Reviewed Preprint
v1 • October 22, 2025
Not revised

Reviewed Preprint
v2 • April 2, 2026
Revised by authors

Reviewed Preprint
v3 • June 16, 2026
Revised by authors

The view tolerance of human identity recognition depends on horizontal face information

Alexia Roux-Sibilon[1,2], Hélène Dumont[1], Vincent Bremhorst[3], Christianne Jacobs[1], Valérie Goffaux[1,4]

[1] Psychological Sciences Research Institute (IPSY), UCLouvain, Louvain-la-Neuve, Belgium • [2] Université Clermont-Auvergne, CNRS, LAPSCO, Clermont-Ferrand, France • [3] Statistical Methodology and Computing Service (SMCS), UCLouvain, Louvain-la-Neuve, Belgium • [4] Institute of Neurosciences (IoNS), UCLouvain, Louvain-la-Neuve, Belgium

For correspondence:
valerie.goffaux@uclouvain.be

Funding: See page 21

Reviewing editor: Tessa Dekker, University College London, United Kingdom

eLife Assessment

This **important** study combines behavioural psychophysics with image-based observer modelling to investigate which visual information can support view-tolerant face identity recognition. It offers **convincing** evidence that although diagnostic orientation content about identity varies with viewpoint - more horizontal for frontal views and more vertical for profiles - human recognition remains mainly tuned to horizontal information, identified by a view-tolerant model as carrying the most stable identity cues across viewpoints. Questions remain about how this generalises to ecological scenes and is biologically implemented.

https://doi.org/10.7554/eLife.108495.3.sa3

Abstract

Our ability to recognize objects and people across dramatic changes in physical appearance is a central yet unresolved question in vision science. In particular, the visual information that supports the human ability to recognize face identity across views is not fully specified. Past research suggests horizontally oriented face information plays a key role. To test this hypothesis, we investigated how humans use the visual information physically available across different orientation ranges of the face stimulus to recognize identity in a view-tolerant manner.

Human observers performed an old/new identity recognition task with face stimuli presented under different viewpoints, achieved by rotating the faces in yaw (from left to right profile) and filtered to preserve contrast in selective orientation ranges. We found that human recognition performance remained tuned to the horizontal range of face information irrespective of yaw.

We used a model observer approach to define the information physically available in the stimulus for matching face identity within each viewpoint (view-selective model observer) or across different viewpoints (view-tolerant model observer). The view-selective model indicated that face identity is carried by orientation ranges shifting from horizontal in frontal views to vertical in profile views. In contrast, the view-tolerant model showed that the horizontal range provides the most stable identity cues across views. The horizontally-tuned orientation profile of human recognition performance was predicted by the high diagnosticity of horizontal information in frontal views and the stability of the horizontal identity cues across views. The informativeness and stability of the identity cues carried by horizontal face information are confirmed by additional model observer evidence that this range best predicts the average summary of full-spectrum face appearances across views.

Our findings indicate that the invariant representation of a face, gradually learned through repeated exposure to its natural appearance statistics, relies primarily on horizontal facial information. By identifying the spatial information supporting view-tolerant face recognition in

humans, the present work yields concrete, data-driven constraints for the refinement of theoretical and computational models of visual recognition.

Introduction

The way in which we recognize objects and people is a central but as yet unresolved question in vision science. Our understanding of visual perception is still limited by our ability to account for how we recognize objects and people despite the sometimes radically different images they project onto the retinae, due to varying lighting, distance, depth rotation, etc. (DiCarlo & Cox, 2007). We experience faces under a broad range of depth rotations mainly along the x axis, i.e. from left to right profile (i.e. yaw; Favelle & Palmisano, 2012) due to biomechanical constrains favoring x-axis head rotations and the vantage point on our conspecifics' faces also varying more along the x than the y axis (i.e. when moving around people). While rotating in depth, a given face projects retinal images that are more dissimilar than the ones that different faces under the same viewpoint would project (Adini et al., 1997; Burton et al., 2016; Hill et al., 1997). Yet humans generally have no difficulty in recognizing a familiar face across views, which implies the joint ability to differentiate its identity from others and to generalize it from one view to another, i.e. with tolerance to variations (Ritchie & Burton, 2017).

The tolerance of face identity recognition is stronger for familiar than unfamiliar faces (e.g., Favelle & Palmisano, 2018; Hancock et al., 2000; Hill et al., 1997; Jeffery et al., 2006; Johnston & Edmonds, 2009; Jones et al., 2017; Liu & Chaudhuri, 2002; Newell et al., 1999; O'Toole et al., 1998; Troje & Bülthoff, 1996), suggesting that a core determinant of tolerant recognition is the repeated exposure to the natural statistics of a person's face (e.g., to the variability of a person appearance across different views; Tian & Grill-Spector, 2015; Troje & Bülthoff, 1996; Van Meel & Op de Beeck, 2018; Wallis & Bülthoff, 2001). Such statistical learning is assumed to be the main unsupervised learning route for tolerant face/object recognition in humans and animals (DiCarlo & Cox, 2007; Fiser & Aslin, 2002; Hauser et al., 2001; Huber et al., 2023; Li & DiCarlo, 2008, 2010, 2012; Tian & Grill-Spector, 2015). Since seeing the different views of a given person in succession improves the tolerance of identity recognition, temporal contiguity is presumed to contribute to the statistical learning of face identity (Li & DiCarlo, 2010; Miyashita, 1993; Pitts & McCulloch, 1947; Van Meel & Op de Beeck, 2018; Wallis et al., 2009; Wallis & Bülthoff, 1999).

While some attention has been devoted to the contribution of temporal contiguity in view-tolerant recognition, the *spatial* aspects of the natural statistics supporting view-tolerant face identity recognition are still largely elusive. Face appearance results from the complex interaction between extrinsic viewing conditions and the intrinsic 3D shape and reflectance properties of the face (determined e.g., by viewpoint and lighting; Favelle et al., 2017; Hill & Bruce, 1996; Liu et al., 2000). Burton and colleagues (Burton, 2013; Burton et al., 2005) suggested that as exposure to multiple appearances of a person increases, the accidental, irrelevant variations would be progressively whitened (i.e. averaged out) and reveal the stable cues to identity. Figure 1 simulates such averaging using the pictures of a given identity taken from variable poses, as experienced in natural viewing (Oruc et al., 2019). It can be seen that additionally to the whitening of accidental properties, the resulting average contains a strong horizontal structure; namely it results in the emergence of the so-called (horizontal) bar code of the face (Dakin & Watt, 2009).

We, and others, proposed that the horizontal content of the face stimulus may drive the visual mechanisms engaged for the view-tolerant recognition of face identity (Caldara & Seghier, 2009; Dakin & Watt, 2009; Gilad-Gutnick et al., 2018; Goffaux, 2008; Goffaux & Dakin, 2010). Several lines of empirical evidence indicate that the horizontally-oriented face information conveys optimal cues to identity. First, the visual mechanisms specialized for the identity recognition of faces' frontal view rely preferentially on the horizontal structure of the face image, indicating a better sensitivity to identity in the horizontal range of face information (e.g., Balas et al., 2015; Dumont et al., 2024; Goffaux, 2019; Goffaux & Dakin, 2010; Pachai et al., 2013, 2018) The horizontal dependence of face identity recognition has also been shown to predict

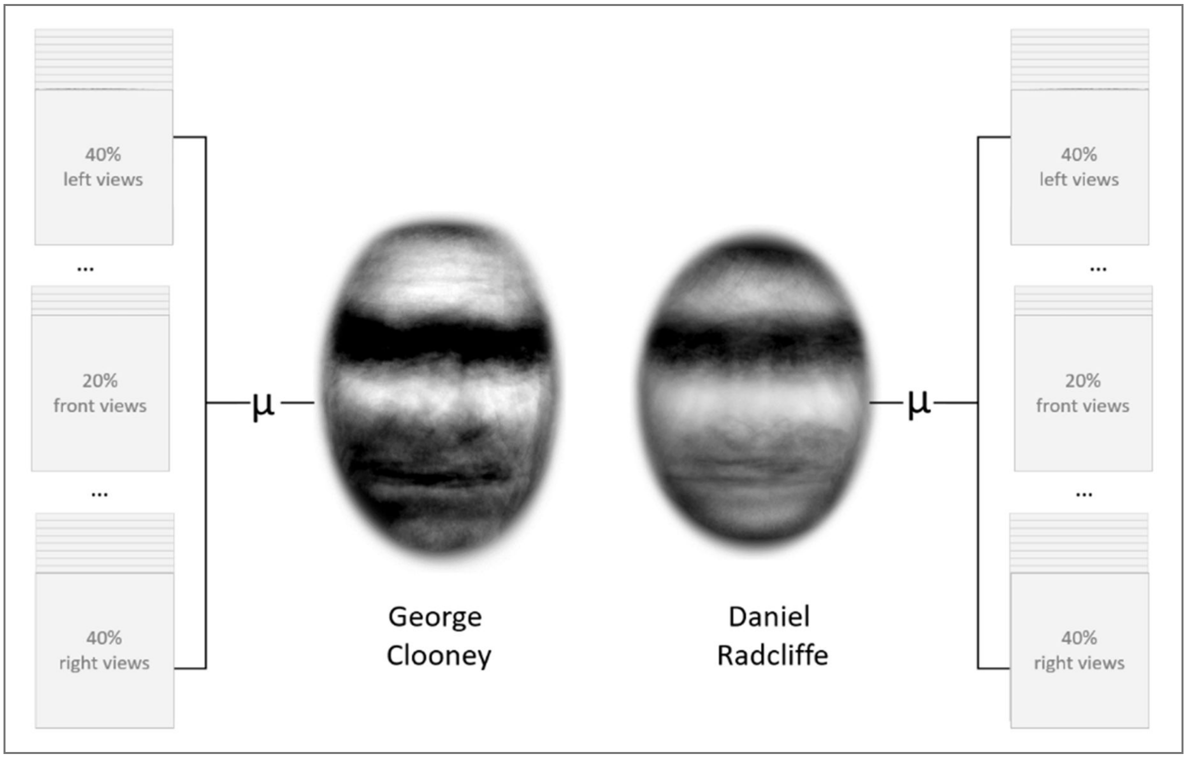

Figure 1. Graphic illustration of the horizontal smearing occurring when averaging multiple views of a face.

The tolerance of human face identity recognition to drastic appearance variations caused by varying lighting, viewpoint, facial expression, etc. has been proposed to emerge through averaging (Burton et al., 2005). With increased exposure to a face, an averaging mechanism would progressively whiten accidental variations in appearance while preserving stable cues to identity . Past illustrations (Burton, 2013; Burton et al., 2005) used varying lighting and expressions but moderate pose variations. Here we show that when pose varies largely from left to right profile, averaging horizontally smears the face suggesting that across encounters with a face, cues at orientations other than horizontal are whitened. As the observer learns the natural statistics of a person's face, it seems plausible that they increasingly rely on horizontal cues for identity recognition. Images of two celebrities (George Clooney and Daniel Radcliffe) were sampled from the internet and sorted into three view categories: frontal, left-, and right-averted. In order to illustrate the horizontal smearing due to view variations, the averages were made of 40% left-averted, 40% right-averted, and 20% frontal views, in line with exposure to face views in natural viewing (Oruc et al., 2019). The luminance and RMS contrast of the averaged faces were set to a luminance of .5 and contrast of .4. Using this procedure, one can appreciate the emergence of the so-called bar code, namely the vertical arrangement of horizontally-oriented cues which carries the natural statistics of the face category and of face individual identity (Dakin & Watt, 2009).

face identification accuracy at the individual observer level (Duncan et al., 2019; Pachai et al., 2013). Furthermore, there is indirect evidence that horizontal face information may optimally drive the tolerance of face identity recognition. For example, while the horizontal dependence manifests from three months of age (de Heering et al., 2016), it strengthens over the lifespan as individuals accumulate experience with the natural statistics of face appearance (Goffaux et al., 2015). Moreover, and considering that the recognition of familiar faces differs from unfamiliar face recognition by its stronger tolerance to retinal variations due to e.g., a change in view (Bruce et al., 1999; Burton et al., 1999, 2010; Hill & Bruce, 1996; Megreya & Burton, 2006; O'Toole et al., 1998; Ramon, 2015), the finding by Pachai and colleagues that the recognition of a face from % to full-frontal view increasingly relies on horizontal information as a function of familiarity provides indirect support to the notion that the tolerance of identity recognition most crucially depends on horizontal face cues (Pachai et al., 2017). In one of the Goffaux and Dakin (2010)'s experiments, the matching of unfamiliar faces from frontal to % view was more largely disrupted by horizontal than vertical noise masking. It is not clear, however, whether this effect reflects view-tolerant recognition or the horizontally-tuned encoding of the frontal view probe. The important question of whether the horizontal range of face information is a privileged informational avenue for view-tolerant identity recognition is therefore still unanswered.

The present study aims at directly investigating the hypothesis that the tolerance of human face identity recognition is supported by the horizontal range of face information. First, we familiarized a sample of human observers with multiple views of a set of face identities. In an old-new recognition task involving face stimuli presented in various views, we demonstrated that humans stay broadly tuned to the horizontal range of face information irrespective of yaw, with a stronger tuning observed for frontal views.

Second, we used a model observer approach to define the information physically available in the stimulus for the matching of face identity within a given viewpoint and across different viewpoints. A model observer is a basic image processing algorithm that cross-correlates a target image with probe images at the pixel level, and selects the probe with the highest correlation, i.e. the most likely match. Combined with orientation filtering, this method provides a formal way to describe how the information most useful for matching identity is distributed in the orientation domain of the face image (e.g., Collin et al., 2014; Gold et al., 1999; Näsänen, 1999; for a similar approach in the spatial frequency domain). The pixelwise image correlation performed by a model observer is not taken as a computationally valid analog of human face processing; on the contrary, it is a means to separate information objectively available in the stimulus from human vision (Collin et al., 2014).

We tested two model observers on the same multiple views of face identities as used in the human old-new recognition task. One so-called *view-tolerant* model observer matched identities across distinct views (e.g., matching a profile view of identity A with the other views of all identities). This model observer necessarily relied on the physical image properties that were most stable across views. We compared the performance of the view-tolerant model to a *view-selective* model observer, which matched identities within the same view (e.g., matching a profile view of identity A with the profile view of all identities), therefore revealing what orientations were physically (irrespective of human vision) the most diagnostic to match faces at each view specifically.

This approach revealed a substantial difference in the orientation distribution of view-specific and view-tolerant information. The view-selective model indicated that, at the image level, the orientation ranges supporting view-selective identity matching differ drastically from one view to another : identity cues were distributed in the horizontal range at frontal views (in line with Goffaux, 2019; Keil, 2008; Pachai et al., 2013) but shifted towards the vertical range as the face turned to profile. In contrast, the view-tolerant model performance indicated that the horizontal structure of the face stimulus provides the most stable cues for matching identity across yaws. Partial correlations between model and human observer performance suggest that the horizontal tuning of human face identity recognition is due not only to the high diagnosticity of horizontal information at frontal view but also to the stability of the horizontal identity cues across views.

Methods

Subjects

Twenty-two healthy young adults took part in this experiment in exchange for monetary compensation (8 euros per hour of testing). They were 14 females and eight males, aged 23.5 (± 3.4) on average (four were left-handed), recruited via Facebook advertisement. They received a written description of the experiment protocol and gave their written informed consent. Participants had normal vision as verified by a Landolt C acuity test (conducted using FRACT; Bach, 1996). Participants wore optical corrections when necessary. The experimental protocol was approved by the local ethical committee (Psychological Sciences Research Institute, UCLouvain).

Stimuli

We selected 30 face identities (15 male, 15 female) from the 3D laser-scanned face database of the Max-Planck Institute for Biological Cybernetics (Tuebingen, Germany; Troje & Bülthoff, 1996). Faces were viewed under seven different viewpoints ranging from −75° (left profile) to +75° (right profile) in steps of 25° (full front faces at 0°; see Figure 2). We first converted all images to a gray scale ranging from 0 to 1 and resized them so that all faces subtended a height of 210 pixels. All images were padded into 400×400 pixels gray canvas and alpha-blended with a view-specific aperture designed to cover the hair and neck of the average of all face images at a given viewpoint (using Adobe Photoshop).

Next, images were normalized to obtain a mean luminance of 0 and root-mean square (RMS) contrast of 1 and submitted to a fast two-dimensional Fourier transform to manipulate orientation content. Since manipulations in the Fourier domain apply to the whole image, they encompass both the face and background pixels. When the image of a face on a plain background is filtered in the Fourier orientation domain, energy belonging to the face therefore smears to the background and vice-versa, resulting in an oriented halo. To minimize this smearing, we applied an iterative phase-scrambling procedure (as in Canoluk et al., 2023; Petras et al., 2019; Roux-Sibilon et al., 2023; Schuurmans et al., 2023), which consists in iteratively phase-scrambling the image, pasting the original face pixels, and phase-scrambling again (50 iterations). By making the power spectra of the face and background pixels more comparable, this procedure minimizes smearing during orientation filtering. The amplitude spectra of the resulting images were then multiplied with wrapped Gaussian filters (standard deviation of 20°) centered on orientations ranging from 0° (vertical) to 157.5° in steps of 22.5°. After inverse-Fourier transform, the filtered images were combined with a view-specific aperture.

Images of faces cropped from their background as used here contain most of their energy in the horizontal range (Goffaux, 2019; Goffaux & Greenwood, 2016; Keil, 2009). Across yaws, we found face energy to range between .11 and .14 on a 0 to 1 grayscale, which is moderate compared to the range of face energy variations we measured across identities (from .08 to .18). To prevent energy from explaining our results, in all images, the luminance and RMS contrast of the face pixels were fixed to 0.55 and 0.15, respectively, and background pixels were uniformly set to 0.55. The percentage of clipped pixel values (below 0 or above 1) per image did not exceed 3%.

We used custom-written scripts in Matlab 2014a (Mathworks Inc, Natick, MA) for stimulus preparation. Stimuli were displayed against a grey background (0.55 luminance across RGB channels) at a viewing distance of 57 cm on a Viewpixx monitor (VPixx Technologies Inc., Saint-Bruno, Canada) with a resolution of 1920 × 1080 pixels and a 70Hz refresh rate using PsychToolbox (Brainard, 1997). With this display, face area subtended 5° (width) by 8° (height) of visual angle (approximately corresponding to conversational distance). At the start of the experiment lighting was switched off and the testing area of the lab was closed off separately with light-draining black curtains.

Procedure

Face identification performance was measured with an old/new recognition task.

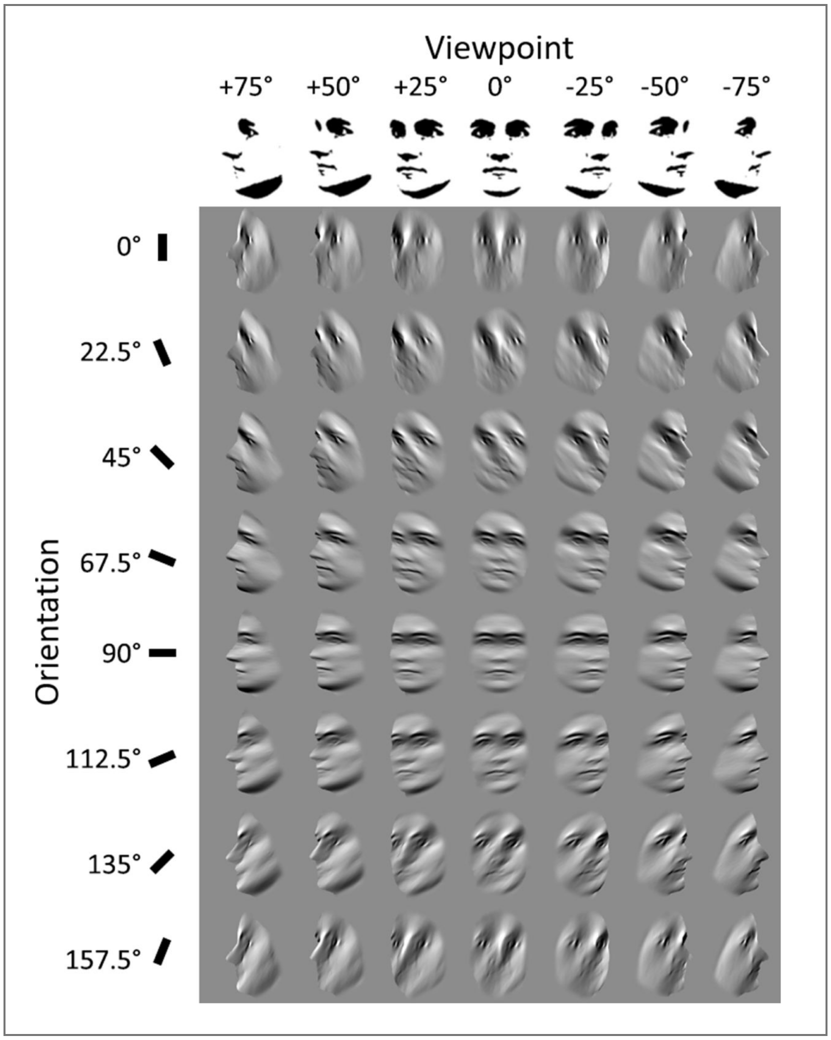

Figure 2. **Stimulus conditions.**

Columns. Each identity was viewed from seven different viewpoints ranging from +75° to −75° in steps of 25°. **Rows**. All images were filtered in the Fourier domain to preserve only a selective range of orientation, from 0° (vertical) to 157.5° in steps of 22.5°.

Participants were familiarized with five female and five male faces. These familiar identities were randomly sampled from the stimulus set of 30 face identities on the participant's first visit and were kept identical throughout the experiment. In the first testing session, we familiarized each participant with their selection of to-be-learnt 'old' faces, presented in their full spectrum version. To engage participants in the face learning procedure, they were asked to remember a face-name pairing, although during the main experiment they were only asked to judge whether the face was seen during familiarization ('old') or new (i.e. incidental learning method as in Liu & Chaudhuri, 2002). Each identity was first shown centrally rotating from left profile (+75°) to right profile (−75°) pseudo-dynamically along with its assigned first name on top of the screen (400ms per frame). The various views of the given face (in succession from +75° to −75° in steps of 25°, i.e. seven views) were then presented one by one for 400ms. After each identity presentation, a recap screen appeared with all the faces learned in previous trials, clustered by identity and shown under the various viewpoints side by side. All learnt identities were randomly presented one-by-one in the second familiarization phase, each under one of the seven viewpoints. The face appeared in isolation for 1000ms. Then the name associated with the face was added and both name and face were shown together for another 2000ms. Participants were then evaluated on their ability to name the learnt so-called 'old' faces at various viewpoints (+75° to −75°, in steps of 25°). A trial started with a 500ms fixation, then the stimulus was presented at screen center along with the name options (the five names of the 'old' faces of the same gender of the shown face), numbered from 1 to 5, at the top of the screen until participant response (maximum of 3000ms). Participants responded by pressing the corresponding key (1 to 5). The fixation turned green in case of a correct response. When response was incorrect, the fixation turned red. In both cases, the correct name appeared along with feedback. Feedback lasted for 500ms. Familiarization and test were looped until naming accuracy reached 80 %.

During the experimental old/new recognition task, faces were presented one by one at screen center and participants were instructed to determine for each of them whether it was a face with which they had been a priori familiarized ('old' face) or not ('new' face). They answered using a button box by pressing '1' for 'old' and '2' for 'new'. The main experiment was divided into 32 runs of 40 trials. A run started with recap screens for all learned identities under the seven viewpoints (from +75° to −75° in 25° steps) along with their associated name. In each main experiment trial, a face stimulus was presented centrally at a specified viewpoint and orientation range (from 0° to 157.5° in 22.5° steps). Viewpoint and orientation range varied randomly from one trial to the other. A trial started with a 1s fixation, next the face stimulus appeared until response or for 3s maximum. At the end of a run, participants received feedback on their average accuracy. To avoid inducing a response bias, there were as many 'old' as 'new' trials, making the 10 'old' faces twice as frequent as the 20 'new' ones, which were each presented only once per condition. There were 40 trials per condition and a total of 1280 trials. Participants first practiced the main experimental task on 20 randomly selected trials with full spectrum stimuli. If they reached 80% accuracy, they could start the main experiment with the filtered stimuli. If practice accuracy was lower than 80%, participants were invited to run the familiarization (learning and test) again. Whenever accuracy in an experiment run dropped below 55% correct, and every third run regardless of performance, participants were presented with the recap screens again. All through the experiment, they were encouraged to respond as accurately and rapidly as possible.

The total experiment lasted for 1.5 hours on average, split into three testing sessions.

Human data analysis

To prevent outlier responses from contaminating the results, we applied a log10 transformation on response latencies at the trial level and excluded trials with latencies at more than 2.5 times the standard deviation above and below the individual mean, in each participant. This procedure resulted in the exclusion of 1.81% of the trials on average.

In order to estimate the orientation dependence of human sensitivity to identity across views, we derived individual *d'* at each viewpoint and orientation. To do so, we determined hit and false alarm rates (Tanner & Birdsall, 1958) from the 'old'/'new' response in each participant, at each orientation and for each viewpoint. Following the log-linear rule (Hautus, 1995), we added a 0.5 correction to both before calculating the z scores. Performance at the 0° filter was duplicated to have circular filter values from 0° to 180°. As expected from previous studies (Dumont et al., 2024; Goffaux & Greenwood, 2016; Pachai et al., 2018), the d’ plotted as a function of orientation in the frontal view condition (yaw = 0°) depicted a bell-shaped function centred on horizontal orientation (i.e. 90°; see Figure 3A, central panel). At the other viewpoints, d’ followed a very similar shape, with maximum sensitivity roughly centred on horizontal orientation. Therefore, fitting the human sensitivity data using a Gaussian model seemed most appropriate as it allows characterizing the parameters of the tuning profile, namely, peak location, peak amplitude, standard deviation and base amplitude, which are directly interpretable in cognitive/functional terms. Moreover, the use of a nonlinear, Gaussian model is motivated by past work that showed that the Gaussian function fits the evolution of recognition performance as a function of orientation (Dakin & Watt, 2009; Goffaux & Greenwood, 2016). Simpler frameworks, i.e. a linear model predicting d’ from the interaction between orientation and viewpoint, would result in an 8 (orientation) * 7 (viewpoint) design that is difficult to analyze and interpret.

The Gaussian model is defined as:

$$f\left(orientation\right) = BaseAmplitude + PeakAmplitude$$
$$.\exp(-\frac{\left(orientation - PeakLocation\right)^2}{2StandardDeviation^2})$$

with *orientation* being the orientation of the filter, ranging from 0 to 180°. The model estimates four free parameters. *Peak Location* is the orientation on which the gaussian curve is centred. *Standard Deviation* is the width of the Gaussian curve (i.e. strength of the tuning). *Base Amplitude* is the height of the Gaussian curve base (i.e. the minimum sensitivity, typically found near vertical orientations). *Peak Amplitude* refers to the height of the Gaussian curve relative to its baseline, reflecting the advantage of horizontal over vertical orientations for horizontally centered Gaussians.

We used the package *brms* (Bürkner, 2018) in R to implement this model in the Bayesian framework, using a multilevel modelling approach (Dumont et al., 2024; Moors et al., 2020; Nalborczyk et al., 2019). The four parameters *Peak Location, Standard Deviation, Base Amplitude,* and *Peak Amplitude* were conjointly estimated by linear predictor terms which included an intercept and the effect of Viewpoint. We also estimated a subject-level intercept of *Standard Deviation* and *Base Amplitude* as random effects, to allow the shape of the gaussian to vary across participants. This multilevel structure provides a more accurate estimation of population-level parameters by accounting for subject variability. The prior distributions of the different parameters of the model were specified based on a compromise between (1) using knowledge from previous research (e.g., we used a normal distribution centred on horizontal orientation − 90° – for *Peak Location,* in line with the well-established horizontal tuning of face identification), (2) keeping unbiased uncertainty when previous research was not informative (e.g., for the effect of Viewpoint on the different parameters), and (3) allowing the convergence of the model. The details of the prior distributions can be found in Supplementary File 1. We ran four Markov Chain Monte-Carlo simulations, with 20,000 iterations including 3,000 warm-up iterations.

Model diagnostics of the model were checked and indicated a good convergence: the potential scale reduction factor (R-hat) was of 1.00 for all parameters, the Bulk Effective Sample Size (ESS) was superior to 10,000 for all but four parameters (Intercept of the Base Amplitude: Bulk ESS = 5504; Intercept of the Peak Amplitude: Bulk ESS = 8494; Subject-level random effect of the Standard Deviation: Bulk ESS = 5965; Subject-level random effect of the Base Amplitude: Bulk ESS = 6308), and the Tail ESS was superior to 10.000 for all but one parameters (Subject-level random effect of the Standard Deviation: Bulk ESS = 5321).

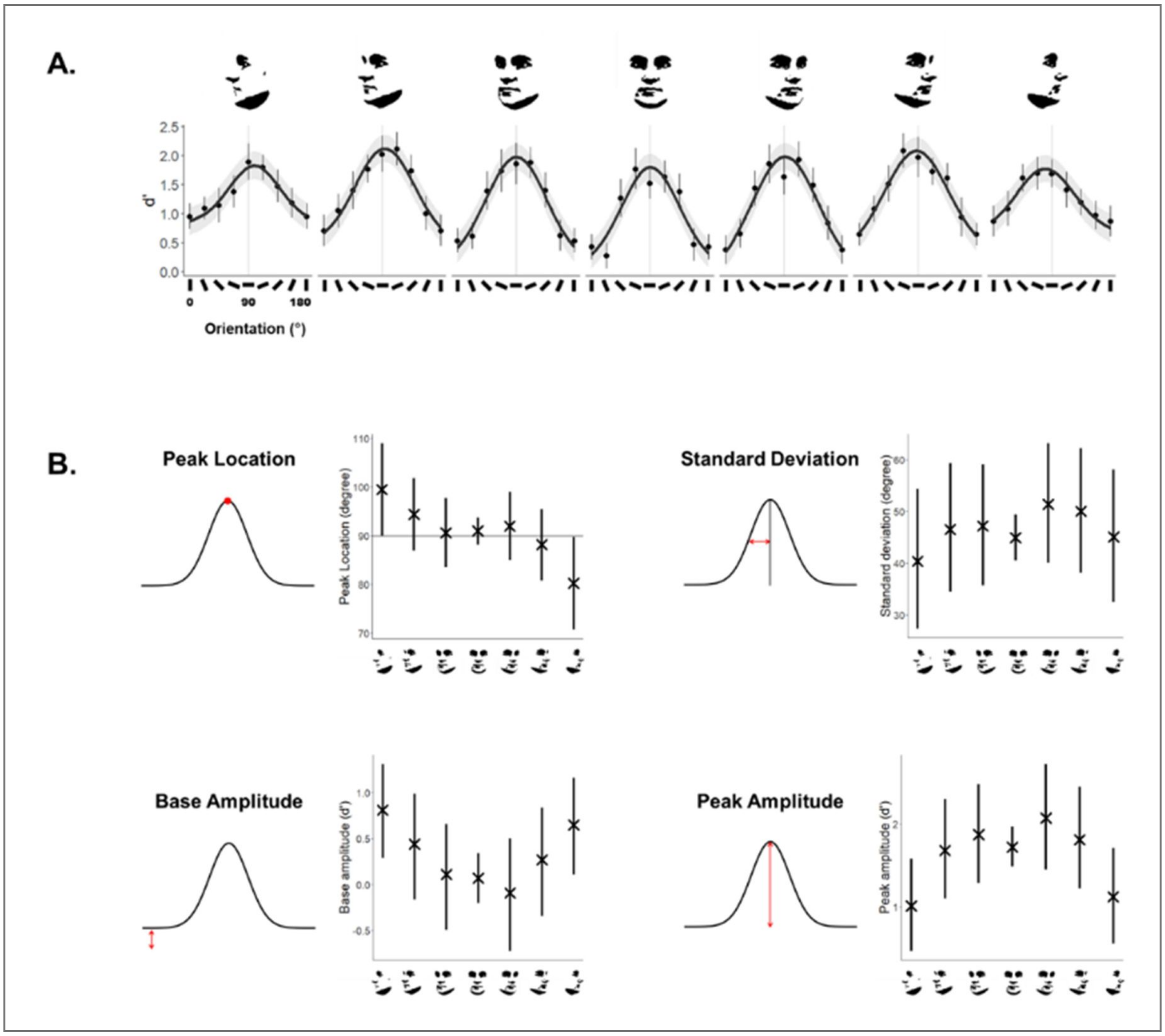

Figure 3.

A. Sensitivity of human observers to facial identity (d’) as a function of the orientation filter (0° to 180° in 22.5° steps), and face viewpoint (yaw: +75° to −75° in 25° steps). Dots and error bars represent mean d’ values and 95% confidence intervals across participants. Solid lines and shaded areas indicate the mean posterior predictions and 95% credible intervals from the Gaussian Bayesian multilevel model. **B.** Population-level mean parameters of the Gaussian Bayesian Multilevel model, plotted with 95% credible intervals as a function of face viewpoint. The 95% credible intervals reflect the uncertainty of the model. They indicate a 95% probability that the true population parameter lies within the specified range, given the observed data.

Image analysis

We investigated how the distribution of energy across orientations in the experimental stimuli varied as a function of viewpoint. The amplitude and phase spectra for the image of each full spectrum face were obtained by means of a two-dimensional fast Fourier Transform and multiplied with wrapped Gaussian filters, with peak orientations centred on orientation values from 0° to 157.5° in steps of 22.5° (20° standard deviation; as in Goffaux & Greenwood, 2016) and with peak spatial frequencies ranging from 1 to 200 cycles per image in 20 logarithmic steps. Amplitude values within each spatial frequency and orientation bin were squared and summed, then averaged across spatial frequencies per orientation bin.

Note that because the Fourier transform represents image energy in a discrete manner, energy at the lowest spatial frequency components can only be reliably sampled at the main cardinal and oblique ranges (i.e. 0°, 45°, 90°, 135°) especially when very narrow orientation filters are used (Hansen & Essock, 2004). However, because orientation filters were broad, and we interpret relative differences and not absolute values of energy distribution across viewpoints, the influence of this issue is minimal. We obtained one mean energy value per orientation band by averaging across identities.

We found that while face images contained most of their energy in the horizontal range, irrespective of viewpoint (Figure 3), the amplitude of the horizontal energy peak decreased as the face viewpoint moved away from frontal. For profile views, there was a plateau covering the horizontal plus the adjacent left and right oblique orientations for the left- and right-pointing profile views, respectively. In comparison, vertical energy was lower than any other range regardless of viewpoint. These findings replicate past evidence that most of the energy in a face image is contained in a range centred over the horizontal angle (Goffaux, 2019; Goffaux & Greenwood, 2016; Keil, 2009).

Yet the distribution of image energy as a function of orientation does not provide any insight about its potential usefulness for face identity recognition; we addressed this using model observers.

Model observers

To systematically quantify the orientation profile of the stimulus information physically available to match identity either (1) within a fixed viewpoint or (2) across distinct viewpoints (i.e. thus requiring the generalization of identity from one view to other views), we designed two different model observers: a view-selective model observer and a view-tolerant model observer, respectively. By characterizing the orientation profile of the information available to recognize face identity in a view-selective and -tolerant manner, this approach enables to investigate objectively whether and how this information is used by humans (Collin et al., 2014; Pachai et al., 2013). Namely, it allows addressing whether the orientation range used during view-tolerant identity recognition by humans is the most informative for the view at stake, or the one most stable across views, at the physical image level.

Model observers matched the same randomly selected 10 “old” faces and 20 “new” faces as the human observers; we ran 22 view-selective and 22 view-tolerant model observers to match the number of tested human subjects. We presented each of the 30 faces once per condition. In each trial, the models computed the pixelwise similarity based on the calculation of a cross-correlation between an orientation-filtered target image (either from the ‘old’ or ‘new’ set) and each of the possible exemplars of the same gender (i.e. probes) filtered at the same orientation as the target image. The probe face with the highest correlation (i.e. the more similar) to the target image was selected as the model response (winner-take-all scheme). Depending on the correspondence between the “winner” probe and the actual target, the responses of the model observer were categorized as hit or false alarm, allowing for the computation of a sensitivity d’ in each orientation and viewpoint condition along a procedure like the one used to compute human performance.

In both model types, targets and probes were of the same orientation range (see Collin et al., 2014; Näsänen, 1999 for a similar method applied to the spatial frequency domain). Targets and probes in the view-selective model observer stemmed from a fixed viewpoint whereas the view-tolerant model observer matched target and probe separated by more than one viewpoint step to force tolerance to viewpoint in this model performance (e.g., a face at +025 of yaw was matched to faces at +075, −025, −050, and −075). This separation of views between target and probes was essential to force the view-tolerant model observer to generalize identity across views when matching identity at the image pixel level. It is not meant to be a computationally-valid analogy of how humans recognize identity in a view-tolerant manner. Since profile views stood at the viewpoint continuum extrema, one extra viewpoint was available for comparison. We therefore decided to drop the mirror profile view to match the number of comparisons across profile and non-profile viewpoints. Performance of the view-tolerant model was averaged across comparison viewpoints.

In a pilot phase, we measured the overall identification performance of each model. Initially, the view-selective model performed at ceiling, yielding a correlation of 1 since there was an exact target-probe match across all trials. To avoid ceiling effects and to keep model performance close to human levels (Supplementary File 2), we thus decreased the signal-to-noise ratio (SNR) of the target and probe images to .125 by combining each with distinct noise patterns (face RMS contrast: .01; noise RMS contrast: .08). Each trial (i.e. target-probe pairing) was iterated ten times with different random noise patterns. Sensitivity d’ of the view-tolerant model was much lower than view-selective model and human sensitivity (Supplementary File 2), even without noise. The view-tolerant model therefore processed fully visible stimuli (SNR of 1). This decreased sensitivity in the view-tolerant compared to the view-selective model is expected, as none of the probes exactly matched the target at the pixel level due to viewpoint differences. In contrast to humans who rely on internally stored representations to match identity across views, the model observer lacks such internal representations and entirely relies on (less efficient) pixelwise comparisons.

The main objective of running these model observers is to interpret human orientation sensitivity profiles considering the available information in the stimulus; following this reasoning, model observer performance is only interesting when compared to human performance. Therefore, we investigated which model observer best predicted the orientation dependence profile of human face recognition using partial Pearson correlation analyses. We controlled for the variance in image energy across orientations and viewpoints in all partial correlation analyses to exclude the possibility that view-selective/tolerant model performance was a mere epiphenomenon of absolute oriented energy. The Fisher Z-transform of perfect correlations leading to infinite values, we replaced all −1 and 1 partial correlation coefficients to −.99 and .99, respectively, before applying Fisher Z-transformation and computing 95% confidence intervals (see Table 1). We (partial) correlated the orientation profiles of human and each model observer at each viewpoint separately (eight orientation vectors). For each specific viewpoint, the orientation sensitivity profile of each human observer was correlated to the average orientation sensitivity profile of either model observer, while controlling for the variance explained by (1) the average orientation sensitivity profile of the alternate model observer and (2) the average profile of orientation energy. The variability in human individual orientation profiles was taken as an estimate of the maximally achievable data-model correlation. We computed the crosscorrelation between each human individual orientation sensitivity profile and the average sensitivity profile of the remaining participants. The maximally achievable correlation was the mean of these individual-to-group correlations.

Table 1. **Posterior mean and 95% credible interval for each parameter of the Gaussian model, at each viewpoint.**

Parameter	Viewpoint	Estimate (posterior mean)	95% credible interval – lower bound	95% credible interval – upper bound
Peak Location	-75 (left profile)	99.5	90.02	108.98
	-50	94.4	86.98	101.87
	-25	90.62	83.58	97.73

	0 (full front view)	90.99	88.22	93.78
	25	91.98	85.00	99.01
	50	88.19	80.82	95.51
	75 (right profile)	80.24	70.73	89.76
Standard Deviation	-75 (left profile)	40.37	27.40	54.4
	-50	46.51	34.51	59.29
	-25	47.2	35.75	59.07
	0 (full front view)	44.94	40.55	49.41
	25	51.36	40.04	63.14
	50	50.02	38.18	62.23
	75 (right profile)	45.06	32.5	58.12
Base Amplitude	-75 (left profile)	0.81	0.29	1.31
	-50	0.44	-0.16	0.99
	-25	0.11	-0.49	0.66
	0 (full front view)	0.07	-0.2	0.34
	25	-0.09	-0.72	0.5
	50	0.27	-0.34	0.84
	75 (right profile)	0.65	0.11	1.16
Peak Amplitude	-75 (left profile)	1.01	0.47	1.58
	-50	1.68	1.10	2.30
	-25	1.87	1.29	2.48
	0 (full front view)	1.72	1.49	1.97
	25	2.07	1.45	2.72
	50	1.81	1.22	2.45
	75 (right profile)	1.12	0.56	1.71

Results - Human observers

Upon visual inspection, group-averaged d’ plotted as a function of orientation depicted a bell-shaped function roughly centred on horizontal orientation in the frontal view condition as in the other Viewpoint conditions (Figure 3A). The fitting confirmed that sensitivity to identity shows a roughly similar Gaussian orientation tuning profile across viewpoints (see bell-shaped curves on Figure 3A).

The relative stability of the human orientation sensitivity profile was corroborated by the stable and significantly positive correlations of the orientation tuning profile across viewpoints (rs > .67, ps < .05; Figure 4).

Despite human sensitivity for face identity always being best around horizontal orientations and worse around vertical ones, there were notable fluctuations in the orientation sensitivity profile across viewpoints. To better grasp these variations, we plotted the population-level estimates of the four parameters of the Gaussian curve as a function of viewpoint in Figure 3B. Population-level estimates and 95% credible intervals can also be found in Table 2.

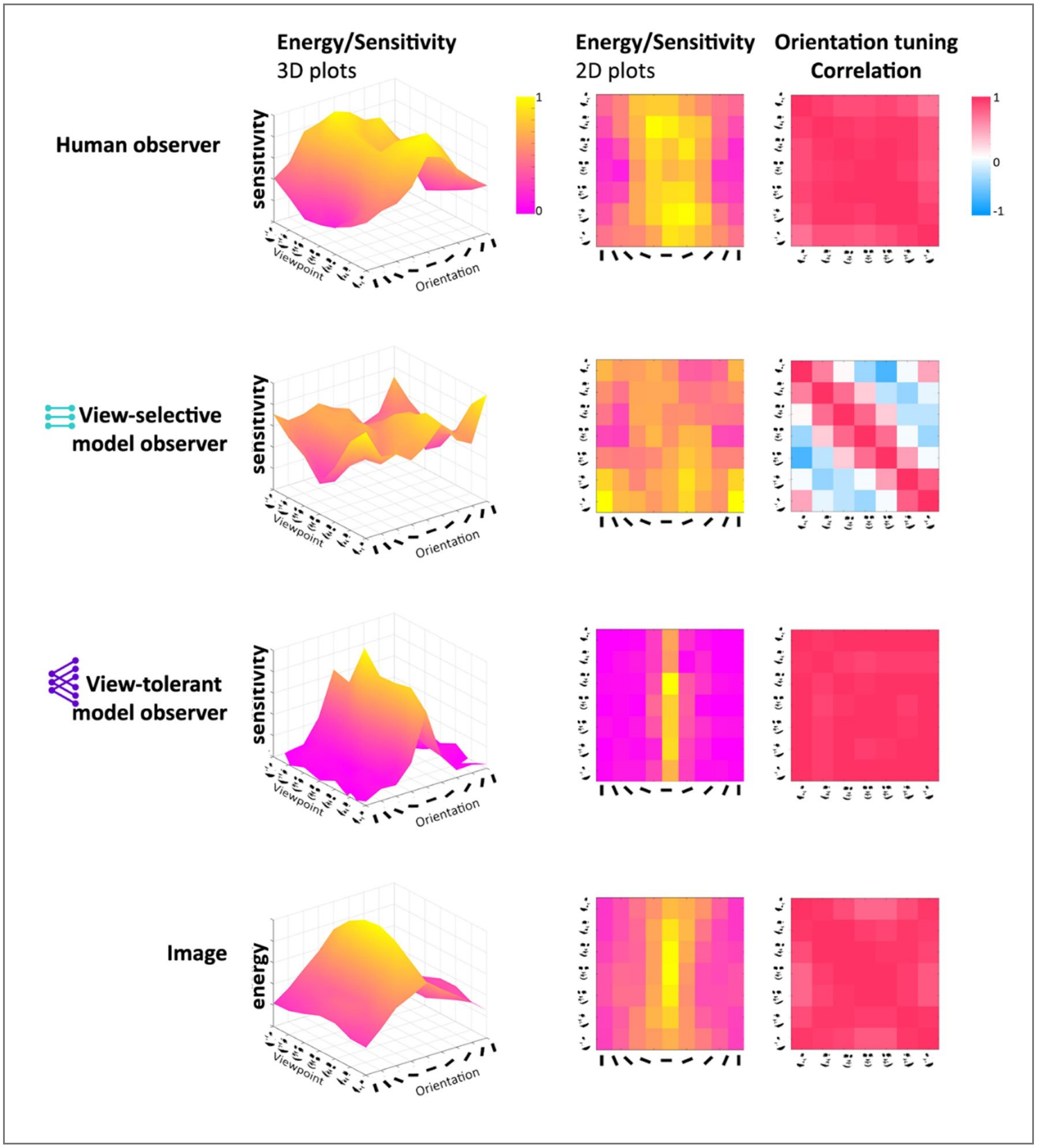

Figure 4. **Sensitivity (i.e. performance in the recognition task) of human and model observers and image energy across viewpoints and orientations.**

Left column. 3D surf plots of the normalized energy/sensitivity across orientation and viewpoints. *Middle column*. Matrix representations of the normalized energy/sensitivity across orientation and viewpoints. *Right column*. Matrix representations of the Pearson correlation (non Fisher Z-transformed) of the normalized orientation distributions of energy/sensitivity across viewpoints.

Table 2. **Mean and 95% confidence interval of the (Fisher Z-transformed) partial correlation coefficients between human and each model orientation d′ profiles while controlling for the variance in image energy and alternate model.**

Model	Viewpoint-specific		Viewpoint-tolerant	
	mean	95% ci	mean	95% ci
+075	0.06	[-0.09 0.2]	0.49	[0.30 0.67]
+050	0.04	[-0.10 0.18]	0.5	[0.34 0.66]
+025	0.56	[0.42 0.71]	0.3	[0.17 0.42]
000	0.79	[0.62 0.97]	0.19	[0.09 0.30]
-025	0.36	[0.18 0.54]	0.34	[0.17 0.51]
-050	0.1	[-0.06 0.25]	0.59	[0.45 0.72]
-075	-0.05	[-0.17 0.08]	0.56	[0.35 0.77]

Peak Location is estimated to lie close to 90° (grey horizontal line in Figure 3B) at all viewpoints except at the two profile views where peak location shifts toward adjacent obliques orientations (yaw = −75/75; see Figure 3A). Specifically, peak human sensitivity tended to shift towards left and right obliques for the most extreme leftward and rightward deviations in viewpoint, respectively. *Peak amplitude,* which corresponds to the difference in sensitivity between the vertical and horizontal ranges, is relatively stable across all viewpoints except for the two profile views where it is lower. *Base amplitude,* reflecting sensitivity in the vertical range, is highest for the two profile views and progressively decreases towards the full-frontal view. This pattern suggests that vertically-oriented face content is more diagnostic for profile than for other viewpoints. The variation of *Standard Deviation* across viewpoints can hardly be interpreted because of the high uncertainty of the estimations; the 95% credible intervals of the *Standard Deviation* for the different viewpoints mostly overlap.

Results - View-selective model observer

The performance of the model observer matching faces at specific viewpoints indicates that the most objectively diagnostic orientation range varies as a function of viewpoint (Figure 4).

Within the frontal view, optimal cues to identity are conveyed by orientations close to the horizontal angle (between 90° and 112.5°). The striking similarity of the orientation tuning profile of the view-selective model observer with the human performance indicates that at frontal view, the human visual system makes an efficient use of the information physically available in the stimulus. The peak of model performance shifted to the left and right oblique orientations closest to horizontal angle (67.5° and 112.5°) for leftward and rightward deviations in viewpoint, respectively. Additionally, the view-selective model observer increasingly relied on vertical cues as the face is turned towards its profile, in line with human performance.

We found only (significant) weakly positive correlations of the orientation sensitivity profiles across adjacent viewpoints (e.g., +075 and +050) confirming the orientation tuning variations across views (Figure 4).

Except for the frontal view, the orientation tuning of the view-selective model observer differs from human orientation profiles, which kept tuned to horizontal and adjacent oblique ranges irrespective of viewpoint.

Results - View-tolerant model observer

The model observer matching face identity across viewpoints performed in a drastically different way from the view-selective model observer. Recognition performance peaked sharply in horizontal angles and dropped abruptly at other ranges (Figure 4). Such orientation profile was relatively stable across viewpoints as shown by the homogeneous matrix of large positive and significant Pearson correlation coefficients for sensitivity profiles across viewpoints (rs > .88, ps < .002).

Such horizontal tuning was sharper than the one observed for humans with a much more severe drop of sensitivity in the vertical orientation range. The horizontal tuning of human performance is much shallower even at frontal view. It shows a similar dip in sensitivity for the vertical range but tends to decrease when moving away from the frontal view, which does not happen for the view-tolerant model observer that keeps sharply horizontally-tuned.

Results - Human Versus Model observers

What model observer best predicts the orientation tuning profile of human face recognition? Is it the view-selective model taking advantage of the identity cues that are optimal at selective viewpoints, or the view-tolerant model, which relies on the identity cues that are the most stable across viewpoints? To address this question directly, human orientation d’ profiles were correlated at the individual level with each model orientation d’ profile, when partialling out the variance explained by the alternate model and image energy. We submitted the so-obtained individual Fisher Z-transformed partial correlation coefficients to a repeated-measure ANOVA with Model and Viewpoint as within-subject factors.

These partial correlations showed that the orientation tuning profiles of human and view-selective model observers correlated strongly for frontal and near-frontal views. However, partial correlations dropped steeply as the views deviated further from frontal (Figure 5; Table 2). The correlation of orientation tuning profiles between human and view-tolerant model observers was lower overall but significant across all viewpoints.

The repeated-measure ANOVA confirmed that the partial correlations between human and model orientation sensitivity profiles differed depending on Model (F(1,21)= 5.21, p = .033, η^2= .017) and Viewpoint (F(3.85,80.9)= 11.064, p < .001, η^2= .11). The interaction between Model and Viewpoint was also significant (F(4.2,88.35)= 13.035, p < .001, η^2= .23).

We explored the impact of viewpoint on human-model correlation for the view-selective and viewtolerant models separately using a repeated-measure ANOVA with Viewpoint as a within-subject factor. For the view-selective model, this analysis revealed a robust effect of Viewpoint on the humanmodel partial correlation (F(3.86,80.99)= 22.795, p < .001, η^2= .52). For the view-tolerant model, the effect of Viewpoint on the human-model partial correlation was not significant (F(6,126)= 1.2, p= .32, η^2= .054). This confirms the relatively stable human-model partial correlation coefficients across viewpoints observed for the view-tolerant model and the fluctuant profile of human-model partial correlation coefficients for the view-specific model (peaking at frontal views and decreasing as moving toward profile views).

Furthermore, at each viewpoint, we examined which model observer best correlated with human orientation sensitivity profile (using Holm-corrected post-hoc tests on Fisher Z-transformed partial correlation coefficients; Table 3). The view-tolerant model predicted a significantly larger portion of the variance in human orientation sensitivity profile at +/-050 and +/-075 viewpoints, this difference was marginal for the +050 viewpoint. It is only for the identification of frontal views of faces that the view-selective model correlated best with human data. Correlations were of similar strength for the +075, +025, −025 (and +050) viewpoints (Table 3).

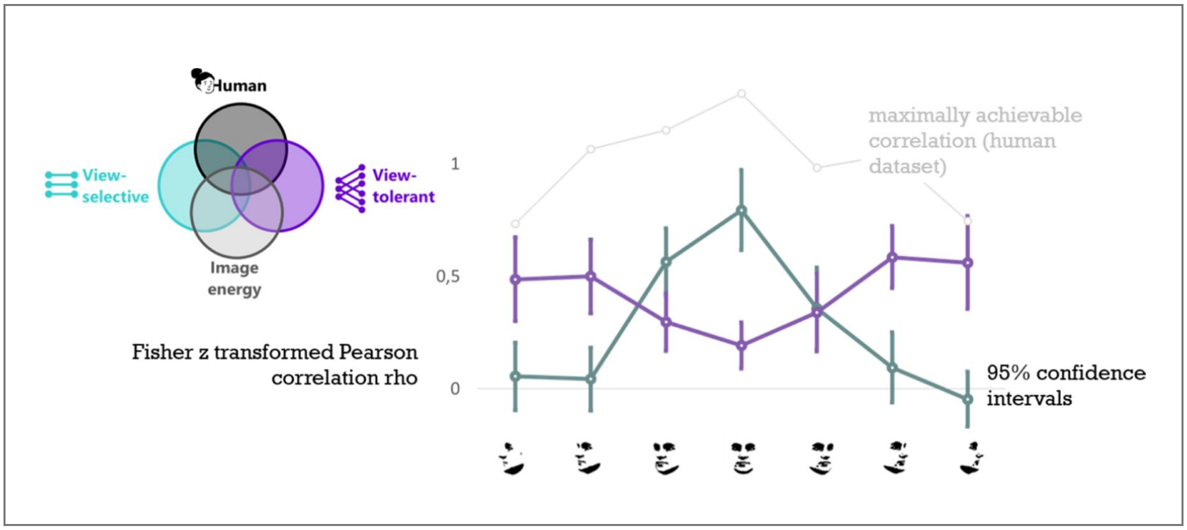

Figure 5. **Fisher Z-transformed Pearson partial correlation of the orientation sensitivity profiles between humans and each model, while controlling for the alternate model and image energy profiles.**

Error bars show the 95% confidence intervals. The faded grey line depicts the maximally achievable correlation for the separate viewpoint conditions in the human dataset (see Methods for details).

Table 3. **Difference of the human-model partial correlation coefficients between viewpoint-selective and view-tolerant model observers.**

Positive t values indicate a stronger correlation with the viewpoint-selective model and negative t values stronger correlation with the view-tolerant model.

		95% CI for Mean Difference			
View	Mean-Difference	Lower	Upper	t	p Holm
+075	-0.47	-0.86	-0.07	-4.20	0.002
+050	-0.43	-0.82	-0.03	-3.82	0.01
+025	0.23	-0.16	0.63	2.08	1
+000	0.69	0.30	1.09	6.21	< .001
-025	0.03	-0.37	0.42	0.23	1
-050	-0.61	-1.00	-0.21	-5.44	< .001
-075	-0.78	-1.18	-0.39	-7.02	< .001

Results - Is model observers' performance predicted by horizontal energy predominance?

In the above correlation analyses, we controlled the variance in image energy to yield a clean measure of the functional link between human and model observer performance. Here, we explore the possibility that the horizontal reliance of either model observer performance scales with the energy predominance of this orientation range in the stimulus image.

To analyze the influence of horizontal predominance on the model's performance, we computed for each image the horizontal minus vertical difference in terms of energy and model observer sensitivity. For each identity and at each viewpoint, energy and model sensitivity difference values were submitted to a Pearson correlation. We found a modest positive correlation between view-selective model observer and image energy (Pearson r = .24, $p < .0005$). There was no similar correlation for the view-tolerant model (Pearson rho = .025, p = .72) despite energy and view-tolerant model performance similarly peaking in the horizontal range across viewpoints.

What this analysis shows is that, in a given face, the stronger the predominance in horizontal energy (relative to the vertical energy), the more diagnostic the horizontally-oriented identity cues for view-selective recognition. However, horizontal energy predominance in a given face does not account for this range carrying the most stable cues across viewpoints.

This image-level analysis did not include human data since human sensitivity necessarily aggregates performance across trials. However, it would likely show a similar detachment from horizontal energy predominance as the view-tolerant model. Namely, humans are expected to rely on horizontal cues to match faces across views no matter the amount of horizontal energy predominance in the face at stake (e.g., George Clooney versus Brad Pitt).

Discussion

Tolerant face identity recognition relies on the ability to extract the idiosyncratic identity cues of a face from its highly variable appearance (e.g., Burton et al., 2016; Kramer et al., 2018). The spatial information supporting the tolerance of visual recognition is a central and debated topic in the field of visual and computational neurosciences (e.g., Andrews et al., 2023; DiCarlo & Cox, 2007). By showing that the information supporting tolerance in face identity recognition is image-computable, i.e. that it can be objectively defined in the orientation domain of the face image, the present work makes a decisive advance on this question. Our finding that view-tolerant face identity recognition is driven by the horizontal range of face information yields concrete, image-driven constraints for the development of theoretical and computational models of visual recognition.

Human participants performed an old/new identity recognition task on face stimuli presented under a variable view (ranging from left to right profile) and filtered to contain a restricted range of oriented content (ranging from 0° to 157.5° in steps of 22.5°; see Figure 2). For each view, recognition performance followed a Gaussian profile with a peak in the horizontal range and declining progressively when shifting towards the vertical range (Figure 3). Yet, while identity recognition stayed broadly tuned to the horizontal range irrespective of the vantage view, there were moderate but notable fluctuations in the tuning profiles. First, the peak location of the Gaussian tuning profile slightly and gradually shifted away from horizontal towards right and left adjacent obliques as the face turned to left or right profile, respectively. Second, the base amplitude of the Gaussian increased drastically towards profile views. The U-shaped profile of base amplitude as a function of viewpoint shows the increasing contribution of vertically oriented information as the face moves away from the frontal view (see Figure 3B). In other words, while human identity recognition stays tuned to the horizontal range irrespective of viewpoint, it tends to increasingly rely on vertical (and close-to-vertical) orientations as the face turns to profile.

As the vantage point of a face shifts away from frontal view, morphological features related to the 3D structure of the face become more apparent (e.g., nose and cheek protuberances, jaw line, nose bridge, eyebrow head; see e.g., Stephan & Caine, 2007 for evidence that the nose gains in informativeness from frontal to profile view). Our finding of moderate but systematic fluctuations in the tuning peak location and that vertical ranges gain importance in non-frontal views suggests that non-horizontal orientations may facilitate access to such features.

In contrast, other sources of information are lost when shifting towards profile views, such as the bilateral symmetric organization of the face as well as the 2D shape of features and their configuration along the x axis (e.g., size of eyebrows, interocular distance; Royer et al., 2016; Troje & Bülthoff, 1996). In a way, it is surprising that this shift in accessible information did not manifest in a more substantial variation of the orientation tuning profile peak location. Instead, the relatively stable preference for horizontal information across views suggests that it is not solely due to this range facilitating access to the 2D shape properties and bilateral symmetry (Dakin & Watt, 2009) but aligns with our recent evidence (Dumont et al., 2024) that the horizontal face information yields not only 2D-shape but also surface cues to identity. The stable horizontal tuning may also be due to identity recognition relying mostly on the eye region irrespective of view. Indeed, despite the eye appearance is strongly affected by changes in view, the identity information extracted from this region is the most diagnostic irrespective of view (Stephan & Caine, 2007; but see (Royer et al., 2016) for conflicting results) and stays best defined in the horizontal range.

Using a model observer approach, we investigated how the human observer makes use of the identity cues physically available in the image across orientations and views. It is indeed important to compare the human tuning profile to an image-based quantification of the information that is available in the stimulus in order to characterize more formally the human sampling specificities at stake when recognizing face identity. We designed two model observers to measure the information available in the stimulus to match (i.e. cross-correlate) face identity within and across views and disentangle stimulus information available for view-selective and view-tolerant recognition, respectively.

Let's first summarize the findings related to the view-selective model observer. It showed that the diagnostic orientation ranges for matching face identity in a view-specific manner vary greatly across views (Figure 4). The view-selective model observer was horizontally-tuned for frontal views of faces, in a manner strikingly similar to the human performance profile. This suggests that human observers make a close-to-optimal use of orientation information in face images when identifying frontal face views. Akin to human recognition, the view-selective model progressively increased its reliance on vertically-oriented cues when the face moved from frontal to profile. However – and in contrast to human performance – this came at the expense of the horizontal tuning, which vanished completely. What these findings suggest is that while the horizontal range conveys the optimal cues to identity at frontal view, it loses its informativeness in non-frontal views; at the image pixel level, face identity at non-frontal views is indeed predominantly carried by non-horizontal ranges of information. Therefore, view-specific informativeness does not account for the generalization of the horizontal tuning profile of human identity recognition across views.

In contrast to the view-selective model observer, the view-tolerant model observer kept sharply tuned to the horizontal range irrespective of view (Figure 4). The horizontal range resulted in the highest cross-correlation among the different views of a given face indicating that this range yields the identity cues that are the most stable across views, i.e. those that enable binding different face views into a unique representation of identity (Burton, 2013b). This physical property of the face image likely explains why human identity recognition keeps horizontally-tuned across views (see also Goffaux & Dakin, 2010).

For a direct comparison of human and model performance, we quantified the variance shared between each model observer and human orientation tuning profile while controlling image energy and the alternate model (Figure 5). These analyses confirmed that the view-selective model best predicted the orientation tuning of human identity recognition at frontal and close-to-frontal views, that are typically experienced during face-to-face conversations, but not at profile and close-to-profile views. In contrast, the variance explained by the view-tolerant model was relatively stable across viewpoints. The particularly strong horizontal tuning of human identity recognition is thus likely due to the visual system extracting a representation that simultaneously prioritizes the orientation range conveying the cues that are the richest at frontal view and the most stable across views.

Tolerant representation of identity presumably builds up through repeated exposure to the physical variability of face appearance (Burton, 2013; Burton et al., 2005). Over encounters, the idiosyncratic properties of a face are reinforced in its internal representation at the expense of uninformative accidental properties, the representation of which gradually attenuate (Figure 1; Tian & Grill-Spector, 2015; Troje & Bülthoff, 1996; Van Meel & Op de Beeck, 2018; Wallis & Bülthoff, 2001). The stable identity cues are horizontal, which entails that this range should best predict the average face appearance across views. Following a suggestion from a reviewer of this paper, we addressed this hypothesis by designing an additional model observer, which matched the experimental stimuli (i.e. orientation-filtered images of faces seen under different views) to the average of the different views of each face identity in full spectrum (Figure 6; see details in the Supplementary File 3). In line with our hypothesis, this 'view-average' model observer performed best with horizontally filtered faces. These additional results support our suggestion that the average summary of a face identity contains a strong horizontal structure, a so-called (horizontal) bar code (Figure 1), which drives the view-tolerant recognition of face identity (Caldara & Seghier, 2009; Dakin & Watt, 2009; Gilad-Gutnick et al., 2018; Goffaux, 2008; Goffaux & Dakin, 2010).

Our stimuli, originally designed by Troje & Bülthoff, 1996, are 3D laser scans of white individuals aged between 20 and 40 years, posing with a neutral expression and shot under a fixed illumination. Ears and a small portion of the neck were visible while the hair region was removed. All face images had a normalized skin color, and we further converted them to grayscales. The use of such a controlled set restricted the range of within- and between-identity variations compared to what is experienced in natural settings. However, the skin color normalization and grayscale conversion, while limiting the range of face variability, did not eliminate the contribution of surface pigmentation in our study as the grayscale 3D laser scanned faces used here contained natural variations in crucial surface cues such as skin albedo and texture. For these reasons, we believe that the present findings generalize to more ecological viewing conditions.

The horizontal tuning of human face recognition was found to be relatively broad across views, compared with the sharp tuning of the view-tolerant model observer. This tuning breadth may serve to retain information of the complex manifold of a given face appearance variability across views. Indeed, each face identity has its own idiosyncratic, way to vary across expression,

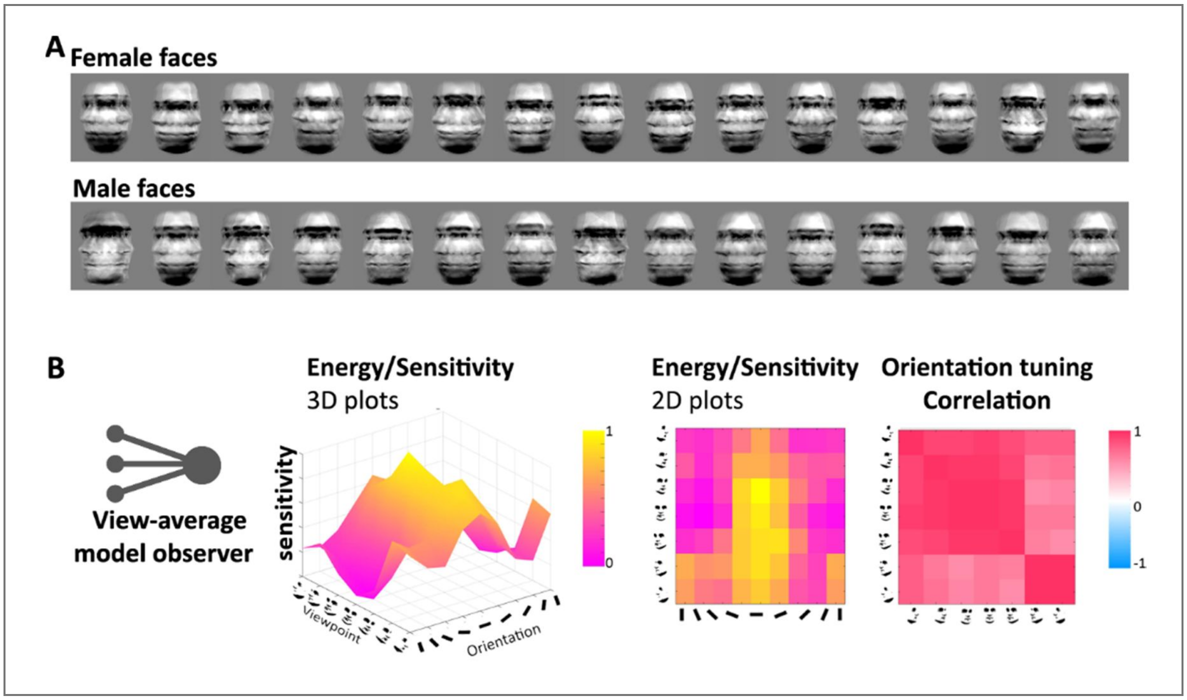

Figure 6. **View-average model observer.**

A. Image averages of the different views of each face identity of the stimulus set. B. Sensitivity of the view-average model observer across views and orientations. *From left to right:* 3D surf plots of the normalized sensitivity across orientation and viewpoints, matrix representations of the normalized sensitivity across orientation and viewpoints, and matrix representations of the Pearson correlation of the normalized orientation distributions of sensitivity across viewpoints.

illumination, view, etc. Such idiosyncratic within-person variability has been proposed to drive familiar face identity recognition as much as the stable (horizontal) identity information (e.g., (Burton et al., 2016; Ritchie & Burton, 2017). Moreover, accidental properties such as head and gaze orientation carry important cues for the regulation of social interactions. Our past evidence shows that the fine discrimination of gaze direction is best supported in the vertical range (Goffaux, 2019). The vertical range also likely carries most of the information about the head direction. Thus, the broad horizontal tuning of identity recognition by humans may allow for the integration of such accidental properties of a face with the (more) stable identity (Or & Wilson, 2010). In other words, for functional social interactions, it may be advantageous to retain the dynamic and variable signals emitted by a face as much as its invariant aspects, which would entail a relatively broad tuning to orientation.

Effect of lighting are even more disruptive than view changes (Adini et al., 1997; Braje et al., 1998; Favelle et al., 2017; Hill & Bruce, 1996; A. Johnston et al., 2013; Tarr et al., 1998, 2008). Image representations that emphasized the horizontal features were found to be less sensitive to changes in the direction of illumination (Adini et al., 1997). Future research should test whether tolerance of human face identity recognition to lighting is also supported by the horizontal range.

To conclude, this study demonstrates that the horizontal range carries the richest identity cues in frontal views, and the most stable across views. The orientation tuning profile of human identity recognition aligns with this combination of high diagnostic value in frontal views and cross-view stability. Taken together, this body of evidence suggests that the invariant representation of face identity, gradually learned through repeated exposure to its natural appearance statistics, relies heavily on horizontal facial information (Figure 1; (Burton et al., 2016; Dakin & Watt, 2009; Ritchie & Burton, 2017).

Data availability

The stimuli and experimental stimulation scripts, the human and model observer data is publicly available at *https://osf.io/fe8s9/*.

Acknowledgements

We thank Julie Juaneda, Zoe Strapazzon, Hana Zjakic, and Stien Van de Plas for their help with the data collection. H.D. is supported by the Belgian National Fund for Scientific Research (FRS-FNRS). V.G. is a research associate of the FRS-FNRS.

Additional information

Contributor Roles

A.R.-S.: Formal analysis, Conceptualization, Visualization, , Writing - review & editing;

H.D.: Formal analysis, Data curation, Visualization;

V.B.: Formal analysis;

C.J.: Methodology, Investigation;

V.G.: Conceptualization, Methodology, Formal analysis, Funding acquisition, Visualization, writing - original draft, Writing - review & editing, Project administration, Supervision.

Funding

Funder	Grant reference number	Author
Fonds De La Recherche Scientifique - FNRS (FNRS)	Aspirant	Helene Dumont
Fonds De La Recherche Scientifique - FNRS (FNRS)	Research Associate	Valerie Goffaux

Author ORCID iDs

Alexia Roux-Sibilon: https://orcid.org/0000-0002-3810-0109
Valérie Goffaux: https://orcid.org/0000-0003-3182-7011

Additional files

Supplementary file 1.

Supplementary file 2.

Supplementary file 3.

References

Adini Y., Moses Y., Ullman S. (1997) Face recognition: The problem of compensating for changes in illumination direction. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **19**:721-732 https://doi.org/10.1109/34.598229

Andrews T. J., Rogers D., Mileva M., Watson D. M., Wang A., Burton A. M. (2023) A narrow band of image dimensions is critical for face recognition. *Vision Research* **212**:108297 https://doi.org/10.1016/j.visres.2023.108297 | PubMed

Balas, Schmidt J., Saville A. (2015) A face detection bias for horizontal orientations develops in middle childhood. *Frontiers in Psychology* **6** https://doi.org/10.3389/fpsyg.2015.00772 | PubMed

Brainard D. H. (1997) The Psychophysics Toolbox. *Spatial Vision* **10**:433-436 https://doi.org/10.1163/156856897X00357 | PubMed

Braje W. L., Kersten D., Tarr M. J., Troje N. F. (1998) Illumination effects in face recognition. *Psychobiology* **26**:371-380 https://doi.org/10.3758/BF03330623

Bruce V., Henderson Z., Greenwood K., Hancock P. J. B., Burton A. M., Miller P. (1999) Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied* **5**:339-360 https://doi.org/10.1037/1076-898X.5.4.339

Bürkner P.-C. (2018) Advanced Bayesian Multilevel Modeling with the R Package brms. *R Journal* **10**:395-411 https://doi.org/10.32614/RJ-2018-017

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

Burton A. M., Bruce V., Hancock P. j. b. (1999) From Pixels to People: A Model of Familiar Face Recognition. *Cognitive Science* **23**:1-31 https://doi.org/10.1207/s15516709cog2301_1

Burton A. M., Jenkins R., Hancock P. J. B., White D. (2005) Robust representations for face recognition: The power of averages. *Cognitive Psychology* **51**:256-284 https://doi.org/10.1016/j.cogpsych.2005.06.003 | PubMed

Burton A. M., Kramer R. S. S., Ritchie K. L., Jenkins R. (2016) Identity From Variation: Representations of Faces Derived From Multiple Instances. *Cognitive Science* **40**:202-223 https://doi.org/10.1111/cogs.12231 | PubMed

Burton A. M., White D., McNeill A. (2010) The Glasgow Face Matching Test. *Behavior Research Methods* **42**:286-291 https://doi.org/10.3758/BRM.42.1.286 | PubMed

Caldara R., Seghier M. L. (2009) The Fusiform Face Area responds automatically to statistical regularities optimal for face categorization. *Human Brain Mapping* **30**:1615-1625 https://doi.org/10.1002/hbm.20626 | PubMed

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

Collin C. A., Rainville S., Watier N., Boutet I. (2014) Configural and Featural Discriminations Use the Same Spatial Frequencies: A Model Observer versus Human Observer Analysis. *Perception* **43**:509-526 https://doi.org/10.1068/p7531 | PubMed

Dakin S. C., Watt R. J. (2009) Biological 'bar codes' in human faces. *Journal of Vision* **9**:2.1-10 https://doi.org/10.1167/9.4.2 | PubMed

de Heering A., Goffaux V., Dollion N., Godard O., Durand K., Baudouin J.-Y. (2016) Three-month-old infants' sensitivity to horizontal information within faces. *Developmental Psychobiology* **58**:536-542 https://doi.org/10.1002/dev.21396 | PubMed

DiCarlo J. J., Cox D. D. (2007) Untangling invariant object recognition. *Trends in Cognitive Sciences* **11**:333-341 https://doi.org/10.1016/j.tics.2007.06.010 | PubMed

Dumont H., Roux-Sibilon A., Goffaux V. (2024) Horizontal face information is the main gateway to the shape and surface cues to familiar face identity. *PLOS One* **19**:e0311225 https://doi.org/10.1371/journal.pone.0311225 | PubMed

Duncan J., Royer J., Dugas G., Blais C., Fiset D. (2019) Revisiting the link between horizontal tuning and face processing ability with independent measures. *Journal of Experimental Psychology: Human Perception and Performance* **45**:1429-1435 https://doi.org/10.1037/xhp0000684 | PubMed

Favelle S., Hill H., Claes P. (2017) About Face: Matching Unfamiliar Faces Across Rotations of View and Lighting. *I-Perception* **8**:2041669517744221 https://doi.org/10.1177/2041669517744221 | PubMed

Favelle S., Palmisano S. (2012) The Face Inversion Effect Following Pitch and Yaw Rotations: Investigating the Boundaries of Holistic Processing. *Frontiers in Psychology* **3** https://doi.org/10.3389/fpsyg.2012.00563 | PubMed

Favelle S., Palmisano S. (2018) View specific generalisation effects in face recognition: Front and yaw comparison views are better than pitch. *PLOS One* **13**:e0209927 https://doi.org/10.1371/journal.pone.0209927 | PubMed

Fiser J., Aslin R. N. (2002) Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **28**:458-467 https://doi.org/10.1037/0278-7393.28.3.458

Gilad-Gutnick S., Harmatz E. S., Tsourides K., Yovel G., Sinha P. (2018) Recognizing Facial Slivers. *Journal of Cognitive Neuroscience* **30**:951-962 https://doi.org/10.1162/jocn_a_01265 | PubMed

Goffaux V. (2008) The horizontal and vertical relations in upright faces are transmitted by different spatial frequency ranges. *Acta Psychologica* **128**:119-126 https://doi.org/10.1016/j.actpsy.2007.11.005 | PubMed

Goffaux V. (2019) Fixed or flexible? Orientation preference in identity and gaze processing in humans. *PLOS One* **14**:e0210503 https://doi.org/10.1371/journal.pone.0210503 | PubMed

Goffaux V., Dakin S. C. (2010) Horizontal information drives the behavioral signatures of face processing. *Frontiers in Psychology* **1**:143 https://doi.org/10.3389/fpsyg.2010.00143 | PubMed

Goffaux V., Greenwood J. A. (2016) The orientation selectivity of face identification. *Scientific Reports* **6**:34204 https://doi.org/10.1038/srep34204 | PubMed

Goffaux V., Poncin A., Schiltz C. (2015) Selectivity of Face Perception to Horizontal Information over Lifespan (from 6 to 74 Year Old). *PLOS One* **10**:e0138812 https://doi.org/10.1371/journal.pone.0138812 | PubMed

Gold J., Bennett P. J., Sekuler A. B. (1999) Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research* **39**:3537-3560 https://doi.org/10.1016/S0042-6989(99)00080-2 | PubMed

Hancock P. J. B., Bruce V., Burton A. M. (2000) Recognition of unfamiliar faces. *Trends in Cognitive Sciences* **4**:330-337 https://doi.org/10.1016/S1364-6613(00)01519-9 | PubMed

Hansen B. C., Essock E. A. (2004) A horizontal bias in human visual processing of orientation and its correspondence to the structural components of natural scenes. *Journal of Vision* **4**:5 https://doi.org/10.1167/4.12.5 | PubMed

Hauser M. D., Newport E. L., Aslin R. N. (2001) Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition* **78**:B53-B64 https://doi.org/10.1016/S0010-0277(00)00132-3 | PubMed

Hautus M. J. (1995) Corrections for extreme proportions and their biasing effects on estimated values of d'. *Behavior Research Methods, Instruments, & Computers* **27**:46-51 https://doi.org/10.3758/bf03203619

Hill H., Bruce V. (1996) Effects of lighting on the perception of facial surfaces. *Journal of Experimental Psychology: Human Perception and Performance* **22**:986-1004 https://doi.org/10.1037/0096-1523.22.4.986

Hill H., Schyns P. G., Akamatsu S. (1997) Information and viewpoint dependence in face recognition. *Cognition* **62**:201-222 https://doi.org/10.1016/S0010-0277(96)00785-8 | PubMed

Huber L. S., Geirhos R., Wichmann F. A. (2023) The developmental trajectory of object recognition robustness: Children are like small adults but unlike big deep neural networks. *Journal of Vision* **23**:4 https://doi.org/10.1167/jov.23.7.4 | PubMed

Jeffery L., Rhodes G., Busey T. (2006) View-Specific Coding of Face Shape. *Psychological Science* **17**:501-505 https://doi.org/10.1111/j.1467-9280.2006.01735.x | PubMed

Johnston A., Hill H., Carman N. (2013) Recognising Faces: Effects of Lighting Direction, Inversion, and Brightness Reversal. *Perception* **42**:1227-1237 https://doi.org/10.1068/p210365n | PubMed

Johnston R. A., Edmonds A. J. (2009) Familiar and unfamiliar face recognition: A review. *Memory (Hove, England)* **17**:577-596 https://doi.org/10.1080/09658210902976969 | PubMed

Jones S. P., Dwyer D. M., Lewis M. B. (2017) The Utility of Multiple Synthesized Views in the Recognition of Unfamiliar Faces. *Quarterly Journal of Experimental Psychology* **70**:906-918 https://doi.org/10.1080/17470218.2016.1158302 | PubMed

Keil M. S. (2008) Does face image statistics predict a preferred spatial frequency for human face processing?. *Proceedings of the Royal Society B: Biological Sciences* **275**:2095-2100 https://doi.org/10.1098/rspb.2008.0486 | PubMed

Keil M. S. (2009) ‘I Look in Your Eyes, Honey’: Internal Face Features Induce Spatial Frequency Preference for Human Face Processing. *PLOS Computational Biology* **5**:e1000329 https://doi.org/10.1371/journal.pcbi.1000329

Kramer R. S. S., Young A. W., Burton A. M. (2018) Understanding face familiarity. *Cognition* **172**:46-58 https://doi.org/10.1016/j.cognition.2017.12.005 | PubMed

Li N., DiCarlo J. J. (2008) Unsupervised Natural Experience Rapidly Alters Invariant Object Representation in Visual Cortex. *Science* **321**:1502-1507 https://doi.org/10.1126/science.1160028 | PubMed

Li N., DiCarlo J. J. (2010) Unsupervised Natural Visual Experience Rapidly Reshapes Size-Invariant Object Representation in Inferior Temporal Cortex. *Neuron* **67**:1062-1075 https://doi.org/10.1016/j.neuron.2010.08.029 | PubMed

Li N., DiCarlo J. J. (2012) Neuronal Learning of Invariant Object Representation in the Ventral Visual Stream Is Not Dependent on Reward. *The Journal of Neuroscience* **32**:6611-6620 https://doi.org/10.1523/JNEUROSCI.3786-11.2012 | PubMed

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**:729-743 https://doi.org/10.1068/p3065 | PubMed

Liu & Chaudhuri (2002) Reassessing the 3/4 view effect in face recognition. *Cognition* **83**:31-48 https://doi.org/10.1016/S0010-0277(01)00164-0 | PubMed

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

Miyashita Y. (1993) Inferior Temporal Cortex: Where Visual Perception Meets Memory. *Annual Review of Neuroscience* **16**:245-263 https://doi.org/10.1146/annurev.ne.16.030193.001333 | PubMed

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

Nalborczyk L., Batailler C., Lœvenbruck H., Vilain A., Bürkner P.-C. (2019) An Introduction to Bayesian Multilevel Models Using brms: A Case Study of Gender Effects on Vowel Variability in Standard Indonesian. *Journal of Speech, Language, and Hearing Research: JSLHR* **62**:1225-1242 https://doi.org/10.1044/2018_JSLHR-S-18-0006 | PubMed

Näsänen R. (1999) Spatial frequency bandwidth used in the recognition of facial images. *Vision Research* **39**:3824-3833 https://doi.org/10.1016/S0042-6989(99)00096-6 | PubMed

Newell F. N., Chiroro P., Valentine T. (1999) Recognizing Unfamiliar Faces: The Effects of Distinctiveness and View. *The Quarterly Journal of Experimental Psychology Section A* **52**:509-534 https://doi.org/10.1080/713755813 | PubMed

Or C. C.-F., Wilson H. R. (2010) Face recognition: Are viewpoint and identity processed after face detection?. *Vision Research* **50**:1581-1589 https://doi.org/10.1016/j.visres.2010.05.016 | PubMed

Oruc I., Shafai F., Murthy S., Lages P., Ton T. (2019) The adult face-diet: A naturalistic observation study. *Vision Research* **157**:222-229 https://doi.org/10.1016/j.visres.2018.01.001 | PubMed

O'Toole A. J., Edelman S., Bülthoff H. H. (1998) Stimulus-specific effects in face recognition over changes in viewpoint. *Vision Research* **38**:2351-2363 https://doi.org/10.1016/S0042-6989(98)00042-X | PubMed

Pachai M., Bennett P. J., Sekuler A. B. (2018) The Bandwidth of Diagnostic Horizontal Structure for Face Identification. *Perception* **47**:397-413 https://doi.org/10.1177/0301006618754479 | PubMed

Pachai M., Sekuler A. B., Bennett P. J., Schyns P. G., Ramon M. (2017) Personal familiarity enhances sensitivity to horizontal structure during processing of face identity. *Journal of Vision* **17**:5 https://doi.org/10.1167/17.6.5 | PubMed

Pachai M., Sekuler A., Bennett P. (2013) Sensitivity to Information Conveyed by Horizontal Contours is Correlated with Face Identification Accuracy. *Frontiers in Psychology* **4** https://doi.org/10.3389/fpsyg.2013.00074 | PubMed

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

Pitts W., McCulloch W. S. (1947) How we know universals the perception of auditory and visual forms. *The Bulletin of Mathematical Biophysics* **9**:127-147 https://doi.org/10.1007/BF02478291 | PubMed

Ramon M. (2015) Perception of global facial geometry is modulated through experience. *PeerJ* **3**:e850 https://doi.org/10.7717/peerj.850 | PubMed

Ritchie K. L., Burton A. M. (2017) Learning faces from variability. *Quarterly Journal of Experimental Psychology* **70**:897-905 https://doi.org/10.1080/17470218.2015.1136656 | PubMed

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

Royer J., Blais C., Barnabé-Lortie V., Carré M., Leclerc J., Fiset D. (2016) Efficient visual information for unfamiliar face matching despite viewpoint variations: It's not in the eyes!. *Vision Research* **123**:33-40 https://doi.org/10.1016/j.visres.2016.04.004 | PubMed

Schuurmans J. P., Bennett M. A., Petras K., Goffaux V. (2023) Backward masking reveals coarse-to-fine dynamics in human V1. *NeuroImage* **274**:120139 https://doi.org/10.1016/j.neuroimage.2023.120139 | PubMed

Stephan B. C. M., Caine D. (2007) What is in a View? The Role of Featural Information in the Recognition of Unfamiliar Faces across Viewpoint Transformation. *Perception* **36**:189-198 https://doi.org/10.1068/p5627 | PubMed

Tanner W. P., Birdsall T. G. (1958) Definitions of d ' and η as Psychophysical Measures. *The Journal of the Acoustical Society of America* **30**:922-928 https://doi.org/10.1121/1.1909408

Tarr M. J., Georghiades A. S., Jackson C. D. (2008) Identifying faces across variations in lighting: Psychophysics and computation. *ACM Transactions on Applied Perception* **5**:1-25 https://doi.org/10.1145/1279920.1279924

Tarr M. J., Kersten D., Bülthoff H. H. (1998) Why the visual recognition system might encode the effects of illumination. *Vision Research* **38**:2259-2275 https://doi.org/10.1016/S0042-6989(98)00041-8 | PubMed

Tian M., Grill-Spector K. (2015) Spatiotemporal information during unsupervised learning enhances viewpoint invariant object recognition. *Journal of Vision* **15**:7 https://doi.org/10.1167/15.6.7 | PubMed

Troje N. F., Bülthoff H. H. (1996) Face recognition under varying poses: The role of texture and shape. *Vision Research* **36**:1761-1771 https://doi.org/10.1016/0042-6989(95)00230-8 | PubMed

Van Meel C., Op de Beeck H. P. (2018) Temporal Contiguity Training Influences Behavioral and Neural Measures of Viewpoint Tolerance. *Frontiers in Human Neuroscience* **12** https://doi.org/10.3389/fnhum.2018.00013 | PubMed

Wallis G., Backus B. T., Langer M., Huebner G., Bulthoff H. (2009) Learning illumination- and orientation-invariant representations of objects throughtemporal association. *Journal of Vision* **9**:6-6 https://doi.org/10.1167/9.7.6 | PubMed

Wallis G., Bülthoff H. (1999) Learning to recognize objects. *Trends in Cognitive Sciences* **3**:22-31 https://doi.org/10.1016/S1364-6613(98)01261-3 | PubMed

Wallis G., Bülthoff H. H. (2001) Effects of temporal association on recognition memory. *Proceedings of the National Academy of Sciences* **98**:4800-4804 https://doi.org/10.1073/pnas.071028598 | PubMed

Peer reviews

Reviewer #1 (Public review):

[Editors' note: this version has been assessed by the Reviewing Editor without further input from the original reviewers. The authors have addressed the comments raised in the previous round of review.]

Summary:

The authors describe the results of a single study designed to investigate the extent to which horizontal orientation energy plays a key role in supporting view-invariant face recognition. The authors collected behavioral data from adult observers who were asked to complete an old/new face matching task by learning broad-spectrum faces (not orientation filtered) during a familiarization phase and subsequently trying to label filtered faces as previously seen or novel at test. This data revealed a clear bias favoring the use of horizontal orientation energy across viewpoint changes in the target images. The authors then compared different ideal observer models (cross-correlations between target and probe stimuli) to examine how this profile might be reflected in the image-level appearance of their filtered images. This revealed that a model looking for the best matching face within a viewpoint differed substantially from human data, exhibiting a vertical orientation bias for extreme profiles. However, a model forced to match targets to probes at different viewing angles exhibited a consistent horizontal bias in much the same manner as human observers.

Strengths:

I think the question is an important one: The horizontal orientation bias is a great example of a low-level image property being linked to high-level recognition outcomes and understanding the nature of that connection is important. I found the old/new task to be a straightforward task that was implemented ably and that has the benefit of being simple for participants to carry out and simple to analyze. I particularly appreciated that the authors chose to describe human data via a lower-dimensional model (their Gaussian fits to individual data) for further analysis. This was a nice way to express the nature of the tuning function favoring horizontal orientation bias in a way that makes key parameters explicit. Broadly speaking, I also thought that the model comparison they include between the view-selective and view-tolerant models was a great next step. This analysis has the potential to reveal some good insights into how this bias emerges and ask fine-grained questions about the parameters in their model fits to the behavioral data.

Weaknesses:

I'll start with what I think is the biggest difficulty I had with the paper. Much as I liked the model comparison analysis, I also don't quite know what to make of the view-tolerant model. As I understand the authors' description, the key feature of this model is that it does not get to compare target and probe at the same yaw angle, but must instead pick a best match from candidates that are at different yaws. While it is interesting to see that this leads to a very different orientation profile, it also isn't obvious to me why such a comparison would be reflective of what the visual system is probably doing. I can see that the view-specific model is more or less assuming something like an exemplar representation of each face: You have the opportunity to compare a new image to a whole library of viewpoints and presumably it isn't hard to start with some kind of first pass that identifies the best matching view first before trying to identify/match the individual in question. What I don't get about the view-tolerant model is that it seems almost like an anti-exemplar model: You specifically lack the best viewpoint in the library but have to make do with the other options. I sort of understand the reasoning that this enforces tolerance of viewpoint variability, but I'm not clear on whether or not this is a version of face familiarity and recognition that the authors think has an analog in human visual processing.

I do think that this model is interesting in terms of the differential tuning it exhibits, but don't find it easy to align with any theoretical perspective on face recognition. Specifically, do the authors think there is a stage of face processing in which tolerance as they've operationalized it in the model is extant? What I'm looking for is a concrete description of the circumstances that the authors are saying lead to this kind of model potentially being a meaningful analog of face recognition. For example, is the idea that one may become familiar with a face in some very limited set of viewpoints and then be presented with that face in other views?

Alternatively, if the authors prefer to say that they simply thought this was a nice exercise in terms of identifying a different model and that it may not be a meaningful proxy for face recognition. I think that's fine, to be clear! I just still don't see anything in the text that convinces me of the ecological validity of this version of view-tolerance.

https://doi.org/10.7554/eLife.108495.3.sa2

Reviewer #2 (Public review):

This study investigates the visual information that is used for the recognition of faces. This is an important question in vision research and is critical for social interactions more generally. The authors ask whether our ability to recognise faces, across different viewpoints, varies as a function of the orientation information available in the image. Consistent with previous findings from this group and others, they find that horizontally filtered faces were recognised better than vertically filtered faces. Next, they probe the mechanism underlying this pattern of data by designing two model observers. The first was optimised for faces at a specific viewpoint (view-selective). The second was generalised across viewpoints (view-tolerant). In contrast to the human data, the view-specific model shows that the information that is useful for identity judgements varies according to viewpoint. For example, frontal face identities are again optimally discriminated with horizontal orientation information, but profiles are optimally discriminated with more vertical orientation information. These findings show human face recognition is biased toward horizontal orientation information, even though this may be suboptimal for the recognition of profile views of the face.

One issue in the design of this study was the lowering of the signal-to-noise ratio in the view-selective observer. This decision was taken to avoid ceiling effects. However, it is not clear how this affects the similarity with the human observers.

Another issue is the decision to normalise image energy across orientations and viewpoints. I can see the logic in wanting to control for these effects, but this does reflect natural variation in image properties. So, again, I wonder what the results would look like without this step.

Despite the bias toward horizontal orientations in human observers, there were some differences in the orientation preference at each viewpoint. For example, frontal faces were biased to horizontal (90 deg) but other viewpoints had biases that were slightly off horizontal (e.g. right profile: 80 deg, left profile: 100 deg). This does seem to show that differences in statistical information at different viewpoints (more horizontal information for frontal and more vertical information for profile) do influence human perception. It would be good to reflect on this nuance in the data.

Comments on revisions:

I am happy with the response and changes to the comments in my review. The key findings from this study are: (1) that there is bias toward the use of horizontal information across all viewpoints for face recognition in humans using an old-new recognition task. (2) In contrast, the optimal information for matching faces varies as a function of viewpoint. The view-selective model shows horizontal information is dominant for frontal views and vertical information is dominant for profile views.

The data from the view-tolerant model is less easy to interpret as it doesn't fit with any theoretically plausible model of face recognition. It might be a useful model for a face matching task in which participants had to match unfamiliar faces across viewpoints. This might be a possible extension of the current work.

Nonetheless, I still think this is an interesting contribution to the literature.

https://doi.org/10.7554/eLife.108495.3.sa1

Author response

The following is the authors’ response to the previous reviews

> ***Reviewer #1 (Public review):***
>
> *Summary:*
>
> *The authors describe the results of a single study designed to investigate the extent to which horizontal orientation energy plays a key role in supporting view-invariant face recognition. The authors collected behavioral data from adult observers who were asked to complete an old/new face matching task by learning broad-spectrum faces (not orientation filtered) during a familiarization phase and subsequently trying to label filtered faces as previously seen or novel at test. This data revealed a clear bias favoring the use of horizontal orientation energy across viewpoint changes in the target images. The authors then compared different ideal observer models (cross-correlations between target and probe stimuli) to examine how this profile might be reflected in the image-level appearance of their filtered images. This revealed that a model looking for the best matching face within a viewpoint differed substantially from human data, exhibiting a*

> *vertical orientation bias for extreme profiles. However, a model forced to match targets to probes at different viewing angles exhibited a consistent horizontal bias in much the same manner as human observers.*
>
> *Strengths:*
>
> *I think the question is an important one: The horizontal orientation bias is a great example of a low-level image property being linked to high-level recognition outcomes, and understanding the nature of that connection is important. I found the old/new task to be a straightforward task that was implemented ably and that has the benefit of being simple for participants to carry out and simple to analyze. I particularly appreciated that the authors chose to describe human data via a lower-dimensional model (their Gaussian fits to individual data) for further analysis. This was a nice way to express the nature of the tuning function, favoring horizontal orientation bias in a way that makes key parameters explicit. Broadly speaking, I also thought that the model comparison they include between the view-selective and view-tolerant models was a great next step. This analysis has the potential to reveal some good insights into how this bias emerges and ask fine-grained questions about the parameters in their model fits to the behavioral data.*
>
> *Weaknesses:*
>
> *I will start with what I think is the biggest difficulty I had with the paper. Much as I liked the model comparison analysis, I also don't quite know what to make of the view-tolerant model. As I understand the authors' description, the key feature of this model is that it does not get to compare the target and probe at the same yaw angle, but must instead pick a best match from candidates that are at different yaws. While it is interesting to see that this leads to a very different orientation profile, it also isn't obvious to me why such a comparison would be reflective of what the visual system is probably doing. I can see that the view-specific model is more or less assuming something like an exemplar representation of each face: You have the opportunity to compare a new image to a whole library of viewpoints, and presumably it isn't hard to start with some kind of first pass that identifies the best matching view first before trying to identify/match the individual in question. What I don't get about the view-tolerant model is that it seems almost like an anti-exemplar model: You specifically lack the best viewpoint in the library but have to make do with the other options. Again, this is sort of interesting and the very different behavior of the model is neat to discuss, but it doesn't seem easy to align with any theoretical perspective on face recognition. My thinking here is that it might be useful to consider an additional alternate model that doesn't specifically exclude the best-matching viewpoint, but perhaps condenses appearance across views into something like a prototype. I could even see an argument for something like the yaw-averages presented earlier in the manuscript as the basis for such a model, but this might be too much of a stretch. Overall, what I'd like to see is some kind of alternate model that incorporates the existence of the best-match viewpoint somehow, but without the explicit exemplar structure of the view-specific model.*

The design of the view-tolerant model aligned with the requirements of tolerant recognition and revealed the stimulus information enabling to abstract identity away from variations in face appearance. However, it did not involve the notion that such ability may depend on a prototype or summary representation of face identity built up through varied encounters (Burton, Jenkins, & Schweinberger, 2011; Burton et al., 2016; Jenkins et al., 2011; Menon, Kemp, & White, 2018; Mike Burton, 2013).

We agree with the Reviewer that the average of the different views of a face is a good proxy of its central tendency (i.e., stable identity properties; Figure 1). We thus followed their suggestion and included an additional model observer that compared specific views to full-spectrum view-averaged identities. The examination of the orientation tuning profile of this so-called view-average model observer confirmed the crucial contribution of horizontal identity cues to view-invariant recognition as the horizontal range best predicted the average summary of full-spectrum face appearances across views. This additional model observer is now presented in the Discussion and Supplementary files 2 and 3.

> *Besides this larger issue, I would also like to see some more details about the nature of the cross-correlation that is the basis for this model comparison. I mostly think I get what is happening, but I think the authors could expand more on the nature of their noise model to make more explicit what is happening before these cross-correlations are taken. I infer that there is a noise-addition step to get them off the ceiling, but I felt that I had to read between the lines a bit to determine this.*

In the Methods section, we now provide detailed information about the addition of noise to model observer cross-correlations: 'In a pilot phase, we measured the overall identification performance of each model. Initially, the view-selective model performed at ceiling, yielding a correlation of 1 since there was an exact target-probe match across all trials. To avoid ceiling effects and to keep model performance close to human levels (Supplementary File 2), we thus decreased the signal-to-noise ratio (SNR) of the target and probe images to .125 by combining each with distinct noise patterns (face RMS contrast: .01; noise RMS contrast: .08). Each trial (i.e. target-probe pairing) was iterated ten times with different random noise patterns.'

We also added a supplemental with the graphic illustration of the d' distributions of each model and human observers: 'Sensitivity d' of the view-tolerant model was much lower than view-selective model and human sensitivity (Supplementary File 2), even without noise. The view-tolerant model therefore processed fully visible stimuli (SNR of 1). This decreased sensitivity in the view-tolerant compared to the view-selective model is expected, as none of the probes exactly matched the target at the pixel level due to viewpoint differences. In contrast to humans who rely on internally stored representations to match identity across views, the model observer lacks such internal representations and entirely relies on (less efficient) pixelwise comparisons.'

> *Another thing that I think is worth considering and commenting on is the stimuli themselves and the extent to which this may limit the outcomes of their behavioral task. The use of the 3D laser-scanned faces has some obvious advantages, but also (I think) removes the possibility for pigmentation to contribute to recognition, removes the contribution of varying illumination and expression to appearance variability, and perhaps presents observers with more homogeneous faces than one typically has to worry about. I don't think these negate the current results, but I'd like the authors to expand on their discussion of these factors, particularly pigmentation. Naively, surface color and texture seem like they could offer diagnostic cues to identity that don't rely so critically on horizontal orientations, so removing these may mean that horizontal bias is particularly evident when face shape is the critical cue for recognition.*

Our stimuli were originally designed by Troje and Bulthoff (1996). These are 3D laser scans of white individuals aged between 20 and 40 years, posing with a neutral expression. Different views of the faces were shot under a fixed illumination. Ears and a small portion of the neck were visible while the hair region was removed. All face images had a normalized skin color and we further converted them to grayscales

While we agree that this stimulus set offers a restricted range of within- and between-identity variations compared to what is experienced in natural settings, we believe that the present findings generalize to more ecological viewing conditions. Indeed, past evidence showed that the recognition of face pictures shot under largely variable pose, age, expression, illumination, hair style is tuned to the horizontal range of the face stimulus (Dakin & Watt, 2009; Dumont, Roux-Sibilon, & Goffaux, 2024). In other words, our finding that view-tolerant identity recognition is mainly driven by horizontal face information would likely replicate with the use of a more ecological stimulus set.

Moreover, the skin color normalization and grayscale conversion, while limiting the range of face variability, did not eliminate the contribution of surface pigmentation in our study. It is thus unlikely that our findings exclusively reflect the orientation dependence of face shape processing. Pigmentation refers to all surface reflectance properties (Russell et al., 2006) and hue (color) is only one among others. The grayscaled 3D laser scanned faces used here contained natural variations in crucial surface cues such as skin albedo (i.e., how light or dark the surface appears) and texture (i.e., spatial variation in how light is reflected); they have actually been used to disentangle the role of shape and surface cues to identity recognition (e.g., Jiang et al., 2009; Russell et al., 2007; Russell et al., 2006; Troje & Bulthoff, 1996; Vuong et al., 2005). Moreover, a past study of ours demonstrated that the diagnosticity of the horizontal range of face information is not restricted to face shape cues; the specialized processing of face shape and surface both selectively rely on horizontal information (Dumont, Roux-Sibilon, & Goffaux, 2024).

For these reasons, the present findings are unlikely to be fully determined by shape processing, and we expect them to generalize to more ecological stimulus sets. We discuss these aspects in the revised manuscript.

> ***Reviewer #2 (Public review):***
>
> *This study investigates the visual information that is used for the recognition of faces. This is an important question in vision research and is critical for social interactions more generally. The authors ask whether our ability to recognise faces, across different viewpoints, varies as a function of the orientation information available in the image. Consistent with previous findings from this group and others, they find that horizontally filtered faces were recognised better than vertically filtered faces. Next, they probe the mechanism underlying this pattern of data by designing two model observers. The first was optimised for faces at a specific viewpoint (view-selective). The second was generalised across viewpoints (view-tolerant). In contrast to the human data, the view-specific model shows that the information that is useful for identity judgements varies according to viewpoint. For example, frontal face identities are again optimally discriminated with horizontal orientation information, but profiles are optimally discriminated with more vertical orientation information. These findings show human face recognition is biased toward horizontal orientation information, even though this may be suboptimal for the recognition of profile views of the face.*
>
> *One issue in the design of this study was the lowering of the signal-to-noise ratio in the view-selective observer. This decision was taken to avoid ceiling effects. However, it is not clear how this affects the similarity with the human observers.*

In the Methods section, we now provide detailed information about the addition of noise to model observer cross-correlations: ‘In a pilot phase, we measured the overall identification performance of each model. Initially, the view-selective model performed at ceiling, yielding a correlation of 1 since there was an exact target-probe match across all trials. To avoid ceiling effects and to keep model performance close to human levels (Supplementary File 2), we thus decreased the signal-to-noise ratio (SNR) of the target and probe images to .125 by combining each with distinct noise patterns (face RMS contrast: .01; noise RMS contrast: .08). Each trial (i.e. target-probe pairing) was iterated ten times with different random noise patterns.’

We also added a supplemental with the graphic illustration of the d’ distributions of each model and human observers.

> *Another issue is the decision to normalise image energy across orientations and viewpoints. I can see the logic in wanting to control for these effects, but this does reflect natural variation in image properties. So, again, I wonder what the results would look like without this step.*

All stimuli were matched for luminance and contrast. It is crucial to normalize image energy across orientations as natural image energy is disproportionately distributed across orientations (e.g., Hansen et al., 2003). Images of faces cropped from their background as used here contain most of their energy in the horizontal range (Goffaux & Greenwood, 2016; Keil, 2008, 2009). If not normalized after orientation filtering, such uneven distribution of energy would boost recognition performance in the horizontal range across views. Normalization was performed across our experimental conditions merely to avoid energy from explaining the influence of viewpoint on the orientation tuning profile.

We were not aware of any systematic natural variations of energy across face views. To address this, we measured face average energy (i.e., RMS contrast) in the original stimulus set, i.e., before the application of any image processing or manipulation. Background pixels were excluded from these image analyses. Across yaws, we found energy to range between .11 and .14 on a 0 to 1 grayscale. This is moderate compared to the range of energy variations we measured across identities (from .08 to .18). This suggests that variations in energy across viewpoints are moderate compared to variations related to identity. It is unclear whether these observations are specific to our stimulus set or whether they are generalizable to faces we encounter in everyday life. They, however, indicate that RMS contrast did not substantially vary across views in the present study and suggest that RMS normalization is unlikely to have affected the influence of viewpoint on recognition performance.

In the revised methods section, we explicitly motivate energy normalization: ‘Images of faces cropped from their background as used here contain most of their energy in the horizontal range (Goffaux, 2019; Goffaux & Greenwood, 2016; Keil, 2009). Across yaws, we found face energy to range between .11 and .14 on a 0 to 1 grayscale, which is moderate compared to the range of face energy variations we measured across identities (from .08 to .18). To prevent energy from explaining our results, in all images, the luminance and RMS contrast of the face pixels were fixed to 0.55 and 0.15, respectively, and background pixels were uniformly set to 0.55. The percentage of clipped pixel values (below 0 or above 1) per image did not exceed 3%.’.

> *Despite the bias toward horizontal orientations in human observers, there were some differences in the orientation preference at each viewpoint. For example, frontal faces were biased to horizontal (90 degrees), but other viewpoints had biases that were slightly off horizontal (e.g., right profile: 80 degrees, left profile: 100 degrees). This does seem to show that differences in statistical information at different viewpoints (more horizontal information for frontal and more vertical information for profile) do influence human perception. It would be good to reflect on this nuance in the data.*

Indeed, human performance data indicates that while identity recognition remains tuned to horizontal information, horizontal tuning peak shows some variation across viewpoints. We primarily focused on the first aspect because of its direct relevance to our research objective, but also discussed the second aspect: with yaw rotation, certain non-horizontal morphological features such as the jaw line or nose bridge, etc. may increasingly contribute to identity recognition, whereas at frontal or near frontal views, features are mostly horizontally-oriented (e.g., Keil, 2008, 2009). In the revised Discussion, we directly relate the modest fluctuations of peak location to yaw differences in face feature appearance.

> ***Recommendations for the authors:***

> ***Reviewing Editor Comments:***
>
> *Based on a discussion with the reviewers, we integrated the recommendations and reached a consensus on the eLife assessment. To move from a "solid" to a "compelling/convincing" strength-of-evidence rating, please address the reviewers' comments. Key points are to clarify and test the plausibility of the models (e.g., effects of different noise-addition steps, inclusion/exclusion of specific orientation channels in the view-dependent comparison, and alternative decision criteria), and to address or discuss the limitations of the stimulus set in capturing recognition under more naturalistic scenarios, for example, including texture cues.*
>
> ***Reviewer #1 (Recommendations for the authors):***
>
> *I generally found the paper to be very well-written, so I have only a few minor comments here.*
>
> *(1) I didn't really follow why the estimation of the Gaussian functions described in the text was preferred over a simpler ML framework. Do these approaches differ that much? I see references to prior studies in which these were applied, so I can certainly go check these out, but I could see value in adding just a bit of text to briefly make the case that this is important.*

Employing a simpler linear framework, i.e. a linear model predicting d' from the interaction between orientation and viewpoint, would result in an 8 (orientation) * 7 (viewpoint) design that is difficult to analyze. The interaction term would almost certainly reach significance but its interpretation would be limited. We would either have to rely on numerous local comparisons, which are not particularly informative for our research objectives (e.g., knowing whether d' differs significantly between two adjacent orientations at a given viewpoint is of little relevance), or to use a polynomial contrast approach (testing the linear, quadratic, … up to the 7th order trends), which would also be difficult to interpret. For such complex, approximately Gaussian-shaped data, the highest-order polynomial trend would likely provide the best fit, but without offering meaningful insight.

In contrast, a nonlinear approach appears more appropriate. The Gaussian model we used allows us to characterize the parameters of the tuning profile, namely, peak location, peak amplitude, standard deviation (or bandwidth) and base amplitude. These parameters are not merely statistical parameters. Rather, they are directly interpretable in cognitive/functional terms. The peak location corresponds to the orientation at which the Gaussian curve is centred, i.e. the preferred orientation band for identity recognition. The standard deviation represents the width of the curve, reflecting the strength or selectivity of the tuning. The base amplitude is the height of the Gaussian curve base, indicating the minimum level of sensitivity, typically found near vertical orientation. Finally, the peak amplitude refers to the height of the Gaussian curve relative to its baseline, that is, it captures the advantage of horizontal over vertical orientations.

Moreover, the use of a nonlinear, Gaussian model is motivated by past work that showed that the Gaussian function fits the evolution of recognition performance as a function of orientation (Dakin & Watt, 2009; Goffaux & Greenwood, 2016). Orientation selectivity at primary stages of visual processing has also been modelled using Gaussian (or Difference of Gaussians; Ringach, Hawken, & Shapley, 2003).

We revised the data analysis section to include a justification for our use of a Gaussian model: "Therefore, fitting the human sensitivity data could be fitted using a simple Gaussian model. seemed most appropriate as it allows characterizing the parameters of the tuning profile, namely, peak location, peak amplitude, standard deviation and base amplitude, which are directly interpretable in cognitive/functional terms. Moreover, the use of a nonlinear, Gaussian model is motivated by past work that showed that the Gaussian function fits the evolution of recognition performance as a function of orientation (Dakin & Watt, 2009; Goffaux & Greenwood, 2016). Simpler frameworks, i.e. a linear model predicting d' from the interaction between orientation and viewpoint, would result in an 8 (orientation) * 7 (viewpoint) design that is difficult to analyze and interpret."

> *(2) When reporting the luminance and contrast of your stimuli, please make clear what these units and measures are. This was a case where I had to take a second to assure myself that I knew what the values meant.*

We clarified that the luminance and contrast values reported in the manuscript are on a grey scale ranging from 0 to 1.

> *(3) In your Procedure section, I think describing the familiarization task right away would help the text flow more clearly. At present, you began talking about the old/new task, and I was immediately wondering how familiarization worked!*

The procedure section now starts with the description of the familiarization task.

> *(4) p. 3 - "Culminates" doesn't seem like the right word here.*

We agree and rephrased this way: 'The tolerance of face identity recognition is stronger for familiar than unfamiliar faces'.

> *(5) p. 5 - I think "with the multiple" shouldn't have "the".*

Indeed, we removed the "the".

> ***Reviewer #2 (Recommendations for the authors):***
>
> *I enjoyed reading the manuscript, but thought the Introduction was a bit long. I wasn't sure about the relevance of the section on temporal contiguity. I think this might have been more relevant if this had been a manipulation in the design. So, I wonder if this might be shortened or removed to focus on the key questions. On the other hand, I found the overview of the view-selective and view-tolerant to be a bit brief. There is plenty of detail here, but I found it difficult to break down what was done when I first read it. It might be good to provide an overview in the Discussion too.*

While past research on the contribution of temporal contiguity to face identity recognition brings interesting insights into the nature of the visual experience leading to view-tolerant performance, we agree with the Reviewer that this aspect is not directly at stake here. We reduced the review of this literature in the Introduction.

We clarified the description of the model observers as suggested by the reviewer and made sure to provide an overview of the model observers in the Discussion as well.

References.

Burton, A. M., Jenkins, R., & Schweinberger, S. R. (2011). Mental representations of familiar faces. Br J Psychol, 102(4), 943-958. *https://doi.org/10.1111/j.2044-8295.2011.02039.x*

Burton, A. M., Kramer, R. S., Ritchie, K. L., & Jenkins, R. (2016). Identity From Variation: Representations of Faces Derived From Multiple Instances. Cogn Sci, 40(1), 202-223. *https://doi.org/10.1111/cogs.12231*

Collin, C. A., Rainville, S., Watier, N., & Boutet, I. (2014). Configural and featural discriminations use the same spatial frequencies: a model observer versus human observer analysis. Perception, 43(6), 509-526. *https://doi.org/10.1068/p7531*

Dakin, S. C., & Watt, R. J. (2009). Biological "bar codes" in human faces. J Vis, 9(4), 2 1-10. *https://doi.org/10.1167/9.4.2*

Dumont, H., Roux-Sibilon, A., & Goffaux, V. (2024). Horizontal face information is the main gateway to the shape and surface cues to familiar face identity. PLOS ONE, 19(10), e0311225. *https://doi.org/10.1371/journal.pone.0311225*

Goffaux, V., & Greenwood, J. A. (2016). The orientation selectivity of face identification [Article de recherche] [peer-reviewed]. Scientific Reports, 6(34204), 34204. *https://doi.org/10.1038/srep34204*

Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Identification of band-pass filtered letters and faces by human and ideal observers. Vision Research, 39(21), 3537-3560. *http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10746125*

Hansen, B. C., Essock, E. A., Zheng, Y., & DeFord, J. K. (2003). Perceptual anisotropies in visual processing and their relation to natural image statistics. Network, 14(3), 501-526. *http://www.ncbi.nlm.nih.gov/pubmed/12938769*

Jenkins, R., White, D., Van Montfort, X., & Mike Burton, A. (2011). Variability in photos of the same face. Cognition, 121(3), 313-323. *https://doi.org/10.1016/j.cognition.2011.08.001*

Jiang, F., Dricot, L., Blanz, V., Goebel, R., & Rossion, B. (2009). Neural correlates of shape and surface reflectance information in individual faces. Neuroscience, 163(4), 1078-1091. *https://doi.org/10.1016/j.neuroscience.2009.07.062*

Keil, M. S. (2008). Does face image statistics predict a preferred spatial frequency for human face processing? Proc Biol Sci, 275(1647), 2095-2100. *https://doi.org/10.1098/rspb.2008.0486*

Keil, M. S. (2009). "I look in your eyes, honey": internal face features induce spatial frequency preference for human face processing. PLoS Comput Biol, 5(3), e1000329. *http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19325870*

Menon, N., Kemp, R. I., & White, D. (2018). More than a sum of parts: robust face recognition by integrating variation. R Soc Open Sci, 5(5), 172381. *https://doi.org/10.1098/rsos.172381*

Mike Burton, A. (2013). Why has research in face recognition progressed so slowly? The importance of variability. Quarterly journal of experimental psychology, 66(8), 1467-1485. *https://doi.org/10.1080/17470218.2013.800125*

Näsänen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. Vision Research, 39(23), 3824-3833. *http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10748918*

Oruc, I., Shafai, F., Murthy, S., Lages, P., & Ton, T. (2019). The adult face-diet: A naturalistic observation study. Vision Res, 157, 222-229. *https://doi.org/10.1016/j.visres.2018.01.001*

Ringach, D. L., Hawken, M. J., & Shapley, R. (2003). Dynamics of orientation tuning in macaque V1: the role of global and tuned suppression [Research Support, Non-U.S. Gov't

Research Support, U.S. Gov't, P.H.S.]. Journal of neurophysiology, 90(1), 342-352. *https://doi.org/10.1152/jn.01018.2002*

Russell, R., Biederman, I., Nederhouser, M., & Sinha, P. (2007). The utility of surface reflectance for the recognition of upright and inverted faces. Vision Res, 47(2), 157-165. *https://doi.org/10.1016/j.visres.2006.11.002*

Russell, R., Sinha, P., Biederman, I., & Nederhouser, M. (2006). Is pigmentation important for face recognition? Evidence from contrast negation. Perception, 35(6), 749-759. *https://doi.org/10.1068/p5490*

Troje, N. F., & Bulthoff, H. H. (1996). Face recognition under varying poses: the role of texture and shape. Vision Res, 36(12), 1761-1771. *https://doi.org/10.1016/0042-6989(95)00230-8*

Vuong, Q. C., Peissig, J. J., Harrison, M. C., & Tarr, M. J. (2005). The role of surface pigmentation for recognition revealed by contrast reversal in faces and Greebles. Vision Res, 45(10), 1213-1223. *https://doi.org/10.1016/j.visres.2004.11.015*

https://doi.org/10.7554/eLife.108495.3.sa0