University of Wollongong

Research Online

Faculty of Health and Behavioural Sciences -
Papers (Archive)

Faculty of Science, Medicine and Health

1-1-2011

Face viewpoint effects about three axes: the role of configural and featural processing

Simone K. Favelle

University of Wollongong, skeane@uow.edu.au

Stephen A. Palmisano

University of Wollongong, stephenp@uow.edu.au

Georgina Avery

University of Wollongong

Follow this and additional works at: <https://ro.uow.edu.au/hbspapers>

Part of the [Arts and Humanities Commons](#), [Life Sciences Commons](#), [Medicine and Health Sciences Commons](#), and the [Social and Behavioral Sciences Commons](#)

Recommended Citation

Favelle, Simone K.; Palmisano, Stephen A.; and Avery, Georgina: Face viewpoint effects about three axes: the role of configural and featural processing 2011, 761-784.
<https://ro.uow.edu.au/hbspapers/2094>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

Face viewpoint effects about three axes: the role of configural and featural processing

Abstract

We directly compared recognition for faces following 0° – 75° viewpoint rotation about the yaw, pitch, and roll axes. The aim was to determine the extent to which configural and featural information supported face recognition following rotations about each of these axes. Experiment 1 showed that performance on a sequential-matching task was viewpoint-dependent for all three types of rotation. The best face-recognition accuracy and shortest reaction time was found for roll rotations, then for yaw rotations, and finally the worst accuracy and slowest reaction time was found for pitch rotations. Directional differences in recognition were found for pitch rotations, but not for roll or yaw. Experiment 2 provided evidence that, in all three cases, viewpoint-dependent declines in recognition were primarily driven by the loss of configural information. However, it also appeared that significant featural information was lost following yaw and pitch (but not roll) rotations. Together, these findings show that unfamiliar-face recognition is viewpoint-dependent following rotation about each axis (and in each direction), and that performance is based on the availability of configural and, to a lesser extent, featural information.

Keywords

Face, viewpoint, effects, about, three, axes, role, configural, featural, processing

Disciplines

Arts and Humanities | Life Sciences | Medicine and Health Sciences | Social and Behavioral Sciences

Publication Details

Favelle, S. K., Palmisano, S. & Avery, G. (2011). Face viewpoint effects about three axes: the role of configural and featural processing. *Perception*, 40 (7), 761-784.

Face viewpoint effects about three axes: The role of configural and featural processing

Simone K Favelle, Stephen Palmisano, Georgina Avery

School of Psychology, University of Wollongong, Northfields Avenue, Wollongong 2522, NSW, Australia; e-mail: simone_favelle@uow.edu.au

Received 30 November 2010, in revised form 23 June 2011; published online 18 August 2011

Abstract. We directly compared recognition for faces following 0° – 75° viewpoint rotation about the yaw, pitch, and roll axes. The aim was to determine the extent to which configural and featural information supported face recognition following rotations about each of these axes. Experiment 1 showed that performance on a sequential-matching task was viewpoint-dependent for all three types of rotation. The best face-recognition accuracy and shortest reaction time was found for roll rotations, then for yaw rotations, and finally the worst accuracy and slowest reaction time was found for pitch rotations. Directional differences in recognition were found for pitch rotations, but not for roll or yaw. Experiment 2 provided evidence that, in all three cases, viewpoint-dependent declines in recognition were primarily driven by the loss of configural information. However, it also appeared that significant featural information was lost following yaw and pitch (but not roll) rotations. Together, these findings show that unfamiliar-face recognition is viewpoint-dependent following rotation about each axis (and in each direction), and that performance is based on the availability of configural and, to a lesser extent, featural information.

1 Introduction

Many biological and social aspects of human life rely on the correct recognition and identification of faces. Generally, humans are excellent at identifying others by face and are considered experts in face recognition (see Hancock et al 2000). Naturally, this expertise is applied more successfully to familiar faces than unfamiliar faces (ie faces that may have been seen previously but are not well-known—Bruce et al 2001; Hancock et al 2000; Liu and Chaudhuri 2000; Schwaninger et al 2002). The task of face recognition can be made more difficult by the number of different image changes that faces may undergo. For example, changes in size, lighting, or the observer's position/orientation relative to the face (viewpoint) can dramatically influence the ability to make a recognition judgment. In the case of unfamiliar faces, two different faces viewed front-on can actually appear more similar than the same face viewed from two different viewpoints. Thus, recognition of unfamiliar faces is typically viewpoint-dependent (Hill and Bruce 1996; Hill et al 1997; Jeffery et al 2006; Lee et al 2006; Liu and Chaudhuri 2002; Newell et al 1999; O'Toole et al 1998, 1999; Troje and Bühlhoff 1996).

Viewpoint effects on face recognition have predominantly been examined by using yaw rotations (about the vertical axis) and roll rotations (about the picture plane). Only recently have studies started to examine viewpoint effects for faces following pitch rotation (about the horizontal axis) (Favelle et al 2007; Liu et al 2005; Wallraven et al 2002). This research has shown that there are marked differences in recognition performance following rotations about the pitch and yaw axes. More specifically, and perhaps not surprisingly, face recognition is relatively more difficult in pitch rotations than rotations in yaw. Figure 1a shows images of faces following camera rotation about each of the three axes: yaw, pitch, and roll. Note that in the current study we will refer to our viewpoint changes as 'camera rotations'. This was because our yaw-rotated and pitch-rotated face images were produced by moving a camera (along either a horizontally or vertically oriented semicircular track) relative to the stationary heads of each of our human models (see figure 1b), as opposed to having these models move/rotate

Figure 1. [In colour online, see <http://dx.doi.org/10.1068/p6878>] (a) Examples of the stimuli generated following camera rotation about the three different axes: yaw, pitch, and roll, shown for both left and right, and up and down directions. A full-face view (0°) is included for each axis as a frame of reference. Examples of each direction on each axis refer to rotation to a viewpoint of 75° . (b) A schematic of the camera setup for image capture. Note that during pitch and yaw rotations, both the camera's position in space and its orientation changed with respect to the head. However, in the case of roll rotation, only camera orientation changed (its physical position relative to the head remained the same as the roll increased).

their heads relative to the stationary camera. This distinction is important because it means that, as in the real-world, lighting was always from above the face (even in the case of roll rotations). Also, it means that the terms ‘pitch-up’ and ‘pitch-down’ refer to the camera moving up or down the track and not where the face appears to be looking.⁽¹⁾

Favelle et al (2007) investigated face recognition following yaw and pitch viewpoint rotations using both a sequential matching and an old/new recognition task in separate experiments. A wide range of orientations was investigated, including 0° – 90° for yaw

⁽¹⁾ The labeling of the pitch-up and pitch-down conditions in Favelle et al (2007) mistakenly refers to head rotation and not viewpoint rotation, ie the opposite of the current paper. The pitch-rotated face stimuli are identical in these two papers.

(leftwards yaw), $0^\circ - 90^\circ$ for pitch-down, and $0^\circ - 75^\circ$ for pitch-up.⁽²⁾ Increments of 15° were used, providing a total of 18 different viewpoints for each face. While results obtained with these two tasks showed that face recognition was viewpoint-dependent following both pitch and yaw camera rotations, significant axis- and direction-based effects were observed. Specifically face recognition was found to tolerate yaw camera rotations better than pitch-down camera rotations, which in turn were better tolerated than pitch-up camera rotations.

While the viewpoint-dependent recognition of unfamiliar faces is well documented, less attention has been directed toward identifying the specific information required to support face recognition across rotations. What is the nature of the critical information contained in a view? Previous studies investigating rotations in depth have shown that viewpoint-dependent face recognition appears to be modulated by the availability of shape and texture cues (Troje and Bühlhoff 1996), the eyes (Stephan and Caine 2007), colour (Hill et al 1997), and distinctive marks (Valentin et al 1999). However, whether these findings extend to rotations about different axes, and whether there is critical information that is required to support face recognition across these types of viewpoint changes, remains largely unexplored.

One possible explanation of the differential effects of axis on viewpoint-dependence is that yaw- and pitch-rotated views of faces contain different amounts or types of the visual information needed to successfully recognise the face. View-interpolation models of object/face recognition (eg Bühlhoff and Edelman 1992; Poggio and Edelman 1990) propose that generalisation between views is based on the perceptual similarity of different views or exemplars rather than a transformational process. In principle, this similarity can be defined in a number of ways, ranging from low-level measures (such as the number of pixels corresponding to the visual image of the face) to higher-level characteristics (such as 3-D shape, shadows and lighting, facial features or configuration of features). For example, Favelle et al's (2007) finding of superior face recognition performance following pitch-down camera rotations could simply have been due to these conditions having a greater number of face pixels available than the equivalent pitch-up camera conditions (which can be seen in figure 2, particularly for the 75° pitch rotations). Similarly, there were also more face pixels available following yaw camera rotations than following pitch camera rotations, which could have better supported comparisons to full-face views (of course, recognition in the case of yaw rotation could also have been enhanced by the bilateral similarity of the face or use of virtual views—eg Troje and Bühlhoff 1998).

However, similarity can also be measured in terms of higher-level information specific to faces such as features and their configuration. Featural information includes parts of the face defined by relatively clear boundaries in the image, such as the nose, eyes, and mouth, or distinctive marks. Configural information is derived from a holistic representation of the face, and includes detailed information about the spatial relationships and distances between features, such as the horizontal distance between the eyes or the vertical distance between mouth and nose (Diamond and Carey 1986; Maurer et al 2002; Rakover 2002; Tanaka and Farah 1993). While both featural and configural information is necessary for accurate face recognition (Cabeza and Kato 2000; Collishaw and Hole 2000, 2002), configural processing has been demonstrated to be critical for face recognition, with empirical evidence coming from a number of different paradigms including the face-inversion effect (Robbins and McKone 2007; Valentine 1988; Yin 1969), the part-whole effect (Donnelly and Davidoff 1999; Gauthier and Tarr 1997; Tanaka and Farah 1993), and the composite effect (Gauthier and Tarr 1997; Robbins and McKone 2007; Young et al 1987).

⁽²⁾ Pitch-down camera rotations went only to 75° , as it was physically impossible to capture images of the faces further than this with a digital camera.

Figure 2. [In colour online.] An example of the full set of stimuli created for each axis; yaw (top row), pitch (centre row), and roll (bottom row) at each viewpoint for both left/down (indicated by a negative sign) and right/up (indicated by a positive sign) directions. Note: The full-face (0°) is shown three times only as a point of reference for each row, in the task only one 0° view is created for each face.

The broad aim of this paper is to determine whether there is critical information required to support face recognition after camera rotations about all three axes. Featural and configural information are excellent contenders given the widely acknowledged key role they play in face recognition. McKone (2008) directly tested configural processing with yaw rotated views (front, three-quarter, and profile) using composite and peripheral inversion tasks. Her results showed that, while inversion affected configural processing, depth (yaw) rotation did not. She concluded that configural processing appears to be independent of view and, importantly, unaffected by natural view frequency or loss of ‘best’ local parts (ie the eyes—Stephan and Caine 2007). This may well explain the Favelle et al (2007) finding that face recognition is better after yaw than after pitch rotations, but whether this is because configural processing is affected by pitch rotations is unknown.

Faces are 3-D objects experienced in a 3-D world. To our knowledge, the current study is the first to investigate face recognition following camera rotations about all three-dimensional axes (ie following yaw, pitch, or roll rotations). Favelle et al (2007) recently compared the effects of yaw (0° – 90°) and pitch (0° – 75°), but not roll, camera rotations. Previous face-recognition studies examining roll rotations have been primarily focused on the ‘face inversion effect’ (ie faces rotated 0° or 180° about roll) and have sought to provide evidence for a face-specific processor (see Valentine 1988) or coding strategy (Barton et al 2001; Collishaw and Hole 2000; Freire et al 2000; Leder and Bruce 2000; Murray et al 2000; Yin 1969). However, there have been a few studies of the effects of incremental changes in roll rotation (between 0° and 180°). For example, Valentine and Bruce (1988) had subjects perform a sequential-matching task using unfamiliar faces seen across five different angles of roll rotation (0° , 45° , 90° , 135° , and 180°). They found that task difficulty increased in a linear fashion with the degree of roll rotation, and concluded that there was no qualitative difference between the processing of upright and roll-rotated faces. In contrast, Murray et al (2000) found a discontinuity in the function relating orientation and bizarreness ratings of normal, Thatcherised, and featurally distorted faces (which were rotated in 30° increments in roll from 0° to 180°). They argued that spatial-relational information is disproportionately impaired by inversion, and that this discontinuity occurs at around 90° – 120° .

We used a range of roll camera rotations (up to 75° either side of upright) which, based on the above-mentioned studies, should be within the range of sensitivity of any spatial-relational encoding mechanisms in normal face perception. Not only is this range

more in keeping with our everyday exposure to faces (which may be related to Murray et al's findings), but it will also allow us to make direct comparisons to the recognition performance following equivalent pitch and yaw camera rotations.

In experiment 1 we compared face-recognition performance following yaw, pitch, and roll camera rotations. In principle, rotations about each of the three axes should place different demands on face-recognition mechanisms (Gauthier et al 2002). As such, we can gain greater insight into these mechanisms if we can compare them in a single analysis within a single set of faces. Natural view frequency effects would predict better performance following yaw and pitch rotations than roll rotations (eg Perrett et al 1998), since roll-rotated faces are less frequently encountered than either pitch or yaw-rotated faces (most noticeably with larger angular deviations from the full-face view). However, in terms of low-level similarity information, the number of pixels corresponding to the face remains constant across rotations in roll, but not for pitch or yaw. We might also expect there to be differences in recognition performance based on higher-level information, such as features, configuration, and top–bottom relations. The same configural and featural information should be available for all degrees of roll rotation, even though the recovery of this information may be hindered by transformation processes (eg mental rotation). By contrast, both yaw and pitch rotations will alter the visibility of facial features and consequently the recovery of configural information is likely to become more challenging.

Top-bottom relations refer to the vertical spatial relationships between features, eg distance between nose and mouth. While the visibility of all of the facial features is maintained in roll rotations, top–bottom relations between (and sometimes within) those features become perturbed. Top–bottom relations remain intact in yaw rotations, but are compressed (proportionately) by pitch rotations. Further, experience might also play a role in any pitch-up versus pitch-down directional differences, since we all have considerable practice at viewing faces from below (ie pitch-down) during childhood.

As already discussed, not only are there differences in the amount/type of available featural information following rotation about the three axes, but the use of configural information is likely to be quite different in the roll axis compared to the pitch and yaw axes. In experiment 2 we investigated the role of higher-level configural and featural information in these viewpoint and axis effects by directly manipulating the availability of these two kinds of information. We used scrambling and blurring techniques on the face images to further degrade the available configural and featural information following camera rotations about either the yaw, pitch, or roll axes.

2 Experiment 1

To our knowledge, face recognition has not been directly compared following yaw, pitch, and roll camera rotations in any previous study. We expected face recognition to be viewpoint-dependent for camera rotations about each of these axes (Collishaw and Hole 2002; Favelle et al 2007; Hill and Bruce 1996; Hill et al 1997; Liu and Chaudhuri 2002; Liu et al 2005; O'Toole et al 1999; Schwaninger and Mast 2005; Searcy and Bartlett 1996; Troje and Bülthoff 1996; Valentine and Bruce 1988; Wallraven et al 2002). However, on the basis of the supposition that more 'face' information is available following a roll rotation (eg Murray et al 2000) and work with macaque monkeys showing that face selective neurons are more sensitive to rotations in depth than roll (Perrett et al 1985), we expect recognition for roll-rotated faces to be superior to that for yaw-rotated and pitch-rotated faces. As in earlier studies, it is expected that faces rotated about yaw will be more easily recognised than faces rotated about pitch (Favelle et al 2007; Wallraven et al 2002). While there is evidence for the lateralisation of face-recognition processes (see Rhodes 1985), the set of yaw-rotated images used in the current study was created by mirroring the original (rightwards) yaw-rotated images from

Favelle et al (2007—see section 2.1.3 below). Thus, we do not expect there to be any left/right recognition differences following either yaw or roll rotations. However, recognition performance is expected to be better for pitch-down than for pitch-up camera rotations.

In the current study we also examined the effect of the viewpoint of the encoding face on recognition. Previous research suggests that the view depicted in the encoding phase is critical to the outcome of the task (Logie et al 1987; Schyns and Bülthoff 1994; Troje and Bülthoff 1996). Specifically, yaw-rotated viewpoints elicited superior performance to full-face (ie 0°) views when presented at encoding. Schyns and Bülthoff (1994) and Troje and Bülthoff (1996) have argued that a full-face view provides redundant information that is not easily generalised to alternative viewpoints, particularly in regard to accessing 3-D information. Many studies, including Favelle et al (2007), have demonstrated viewpoint-dependent performance using full-face views at encoding (Liu et al 1999; Newell et al 1999; Valentin et al 1999; Schwaninger and Mast 2005). Thus, in the current study, using the same ecologically lit photographic images as in Favelle et al (2007), we investigated whether viewpoint-dependent effects rely on whether or not the encoding face is shown at 0° .

2.1 Method

2.1.1 Design. The design was a mixed factorial design with three within-subjects factors: axis (yaw, pitch, and roll), direction (positive and negative),⁽³⁾ and viewpoint (0° , 15° , 30° , 45° , 60° , and 75°). There was one between-subjects factor: stimulus presentation order in the sequential matching task (ie stimulus presentation at encoding and test). One group was shown a face at 0° at encoding followed by a face presented at any of the 31 possible viewpoints at test (face 1 at 0°); the second group was shown a face presented at any of the 31 possible viewpoints at encoding followed by a face at 0° at test (face 2 at 0°).

2.1.2 Participants. Seventy undergraduate Psychology students attending the University of Wollongong participated in the experiment in return for course credit. There was an equal number of participants in each group (ie thirty-five per group). The participants were made up of seventeen males and fifty-three females, and each group had similar numbers of males and females (group 1: nine males and twenty-six females; group 2: eight males and twenty-seven females). The average age of participants was 20.5 years (age range: 18–43 years). All participants were tested individually, had normal or corrected-to-normal vision, and none was familiar with the faces used as stimuli.

2.1.3 Stimuli. Stimuli were taken from a database of high-quality digital face images (see Favelle et al 2007). The stimuli consisted of 9 Caucasian female faces. Faces portrayed a neutral expression, and any distinctive features (such as moles or blemishes) as well as hair were removed. Faces were illuminated by four directional light sources located 1 m in front of the model (just above and to the left, just above and to the right, just below and to the left, and just below and to the right of the face)⁽⁴⁾ and an ambient light source located directly above. Lighting was held constant across all viewpoints. In addition to a full-face (0°) view, each face was presented from 10 different viewpoints rotated 15° , 30° , 45° , 60° , and 75° either side of 0° about one of three different axes (see figure 2). In the case of yaw camera rotations, faces were presented at rotations to the left or to the right of 0° . Since the face images were of real people and captured with a digital camera mounted on a semi-circular frame (see Favelle et al 2007), to minimise noise (eg slight variations in camera angle or model expression)

⁽³⁾ Depending on the axis, positive/negative will refer to either left/right (yaw and roll) or up/down (pitch).

⁽⁴⁾ The two lights located just above the face were oriented horizontally (ie at 90° to gravity), whereas the lights located just below the face were oriented 45° below the horizontal.

the leftwards yaw stimuli were created as mirror images of the original rightwards yaw stimuli. In the case of pitch camera rotations, faces were shown either above 0° (pitch-up) or below 0° (pitch-down). The images used for roll axis camera rotations were created by rotating the full-face (0°) images left and right. In total there were 31 different viewpoints generated for each face: full-face 0° , left and right yaw viewpoints of 15° – 75° , pitch-up and pitch-down viewpoints of 15° – 75° , and left and right roll viewpoints of 15° – 75° . These 31 different image manipulations were created for each of the 9 face models. Overall there were 279 images created for the experimental task.

All images were viewed in the centre of the computer screen against a white background. The on-screen height of 0° faces was approximately 16 cm, which with an on-screen width of 10 cm, meant that they subtended a visual area of $14.7^\circ \times 19.2^\circ$. Following yaw rotations, the height of the face image remained constant, but the face width increased as the viewpoint was rotated further away from 0° . For pitch camera rotations the face width remained constant, but the face height decreased as the viewpoint was rotated further away from 0° (for both pitch-up and pitch-down conditions). The smallest image for the camera pitch-up condition was at a viewpoint of 75° , which was 12 cm high and subtended a visual area of $11^\circ \times 9.2^\circ$. The rectangular patterned mask used in the experiment subtended a visual area of $18^\circ \times 22^\circ$ and was composed of various elements taken from the stimuli used in the task.

2.1.4 Apparatus. Full-colour images were presented to participants on a 48 cm flat-screen monitor with a resolution of 1024×768 pixels. Trials were run on a Macintosh G4 computer and RSVP experimental software (version 4.0.5; www.tarrlab.org) guided the trial sequence. Responses were made by key presses on a keyboard placed in front of the participant.

2.1.5 Procedure. The task was a masked sequential-matching task. Participants were randomly assigned to one of two groups. Those in group 1 were presented with a face viewed at 0° as face 1 in the task (face 2 was a face presented at any of the 31 possible viewpoints). Those in group 2 were presented with any of the 31 possible viewpoints as face 1 (face 2 was a face viewed at 0°). Participants were first verbally instructed how to complete the task, with emphasis placed on both speed and accuracy in responding. Written instructions on how to complete the task were also provided on the computer screen. After reading the instructions, participants completed 14 practice trials to familiarise themselves with the task. Stimuli used in the practice trials were different (ie generated from a different set of face models) from the stimuli used in the task. After the practice trials, participants could ask questions about the procedure, should they have any, before continuing with the experiment.

The task consisted of 558 experimental trials (31 viewpoints $\times$ 9 identities each for same and different trials).⁽⁵⁾ In half of the trials the two faces presented were the same, regardless of viewpoint (same trials). In the other half of trials the two faces were different; the different face was randomly selected from the remaining 8 faces (different trials). Trial type was presented in random order. Participants were given 8 self-timed rest periods spaced equally throughout the experiment. The experiment lasted approximately 45 min.

Each trial began with a fixation cross displayed for 500 ms. This was followed by the presentation of face 1 for 250 ms. Then the mask was presented for 500 ms. Face 2 was then presented for 250 ms, followed by a second presentation of the mask for 500 ms.

⁽⁵⁾ While participants were not familiar with any of the faces before testing, it is possible that with only 9 different identities, some learning may have occurred over the course of the experiment. This possible learning was not of particular concern, since becoming familiar with these faces would not have advantaged one condition over another (the axis and degree of face rotation were fully randomised). Indeed, Favelle et al (2007) demonstrated significant differences in performance following yaw and pitch rotations using the same set of faces in a learning task.

Following the second mask the screen remained blank for 2 s or until a response was made by the participant. If a response was not made within this time, the trial ended (ie ‘timed-out’). The interval between trials was 1 s. Participants were required to respond by pressing clearly labelled “same” and “different” keys on a keyboard depending on whether they judged face 1 and face 2 to be the same face or two different faces.

2.2 Results

Participants’ responses were converted into hit (H) and false alarm (FA) rates, where a hit was a correct “same” response to a face during the test phase, and a false alarm was an incorrect “same” response to a “different” face during the test phase. These H and FA rates were converted into *z*-scores and then used to calculate *d'* (see Macmillan and Creelman 2005). The reaction-time data were based only on participants’ correct responses to same trials. Trials that timed-out were not included in the analysis (timed-out trials accounted for 0.9% of total trials). Mean sensitivity (*d'*) and reaction-time

Figure 3. Mean recognition sensitivities (a) and mean reaction times (b) as a function of the presentation order and the type of camera rotation. Encoding face at 0° is represented by ‘1’ and test face at 0° is represented by ‘2’. Rotation could vary in terms of the inclination (0°–75°), axis (yaw, pitch, or roll), and direction (+/– of the rotation). Left/down viewpoint directions are indicated by a negative sign and right/up viewpoint directions are indicated by a positive sign. Error bars represent ±1 SEM.

scores were analysed in separate 2 (presentation order: encoding face shown at 0°; test face shown at 0°) × 3 (axis: yaw, pitch, roll) × 2 (direction: positive, negative) × 5 (viewpoint: 15°, 30°, 45°, 60°, 75°) split-plot ANOVAs where the between-subjects factor was encoding versus test face shown at 0°. The α level was 0.05. All a-posteriori comparisons were Bonferroni-adjusted pairwise comparisons.

2.2.1 Comparison of the encoding face at 0° and the test face at 0°. The ANOVA conducted on the sensitivity data revealed no effect of presentation order ($F < 1$, $p = 0.79$ —see figure 3a). Further, presentation order did not interact with any of the other factors (all F s < 1.5 , $p > 0.2$) except viewpoint ($F_{4,272} = 2.58$, $p = 0.04$, $MSE = 4.16$, $\eta_p^2 = 0.04$). A-posteriori comparisons showed no significant difference between the two levels of presentation order at any of the viewpoints (all p s > 0.08). Thus, the presentation order by viewpoint interaction appears to be based on the smaller overall viewpoint effect found for the encoding face at 0° group compared to the test face at 0° group.

The analysis of reaction-time data revealed no effect of the presentation order ($F = 1.05$, $p = 0.31$ —see figure 3b). There was a significant interaction between presentation order and axis ($F_{2,136} = 6.5$, $p < 0.01$, $MSE = 100\,808$, $\eta_p^2 = 0.11$) based on slower reaction times for pitch rotations in the encoding face at 0° group, and between presentation order, direction, and viewpoint ($F_{2,136} = 2.56$, $p = 0.04$, $MSE = 19\,255$, $\eta_p^2 = 0.14$). Presentation order did not interact with any of the other factors (all F s < 1 , $p > 0.1$).

Overall, these results demonstrated a minimal effect of whether the face shown at encoding was a 0° view or not. Because presentation order interacted with some (although, admittedly, very few) factors, we analysed the data for each level of this between-subjects factor separately. That is, mean sensitivity (d') and reaction-time scores were analysed in separate 3 (axis: yaw, pitch, roll) × 2 (direction: positive, negative) × 5 (viewpoint: 15°, 30°, 45°, 60°, 75°) repeated-measures ANOVAs for each level of the between-subjects factor.

2.2.2 Encoding face shown at 0°. As can be seen in figure 4, sensitivity to matching faces when the encoding face is shown at 0° varies with the axis of rotation and is viewpoint-dependent. The ANOVA conducted on sensitivity data revealed a main effect of axis ($F_{2,68} = 121.39$, $p < 0.001$, $MSE = 2.08$, $\eta_p^2 = 0.78$) and viewpoint ($F_{4,136} = 9.70$, $p < 0.001$, $MSE = 1.55$, $\eta_p^2 = 0.22$). Simple contrasts show that: (i) recognition following

Figure 4. Mean recognition sensitivity for yaw, pitch, and roll axes for both left/right and up/down directions for the encoding face at 0° condition. Error bars represent ± 1 SEM.

roll camera rotation was more accurate than following yaw and pitch camera rotations ($p < 0.001$), and (ii) recognition following yaw camera rotation was more accurate than following pitch camera rotation ($p < 0.001$). There was no main effect of direction ($F < 1$, $p = 0.68$); however, there was an interaction between axis and direction ($F_{2,68} = 5.76$, $p = 0.001$, $MSE = 1.64$, $\eta_p^2 = 0.15$). This interaction appeared to be caused by performance differences in response to pitch-up and pitch-down camera rotations. A-posteriori comparisons demonstrated that directional differences were only significant after pitch camera rotations, with pitch-down rotations demonstrating better recognition accuracy than pitch-up rotations ($p < 0.05$). No other interactions were significant (all $F_s < 1.5$, $p > 0.17$).

The patterns found in the analysis of the reaction-time data reflect those found for the sensitivity data (see figure 5). The ANOVA revealed a main effect of axis ($F_{2,68} = 68.02$, $p < 0.001$, $MSE = 1970996$, $\eta_p^2 = 0.67$) and viewpoint ($F_{4,136} = 8.08$, $p < 0.001$, $MSE = 91955$, $\eta_p^2 = 0.19$). Simple contrasts show that face recognition reaction-time following roll camera rotation was significantly faster than following yaw camera and pitch camera rotation and that recognition reaction time following yaw camera rotation was faster than following pitch camera rotation (both $p_s < 0.001$). There was no main effect of direction of rotation ($F_{2,68} = 2.9$, $p = 0.1$). However, there was an interaction between axis and direction ($F_{2,68} = 7.33$, $p = 0.002$, $MSE = 53668$, $\eta_p^2 = 0.18$). A-posteriori comparisons showed that recognition following pitch-down rotation was significantly faster than at pitch-up rotation ($p < 0.05$), but no directional differences were found for the yaw or roll axes ($p = 0.10$ and $p = 0.36$, respectively). No other interactions were significant (all $F_s < 1.9$, $p > 0.12$).

Figure 5. Mean reaction time for yaw, pitch, and roll axes for both left/right and up/down directions for the encoding face at 0° condition. Error bars represent ± 1 SEM.

2.2.3 Test face shown at 0° . The sensitivity data for the condition in which the test face was always shown at 0° is shown in figure 6. The ANOVA conducted on these data revealed significant main effects of axis ($F_{2,68} = 130.69$, $p < 0.001$, $MSE = 2.13$, $\eta_p^2 = 0.79$) and viewpoint ($F_{4,136} = 24.38$, $p < 0.001$, $MSE = 1.68$, $\eta_p^2 = 0.42$). Simple contrasts show that: (i) face-recognition accuracy following roll camera rotation was significantly greater than for yaw and pitch camera rotations; an (ii) face recognition following yaw camera rotation was more accurate than that following pitch camera rotation (both $p_s < 0.001$). There was no main effect of direction ($F_{2,68} = 2.3$, $p = 0.14$), nor interaction between direction and viewpoint ($F = 1.63$, $p = 0.17$); however,

Figure 6. Mean recognition sensitivity for yaw, pitch, and roll axes for both left/right and up/down directions for the test face at 0° condition. Error bars represent one ± 1 SEM.

there was a significant interaction between axis and direction ($F_{2,68} = 4.17, p < 0.05, \text{MSE} = 0.17, \eta_p^2 = 0.11$). The interaction between axis and direction again appeared to be driven by differences in responding to pitch-up and pitch-down rotations. A-posteriori comparisons showed that face-recognition accuracy was superior after pitch-down, as opposed to pitch-up, camera rotation ($p < 0.01$), but no directional differences were found for the yaw or roll axes (both $ps > 0.17$). There was also an interaction between axis and viewpoint ($F_{8,272} = 4.07, p < 0.001, \text{MSE} = 1.59, \eta_p^2 = 0.11$). While a-posteriori comparisons showed viewpoint-dependent performance for yaw, pitch, and roll, the interaction between axis and viewpoint appeared to be based on the relatively smaller viewpoint effect found for the roll axis compared to pitch and yaw.

A significant three-way interaction was found between axis, direction, and viewpoint ($F_{8,272} = 2.24, p < 0.05, \text{MSE} = 1.30, \eta_p^2 = 0.06$). Referring to figure 6, it appears that this interaction was caused by the finding that, while there were no directional differences for either the roll or yaw axes, the difference between pitch-up and pitch-down rotations did not manifest itself until the 45° viewpoint.

Analysis of the reaction-time data again revealed a similar pattern of results to that of the sensitivity data (see figure 7). The ANOVA revealed significant main effects of axis ($F_{2,68} = 34.25, p < 0.001, \text{MSE} = 518\,007, \eta_p^2 = 0.50$) and viewpoint ($F_{4,136} = 10.02, p < 0.001, \text{MSE} = 100\,469, \eta_p^2 = 0.23$), but not direction ($F = 3.0, p = 0.10$). Simple contrasts showed that face recognition was significantly faster following roll-camera rotation than for equivalent yaw and pitch camera rotations, and that face recognition was faster following yaw camera rotation than for pitch camera rotation (both $ps < 0.001$). Significant interactions were found between axis and direction ($F_{2,68} = 4.50, p < 0.05, \text{MSE} = 24\,406, \eta_p^2 = 0.12$), between axis and viewpoint ($F_{8,272} = 4.20, p = 0.001, \text{MSE} = 54\,281, \eta_p^2 = 0.11$), and between direction and viewpoint ($F_{4,136} = 3.10, p = 0.025, \text{MSE} = 27\,229, \eta_p^2 = 0.08$). The three-way interaction was not significant ($F = 1.9, p = 0.07$). The interaction between axis and direction appeared to be driven by the difference between pitch-up and pitch-down rotations. A-posteriori comparisons showed that face reaction time following pitch-down camera rotation was greater than following pitch-down camera rotation ($p < 0.01$), but no directional differences were found for the yaw or roll axes (both $ps > 0.60$). A-posteriori comparisons of the interaction between axis and viewpoint show viewpoint-dependent performance for the pitch and roll axes but only a marginal effect of viewpoint for yaw.

Figure 7. Mean reaction times for yaw, pitch, and roll axes for both left/right and up/down directions for the test face at 0° condition. Error bars represent ± 1 SEM.

The interaction between direction and viewpoint was most likely due to the relatively large increase in reaction time at the 60° and 75° viewpoints in the pitch-up direction (refer to figure 6). A-posteriori comparisons showed no directional differences for 15°, 30°, or 45° (all $ps > 0.05$), but for viewpoints 60° and 75° reaction time was significantly slower for the right/down direction (both $ps < 0.05$).

2.3 Discussion

The results of experiment 1 demonstrated a viewpoint-dependent decline in face recognition following camera rotations about each of the yaw, pitch, and roll axes. As predicted, the level of viewpoint-dependence varied depending on the axis of camera rotation. Overall, the best face recognition was found after camera rotations about the roll axis, followed by camera rotations about the yaw axis, with the worst performance found for camera rotations about the pitch axis. The results support previous findings that face recognition is more accurate and faster for yaw than for pitch rotations (Favelle et al 2007; Liu et al 2005; Wallraven et al 2002) and that recognition after pitch-down camera rotations is faster and more accurate than after pitch-up camera rotations (Favelle et al 2007; Liu et al 2005, experiment 3). As expected, no directional differences were found between either left and right yaw rotations or left and right roll rotations. Taken together, the current findings are consistent with the proposal that face recognition is determined by the axis of rotation and the amount of specific face information that is available or retained when the observer changes his/her viewpoint.

We found viewpoint-dependent performance irrespective of whether the full face was shown in the encoding or test phase. The choice of encoding view was shown to have a modest effect on face recognition—there were smaller overall viewpoint costs when the encoding face was shown at 0°. This result appears contrary to the findings of Troje and Bühlhoff (1996), but in line with others (eg Liu et al 1999; Valentin et al 1999). Our finding suggests that recognition performance is not based solely on the view shown at encoding, but rather it is the relationship between encoding and test views and the information available for generalisation across views (ie extracted and utilised from multiple views) that are of most importance. There was an overall shift towards slower reaction times when the encoding face was shown at 0°, particularly for pitch rotations. Since there was no similar effect on sensitivity, this could

represent a general increase in the time taken to utilise information from a full-face view so as to recognise a face from a different viewpoint (rather than a difference in the quality of the information available from a full-face view). In the future, this possibility could be investigated with an experimental design similar to that of Troje and Bühlhoff (1996), where different views are compared at both encoding and test (as opposed to keeping either the encoding or test face view constant as in the current study).

The superior face-recognition performance following roll camera rotation suggests that it is not just angular difference between views, but the greater loss in terms of the available face information that is responsible for the reduced performance following yaw and pitch camera rotations. It has been previously shown that configural and featural information are both preserved after roll face rotations of up to 75° (Lewis 2001; Lewis and Glenister 2003; Murray et al 2000). This suggests that any decrement in face recognition following yaw and pitch camera rotations (relative to roll) must be at least partially due to the loss of one or both types of face information. In experiment 2 we directly investigated the contribution of configural and featural face information to these axis-of-rotation-based viewpoint effects. Using face-image scrambling and blurring, we directly manipulated the availability of both of these types of information following camera rotations about all three axes.

3 Experiment 2

Recent research by McKone (2008) suggests that, while feature-based face processing is affected by changes in view, configural-based face processing is not. McKone has argued that: (i) view effects are driven by featural processing, and (ii) one of the roles of configural processing is to allow reliable identification in situations where key local information is not available. But while McKone's (2008) findings suggest that configural processing is not affected by yaw rotations, it is unclear whether it is affected by pitch rotations or acute roll rotations ($0^\circ - 75^\circ$).

In experiment 2, we examined the specific roles that configural and featural information play in face recognition following yaw, pitch, and roll camera rotations. We used a technique similar to that reported by Schwaninger et al (2002), in which face images were either scrambled or blurred (also see Hayward et al 2008; Lobmaier and Mast 2007). While this is not an absolute manipulation of configural and featural information, it does place a strong, selective emphasis on the two different types of information. In the scrambled condition, components of the face were cut out and re-arranged. Since this primarily disrupts configural (and holistic) information, in this case faces must be recognised primarily on the basis of their featural information. In the blurred condition, a low-pass filter was applied to intact faces. While this removes much of the fine detail required for featural processing, it leaves first-order and second-order configural information for face recognition. We applied this technique to 0° , 45° , and 75° rotated face images about the yaw, pitch, and roll axes. The images were used in a sequential-matching task similar to that used in experiment 1.

Schwaninger et al (2002) used an old/new face recognition task and found better performance for blurred than scrambled faces, regardless of whether the faces were familiar or unfamiliar. Hayward et al (2008) found a similar pattern of results in own-race and other-race face recognition. In contrast, Lobmaier and Mast (2007) used a sequential-matching task with longer stimulus presentation times (1000 ms compared to the 250 ms used in the current study) and found that accuracy for matching unfamiliar scrambled faces to intact faces was greater than for matching unfamiliar blurred faces to intact faces. When faces were familiar, there was no difference between scrambled and blurred conditions in terms of matching accuracy. Thus, while the Schwaninger et al (2002) and Hayward et al (2008) studies suggest that configural information is predominant in both familiar-face and unfamiliar-face recognition, there is some evidence

that for unfamiliar faces featural processing may play a relatively larger role than configural processing. Since we are investigating unfamiliar-face recognition in a matching task, we might expect a pattern of results similar to Lobmaier and Mast's (2007). However, since it is widely held that configural information is key to face recognition, we might expect that performance in matching blurred to intact faces will be better than in matching scrambled to intact faces. Whatever the pattern of results, how it holds up over changes in viewpoint is unclear. McKone (2008) found evidence that configural (but not featural) processing is insensitive to changes in yaw viewpoint. Thus, in the current experiment we might expect that, for yaw camera rotations, blurred faces will show little to no effect of view with scrambled faces showing viewpoint-dependence. If configural processing is insensitive to view irrespective of the axis of rotation, then we would expect this pattern to be replicated for the pitch and roll axes.

However, if pitch camera rotations produced the greatest viewpoint-dependent declines in face recognition in experiment 1 because they caused the greatest disruption to configural processing, then the blurred-face stimuli should show a sensitivity to view (as this manipulation leaves configural information relatively intact). Alternatively, if the differential effects of yaw, pitch, and roll camera rotations on face recognition in experiment 1 were caused by different rates of loss of featural information (McKone 2008), then we would expect these view effects to persist with scrambled faces, but not with blurred faces.

3.1 Method

3.1.1 Design. The design was a mixed factorial design with three within-subjects factors: axis (yaw, pitch, and roll), direction (positive and negative)⁽⁶⁾ and viewpoint (0°, 45°, and 75°). There was one between-subjects factor: image distortion (scrambled, blurred, or scrambled-blurred).

3.1.2 Participants. Sixty-six undergraduate Psychology students attending the University of Wollongong participated in the experiment in return for course credit. There was an equal number of participants in each group (ie twenty-two per group). The participants were made up of twenty males and forty-four females (two participants did not enter sex information), and each group had similar numbers of males and females (blur group: five males, fifteen females, and two no response; scramble group: six males and sixteen females; scramble and blur group: nine males and thirteen females). The average age of participants was 21.9 years (age range: 17–52 years, one participant did not enter age information). All participants had normal or corrected-to-normal vision and none was familiar with the faces used as stimuli.

3.1.3 Stimuli. Stimuli were created using a subset of the images from experiment 1 (only the 0°, 45°, and 75° views in both directions and about all three axes). The neck and under-chin areas were removed from all images where present (since this is not a feature of the face). For each of the 9 faces, there were 13 views. All faces were processed with Adobe Photoshop. Scrambled faces were created by cutting each of the 13 intact views into components using the polygonal lasso tool with a 2 pixel feather. The number of components differed depending on the amount of the face visible due to view. 0°, 45°, and 75° views in either direction about the roll axis were each cut into 7 components (mouth and chin, nose, eyes, cheeks, and forehead). 45° views in either direction about the yaw axis were cut into 6 components (mouth and chin, nose, eyes, cheek, and forehead) and 75° views in either direction about the yaw axis were cut into 5 components (mouth and chin, nose, eye, cheek, and forehead). 45° views in either direction about the pitch axis were cut into 7 components (mouth and chin,

⁽⁶⁾ Depending on the axis of camera rotation, positive/negative will refer to either left/right (yaw and roll) or up/down (pitch).

nose, eyes, cheeks, and forehead), 75° views about the pitch-up camera axis were cut into 6 components (mouth and chin, nose, eyes, and cheeks), and 75° views about the pitch-down camera axis were cut into 4 components (nose, eyes, and forehead). Three different scrambling versions for each view were used and appeared randomly. Each version was arranged such that natural relations to neighbouring components were disrupted. The components were distributed such that the image size of the scrambled faces were only slightly larger than that of the intact faces (see figure 8a).

Figure 8. [In colour online.] Sample stimuli for the (a) scrambled, (b) blurred, and (c) scrambled-blurred conditions for both left/down (indicated by a negative sign) and right/up (indicated by a positive sign) directions.

The blurred faces were created by the same technique as that used by Hayward et al (2008), by taking each of the 13 views and repeatedly applying a Gaussian filter (see figure 8b). The filter had a radius of 3 pixels and a standard deviation of 3 pixels, and was applied four times. With an average stimulus width of 10 cm and a viewing distance of 40 cm, the blurring cut off spatial frequencies above 21 cycles per face width or 1.4 cycles per visual degree. Scrambled-blurred faces were made in the same way as the blurred faces, but were based on the scrambled faces rather than the intact faces (see figure 8c). The spatial arrangement of each scrambled-blurred face was matched to its scrambled face complement. In total, there were 360 images created for the experiment (9 intact, full-spectrum faces; and 117 each of scrambled, blurred, and scrambled-blurred faces).

All images were viewed in the centre of the computer screen against a white background. The on-screen height of 0° faces was approximately 16 cm with a width of 10 cm, which produced a visual area of $14.7^\circ \times 19.2^\circ$. The rectangular patterned mask used in the experiment subtended a visual area of $18^\circ \times 22^\circ$ and was composed of various elements taken from the stimuli used in the task.

3.1.4 Apparatus. Full-colour images were presented to participants on a 48 cm flat-screen monitor with a resolution of 1024×768 pixels. Trials were run on a Macintosh G4 computer and RSVP experimental software (version 4.0.5; www.tarlab.org) guided the trial sequence. Responses were made by key presses on a keyboard placed in front of the participant.

3.1.5 Procedure. Participants completed the experiment in a dimly lit room in undisturbed conditions. Participants were randomly assigned to one of three groups. Each group was tested in one experimental condition: scrambled, blurred, or scrambled-blurred. The task was a sequential-matching task in which face 1 was presented first followed by a mask, and then face 2 was presented followed by the mask again. Face 1 was always presented in a full-face view (0°), intact, and in the full spatial spectrum. Face 2 was presented scrambled, blurred, or scrambled-blurred (depending on group membership) and at any one of the 13 possible viewpoints. Participants were first verbally instructed how to complete the task, with emphasis placed on both speed and accuracy in responding. Written instructions how to complete the task were also provided on the computer screen. After reading the instructions, participants completed 14 practice trials to familiarise themselves with the task. Stimuli used in the practice trials were different from the stimuli used in the task. After the practice trials, participants were given a chance to ask any questions about the procedure before continuing on with the experiment.

The task consisted of 234 experimental trials ($13 \text{ viewpoints} \times 9 \text{ identities}$, each for “same” and “different” trials). In half of the trials the two faces presented were the same, regardless of viewpoint or image distortion (same trials). In the other half of trials the two faces were different; the different face was randomly selected from the remaining 8 faces (different trials). Trial type was presented in random order. Participants were given 5 self-timed rest periods spaced equally throughout the experiment. The experiment lasted approximately 25 min. The trial sequence was identical to that in experiment 1.

3.2 Results

Data analysis was conducted on d' and reaction time (these reaction-time data were based only on participants' correct responses to same trials). Trials that timed-out were not included in the analysis (timed-out trials accounted for 0.01% of total trials). Viewpoint costs (collapsed across direction) were analysed using difference scores in both sensitivity and reaction time between 0° and 45° views and between 45° and 75° views as a measure of performance. As expected, performance was worse in the scrambled-blurred group compared to both the blurred-only group and scrambled-only group. Planned t -tests on the data for this scrambled-blurred group showed that: (i) yaw -45° , $+75^\circ$; pitch -75° , -45° , $+45^\circ$; and roll -75° , -45° were not different from chance (all t s < 1.7); (ii) yaw -75° and pitch 75° were significantly below chance (both t s < -2.2 , $p < 0.05$); and (iii) 0° , yaw $+45^\circ$, and roll $+45^\circ$, $+75^\circ$ were just above chance (all t s > 3.0 , $p < 0.05$). The remaining analysis only investigated the critical comparisons between the blurred group and scrambled group. Two separate 3 axis (yaw, pitch, and roll) $\times$ 2 view difference ($0^\circ - 45^\circ$, $45^\circ - 75^\circ$) $\times$ 2 image alteration (blur and scramble, between subjects) mixed-design ANOVAs were conducted on the d' and reaction-time data. The α level was 0.05. All a-posteriori comparisons were Bonferroni-adjusted pairwise comparisons.

3.2.1 Signal detection analysis. The sensitivity results are shown in figure 9. Analysis of the d' data showed better performance in the blurred condition than the scrambled condition at 0° rotation ($t_{42} = 6.95$, $p < 0.001$), replicating the findings of Schwaninger et al (2002) and not those of Lobmaier and Mast (2007). On the basis of the results of experiment 1 and Favelle et al (2007) we expected performance to be better for pitch-down than pitch-up camera rotations. Planned t -tests showed sensitivity to be greater for pitch-down ($M = 1.18$, $SEM = 0.17$) than pitch-up ($M = 0.75$, $SEM = 0.15$) rotations at 45° ($t_{21} = 1.87$, $p < 0.05$) but not at 75° ($t = 0.5$, $p > 0.3$) in the blurred face condition ($M_{\text{pitch down}} = 0.09$, $SEM_{\text{pitch down}} = 0.15$; $M_{\text{pitch up}} = -0.03$, $SEM_{\text{pitch up}} = 0.16$). There were no significant differences between pitch-down and pitch-up rotations at either 45° or 75° in the scrambled-face condition (both $ts < 0.85$, $p > 0.2$).⁽⁷⁾

Figure 9. Mean recognition sensitivity (d') for viewpoint differences in the blurred (a) and scrambled (b) conditions as a function of axis. Error bars represent ± 1 SEM.

The mixed-design ANOVA showed main effects of image alteration ($F_{1,42} = 29.11$, $p < 0.01$, $MSE = 1.05$, $\eta_p^2 = 0.41$), axis ($F_{2,84} = 61.0$, $p < 0.01$, $MSE = 0.19$, $\eta_p^2 = 0.59$), and viewpoint differences ($F_{1,42} = 24.99$, $p < 0.01$, $MSE = 1.27$, $\eta_p^2 = 0.37$). Viewpoint costs were: (i) greater for blurred than for scrambled images, and (ii) greater between 0° and 45° views than between 45° and 75° views. A-posteriori comparisons showed that viewpoint costs were significantly different for the three axes (all $ps < 0.01$), with the greatest view costs found for pitch, followed by yaw and then roll camera rotations. The between-subjects factor, image alteration, interacted with both rotation axis ($F_{2,84} = 9.78$, $p < 0.01$, $MSE = 0.16$, $\eta_p^2 = 0.19$) and view difference ($F_{1,42} = 12.35$, $p < 0.01$, $MSE = 1.27$, $\eta_p^2 = 0.23$). A-posteriori comparisons showed that the difference between the rotation axes depended on the type of image alteration. For blurred images, the viewpoint costs were significantly different for all axes (all $ps < 0.01$), with the greatest view costs for pitch, followed by yaw, and then roll camera rotations. For scrambled images, there was no difference in cost between pitch and yaw camera rotations ($p = 0.42$), and both of these were greater than the cost for roll camera rotations (both $ps < 0.01$). The interaction between image alteration and view difference was based on the 0° to 45° cost being greater for blurred than scrambled images ($p < 0.01$), with there being no difference in the 45° to 75° cost between blurred and

⁽⁷⁾ While sensitivity was not different from chance following 75° pitch-down or pitch-up rotations of blurred or scrambled faces (all $ts < 0.9$, $p > 0.39$), the proportions of correct responses following 75° pitch-down and pitch-up rotations of blurred faces (0.61 and 0.60, respectively) were both above chance (both $ts > 2.4$, $p < 0.05$).

scrambled images ($p = 0.13$). Further, for blurred images, the 0° to 45° cost was greater than the 45° to 75° cost ($p < 0.01$), suggesting an uneven decline in performance. For scrambled images, the decline appeared more steady with the 0° to 45° cost not different from the 45° to 75° cost ($p = 0.3$).

3.2.2 Reaction-time analysis. Figure 10 shows the reaction times as a function of the viewpoint differences for experiment 2. Analysis of the reaction-time data showed faster reaction times in the blurred condition than in the scrambled condition for the 0° full-face view ($t_{42} = 2.2$, $p < 0.05$). The mixed-design ANOVA showed a main effect of camera rotation axis only ($F_{2,84} = 24.1$, $p < 0.01$, $\text{MSE} = 84454$, $\eta_p^2 = 0.37$). No other main effects or interactions were significant. A-posteriori comparisons reflected the pattern in the d' data: viewpoint costs were significantly different for all axes (all p s < 0.01) with the greatest view costs for pitch, followed by yaw, and then roll camera rotations. While the reaction-time data are somewhat noisy, these results suggest that the effects observed for sensitivity are not due to a speed-accuracy trade-off.

Figure 10. Mean reaction time for viewpoint differences in the blurred (a) and scrambled (b) conditions as a function of axis. Error bars represent ± 1 SEM.

4 General discussion

This study was the first to directly compare recognition performance for faces viewed following (camera) rotation about the yaw, pitch, and roll axes. In experiment 1, viewpoint-dependent declines in face recognition were observed following camera rotations about all three axes. Viewpoint costs were greatest for pitch camera rotations, less for yaw camera rotations, and smallest for roll camera rotations. The only directional difference found was for pitch camera rotations—face recognition was less sensitive after pitch-up than after pitch-down camera rotations (left and right rotations were not significantly different for either roll or yaw). While the results of experiment 1 replicate those of Favelle et al (2007), our aim with experiment 2 was to determine whether there was critical information required to support face recognition across viewpoint changes in all three-dimensional axes. We proposed that differences in face recognition following yaw, pitch, and roll camera rotations could be explained by taking into account both the degree of feature occlusion and the amount of disruption to configural information following camera rotation in each case. For example, it was proposed that face recognition was best following roll camera rotations because, unlike rotations about the other two axes, they did not occlude any facial features and generated little or no disruption to the available configural information.

We tested the above-mentioned proposal in experiment 2 using the scrambled-blurred paradigm. We found that recognition performance was better overall for blurred (or feature-information degraded) faces than for scrambled (or configuration degraded) faces (Hayward et al 2008; Schwaninger et al 2002), which supports the widely reported predominance of configural processing in face recognition (Diamond and Carey 1986; Maurer et al 2002; Robbins and McKone 2007; Searcy and Bartlett 1996; Tanaka and Farah 1993). Results also revealed large viewpoint-dependent declines in blurred-face recognition following camera rotations about the yaw, pitch, and roll axes. Since the recognition of these blurred faces was based primarily on configural information, this shows that there were significant losses in configural information following rotations about all three axes. Specifically, pitch camera rotations appeared to result in greater costs to configural processing than yaw camera rotations, and roll camera rotations appeared to result in the smallest cost to configural processing.

Compared to blurred-face recognition, we found comparatively smaller viewpoint-dependent declines in scrambled-face recognition. Since the recognition of these scrambled faces was based primarily on featural information, the results showed that there were significant losses in this type of information following pitch and yaw camera rotations. While the magnitude of this decline in featural information processing was similar for pitch and yaw camera rotations, roll camera rotations appeared to produce little to no viewpoint-dependent decline. Performance was above chance in all scrambled conditions, except for 75° pitch camera rotations, suggesting that featural processing does play a role in face recognition following roll and yaw rotations, but not following extreme rotations in pitch. That is, when available, featural information can be and is used to recognise faces (Schwaninger et al 2002).

Taken together, these patterns of viewpoint decline for blurred and scrambled faces suggest that, while face recognition relies more heavily on configural processing (compared to featural processing), configural processing may actually be more prone to viewpoint-dependent declines following rotation about all three axes. This finding may not be surprising, since configural processing requires a more holistic representation of the face. That is, the front view of a face may be a canonical view for configural processing (since all major features are available) which is likely to be degraded by rotation when 'face' information and features are lost. Featural processing, on the other hand, can be done on the basis of individual features and so is more robust to rotation, at least when a subset of the features is still visible. At first glance, these results appear to contradict the findings of McKone (2008), who found that configural processing was insensitive to changes in yaw axis views. However, task differences between the two studies may account for the apparently discrepant effects of yaw rotation on configural processing. First, McKone employed a much stricter technique to isolate configural processing (eg the composite task). In the current study, a blurring manipulation was used to examine configural processing. While this manipulation should have at least severely limited the accessibility of featural information, some residual featural information (such as colour and skin tone) probably remained in these stimuli. Further research using stricter techniques to isolate configural information in faces rotated about all three axes would resolve this issue.

Perhaps a more critical task difference is that McKone (2008) used a task in which participants were trained to reliably learn faces (top-halves only) as opposed to sequential matching of unfamiliar faces used in the current study. McKone found very low (less than 9%) naming error rates across all views (0°, 45°, and 90°), whereas we found a viewpoint-dependent decline in matching sensitivity from 0° to 75°. That is, McKone (2008) tested configural information processing across different yaw rotations of familiar faces, whereas we investigated here the extent to which configural information processing is available for unfamiliar faces across different view rotations. We argue that

our findings are compatible with McKone's (2008) since our results also suggest that configural information is critical for successful recognition of faces across changes in view. Indeed the two studies taken together seem to suggest that learning configural information across views is critical for learning faces. How it is that we learn to extract and/or process configural information across changes in view is unclear. It is possible that upright faces are encoded as three-dimensional models and so are responsive to all rotated views of a face. However, given that unfamiliar-face recognition is viewpoint-dependent and that the current results suggest that both configural and featural information is susceptible to changes in viewpoint, this option seems unlikely. It may be that the descriptions of the changes to configural information are also view-specific or that we encode the configurations of facial features at multiple viewpoints. Lobmaier and Mast (2007) have shown that configural information plays a larger role in familiar-face than in unfamiliar-face recognition. It may be that learning to encode configural information from different views of a face is an initial part of the face familiarisation process, leading to the robust and viewpoint-invariant recognition performance often found with familiar faces. At any rate, how our use of configural cues across viewpoint changes with increased familiarity with a face may provide important insights into the familiarisation process itself and is certainly worthy of future research.

Findings from experiment 1 and Favelle et al (2007) showed better face-recognition performance for yaw than for pitch camera rotations. The results of experiment 2 suggest that, while the loss of featural information plays a role, it is the greater disruption to configural processing following pitch camera rotation that is responsible for these axis-based view differences in face recognition. Why might pitch rotations disrupt configural processing to a greater degree than yaw? One explanation could be that since configural processing requires a more holistic representation of the face, the fact that pitch camera rotations result in greater foreshortening and occlusion of features means that there is less 'face' information available in order to extract configural information. Another way in which pitch and yaw camera rotations might differ in the degree of disruption to configural processing is in differential distortion to the vertical (eg distance between the mouth and nose) and horizontal (eg interocular distance) relations. Goffaux (2008) investigated the optimal spatial-frequency range for vertical and horizontal spatial relations in face. She found that vertical relations were processed best in intermediate spatial frequencies, which have been shown to carry information useful for face individuation (Näsänen 1999), whereas horizontal relations were processed best in high spatial frequencies, which are involved in processing local features. All vertical relations are dramatically distorted by pitch camera rotations, but by not yaw camera rotations. Thus, our findings of a greater cost to face recognition following pitch camera rotations may be due to face individuation relying more heavily on vertical relations (which would have been available even in our blurred face stimuli—which only cut off high spatial frequencies).

Findings from experiment 1 and Favelle et al (2007) also showed better face recognition performance after pitch-down than after pitch-up camera rotations. Experiment 2 showed better performance for pitch-down than pitch-up blurred faces at 45°, but no difference at 75°. There was no difference between pitch-up and pitch-down in the scrambled-faces condition. This suggests that more configural information is preserved after pitch-down than after pitch-up camera rotations (at least up to 45°), which in turn leads to better face recognition. However, view familiarity could also explain the directional differences in performance following pitch camera rotation. While pitch-down and pitch-up views are relatively common in everyday experience (eg nodding heads or interactions between one person sitting and the other standing), as children we are more exposed to pitch-down than pitch-up camera views. This extensive experience may at least partially account for the better performance for these views. Physiological data

show that there are face-selective cells that respond maximally to yaw-rotated views with fewer cells responding to pitch views (Perrett et al 1985). Perrett and colleagues (1998) proposed a mechanism whereby cells become tuned to a particular face view via repeated exposure. Thus, people take longer to recognise familiar faces in unusual orientations because the neural representation of the face is less strongly activated.

The present study has focused on the role of configural and featural information in face-viewpoint effects. However, there are other factors that might be considered, including visibility of specific facial features and face lighting (eg direction). While we have manipulated the availability of overall featural information by image blurring in experiment 2, there is still the possibility that the processing of specific features plays an important role in recognising intact, full-spectrum faces. For example, many studies have demonstrated the importance of eyes in face recognition (eg Davies et al 1977; Ellis et al 1979; Young et al 1985; Haig 1986). Looking at figure 2, much of the eye information has disappeared at 45° pitch-up camera rotations, but in pitch-down camera rotations there is at least partial eye information available at rotations of 75°. This could explain the poorer recognition performance for pitch-up compared to pitch-down camera rotations. By contrast, yaw camera rotations only ever resulted in the (partial) loss of information about one eye. Thus, the relatively better performance following yaw camera rotations may be a consequence of more eye information being available.

In experiment 1, we directly compared the effects of roll, pitch, and yaw camera rotations on face recognition, and found that performance was always best following the roll camera rotations. This finding was further supported by the results of experiment 2, which showed that roll camera rotation was less disruptive to both configural and featural processing than yaw and pitch camera rotations. It is possible that this superior performance following roll camera rotations was at least partly a consequence of the lighting conditions. The face images were created to simulate roll changes in observer viewpoint and, as such, the lighting was held constant (faces were lit from above the head). As a result, there was no change in the pattern of lighting across the face following roll camera rotations. By contrast, there were noticeable changes in the pattern of lighting following pitch and yaw camera rotations, which were actual photographs of the same upright observer's face taken from different positions in space (see figures 1 and 2). While these different lighting patterns were ecological (consistent with viewpoint, as opposed to face, rotation), it is possible that this constant face-centred pattern of lighting in roll rotation conditions was used as an additional cue to recognition. However, since Favelle and Hill (2009) have recently demonstrated that pitch and yaw head rotation effects are stronger than lighting rotation effects, we believe it is unlikely that the particular lighting conditions used in our study had much influence on the current pattern of results.

The present study is the first to examine the effects of yaw, pitch, and roll camera rotations on face recognition. It also examined the specific effects of these three types of rotation on configural and featural face processing. The range of viewpoints examined in this study (0°–75°) was chosen to approximately reflect our everyday experience with faces. Following roll, pitch, and yaw camera rotations within this range: (i) face recognition was found to be viewpoint-dependent, and (ii) configural face processing was found to be superior to featural face processing. Configural face processing was best after roll camera rotations, followed by yaw camera rotations, and was found to be weakest for pitch camera rotations, pitch-up camera rotations in particular. Featural face processing showed somewhat similar (but lesser) patterns of viewpoint-dependent decline following pitch and yaw (but not roll) camera rotations. These results suggest that while both configural and featural processing is used to recognise faces across different views, the use of configural processing (at least when sufficient 'face' information is still available) is critical to accurate face recognition.

Acknowledgments. This research was supported by an Australian Research Council Discovery Grant DP0986420. We thank our reviewers for their helpful comments and feedback.

References

- Barton J J S, Keenan J P, Bass T, 2001 "Discrimination of spatial relations and features in faces: Effects of inversion and viewing duration" *British Journal of Psychology* **92** 527–549
- Bruce V, Henderson Z, Newman C, Burton A M, 2001 "Matching identities of familiar and unfamiliar faces caught on CCTV images" *Journal of Experimental Psychology: Applied* **7** 207–218
- Bülthoff H H, Edelman S, 1992 "Psychophysical support for a 2-D view interpolation theory of object recognition" *Proceedings of the National Academy of Sciences of the USA* **89** 60–64
- Cabeza R, Kato T, 2000 "Features are also important: Contributions of featural and configural processing to face recognition" *Psychological Science* **11** 429–433
- Collishaw S M, Hole G J, 2000 "Featural and configural processes in the recognition of faces of different familiarity" *Perception* **29** 893–909
- Collishaw S M, Hole G J, 2002 "Is there a linear or a nonlinear relationship between rotation and configural processing of faces?" **31** 287–296
- Davies G, Ellis H, Shepherd J, 1977 "Cue saliency in faces as assessed by the 'Photofit' technique" *Perception* **6** 263–269
- Diamond R, Carey S, 1986 "Why faces are and are not special: An effect of expertise" *Journal of Experimental Psychology: General* **115** 107–117
- Donnelly N, Davidoff J, 1999 "The mental representations of faces and houses: Issues concerning parts and wholes" *Visual Cognition* **6** 319–343
- Ellis H D, Shepherd J W, Davies G M, 1979 "Identification of familiar and unfamiliar faces from internal and external features: some implications for theories of face recognition" *Perception* **8** 431–439
- Favelle S, Hill H, 2009 "Matching faces across rotations in view and lighting" paper presented at the 9th Conference of the Australasian Society for Cognitive Science, Sydney, Australia, 30 September–2 October 2009
- Favelle S, Palmisano S, Maloney R, 2007 "Things are looking up: differential decline in face recognition following pitch and yaw rotation" *Perception* **36** 1334–1352
- Freire A, Lee K, Symons L A, 2000 "The face-inversion effect as a deficit in the encoding of configural information: Direct evidence" *Perception* **29** 159–170
- Gauthier I, Hayward W G, Tarr M J, Anderson A, Skudlarski P, Gore J C, 2002 "BOLD activity during mental rotation and viewpoint-dependent object recognition" *Neuron* **34** 161–171
- Gauthier I, Tarr M J, 1997 "Becoming a 'Greeble' expert: Exploring mechanisms for face recognition" *Vision Research* **37** 1673–1682
- Goffaux V, 2008 "The horizontal and vertical relations in upright faces are transmitted by different frequency ranges" *Acta Psychologica* **128** 119–126
- Haig N, 1986 "Exploring recognition with interchanged facial features" *Perception* **15** 235–247
- Hancock P J B, Bruce V, Burton A M, 2000 "Recognition of unfamiliar faces" *Trends in Cognitive Sciences* **4** 330–337
- Hayward W G, Rhodes G, Schwaninger A, 2008 "An own-race advantage for components as well as configurations in face recognition" *Cognition* **106** 1017–1027
- 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
- Hill H, Schyns P G, Akamatsu S, 1997 "Information and viewpoint dependence in face recognition" *Cognition* **62** 201–222
- Jeffery L, Rhodes G, Busey T, 2006 "View-specific coding of face shape" *Psychological Science* **17** 501–505
- Leder H, Bruce V, 2000 "When inverted faces are recognized: the role of configural information in face recognition" *Quarterly Journal of Experimental Psychology A* **53** 513–536
- Lee Y, Matsumiya K, Wilson H R, 2006 "Size-invariant but viewpoint-dependent representation of faces" *Vision Research* **46** 1901–1910
- Lewis M B, 2001 "The lady's not for turning: Rotation of the Thatcher illusion" *Perception* **30** 769–774
- Lewis M B, Glenister T E, 2003 "A sideways look at configural encoding: Two different effects of face rotation" *Perception* **32** 7–14
- Liu C H, Chaudhuri A, 2000 "Recognition of unfamiliar faces: three kinds of effects" *Trends in Cognitive Sciences* **14** 445–446

- Liu C H, Chaudhuri A, 2002 "Reassessing the $\frac{3}{4}$ view effect in face recognition" *Cognition* **83** 31–48
- Liu C H, Collin C A, Farivar R, Chaudhuri A, 2005 "Recognizing faces defined by texture gradients" *Perception & Psychophysics* **67** 158–167
- Liu C H, Collin C A, Burton A M, Chaudhuri A, 1999 "Lighting direction affects recognition of untextured faces in photographic positive and negative" *Vision Research* **39** 4003–4009
- Lobmaier J S, Mast F W, 2007 "Perception of novel faces: The parts have it!" *Perception* **36** 1660–1673
- Logie R H, Baddeley A D, Woodhead M M, 1987 "Face recognition, pose and ecological validity" *Applied Cognitive Psychology* **1** 53–69
- McKone E, 2008 "Configural processing and face viewpoint" *Journal of Experimental Psychology: Human Perception and Performance* **34** 310–327
- Macmillan N A, Creelman C D, 2005 *Detection Theory: A User's Guide* 2nd edition (London: Lawrence Erlbaum Associates)
- Maurer D, Le Grand R, Mondloch C J, 2002 "The many faces of configural processing" *Trends in Cognitive Sciences* **6** 255–260
- Murray J, Yong E, Rhodes G, 2000 "Revisiting the perception of upside-down faces" *Psychological Science* **11** 492–496
- Näsänen R, 1999 "Spatial frequency bandwidth used in the recognition of facial images" *Vision Research* **39** 3824–3833
- Newell F N, Chiroro P, Valentine T, 1999 "Recognizing unfamiliar faces: The effects of distinctiveness and view" *Quarterly Journal of Experimental Psychology A* **52** 509–534
- 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
- O'Toole A J, Vetter T, Blanz V, 1999 "Three-dimensional shape and two-dimensional surface reflectance contributions to face recognition: and application of three-dimensional morphing" *Vision Research* **39** 3145–3155
- Perrett D I, Oram M W, Ashbridge E, 1998 "Evidence accumulation in cell populations responsive to faces: an account of generalisation of recognition without mental transformations" *Cognition* **67** 111–145
- Perrett D, Smith P, Potter D, Mistlin A, Head A, Milner A, Jeeves M, 1985 "Visual cells in the temporal cortex sensitive to face view and gaze direction" *Proceedings of the Royal Society of London, Series B* **223** 293–317
- Poggio T, Edelman S, 1990 "A network that learns to recognize three-dimensional objects" *Nature* **343** 263–266
- Rakover S S, 2002 "Featural vs. configurational information in faces: A conceptual and empirical analysis" *British Journal of Psychology* **93** 1–30
- Rhodes G, 1985 "Lateralized processes in face recognition" *British Journal of Psychology* **76** 249–271
- Robbins R, McKone E, 2007 "No face-like processing for objects-of-expertise in three behavioural tasks" *Cognition* **103** 34–79
- Schwaninger A, Lobmaier J S, Collishaw S M, 2002 "Role of featural and configural information in familiar and unfamiliar face recognition" *Lecture Notes in Computer Science* **2525** 643–650
- Schwaninger A, Mast F W, 2005 "The face-inversion effect can be explained by the capacity limitations of an orientation normalization mechanism" *Japanese Psychological Research* **47** 216–222
- Schyns P G, Bülthoff H H, 1994 "Viewpoint dependence and face recognition", in *Proceedings of the XVI Meeting of the Cognitive Science Society* Eds A Ram, K Eiselt (Hillsdale, NJ: Lawrence Erlbaum Associates) pp 789–793
- Searcy J H, Bartlett J C, 1996 "Inversion and processing of component and spatial-relational information in faces" *Journal of Experimental Psychology: Human Perception and Performance* **22** 904–915
- 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
- Tanaka J W, Farah M, 1993 "Parts and wholes in face recognition" *Quarterly Journal of Experimental Psychology* **46** 225–245
- Troje N F, Bülthoff H H, 1996 "Face recognition under varying poses: The role of texture and shape" *Vision Research* **36** 1761–1771
- Troje N F, Bülthoff H H, 1998 "How is bilateral symmetry of human faces used for recognition of novel views?" *Vision Research* **38** 79–89
- Valentin D, Abdi H, Edelman B, 1999 "From rotation to disfiguration: testing a dual-strategy model for recognition of faces across view angles" *Perception* **28** 817–824

-
- Valentine T, 1988 "Upside down faces: A review of the effect of inversion upon face recognition" *British Journal of Psychology* **79** 471–491
- Valentine T, Bruce V, 1988 "Mental rotation of faces" *Memory & Cognition* **16** 556–566
- Wallraven C, Schwaninger A, Schuhmacher S, Bülthoff H H, 2002 "View-based recognition of faces in man and machine: Re-visiting inter-extra-ortho" *Lecture Notes in Computer Science* **2525** 651–660
- Yin R K, 1969 "Looking at upside-down faces" *Journal of Experimental Psychology* **81** 141–145
- Young A W, Hay D C, McWeeny K H, Flude B M, Ellis A W, 1985 "Matching familiar and unfamiliar faces on internal and external features" *Perception* **14** 737–746
- Young A W, Hellawell D J, Hay D C, 1987 "Configural information in face perception" *Perception* **16** 747–759

ISSN 0301-0066 (print)

ISSN 1468-4233 (electronic)

PERCEPTION

VOLUME 40 2011

www.perceptionweb.com

Conditions of use. This article may be downloaded from the Perception website for personal research by members of subscribing organisations. Authors are entitled to distribute their own article (in printed form or by e-mail) to up to 50 people. This PDF may not be placed on any website (or other online distribution system) without permission of the publisher.