Camera-to-subject distance affects face configuration and perceived identity

Eilidh Noyes & Rob Jenkins

Department of Psychology, University of York, UK

Correspondence concerning this article should be addressed to:

Rob Jenkins, Department of Psychology, University of York, York, YO10 5DD, UK

Email: rob.jenkins@york.ac.uk

Phone: +44(0)1904 323144

Running Head: Camera distance affects face configuration

Keywords: Face recognition, face matching, identification, configural processing, camera, passport

Abstract

Face identification is reliable for viewers who are familiar with the face, and unreliable for viewers who are not. One account of this contrast is that people become good at recognising a face by learning its configuration—the specific pattern of feature-to-feature measurements. In practice, these measurements differ across photos of the same face because objects appear more flat or convex depending on their distance from the camera. Here we connect this optical understanding to face configuration and identification accuracy. Changing camera-to-subject distance (0.32 m versus 2.70 m) impaired perceptual matching of unfamiliar faces, even though the images were presented at the same size. Familiar face matching was accurate across conditions. Reinstating valid distance cues mitigated the performance cost, suggesting that perceptual constancy compensates for distance-related changes in optical face shape. Acknowledging these distance effects could reduce identification errors in applied settings such as passport control.

Introduction

Recognising a person's face involves mapping an image onto an identity (Bruce & Young, 1986). Completing this mapping reliably is a challenge, because images of the same face can be as varied as images of different faces (Jenkins et al., 2011). One way of characterising this challenge is as a signal-to-noise problem: the signal is the 'true' appearance of the person's face, the noise consists of deviations from that appearance, and the task of the visual system is to extract the signal from the noise.

For some time, the search for an identity signal has centred on *configural* information—the idiosyncratic spatial layout of a person's face, typically defined in terms of 'metric distances between features' (Tanaka & Gordon, 2011). The proposal is that each face has a unique configuration, which viewers come to learn, and that knowledge of that configuration allows the viewer to recognise that particular person (Diamond & Carey, 1986; Maurer, Le Grand, & Mondloch, 2002). A parallel research effort has documented effects of different types of noise on identification accuracy. For example, many studies have measured the impact of viewing angle, facial expression, and lighting conditions on observers' performance (Bruce, 1982; Hill & Bruce, 1996; Johnston, Hill, & Carman, 1992; Troje, & Bülthoff, 1996; Young et al., 1986).

Relatively few studies have examined effects of viewing *distance*. Those studies have typically been motivated by forensic questions concerning eyewitness testimony (De Jong et al., 2005; Greene & Fraser, 2002; Hahn, O'Toole, & Phillips, 2016; Lampinen et al., 2014; Lindsay et al., 2008; Loftus & Harley, 2005; Wagenaar & van der Schrier, 1996). As such, they have focused almost exclusively on two inter-related questions, (i) What is the maximum distance at which we can recognise a face?, and (ii) What is the minimum information required to recognise a face? These questions arise directly from the optics of the situation: more distant objects project smaller retinal images (with implications for spatial frequency content; Loftus & Harley, 2005). But optics gives us another reason to take viewing distance seriously: changes in viewing distance affect configural information in the face image.

Perhaps the most compelling demonstration of this configural change comes not from face recognition research, but from analyses of perspective in portraiture. Harper and Latto (2001) photographed models' faces at different camera-to-subject distances (0.32 m, 0.71 m, 2.70 m), and rescaled the faces to the same interocular distance. As Figure 1 illustrates, faces look convex when close, and flatter from afar. In other words, the same face appears to have quite distinct shapes when viewed from different distances. Indeed, participants in Harper and Latto's (2001) study gave higher weight estimates for the models as camera-to-subject distance increased.

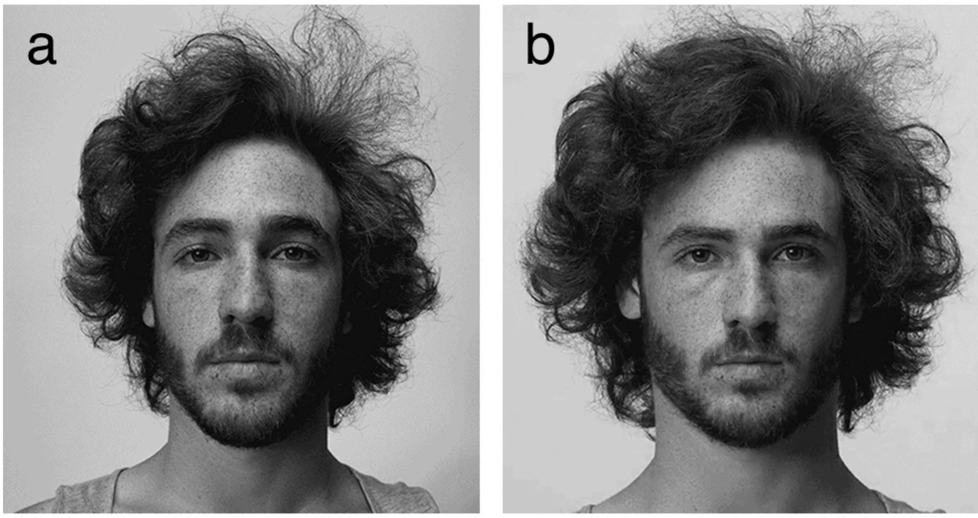

Figure 1. Two photographs of the same face taken from different viewing distances: (a) ~0.20 m; (b) ~3.00 m. Photos are shown rescaled to the same interocular distance. © Dan Vojtěch 2016. Reproduced with permission.

More recent work (Bryan, Perona & Adolphs, 2012) has shown that social inferences from faces also change with camera-to-subject distance. Viewers' ratings of trustworthiness, competence, and attractiveness were all lower for photos that were taken closer (0.45 m) than for photos that were taken further away (1.35 m), presumably because such inferences rely partly on shape cues (Oosterhof & Todorov, 2008; Zebrowitz, 2011).

Taken together, these studies confirm that (i) face shape in the image changes with camera-to-subject distance, and (ii) these shape changes are large enough to have psychological consequences, even when the difference in viewing distance is small (e.g. 1–2 metres). Despite the clarity of these findings, there has been almost no attempt to pursue their implications for the important issue of face identification (see Liu, 2003, for an exception). This oversight is perhaps surprising, given the emphasis on configural information in the face recognition literature. If viewing distance alters configural information, and configural information is key to face identification, it follows that viewing distance should affect face identification. That is the argument that we examine in the present studies. We begin with a direct test of the first premise, that viewing distance alters configural information.

Study 1. Camera-to-subject distance affects feature-to-feature measurements

The purpose of this study was to relate changes in camera-to-subject distance to changes in facial configuration. The apparent size of an object clearly varies with viewing distance, in the sense that the size of the retinal image changes. Linear changes in the size of a 2D face image (as when a photograph is rescaled) do not affect configural layout because they do not affect the relative distances between features. Consistent with the conservation of configural layout over size changes, behavioural and neuroimaging studies have found that face recognition is unaffected by linear rescaling (Andrews & Ewbank, 2004; Bruce et al., 1994; Bindemann et al., 2008; Grill-Spector et al. 1999). For 3D objects (e.g. live faces as opposed to face photographs), the optical situation is very different. Changes in camera-to-subject distance produce non-linear changes in the image, such that different parts of the image are affected to differing degrees (Latto & Harper, 2007; Pirenne, 1970). For convex objects such as faces, distant viewing leads to flatter appearance, whereas closer viewing leads to more convex appearance (see Figure 1). To tie this optical transformation directly to the notion of configuration in the face perception literature, we measured distances between key facial features in photos that were taken at different viewing distances. The expectation was that, as a reflection of the flat-to-convex transformation, measures nearer the edge of the face would be compressed relative to measures nearer the centre of the face.

Photographic procedure

The images used for all of these studies were face photographs of 18 consenting undergraduates at the University of York. These volunteer models were photographed in two separate sessions, one week apart. In each session, each model was photographed at two distances—*Near* (camera-to-subject distance = 0.32 m) and *Far* (camera-to-subject distance = 2.70 m), following Harper and Latto (2001). This regime resulted in four photographs for each of the 18 models: Week 1 *Near*, Week 1 *Far*, Week 2 *Near*, and Week 2 *Far* (72 photos in total). All models were photographed with a neutral expression using an Apple iPhone 5s on default settings. Photos were then cropped around the head to remove extraneous background. For anthropometric analysis, each image was scaled to an interocular distance of 150 pixels, with aspect ratio preserved.

Anthropometric analysis

We follow Burton et al. (2015) in extracting from the literature those feature-to-feature distances that have been offered as specific examples: distance between the corner of the eye and the edge of the nose (left and right; Leder & Carbon, 2006), distance between the corner of the nose and the corner of the mouth (left and right; Leder & Bruce, 2000), and distance between the nose and the mouth (Leder & Carbon, 2006; see Burton et al., 2015, for precise anatomical definitions). This resulted in five measurements in total for each photograph, which were made using the Ruler tool in Adobe PhotoShop. Figure 2 shows these five measurements for *Near* and *Far* photos of one volunteer model.

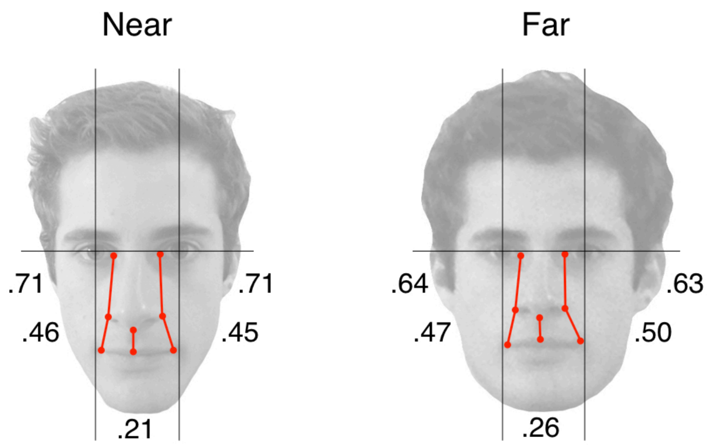

Figure 2. Configural information (distances between facial features) for *Near* (left) and *Far* (right) photos of the same person. Photo size is standardised for interocular distance. Red lines show distances between facial features. All distances are expressed as proportions of standardised interocular distance (see Burton et al., 2015).

Results and Discussion

For each of the five feature-to-feature metrics, we conducted a 2 x 2 repeated-measures ANOVA with the factors of *Photographic Session* (*Week 1* versus *Week 2*) and *Camera-to-Subject Distance* (*Near* versus *Far*). Results of these analyses are summarised in Table 1.

Table 1. Means across eighteen faces for five different feature-to-feature measurements, separated by photographic condition. Measurements are shown as proportions of interocular distance. *Wk 1* and *Wk 2* refer to the two photographic sessions. *Near* and *Far* refer to the two camera-to-subject distances. *Diff* = Difference; *ES* = Effect Size (partial eta squared).

	Photographic Session						**Camera-to-Subject Distance**					
Measurement	Wk 1	Wk 2	Diff	*F*	*p*	*ES*	Near	Far	Diff	*F*	*p*	*ES*
Ear–Nose (Left)	.57	.57	.00	.05	>.05	.00	.58	.56	-.02	1.67	>.05	.09
Ear–Nose (Right)	.58	.57	.01	.20	>.05	.01	.59	.57	-.02	3.53	>.05	.17
Nose–Mouth (Left)	.42	.41	-.01	1.58	>.05	.09	.38	.45	.07	39.82	<.001	.70
Nose–Mouth (Right)	.42	.41	-.01	1.23	>.05	.07	.38	.45	.07	42.97	<.001	.72
Nose–Mouth (Centre)	.22	.23	.01	1.10	>.05	.06	.20	.24	.04	17.26	<.001	.50

As can be seen from Table 1, *Photographic Session* had no significant effect on any of the measurements ($p > .1$ for all), indicating that incidental changes in viewpoint and expression for *Week 1* versus *Week 2* were negligible. In this context, *Camera-to-Subject Distance* systematically affected some measures but not others. The relatively peripheral nose-to-mouth measurements were larger for *Far* images than for *Near* images, whereas the more central eye-to-nose measurements were statistically equivalent at the two camera distances we compared. This pattern in the anthropometric data corroborates the flatter appearance of the *Far* images and the more convex appearance of the *Near* images, and is consistent with the differential weight estimates in previous studies (Harper & Latto, 2001). More importantly for the current study, it confirms the non-linear effect of camera-to-subject distance on configural information: some feature-to-feature measurements changed substantially and others did not (see Smith, 2016, for a computational perspective). We next used a paired matching task to assess the implications of these configural changes for perception of facial identity.

Study 2. Distance-related changes in configuration impair unfamiliar face matching

Many previous studies have shown that unfamiliar face matching is a difficult task. One influential measure of this ability is the Glasgow Face Matching Test (GFMT). In this task, participants are presented with pairs of face photographs that were taken with different cameras (Burton et al. 2010). For each pair, the participant's task is to decide whether the two photos show the same person (*Same* trials; 50% prevalence) or two different people (*Different* trials; 50% prevalence). Despite the simplicity of this task, error rates are high when the faces are unfamiliar, because the viewer does not have enough information to distinguish image changes from identity changes. When viewers match familiar faces, errors are virtually absent, presumably because the range of possible appearances for those faces is better understood (Jenkins & Burton, 2011; Jenkins, McLachlan, & Renaud, 2014).

In the current study, we extended the standard paired matching design by adding camera-to-subject distance as an experimental factor. All of the images were presented at the same standard size. Because the configural information in a face image changes with camera-to-subject distance (Study 1), manipulating distance allowed us to make very specific predictions. For viewers who are *unfamiliar* with the faces concerned, a change in camera-to-subject distance should impair performance on *Same Identity* trials, because it generates dissimilar images. At the same time, it should *improve* performance on *Different Identity* trials, (for the same reason). If identity judgments by familiar viewers rely on configural information, then their performance should be similarly affected. However, given that familiar viewers readily see through changes in viewpoint, lighting, facial expression, and other factors (e.g. Burton et al., 1999; Jenkins & Burton, 2011; Jenkins et al., 2011), we anticipated that familiar viewers might similarly see through changes arising from camera-to-subject distance, such that their performance would be unaffected by this manipulation.

Method

Participants

Forty-five psychology undergraduates at the University of York participated in exchange for payment or course credit. Twenty-three of these participants were first-year students who arrived at the University of York after our photographic models had left, and so had never seen the faces in the stimulus set (verified post-test; see Procedure section below). We refer to these participants as *Unfamiliar* viewers (19 female, 4 male; mean age = 18.7 years). The remaining 22 participants were other students from the same year group as our photographic models, and had spent over two years studying on the same course (19 female, 3 male; mean age = 22.1 years). We refer to these participants as *Familiar* viewers because they had seen the faces in the stimulus set routinely over those two years (again, verified post-test). All participants reported normal or corrected-to-normal vision at the point of recruitment.

Stimuli and Design

144 face pairs were constructed from the 72 photographs used in Study 1. As in Study 1, all images were scaled to an interocular distance of 150 pixels, preserving aspect ratio. To achieve full counterbalancing, we crossed the within-subjects factors of *Identity* (*Same*, *Different*) and camera-to-subject *Distance* (*Same*, *Different*), so that each model's face was paired with (i) another photo of the same person, taken at the same distance (*Same Identity, Same Distance* condition), (ii) another photo of the same person, taken at a different distance (*Same Identity, Different Distance* condition), (iii) a photo of a different person, taken at the same distance (*Different Identity, Same Distance* condition), and (iv) a photo of a different person, taken at a different distance (*Different Identity, Different Distance* condition). Each pair comprised two photos that were taken in different photographic sessions (*Week 1* and *Week 2*) so that viewers were never comparing identical images (Bruce, 1982). For different identity trials, the foil was the person in the stimulus set who most closely resembled the target (White et al., 2014; Jenkins & Kerr, 2013). Figure 3 summarises the experimental conditions.

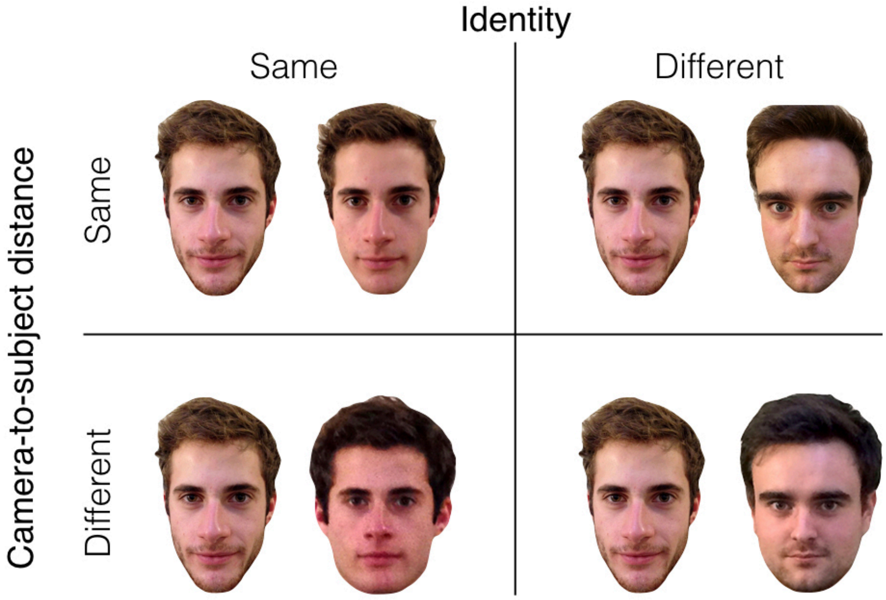

Figure 3. Summary of the experimental conditions in Study 2 with example stimuli. The left column shows *Same Identity* pairs and the right column shows *Different Identity* pairs. The top row shows *Same Distance* pairs and the bottom row shows *Different Distance* pairs. *Near* and *Far* images were presented at the same size (same interocular distance). Each image pair contained one *Week 1* photo and one *Week 2* photo. This design ensured that the *Same Identity, Same Distance* condition was not reduced to image matching.

Procedure

The two groups of participants (*Familiar* and *Unfamiliar*) were tested in separate classroom sessions on different days. Each session began with verbal instruction followed by four example trials using images that were not used in the main experiment. Face pairs consisted of two images presented at the same size (i.e. matched for interocular distance). These displays were projected at a size of approximately 180 cm x 100 cm using a computer controlled data projector. Participants viewed the projection screen at a distance of between 3 and 5 metres. At this range, the image pairs subtended a visual angle of approximately 11–19° vertically and 20–33° horizontally.

Each of the 144 face pairs was presented for 5 seconds, and successive pairs were separated by a 3-second interval that was filled with a visual countdown. During this 8-second cycle, each participant was required to decide whether the two images showed the same face or two different faces, and to record this decision by circling 'Same' or 'Different' on a printed answer sheet. Participants worked through the sheet individually and in silence.

Following the matching task, participants were presented with a photographic array showing all 18 experimental models, and were asked to indicate any faces that were already familiar to them before the experiment. The entire testing session took approximately 30 minutes to complete.

Results and Discussion

To ensure that the familiarity manipulation was not compromised by items from the opposite category, any faces that were unknown to an observer in the *Familiar* group (< 10%) or known to an observer in the *Unfamiliar* group (0%) were excluded from analysis. Participants' accuracy scores from the remaining trials were submitted to a 2 x 2 x 2 mixed ANOVA with the within-subjects factors of *Identity* (*Same*, *Different*) and camera-to-subject *Distance* (*Near*, *Far*), and the between-subjects factor of *Familiarity* (*Familiar*, *Unfamiliar*). Results of this analysis are summarised in Figure 4.

As expected, overall accuracy was significantly lower for *Unfamiliar* viewers (M = 85.71%, SE = .80, CI = 84.11 – 87.31) than for *Familiar* viewers (M = 97.89%, SE = .81, CI = 96.25 – 99.53), [$F(1,43) = 114.8$, $p < .001$, $\eta^2 = .73$]. There was a significant main effect of *Identity*, with slightly higher accuracy for *Different Identity* trials (M = 93.36%, SE = .91, CI = 91.53 – 95.20) than for *Same Identity* trials (M = 90.25%, SE = .87, CI = 88.49 – 92.00), [$F(1,43) = 5.17$, $p = .03$, $\eta^2 = .11$]. However there was no interaction between *Identity* and *Familiarity* [$F(1,43) = 1.53$, $p = .22$, $\eta^2 = .03$].

More importantly for this study, there was also a significant main effect of *Distance*, with participants performing better overall for *Same Distance* trials (M = 94.92%, SE = .55, CI = 93.82 – 96.03) than for *Different Distance* trials (M = 88.69%, SE = .85, CI = 86.97 – 90.38), [$F(1,43) = 51.90$, $p < .001$, $\eta^2 = .547$]. Moreover, the interaction between *Distance* and *Familiarity* [$F(1,43) = 21.80$, p < .001, $\eta^2 = .34$] and the interaction between *Distance* and *Identity* [$F(1,43) = 133.64$, $p < .001$, $\eta^2 = .76$] were highly robust. Finally, there was a significant three-way interaction between *Identity, Distance* and *Viewer Familiarity* $F(1,43)$

= 117.45, $p < .001$, $\eta^2 = .73$. To break down this three-way interaction, we next carried out separate 2 x 2 within-subjects ANOVAs for the *Familiar* and *Unfamiliar* groups.

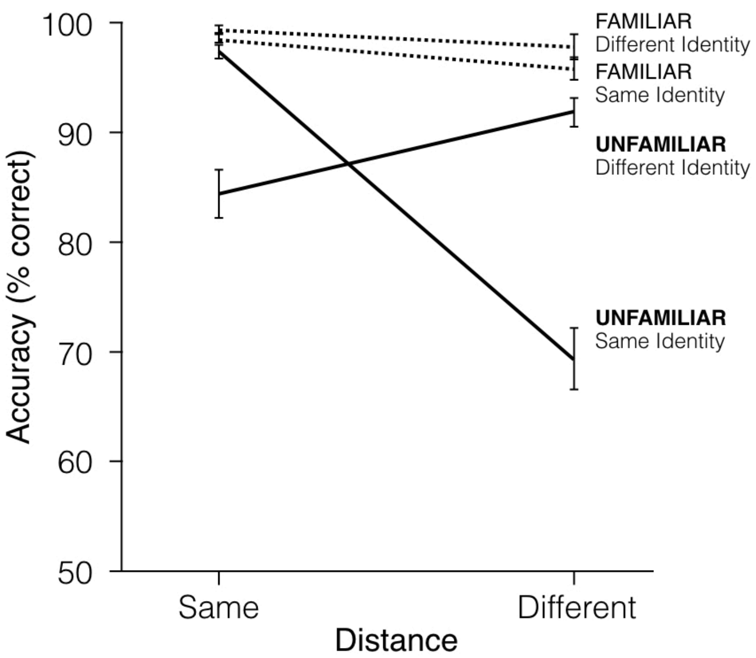

Figure 4. Mean accuracy scores for the conditions in Study 2. Dotted lines show *Familiar* viewers, solid lines show *Unfamiliar* viewers. Error bars show SEM.

Familiar viewers

As can be seen from Figure 4, *Familiar* viewers performed with very high accuracy in both *Same Distance* and *Different Distance* conditions. Accuracy was significantly higher for *Same Distance* image pairs (M = 98.84% correct, SD = 1.37) than for *Different Distance* image pairs (M = 96.70% correct, SD = 4.17), [$F(1,21) = 5.49$, $p < .05$, $\eta^2 = .21$], although the magnitude of this effect was small. There was no significant effect of *Identity* [$F(1,21) = 3.41$, $p = .08$, $\eta^2 = .14$] and no interaction between these two factors $p > .05$.

Unfamiliar viewers

Unfamiliar viewers performed significantly better for *Same Distance* pairs (M = 90.85% correct, SD = 4.97) than for *Different Distance* (M = 80.54%, SD = 6.83) [$F(2,22) = 51.20$, $p < .001$, $\eta^2 = .67$]. As with the *Familiar* group, there was no significant main effect of *Identity* [$F(1,22) = 3.48$, $p = .08$, $\eta^2 = .14$]. Critically however, there was a strong interaction between

Identity and *Distance* [$F(1,22) = 155.13$, $p < .001$, $\eta^2 = .88$]. For *Same Identity* pairs, accuracy was much lower for *Different Distance* images ($M = 69.29\%$, SD = 13.41) than for *Same Distance* ($M = 97.33\%$, SD = 3.09)—a cost of 28%. However, for *Different Identity* pairs the opposite pattern emerged: accuracy was higher for *Different Distance* ($M = 91.86\%$, SD = 6.29) images than for *Same Distance* images ($M = 84.37\%$, SD = 10.64), though this effect was smaller (8%). Simple main effects confirmed that there was a significant effect of *Distance* for *Same Identity* pairs [$F(1,44) = 191.92$, $p < .001$, $\eta^2 = .81$], and for *Different Identity* pairs [$F(1,44) = 13.69$, $p < .001$, $\eta^2 = .31$]. There was also a significant effect of *Identity* for both *Same Distance* images [$F(1,44) = 19.34$, $p < .001$, $\eta^2 = .31$] and *Different Distance* images [$F(1,44) = 58.73$, $p < .001$, $\eta^2 = .57$].

Changing camera-to-subject distance severely impaired face-matching performance for *Unfamiliar* viewers. In contrast, *Familiar* viewers were barely affected, and performed highly accurately in all conditions. The pattern for *Unfamiliar* viewers was exactly in line with our predictions: changing camera-to-subject distance drastically reduced accuracy on *Same Identity* trials, but increased accuracy on *Different Identity* trials.

Previous work has established the pictorial nature of unfamiliar face matching (Burton & Jenkins, 2011; Hancock, Bruce, & Burton, 2000), in which identity judgements are easily swayed by superficial characteristics of the comparison image. Changes in viewing angle (Bruce, 1982), facial expression (Young et al., 1986), and lighting (Johnston, Hill, & Carman, 1992) have all been found to reduce accuracy in unfamiliar face matching. The GFMT (Burton, White, & McNeill, 2010) reveals a performance cost for a change in camera type, in the context of camera-to-subject distance being held constant. Here we demonstrate the converse effect: changing camera-to-subject distance impairs performance, even when the camera stays the same.

The observation that, for the *Familiar* group, accuracy was high across all conditions is consistent with the general robustness of familiar face matching against image change (Burton et al, 1999; Jenkins & Burton, 2011, Jenkins & Kerr, 2013). However, the fact that this particular image change was a change in facial configuration suggests that whatever underpinned the accurate performance of familiar viewers, it can not have been configural information. We return to this issue in the general discussion.

The results of this study raise an interesting question: if faces look more flat or convex at different viewing distances, why do we not notice these changes in daily life? The next experiment addresses this question by examining perceptual constancy for face shape.

Study 3. Perceptual constancy for face shape

So far, we have seen that changing camera-to-subject distance changes facial configuration (Study 1), and that this manipulation impairs unfamiliar face matching (Study 2). Yet these findings seem at odds with day-to-day experience. In daily life, we tend not to perceive distance-related changes in face shape. Faces do not appear to become more convex as they approach. How are we to account for this discrepancy between laboratory findings and real world experience?

Vision research has identified many image transformations that go unnoticed. A banana may be illuminated by light of various wavelengths, yet it will still look yellow (even when different colour receptors are stimulated; e.g. Webster & Mollon, 1995). A door may be viewed from various angles, yet it will still look rectangular (even when its retinal projection is trapezoidal; e.g. Pizlo, 1994). These examples illustrate *perceptual constancy*—the feature of the visual system that preserves perceived properties of an object across changes in sensory input. Perceptual constancy has been intensively studied (Walsh & Kulikovsky, 1998), but mainly for low-level visual features such as colour and shape.

In this final experiment, we test the possibility that perceptual constancy in high-level vision might compensate for distance-related changes in facial configuration. Perceptual constancy for face shape could account for the perceived stability of facial appearance across viewing distances in everyday life. In live viewing situations, the distance of a face can be estimated from a variety of depth cues including image size and height in picture plane. In principle, these depth cues could allow a constancy mechanism to correct for distance-related configural changes and deliver the perceptual stability that we experience. When viewing *photographs* of faces, the situation is very different. Depth cues such as size and height in picture plane are routinely altered in photographic presentations, so that they do not convey camera-to-subject distance as it was when the photo was taken. We reasoned that whenever such alterations have been made, a constancy mechanism that is based on these cues will be maladaptive. Instead of stabilising perceived face shape, its effect will be to inflate apparent differences.

To test perceptual constancy for face shape, we manipulated congruency between the actual camera-to-subject distance at the time of image capture, and the distance implied by depth cues at the time of presentation. For *Congruent* trials, on-screen size was consistent with the original camera-to-subject distance, so that the *Near* image was larger than the *Far* image (in the correct proportion). For *Incongruent* trials, on-screen size conflicted with camera-to-subject distance, so that the *Far* image was larger than the *Near* image.

Our hypothesis demands a highly specific pattern of results: For *Same Identity* pairs, participants should perform more accurately on *Congruent* trials than on *Incongruent* trials. Valid distance cues should allow a constancy mechanism to compensate for viewing distance normally, such that *Near* and *Far* images look more similar than they would otherwise. For *Different Identity* pairs, participants should perform more accurately on *Incongruent* trials than on *Congruent* trials. Image differences due to the change in identity will be augmented by perceptual differences caused by compensating for distance in the wrong direction. If there is no such compensation for distance, then our manipulation of distance cues should not affect performance.

Method

Participants

Thirty volunteers at the University of York (24 female, 6 male; mean age = 19.2) participated in exchange for payment or course credit. All participants were students who had enrolled after our photographic models had graduated, and were thus unfamiliar with the faces in the stimulus set (verified post-test). As with Study 2, all participants reported normal or corrected-to-normal vision.

Design

The manipulation of identity (*Same Identity*, *Different Identity*) was exactly the same as in

Study 2. However, the previous manipulation of camera-to-subject distance was replaced with a manipulation of distance cue congruency as follows. Each face pair consisted of a *Near* photo and a *Far* photo, one presented at a large size (700 pixels high x 525 pixels wide) and the other presented at a small size (100 pixels high x 75 pixels wide). In *Congruent* pairs (50%), the *Near* image was large and the *Far* image was small, consistent with the normal optical situation. To preserve the veridical relationship between camera-to-subject distance and image size, we simply cropped the faces from the original photographs and presented them without rescaling. In *Incongruent* pairs (50%), the *Far* photo was large and the *Near* image was small—a reversal of the normal optical situation. We implemented this reversal by simply exchanging the original heights of the *Near* and *Far* face images while preserving their aspect ratios. To strengthen the impression of distance, we added three further depth cues to the displays.

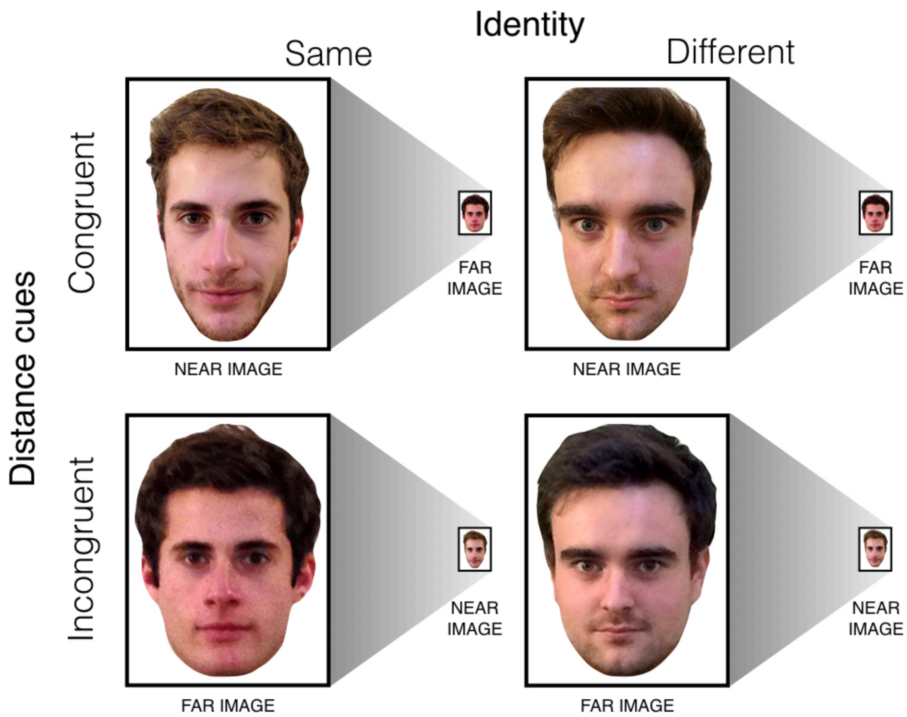

Figure 5. Summary of the experimental conditions in Study 3 with example stimuli. The left column shows *Same Identity* pairs and the right column shows *Different Identity* pairs. The top row shows *Congruent* distance cues and the bottom row shows *Incongruent* distance cues. Image labels are for exposition only and were not presented in the experiment.

The small image was presented higher in the picture plane than the large image, the two images were connected with perspective lines, and the area between the perspective lines was shaded with a receding contrast gradient. Example stimuli can be seen in Figure 5. Combining the 4 photos of each of the 18 models in *Same Identity* and *Different Identity* pairs using *Congruent* and *Incongruent* distance cues resulted in a total of 144 face pairs.

Procedure

The testing procedure was the same as in Study 2, except that all of the new participants were unfamiliar with the faces concerned, and the new stimulus displays incorporated congruent or incongruent distance cues as described in the preceding section. For each face pair, participants were asked to decide whether the two face images showed the same person or two different people. Participants viewed all 144 face pairs in a random order.

Results & Discussion

As expected, none of the participants reported being familiar with any of the stimulus faces in the familiarity check. Participants' accuracy scores were submitted to a 2 x 2 within-subjects ANOVA with the factors of *Distance* (*Congruent*, *Incongruent*) and *Identity* (*Same*, *Different*). Results of this analysis are summarised in Figure 6.

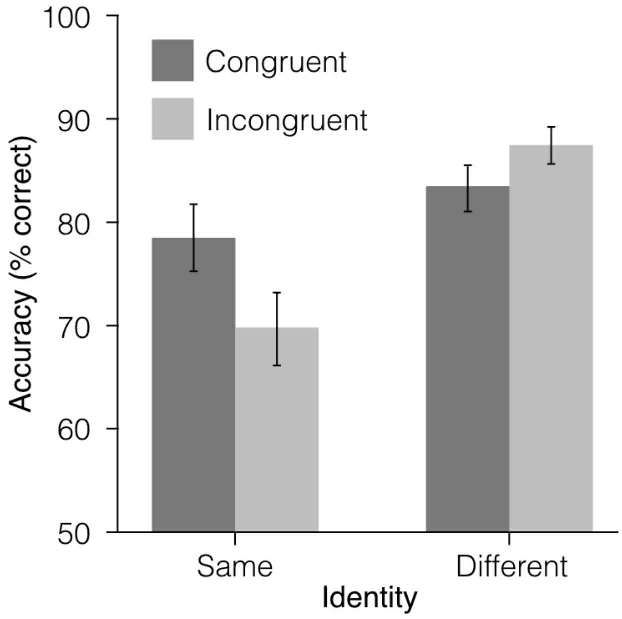

Figure 6. Mean accuracy scores for the conditions in Study 3. Dotted lines show *Familiar* viewers, solid lines show *Unfamiliar* viewers. Error bars show SEM.

The analysis revealed no significant main effect of *Distance Cues*. Overall accuracy was comparable for *Congruent* trials (M = 80.82%) and *Incongruent* trials (M = 78.49%) [$F(1,27) = 3.29$, $p = .08$, $\eta^2 = .11$]. However, there was a significant main effect of *Identity*, with higher accuracy for *Different* trials (M = 85.35%) than for *Same* trials (M = 73.97%) [$F(1,27) = 8.13$, $p = .008$, $\eta^2 = .23$]. More importantly, there was a highly reliable interaction between *Distance Cues* and *Identity* [$F(1,27) = 23.22$, $p < .001$, $\eta^2 = .46$], indicating that the distance manipulation had differential effects for *Same Identity* and *Different Identity* trials.

Simple main effects confirmed that for *Same Identity* pairs, participants were significantly more accurate when distance cues were *Congruent* (M = 78.38%, SE = 3.3) than when distance cues were *Incongruent* (M = 69.57%, SE = 3.6), [$F(1,54) = 22.44$, $p <.001$, $\eta^2 = .29$]. For *Different Identity* trials, there was a smaller effect in the opposite direction: participants were more accurate when distance cues were *Incongruent* (M = 87.42%, SE = 1.8) than when they were *Congruent* (M = 83.27%, SE = 2.18), [$F(1,54) = 4.97$, $p < .01$, $\eta^2 = .08$].

This pattern of results is exactly as one would expect if perceptual constancy compensates for distance-related changes in face shape. Given two photos of the same person, taken at different distances, observers found it easier to see that they showed the same person when congruent distance cues were present (78.4%) than when incongruent distance cues were present (69.6%). When the two photos showed different people, the effect was smaller and in the opposite direction: observers found it easier to tell the faces apart when incongruent distance cues were used. In both situations, the congruency of the distance cues affected the perceptual similarity of the face images. All of these findings can be explained by supposing that the perceptual system compensates for effects of viewing distance on face shape.

General Discussion

Many previous studies of face perception have concluded that configural information—specifically the pattern of distances between facial features—plays an important role in face identification (Maurer, Le Grand & Mondloch, 2002; Tanaka & Gordon, 2011). At the same time, principles of optics dictate that configural information changes with viewing distance (Harper & Latto, 2001). Here we sought to bring these two ideas together by asking whether changes in camera-to-subject distance affect identification accuracy.

In Study 1, we quantified the physical problem by comparing feature-to-feature measurements in photos of the same identities taken at two different distances. The analysis confirmed that our change in camera-to-subject distance affected some feature-to-feature measurements more than others. That is, it altered the configural information in the image. If face photographs are to be used in identification, this physical variability matters. Although our specific camera distances (0.32 m versus 2.70 m) were informed by Harper and Latto's (2001) work, they also correspond to the range typically used for passport photographs (Verhoff et al., 2008), as well as selfies (Lee & Kosasih, 2015) and personal photos (De Vignemont & Iannetti., 2015; see Hayduk, 1983). Passport photos are particularly interesting in this regard, as they are so frequently used in establishing identity (White et al., 2014). US Department of State guidance specifies a lower limit on camera-to-subject distance of 4 feet (1.25 m), but no upper limit. Australian Passport Office guidance similarly specifies a lower limit only (1.20 m). EU and UK guidance do not mention camera-to-subject distance at all, but do specify the size of the head in the frame. Each of these specifications allows some distance-related variability in configural information.

As Study 2 shows, distance-related changes in facial configuration can alter perceived identity. For unfamiliar viewers, photos of the same face taken at different distances were often seen as different people. Indeed, the distance manipulation incurred a 28% accuracy cost for same-identity trials (from 97% to 69%, in a task where perfect performance is 100% and chance performance is 50%). Very often, *Same Person, Different Distance* images were too dissimilar to be perceived as the same identity. This is perhaps surprising, given the overall similarity of the images in terms of ambient lighting, viewpoint, expression, and capture device (the high accuracy on *Same Person, Same Distance* trials attests to this general image similarity). However, the failure to integrate *Same Person, Different Distance* images accords with previous demonstrations that viewers underestimate within-person variability in facial appearance (Jenkins et al., 2011)—in this case, within-person variability in spatial layout.

Interestingly, our distance manipulation also affected *Different Person* trials. Photos of different faces were more likely to be seen as different people when they were taken at different distances, though this effect was relatively small (8% accuracy). Evidently, image differences due to a change in identity and image differences due to a change in viewing distance were to some extent additive in driving identity judgements.

Both sides of this interaction reinforce the view that unfamiliar face matching is tied to specific images (Burton & Jenkins, 2011; Hancock, Bruce, & Burton, 2000). Similarities between comparison images promote same-person judgements; differences between comparison images promote different-person judgements. Familiar face matching was clearly much less tied to image similarity. Viewers who knew the faces in this task performed almost perfectly (>96% in all conditions), and were virtually unaffected by a change in camera-to-subject distance (<2% accuracy cost). To put it another way, familiar face matching was virtually unaffected by configural differences between comparison images. This finding is difficult to reconcile with a configural processing account of face recognition (see Burton et al., 2015; Maurer, Le Grand, & Mondloch, 2002). However, it is easy to reconcile with a statistical account. We contend that becoming familiar with a person's face involves learning the ways in which that face can vary, and that the dimensions of variation include distance-related changes in configuration.

Study 3 demonstrates that the visual system compensates for such changes based on available distance cues. Accuracy on *Same Person* trials was nearly 10% higher when distance cues were *Congruent* with camera-to-subject distance (78.4%) than when they were *Incongruent* (69.6%; or absent in Study 2; 69.3%), indicating perceptual constancy for face shape. This constancy effect resolves the otherwise puzzling discrepancy between objective variability at the optical level and subjective stability at the perceptual level. It also reveals an unexpected role for distance cues in an identification task.

As well as their theoretical interest, our findings have a number of implications for applied face recognition. First, facial anthropometry is not a reliable method of identification. It has already been shown that distances between features do not reliably discriminate between different faces (Burton et al., 2015; Kleinberg, Vanezis & Burton, 2007; Towler, White, & Kemp, 2015). The current studies confirm that small changes in camera-to-subject distance reliably alter feature-to-feature measurements in the image, even when expression and pose are held constant. Second, identification photographs could be better standardised by specifying lower and upper bounds for camera-to-subject distance. Internationally-agreed standards for passport photographs would reduce the range of distance-related variability in facial appearance, thus reducing one source of identification error. Third, when camera-to-subject distance differs between photos, it may be beneficial to preserve available distance cues. Perceptual constancy for configuration appears to support unfamiliar face matching. But this mechanism can not compensate for viewing distance unless viewing distance can be estimated. In the current study we combined four distance cues—image size, height in the picture plane, perspective lines, and shading. A practical goal for future research would be to establish which distance cues are necessary and sufficient to optimise performance, and the range of conditions over which they are effective.

For now we show that facial configuration in photographs varies with camera-to-subject distance. These changes in configuration impair identification of unfamiliar faces, but not familiar faces. Valid distance cues mitigate the performance cost, suggesting that perceptual constancy compensates for distance-related changes in face shape. Understanding these distance effects could reduce identification errors in applied settings.

References

Andrews, T. J., & Ewbank, M. P. (2004). Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *NeuroImage*, *23*, 905–913.

Bindemann, M., Burton, A. M., Leuthold, H., & Schweinberger, S. R. (2008). Brain potential correlates of face recognition: Geometric distortions and the N250r brain response to stimulus repetitions. *Psychophysiology*, *45*, 535–544.

Bruce, V. (1982). Changing faces: visual and non-visual coding in face recognition. *British Journal of Psychology*, *73*, 105–116.

Bruce, V., Burton, A. M., Carson, D., Hanna, E., & Mason, O. (1994). Repetition priming of face recognition. *Attention and performance*, *15*, 179–201.

Bruce, V., & Young, A. W. (1986). Understanding face recognition. *British Journal of psychology*, *77*(3), 305–327. doi: 10.1111/j.2044-8295.1986.tb02199.x

Bryan, R., Perona, P. & Adolphs, R. (2012). Perspective distortions from interpersonal distance is an implicit visual cue for social judgments of faces. *PLoS ONE 7*(9): e45301.

Burton, A. M., & Jenkins, R. (2011). Unfamiliar face perception. In *Oxford Handbook of Face Perception*, eds. A. J. Calder, G. Rhodes, M. Johnson, J. Haxby, pp. 287–386. Oxford, UK: Oxford University Press.

Burton, A. M., Schweinberger, S. R., Jenkins, R., & Kaufmann, J. M. (2015). Arguments against a configural processing account of familiar face recognition. *Perspectives on Psychological Science*, *10*(4), 482–496.

Burton, A. M., White, D., & McNeill, A. (2010). The Glasgow face matching test. *Behavior Research Methods*, *42*, 286–291.

Burton, A. M., Wilson, S., Cowan, M., Bruce, V. (1999) Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science, 10*, 243–248.

De Jong, M., Wagenaar, W. A., Wolters, G., & Verstijnen, I. M. (2005). Familiar face recognition as a fullnction of distance and illumination: A practical tool for use in the courtroom. *Psychology, Crime & Law, 11*, 87–97.

De Vignemont, F., & Iannetti, G. D. (2015). How many peripersonal spaces?. *Neuropsychologia*, *70*, 327-334.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115(2)*, 107-117.

Greene, E., & Fraser, S. C. (2002). Observation distance and recognition of photographs of celebrities' faces. *Perceptual & Motor Skills, 95*, 637–651.

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, *24*, 187–203.

Hancock, P. J., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences*, *4*, 330–337.

Hahn, C. A., O'Toole, A. J., & Phillips, P. J. (2016). Dissecting the time course of person recognition in natural viewing environments. *British Journal of Psychology, 107(1),* 117-134.

Harper, B. & Latto, R. (2001). Cyclopean vision, size estimation, and presence in orthostereoscopic images. *Presence-Teleoperators and Virtual Environments, 10(3),* 312-330. doi: 10.1162/105474601300343630

Hayduk, L. A. (1983). Personal space: Where we now stand. *Psychological Bulletin*, *94*, 293–335.

Hill, H., & Bruce, V. (1996). The effects of lighting on the perception of facial surfaces. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 986–1004.

Jenkins, R., & Burton, A. M. (2011). Stable face representations. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1571), 1671–1683.

Jenkins, R., & Kerr, C. (2013). Identifiable images of bystanders extracted from corneal reflections. *PLoS ONE, 8*(12), e83325.

Jenkins, R., McLachlan, J. L., & Renaud, K. (2014). Facelock: familiarity-based graphical authentication. *PeerJ, 2*, e444.

Jenkins, R., White, D., Van Montfort, X., & Burton, A.M. (2011). Variability in photos of the same face. *Cognition*, *121*, 313–23.

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

Kleinberg, K. F., Vanezis, P., & Burton, A. M. (2007). Failure of anthropometry as a facial identification technique using high-quality photographs. *Journal of Forensic Sciences*, *52*, 779–783.

Lampinen, J. M., Erickson, W. B., Moore, K. N., & Hittson, A. (2014). Effects of distance on face recognition: implications for eyewitness identification. *Psychonomic Bulletin & Review, 21(6)*, 1489-1494.

Latto, R., & Harper, B. (2007). The non-realistic nature of photography: Further reasons why turner was wrong. *Leonardo*, *40*, 243–247.

Leder, H. & Bruce, V. (2000). When inverted faces are recognized: the role of configural face information in face recognition. *The Quarterly Journal of Experimental Psychology.Section A. 53,* 513–536.

Leder, H., & Carbon, C.C. (2006). Face-specific configural processing of relational information. *British Journal of Psychology, 97(1)*, 19–29.

Lee, T. P., & Kosasih, W. (2015). Exploring the Distance, Location, and Style of "Selfies" with the Self-Pano Mobile Application. In *New Contributions in Information Systems and Technologies* (pp. 875–884). Springer International Publishing.

Lindsay, R. C. L., Semmler, C., Weber, N., Brewer, N., & Lindsay, M. R. (2008). How variations in distance affect eyewitness reports and identification accuracy. *Law and Human Behavior*, *32*, 526–535.

Liu, C.H. (2003). Is face recognition in pictures affected by the center of projection? *IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, (Los Alamitos, CA: IEEE Computer Society), 53–59.

Loftus, G. R., & Harley, E. M. (2005). Why is it easier to identify someone close than far away? *Psychonomic Bulletin and Review, 12*, 43–65.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*, 255–260.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*, 11087–11092.

Pirenne, M. H. (1970) Optics, Painting, & Photography. *New York*: Cambridge University Press.

Pizlo, Z. (1994). A theory of shape constancy based on perspective invariants. *Vision Research, 34*, 1637–1658.

Smith, W. A. (2016). The perspective face shape ambiguity. In *Perspectives in Shape Analysis* (pp. 299–319). Springer International Publishing.

Tanaka, J. W., & Gordon, I. (2011). Features, configuration, and holistic face processing. *The Oxford Handbook of Face Perception*, 177–194.

Towler, A., White, D., & Kemp, R. I. (2014). Evaluating training methods for facial image comparison: The face shape strategy does not work. *Perception, 43*, 214–218.

Troje, N. F., & Bülthoff, H. H. (1996). Face recognition under varying poses: The role of texture andnd shape. *Vision Research*, *36*, 1761–1771.

Verhoff, M. A., Witzel, C., Kreutz, K., & Ramsthaler, F. (2008). The ideal subject distance for passport pictures. *Forensic Science International*, *178*, 153–156.

Wagenaar, W. A., & van der Schrier, J. H. (1996). Face recognition as a function of distance and illumination: A practical tool for use in the courtroom. *Psychology, Crime & Law, 2*, 321–332.

Walsh, V., & Kulikowski, J. (1998). *Perceptual constancy: Why things look as they do*. Cambridge University Press.

Webster, M. & Mollon, J. (1995). Colour constancy influenced by contrast adaptation. *Nature, 373*, 694–698.

White, D., Kemp, R. I., Jenkins, R., Matheson, M., & Burton, A. M. (2014). Passport officers' errors in face matching. *PLoS ONE*, *9*(8), e103510.

Young, A. W., McWeeny, K. H., Hay, D. C., & Ellis, A. W. (1986). Matching familiar and unfamiliar faces on identity and expression. *Psychological Research*, *48*, 63–68.

Zebrowitz L. A. (2011). Ecological and social approaches to face perception. In *Oxford Handbook of Face Perception*, eds. A. J. Calder, G. Rhodes, M. Johnson, J. Haxby, pp. 31–50. Oxford, UK: Oxford University Press.