088

[a] The intercept is the Upright lit from Above condition (original). See Fig. 16(c–f).

shape, and not necessarily the image features that may be important in detecting and recognising faces.

Our task assessed the effects of lighting direction, image rotation and polarity inversion on the perception of 3D surface shape. It has been proposed that this type of 3D representation may be used to provide face recognition with invariance to changes in viewing condition, such as lighting conditions and viewing direction, by creating a representation based on the shape of the face itself, rather than properties of the retinal image. However, our results suggest that the effects of viewing conditions on face recognition are not mediated by changes in perceived 3D surface shape.

It has been proposed that this type of 3D representation may be used to provide face recognition with invariance to changes in viewing condition, such as lighting conditions and viewing direction, by creating a representation based on the shape of the face itself, rather than properties of the retinal image. However, our results suggest that the effects of viewing conditions on face recognition are not mediated by changes in perceived 3D surface shape.

The problem of face recognition has been characterised using the abstract concept of "face space" (Valentine, 1991). In this multidimensional space, each dimension corresponds to a facial feature. Each face is represented by a point in this space, and the similarity or difference between two faces corresponds to the distance between them. Irrespective of the specific nature of these dimensions, robust face recognition requires a representation in which similarity is invariant across changes in viewing conditions, rather than the stronger requirement of the representation itself exhibiting invariance (Blank & Yovel, 2011). Consistent with this, insights from deep learning studies show that representation that supports accurate face recognition may contain image-level information, but organised to allow for invariance across for viewpoint and illumination (Hill et al., 2019; O'Toole & Castillo, 2021). More specifically, lip-thickness, eye-colour, eye-shape and eyebrow-thickness and hair have all been proposed as critical features for face recognition (Abudarham, Shkiller, & Yovel, 2019; Abudarham & Yovel, 2016).

Our results are based on a small number of faces, and did not assess the task of face recognition directly. However, in all cases the effects of viewing condition could be accounted for by a simple affine transformation. This transformation would not be expected to influence the estimation of these critical facial features, since these transformation reflect the structural ambiguity inherent in pictorial space (Belhumeur et al., 1999; Georghiades et al., 2001; Koenderink & Van Doorn, 2003).

The gauge figure technique requires many local slant and tilt settings to be made on a single face, so that its global shape can then be recovered. We are therefore limited to using a small number of stimuli (one or two) across our experiments, as is typical in research of this type. Idiosyncratic differences in features between faces will naturally introduce convexities and concavities that will be reflected in the perception of local shape and texture — key features of interest in the present study that could be controlled to a degree by using a constant face identity. Our study was concerned with the effects of viewing condition on the perception of face shape, rather than assessing sensitivity to differences between faces. We found very similar results across the three stimuli that we used, and anticipate that our observations will be generalisable to other faces.

5. Conclusion

While there were discrepancies when comparing the true depth to the responses across the viewing conditions (lighting direction, image rotation and polarity inversion), they could be explained by applying an affine transformation. Changing the viewing conditions distorted incidental but not intrinsic properties of pictorial relief in faces.

CRediT authorship contribution statement

Jordi M. Asher: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Paul B. Hibbard:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Abigail L.M. Webb:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization.

Acknowledgements

We extend our gratitude to Kai Tajima, Shanel-Oriene Kalombo Nzeba and Chidera Elvis Mbah for their meticulous effort and commitment in diligently collecting the data for this research.

Data availability

Data will be made available on request.

References

3DScanStore (2020). Free 3D head model. URL https://www.3dscanstore.com/blog/Free-3D-Head-Model.

Abudarham, N., Shkiller, L., & Yovel, G. (2019). Critical features for face recognition. *Cognition*, *182*, 73–83.

Abudarham, N., & Yovel, G. (2016). Reverse engineering the face space: Discovering the critical features for face identification. *Journal of Vision*, *16*(3), 40–40.

Asher, J., Hibbard, P., & Webb, A. (2024). Supplementary materials: Perceived intrinsic 3D shape of faces is robust to changes in lighting direction, image rotation and polarity inversion. https://osf.io/fcz7h/.

Belhumeur, P. N., Kriegman, D. J., & Yuille, A. L. (1999). The bas-relief ambiguity. *International Journal of Computer Vision*, *35*(1), 33–44.

Berbaum, K., Bever, T., & Chung, C. S. (1983). Light source position in the perception of object shape. *Perception*, *12*(4), 411–416.

Blank, I., & Yovel, G. (2011). The structure of face–space is tolerant to lighting and viewpoint transformations. *Journal of Vision*, *11*(8), 15–15.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Bruce, V., & Langton, S. (1994). The use of pigmentation and shading information in recognising the sex and identities of faces. *Perception*, *23*(7), 803–822.

Burge, J., McCann, B. C., & Geisler, W. S. (2016). Estimating 3D tilt from local image cues in natural scenes. *Journal of Vision*, *16*(13), 2–2.

Favelle, S., Hill, H., & Claes, P. (2017). About face: matching unfamiliar faces across rotations of view and lighting. *i-Perception*, *8*(6), Article 2041669517744221.

Georghiades, A. S., Belhumeur, P. N., & Kriegman, D. J. (2001). From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *23*(6), 643–660.

Gibson, J. J. (1950). *The perception of the visual world*. Houghton Mifflin.

Henderson, A. J., Holzleitner, I. J., Talamas, S. N., & Perrett, D. I. (2016). Perception of health from facial cues. *Philosophical Transactions of the Royal Society, Series B (Biological Sciences)*, *371*(1693), Article 20150380.

Hibbard, P. B., Goutcher, R., Hornsey, R. L., Hunter, D. W., & Scarfe, P. (2023). Luminance contrast provides metric depth information. *Royal Society Open Science*, *10*(2), Article 220567.

Hibbard, P. B., Hornsey, R. L., & Asher, J. M. (2023). Binocular information improves the reliability and consistency of pictorial relief. *Vision*, *7*(1).

Hill, M. Q., Parde, C. J., Castillo, C. D., Colon, Y. I., Ranjan, R., Chen, J.-C., et al. (2019). Deep convolutional neural networks in the face of caricature. *Nature Machine Intelligence*, *1*(11), 522–529.

Hole, G. J., George, P. A., Eaves, K., & Rasek, A. (2002). Effects of geometric distortions on face-recognition performance. *Perception*, *31*(10), 1221–1240.

Hunter, F., Biver, S., Fuqua, P., & Reid, R. (2021). *Light—science & magic: An introduction to photographic lighting*. Routledge.

Johnston, A., Hill, H., & Carman, N. (1992). Recognising faces: Effects of lighting direction, inversion, and brightness reversal. *Perception*, *21*(3), 365–375.

Kemp, R., Pike, G., White, P., & Musselman, A. (1996). Perception and recognition of normal and negative faces: the role of shape from shading and pigmentation cues. *Perception*, *25*(1), 37–52.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C., et al. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1.

Koenderink, J. J. (1998). Pictorial relief. *Philosophical Transactions of the Royal Society of London. Series A. Mathematical, Physical and Engineering Sciences*, *356*(1740), 1071–1086.

Koenderink, J. J., & Van Doorn, A. J. (1995). Relief: Pictorial and otherwise. *Image and Vision Computing*, *13*(5), 321–334.

Koenderink, J., & Van Doorn, A. (2003). Pictorial space. In *Looking into pictures: an interdisciplinary approach to pictorial space* (pp. 239–403). MIT Press.

Koenderink, J. J., Van Doorn, A. J., & Kappers, A. M. (1992). Surface perception in pictures. *Perception & Psychophysics*, *52*, 487–496.

Koenderink, J. J., Van Doorn, A. J., Kappers, A. M., & Todd, J. T. (2001). Ambiguity and the 'mental eye'in pictorial relief. *Perception*, *30*(4), 431–448.

Kumle, L., Võ, M. L.-H., & Draschkow, D. (2021). Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behavior Research Methods*, *53*(6), 2528–2543.

Liu, C. H., Collin, C. A., Burton, A. M., & Chaudhuri, A. (1999). Lighting direction affects recognition of untextured faces in photographic positive and negative. *Vision Research*, *39*(24), 4003–4009.

Liu, C. H., Collin, C. A., & Chaudhuri, A. (2000). Does face recognition rely on encoding of 3-D surface? Examining the role of shape-from-shading and shape-from-stereo. *Perception*, *29*(6), 729–743.

Mamassian, P., & Goutcher, R. (2001). Prior knowledge on the illumination position. *Cognition*, *81*(1), B1–B9.

Mingolla, E., & Todd, J. T. (1986). Perception of solid shape from shading. *Biological Cybernetics*, *53*(3), 137–151.

Nefs, H. T. (2008). Three-dimensional object shape from shading and contour disparities. *Journal of Vision*, *8*(11), 11–11.

O'Toole, A. J., & Castillo, C. D. (2021). Face recognition by humans and machines: three fundamental advances from deep learning. *Annual Review of Vision Science*, *7*(1), 543–570.

O'Toole, A. J., Vetter, T., & Blanz, V. (1999). Three-dimensional shape and two-dimensional surface reflectance contributions to face recognition: An application of three-dimensional morphing. *Vision Research*, *39*(18), 3145–3155.

Palmer, C. J., Goddard, E., & Clifford, C. W. (2022). Face detection from patterns of shading and shadows: The role of overhead illumination in generating the familiar appearance of the human face. *Cognition*, *225*, Article 105172.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Peterson, L. M., Susilo, T., Clifford, C. W., & Palmer, C. J. (2023). Discrimination of facial identity based on simple contrast patterns generated by shading and shadows. *Vision Research*, *212*, Article 108307.

Phillips, F., Todd, J. T., Koenderink, J. J., & Kappers, A. M. (2003). Perceptual representation of visible surfaces. *Perception & Psychophysics*, *65*(5), 747–762.

Ramachandran, V. S. (1988). Perception of shape from shading. *Nature*, *331*(6152), 163–166.

Russell, R., Biederman, I., Nederhouser, M., & Sinha, P. (2007). The utility of surface reflectance for the recognition of upright and inverted faces. *Vision Research*, *47*(2), 157–165.

Russell, R., & Sinha, P. (2007). Real-world face recognition: The importance of surface reflectance properties. *Perception*, *36*(9), 1368–1374.

Russell, R., Sinha, P., Biederman, I., & Nederhouser, M. (2006). Is pigmentation important for face recognition? Evidence from contrast negation. *Perception*, *35*(6), 749–759.

Sandford, A., & Rego, S. (2019). Recognition of deformed familiar faces: Contrast negation and nonglobal stretching. *Perception*, *48*(10), 992–1012.

Sandford, A., Sarker, T., & Bernier, T. (2018). Effects of geometric distortions, Gaussian blur, and contrast negation on recognition of familiar faces. *Visual Cognition*, *26*(3), 207–222.

Sinha, P., Balas, B., Ostrovsky, Y., & Russell, R. (2006). Face recognition by humans: Nineteen results all computer vision researchers should know about. *Proceedings of the IEEE*, *94*(11), 1948–1962.

Stevens, K. A. (1983). Slant-tilt: The visual encoding of surface orientation. *Biological Cybernetics*, *46*(3), 183–195.

Todd, J. T., Koenderink, J. J., Van Doorn, A. J., & Kappers, A. M. (1996). Effects of changing viewing conditions on the perceived structure of smoothly curved surfaces. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(3), 695.

Todd, J. T., & Petrov, A. A. (2022). The many facets of shape. *Journal of Vision*, *22*(1), 1–1.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology*, *43*(2), 161–204.