This book chapter is published in Forensic Facial Identification: Theory and Practice of Identification from Eyewitness, Composites and CCTV.

Full article available here:

https://onlinelibrary.wiley.com/doi/10.1002/9781118469538.ch9

Davis, J. P., & Valentine, T. (2015). Human verification of identity from photographic images. In T. Valentine and J. P. Davis (Eds.), Forensic Facial Identification: Theory and Practice of Identification from Eyewitnesses, Composites and CCTV (pp. 211-238). Chichester: WileyBlackwell. DOI: 10.1002/9781118469538.ch9

This is a pre-print version of the paper

Any questions please e-mail j.p.davis@gre.ac.uk

Chapter 09: Human Verification of Identity from Photographic Images

Josh P Davis[1] and Tim Valentine[2]

[1]University of Greenwich

[2]Goldsmiths, University of London

Case Study: Jean Charles de Menezes (1978 – 22nd July 2005)

On 7th July 2005, four suicide bombers detonated devices on three London underground trains and a bus murdering 56 people and injuring many more. Two weeks later, four more explosive devices failed to explode on the underground. The four perpetrators went into hiding. One – Osman Hussain - had left a gym photograph membership card in his bag containing one of the unexploded bombs. With security forces on high alert, the police retrieved this card, and a few hours later staked out a South London block of flats – the address registered by the gym card owner. A Brazilian, Jean Charles de Menezes who lived in this building, left the communal entrance the following morning. A police surveillance officer who had seen a copy of the gym card, and a poor-quality CCTV image of Osman, watched de Menezes leave and suggested by radio he deserved 'another look'. For the following hour, other officers briefly viewed de Menezes as he travelled on buses to his workplace. His entirely innocent behaviour was deemed suspicious as it inadvertently replicated tactics for evading surveillance, adding to the growing belief that de Menezes was one of the bombers. de Menezes arrived at Stockwell underground station at about the same time as a firearms team. From operational command information, they were convinced de Menezes was carrying a primed explosive device, and he was shot at very close range shortly after he boarded a train.

Several official investigations were held into the wrongful killing of this innocent man - the Metropolitan Police Service were criticised for their operations, communication and surveillance procedures. On 1st October 2007, a criminal prosecution against the force was bought under the Health and Safety at Work Act (1974). The force pleaded 'not guilty', but for "failing to provide for the health, safety, and welfare of Jean Charles de Menezes" it was fined £175,000, with £385,000 legal costs on 1st November 2007 (BBC News, 2007).

Evidence provided by the defence had included presenting to the jury 'chimeric' composite images - created by aligning the left side of Osman's face with the right side of de Menezes' face (Figure 1). The aim was to demonstrate that the two men were highly similar in appearance and that the original errors of misidentification were partly a consequence of this similarity. An expert witness for the prosecution argued that the images had been manipulated so that contrast and colour matched, and had also been stretched in some places to align them – reducing the salience of differences between the two faces (note that the images in Figure 1 have not been manipulated in this manner). There is also evidence that chimeric display will increase the similarity of any two pairs of faces, regardless of manipulations (Strathie, McNeill, & White, 2012). Faces are primarily processed in a holistic or whole face manner (e.g., Tanaka & Farah, 1993; Tanaka & Sengco, 1997), and this technique increases the likelihood that a viewer will view the two halves as a Gestalt or as a single individual. Strathie *et al*. argue that employing this procedure in court should therefore be avoided as it may mislead the jury.

Figure 1 about here

Other research reported in this chapter also informs as to why errors may have been made in this operation. The inquest reports do not make clear whether any of the surveillance teams attempted to simultaneously match de Menezes with Osman's gym card or CCTV image. However, unfamiliar face matching from even high quality photographs to live actors can be error prone (Davis & Valentine, 2009), particularly with passport sized images (Kemp, Towell, & Pike, 1997). With lower quality images taken from a distance, which if digital may be pixelated, face matching performance will be worse (Bindemann, Attard, Leach, & Johnston, 2013). This may have been relevant as even though higher quality images were available, the police were criticised for supplying the surveillance teams with poor quality CCTV images of Osman. Of course, to avoid attracting the attention of de Menezes, the team may have kept the images concealed and made their identification decisions based on memory, further reducing the likelihood of an accurate judgement. Nevertheless, the first view of the two images may have biased all subsequent decisions anyway, particularly if the gym card and CCTV images of Osman appeared dissimilar, as is surprisingly common when viewing images of the same unfamiliar person from different sources (Jenkins, White, van Montfort, & Burton, 2011). Information that a suspect may have changed appearance can induce a liberal response bias in both matching (Davis & Valentine, 2009) and recognition tasks (Charman & Wells, 2007), and being told that two dissimilar images depict the same person may have been interpreted as an implicit instruction that de Menezes' appearance had changed, and to therefore be more 'flexible' when making their identification judgements. Furthermore, there are no details of the face recognition abilities of the surveillance team which may have influenced all the tasks listed above. Individuals with average ability are likely to make more identification errors than so called super-recognisers (Davis, Lander, Evans, & Jansari, in preparation; Davis, Jansari & Lander, 2013).

Finally, although not central to the current chapter, all human decision-making is susceptible to error. The first police radio call of de Menezes needing 'another look' appears to have initiated a *confirmation bias* for the operations team to interpret subsequent information as evidence of guilt (see Cole & Thompson, 2013). If so, evidence pointing towards de Menezes' innocence will have been ignored or misinterpreted, whereas information pointing towards his guilt will have been assimilated into a model of escalating guilt. Indeed, by the time of the shooting, operations team transcripts suggest there was confidence in the belief that de Menezes was a terrorist, and that he had been visually identified as Osman by at least one of his police pursuers.

It has long been known that human memory is fallible (e.g., Borchard, 1932; Munsterberg, 1908), and that this can create legal controversies when a police investigation relies on eyewitness testimony. Evidence from exoneration cases, as well as the vast body of empirical research reported in this volume, has shown that witnesses often make highly confident, but mistaken identifications of someone they wrongly believe to be the perpetrator of a crime. On the other hand, witnesses also regularly fail to identify the true offender from an identity parade or lineup. Recognition can be adversely influenced by both system variables (see Chapter 6) and estimator variables (see Chapter 7) and for many years, these identification errors were attributed purely to the fragility of human memory. However, there are many important forensic and security situations for which no memory is required, and this chapter describes the large body of relatively recent research demonstrating that there is a similar risk of identification error with these tasks. Examples include:

1. Border control or other security check points where officials are required to verify the identity of individuals from photographs on documents such as passports.
2. Closed Circuit Television (CCTV) cameras operators may be required to match a face on a ‘wanted’ list with an individual under surveillance.
3. Police officers may simultaneously compare CCTV or other photographic images collected from the scene of the crime, with images of known suspects in order to gather evidence to secure a prosecution.
4. In some cases, the identification of victims of crime from photographic images may be as important as identifying the culprits. One such highly distressing scenario is when the police attempt to identify minors depicted in footage of

pornographic events displayed on the internet. For this, officers will similarly use comparison photographs to match with the evidential images.

An important distinction is however made in the law, and is supported by empirical research. Viewers familiar with those depicted are, in most circumstances, highly accurate at identity verification tasks (e.g., Bruce, Henderson, Newman, & Burton, 2001; Burton, Wilson, Cowan, & Bruce, 1999). In contrast, unfamiliar face identification is error prone, even when the task requires no memory (e.g., Bruce, Henderson, Greenwood, Hancock, Burton, & Miller, 1999; Burton, White, & McNeill, 2010; Burton *et al*., 1999; Davis & Valentine, 2009; Henderson, Bruce, & Burton, 2001; Kemp, Towell, & Pike, 1997; Megreya & Burton, 2006; 2008). Furthermore, as with eyewitness identification, many of the variables described elsewhere in this volume, negatively influence these procedures. Of course, familiarity can vary from a fleeting encounter to knowing someone for many years, and differences in the level of familiarity will influence identification accuracy. Furthermore, there are large individual differences in the ability to learn new faces. Nevertheless, it is perhaps rather ironic that the officials tasked with determining identity in the scenarios listed above, are extremely unlikely to have ever previously encountered those who are under scrutiny.

The use of photographic images in court

The United Kingdom has led the world in the widespread installation of CCTV in most city and town centres. There are estimated to be between 1.85 (Thompson & Gerrard, 2011) and 5.9 million CCTV cameras (Barrett, 2013) in the UK. These numbers may be dwarfed by alternative sources of photographic evidence including mobile camera phones, social media websites, police mugshots, as well as GPS-enabled smart video cameras carried by cyclists and other vehicles such as police cars. Some offenders, particularly those in youth gangs, upload images of their activities onto websites, as a ‘badge of honour.’ As a consequence of government support for the proliferation of CCTV, many of the principles for the use of such images in legal proceedings were first applied in the UK, and therefore the primary English legislation will be described in the following section (see Chapters 10 & 12 for further details).

Photographs have been admissible in court since 1864 (*R v. Tolson,* 1864; cited in Murphy, 1999), and CCTV footage itself was first admitted in 1982 to provide information about a theft from a retail store (*R v. Fowden and White*, 1982). The use of such evidence for identification purposes in England and Wales was recently summarised in an Attorney General's Reference (2003), with an important distinction between evidence presented by witnesses who are familiar with the defendant, and evidence provided by those who did not know the defendant prior to the police investigation.

Firstly, if identifications have been made after viewing a video or photograph by individuals previously familiar with a defendant, they may give evidence as a witness for the case, even if the footage is no longer available (e.g., *R v. Caldwell and Dixon*, 1993; *R v. Grimer*, 1982; *Taylor v The Chief Constable of Cheshire*, 1987). In these circumstances, identifications have the same status as those from eyewitnesses present at the incident, and witnesses can therefore be cross-examined in court. In many cases, witnesses will be police officers and to protect them, or indeed any witness, from accusations of collusion the Police and Criminal Evidence Act (1984, Codes of Practice, 2011, Code D) additionally prescribes the procedures by which images should be distributed for eliciting reliable identifications. For instance, police intelligence may lead an investigating team to believe an officer may have encountered the suspect previously. They may wish the officer to view crime scene footage to confirm the suspect's identity. Such identifications should be spontaneous and independently acquired - witnesses should not be prompted. Therefore, witnesses are required to provide details as to how they were previously familiar with the suspect, and the circumstances by which they first viewed the images. A failure to adhere to these codes of practice has resulted in the courts rejecting improperly collected identification evidence (see for instance Fort, 2013).

Similar principles designed to protect against evidence contamination operate in Australia. In *Strauss vs. the Police* (2013), an eyewitness who was intoxicated at the time of the crime, and who may not have had a good view of the perpetrator was encouraged by an acquaintance a few hours later to view images from a social website – *Facebook* - in order to identify the eventual defendant. This identification formed a central component of the prosecution's case. The conviction was overturned on appeal. The court ruled that the identification procedure was highly suggestive and therefore unsafe, as the defendant's name had been provided to the witness by the same acquaintance prior to viewing the website images.

The remaining three circumstances in which photographic identification evidence may be admitted in court in England and Wales concerns witnesses who were not previously familiar with the defendant (Attorney General's Reference, 2003):

First, expert practitioners may apply *facial mapping* or *facial comparison* techniques to provide evidence as to whether an individual captured in photographic evidence is the defendant depicted in a confirmed photograph (e.g., *R v. Clarke*, 1995; Davis, Valentine, & Wilkinson, 2012). The generally poor reliability of these techniques is discussed in Chapter 10. When presented in court with professional gravitas, alongside circumstantial evidence, evidence from a facial comparison expert can be highly compelling.

Second, evidence may be admissible if a witness, not previously familiar with the defendant spends "substantial time viewing and analysing photographic images from the scene" (Attorney General's Reference, 2003), thus familiarizing themselves with the accused and gaining a *special ad-hoc* expertise. For instance, in *R v. Clare and Peach* (1995) a police officer viewed black-and-white CCTV footage of a football crowd riot more than 40 times, examining stills and evaluating details in slow motion. He compared this footage with undisputed photographs of the defendants taken the same day. There is evidence that matching performance is partly diagnostic of familiarity, as increased exposure can increase accuracy at this type of face matching task (Clutterbuck & Johnston, 2002), although not necessarily with low-quality images (Davis, 2007; Lee, Wilkinson, Memon, & Houston, 2009). However, even forensic scientists can be susceptible to 'cognitive biases', or tunnel vision encouraging faulty decision-making, so that prior case information can inadvertently influence judgements (Cole & Thompson, 2013). It is possible that individuals repetitively viewing poor-quality photographic images will be vulnerable to such biases by ignoring 'exculpatory' visual cues in some frames that do not match the appearance of the defendant, while giving greater weight to alternative frames as there appears to be a greater similarity of appearance between the defendant and the individual depicted.

Third, "where the photographic image is sufficiently clear, the jury can compare it with the defendant sitting in the dock" (Attorney General's Reference, 2003). The prosecution may explicitly encourage a jury to provide a verdict based on their perception that the perpetrator on video is the defendant. For instance, in *R v. Dodson and Williams* (1984), the jury were shown CCTV stills from a bank raid and invited to compare them with the two defendants. No alternative corroborating identification evidence was submitted, although the court was also presented with undisputed photographs of one of the accused, contemporaneous to the time of the crime – again, to demonstrate a similarity of appearance. Although this judgement applies to cases in which a jury is *directed* to compare the evidential images to the defendant, it is possible that regardless of instructions, an individual juror may base their verdict on the similarity of appearance between the defendant and crime scene images. Indeed, anyone involved in a police investigation or in court may make such identification judgements. Jurors will be previously unfamiliar with the defendant and may be susceptible to identity judgement errors. Unfortunately, face matching accuracy is not improved, even after warnings of the high risk of error (Thompson, Dunkelberger, & Vescio, 2013).

These legal principles will not necessarily apply in different jurisdictions. Indeed, in Australia, evidence from police officers who claim to recognise the defendant is not normally admissible (*Smith v. The Queen,* 2001). Instead, juries are encouraged to make their own identity decisions from viewing images and directly comparing it to the defendant.

The identification of suspects from CCTV

CCTV operators often view and record criminal acts as they occur (see Stainer, Scott-Brown, & Tatler, 2013, and Troscianko, Holmes, & Stillman, 2004, for research on the surveillance strategies of operators). They may concurrently report what they observe to the police in order to direct them to the crime scene. If the suspect is apprehended, the operator acting as an eyewitness could in court theoretically provide a commentary as to what is depicted on crime scene footage. However, in most cases, criminal activity is not directly observed, and therefore images will be collected for retrospective examination in a police station. The police may 'track' suspects if they move in and out of footage captured by different cameras, in order to obtain the best possible images. This is often achieved by focussing on clothing, and sometimes officers capture incidences in which a suspect discards their outer clothing or a disguise in an attempt to evade detection.

Familiar face recognition

The manner in which images are distributed and publicised can have an important influence on the likelihood of identification and subsequent detection. Images are regularly distributed internally by police forces and displayed on 'wanted' websites. In serious cases they may be depicted in the media as part of an information gathering appeal. The aim is that they will be recognised by someone familiar. In London, if a suspect is located and shown CCTV imagery over 70% confess in police interview (Davis, Jansari, & Lander, 2013). Others may admit to being at the scene of the crime, but to having no involvement. In cases in which the suspect disputes the identification entirely, the same witnesses (police or member of the public) who identified that suspect may be asked to provide evidence of identity in court. In some cases, the defence may successfully argue that the quality of the images is inadequate for a reliable identification to be made, and/or that the witness, who is often a police officer, does not possess sufficient prior familiarity with the defendant to make an objective identification.

The identifications made in the scenarios above are by someone familiar with the suspect and most research has found that identifications of individuals well-known to the observer are normally highly reliable, even when image quality is poor (Bruce *et al*., 2001; Burton *et al*., 1999). Nevertheless, familiar face recognition performance is rarely 100%, even by those with the best face recognition ability (Davis *et al*., 2013). Readers will be aware that photographs of even the closest family members will occasionally fail to ‘capture their likeness’. Indeed, consequential errors of familiar person identification do occur, even with high-quality images. Close family members of a missing person all mistakenly identified a man depicted in airport CCTV footage as their relative (BBC News, 2003). There may be confounding explanations. Nevertheless, this case does illustrate that recognition of even highly familiar people in high quality photographs is not infallible.

Unfamiliar face matching

Despite the occasional identification difficulties we may encounter when viewing images of those we know from an unusual perspective, it might be thought that when images depict an offender in high quality images, identity verification should be relatively easy, regardless of familiarity. However, a large body of recent research employing a variety of different facial databases and experimental designs has demonstrated that the identification of unfamiliar people is often unreliable, even when there are no memory demands and the high-quality images to match are taken at approximately the same time with no attempt to change appearance (e.g., Bindemann, Avetisyan, & Blackwell, 2010; Bruce *et al*., 1999; 2001; Burton *et al*., 1999; Burton *et al*., 2010; Davis & Valentine, 2009; Henderson *et al*., 2001; Megreya & Bindemann, 2009; Megreya, Bindemann, Havard, & Burton, 2012; Megreya & Burton, 2006, 2007, 2008; Özbek & Bindemann, 2011; Strathie, McNeill, & White, 2012). Error rates inevitably increase when the images to be matched are taken some time apart (Davis & Valentine, 2009; Megreya, Sandford, & Burton, 2013). Two studies directly compared familiar and unfamiliar face identification using poor-quality CCTV images. Burton *et al.* (1999; Experiment 1) displayed a series of video stills of university lecturers. In a subsequent recognition task, students familiar with the lecturers, were far more accurate than those students and police officers who were unfamiliar with those depicted.

Similarly, Bruce *et al.* (2001; Experiment 1) employed a *face matching* task involving no memory, in which participants were presented with a series of pairs of facial images. One image was from a poor quality CCTV system, the other a good quality facial photograph. Participants were asked to decide if they depicted the same person or not. When both images were of the same individual (*match trials*), the correct identification rate (*hits)* of participants familiar with those depicted was approximately 93%. However, when targets were unfamiliar, the hit rate was only 76%. If the targets were presented with a distracter image (*mismatch trials*), accuracy at correctly identifying that different people were depicted was high (*correct rejections;* 91%), but only if the targets were familiar. When unfamiliar with those depicted, correct rejection rates were approximately 55%, not much higher than would be expected if guessing (i.e., 50%).

Pre-processing of images taken under different lighting conditions to reduce discrepancies in their illumination can reduce face matching error rates (Liu, Chen, Han, & Shan, 2013). Furthermore, matching judgements, particularly to poor quality images may be improved when the whole body is available (Rice, Phillips, and O'Toole, 2013). However, unfamiliar face matching errors from facial photographs may be common even when the images to be matched are high quality, close-up and taken in optimum conditions. Bruce *et al.* (1999) designed a task requiring the matching of one of a series of 40 male frontal, high quality facial video stills, with a simultaneously presented frontal, high quality facial photograph of the target person among an array of nine distracters (Figure 2). Error rates in target-present trials were 30%. Two-thirds of these errors involved failures to identify any array image. In the remainder, participants selected the wrong array face, even though the correct image was available. Photographs and videos were from a standardised distance and facial angle, under standardised lighting and taken on the same day, so that the appearance of the actors would not have substantially changed. Indeed, they also posed with neutral facial expressions. Similarly high false negative error rates were found in target-absent trials (30%), in which the target was replaced by an additional distracter. Furthermore, when the single target image was depicted from a 30^0 angle, error rates increased to 40% in target-present trials and 38% in target-absent trials. Even a change in the facial expression of the single target from neutral to smiling significantly increased the likelihood of error, with errors of 34% in target-present and 38% in target-absent trials. These reported values also reflect average performance. In one specific target-present trial with the same pose and expression in both target and array, 80% of participants failed to correctly match the target. These findings demonstrate that minor image format differences across even high quality images can be responsible for high error rates.

Figure 2 about here

Two different experiments by Henderson *et al.* (2001) illustrate that even with reduced task demands; face matching is still error prone. In the first (Experiment 4), participants were asked to identify which of two professional studio portrait photographs depicted a target actor shown in a close-up high quality television broadcast still. One was a picture of the target, one a distracter of similar appearance. Overall 24% of decisions were incorrect. However, in one trial, approximately one third of participants incorrectly responded that a still of a first actor was more similar in appearance to a second actor's photograph, than the first actor's own photograph. In a follow-up experiment (Experiment 5), in which participants were presented with a single pair of images, approximately 45% of participants wrongly reported that two images of the same person were of different people. A further 27.5% incorrectly matched the images of two different actors. Decision confidence ratings were collected in both of the above studies (Bruce *et al*., 1999; Henderson *et al.*, 2001). Confidence was often high, even when participants were incorrect. In addition, Papesh and Goldinger (2014) conducted a study designed to better replicate real security conditions in which the carriers of fraudulent identity documents will be rare. They demonstrated that the likelihood of detecting mismatched pairs of images was proportionally worse when mismatch trials were rare (one-in-ten trials), than when they were relatively common (one-in-two-trials).

Some security officials such as passport officers compare photograph identity documents to a person standing in front of them. Many millions of such decisions are made every day. Indeed, on 31st July 2011, 233,561 passengers passed through one airport in London alone (Heathrow Airport, 2013). Consistent with photograph-to-photograph matching, unfamiliar face matching errors are still common even when the target is physically present (Davis & Valentine, 2009; Kemp *et al*., 1997; Megreya & Burton, 2008). For instance, Kemp *et al.* conducted a study in which 'participant-shoppers' submitted credit cards displaying 2-cm^2 photographic images to experienced supermarket cashiers, who were told they would receive a cash bonus if errors in administering these trial cards were low. However, the cashiers failed to detect 64% of shoppers when they presented a photo identity card containing a facial photo of another person matched for facial appearance. When the distracter was simply of the same ethnicity and gender, errors were reduced, but still high at 34%. When the shoppers presented correct photographs of themselves, there was a low false positive rate of 7%. A possible explanation for this liberal acceptance criterion is that challenging too many legitimate shoppers would be embarrassing for a retail outlet. However, the small photographs may have not been sufficiently detailed to make judgements. Indeed, there is a positive relationship between image size and unfamiliar face recognition accuracy (Loftus & Harley, 2005).

Davis and Valentine (2009) also found high simultaneous matching error rates in both match and mismatch trials across a series of three experiments employing much larger higher-quality images in which approximately 1,200 participants were recruited. In these experiments, participants were required to match a target actor present in person against a good quality full-body or close-up facial moving video image displayed on a large display screen. In Experiment 1, videos were three weeks old and the error rate in match trials was 22%; in mismatch trials it was 17%. These results may partly be a consequence of the delay, as some of the eight actors had changed their hairstyle. However, in another experiment (Experiment 3), the close-up facial videos in one condition were taken 10 min beforehand, and yet 17% of participants failed to correctly respond that the video depicted the actor standing by the screen. Errors increased substantially to 33% with images taken a week beforehand. In a further experiment (Experiment 2), videos sometimes depicted the actors in disguise (no disguise, dark glasses, hat), and all videos were one-year-old to replicate a situation in which a defendant's court appearance may occur a long time after crime scene footage is captured. Some participants were informed of this delay, and providing such information biased responses, particularly when actors were in disguise, by increasing both false negative and false positive error rates. As with the photograph-to-photograph matching studies above, many highly confident participants made incorrect decisions, although overall, confidence was correlated with accuracy.

An important feature of the studies reported above was the use of relatively small databases to locate individuals who might be mistaken for one another. Each array in Bruce *et al.* (1999) was constructed by selecting faces from a database of 200 trainee police officers rated as similar in appearance to the target, whilst Henderson *et al.* (2001) "searched through several hundred actor-agency photographs" (*p.* 463), to locate appropriate distracters. Kemp *et al.* (2007) recruited 56 students from a university course to act as shoppers. Davis and Valentine (2009) selected eight rugby players from a student club with approximately 70 members. It therefore appears comparatively easy to construct experiments in which errors in identification matching occur. There are likely to be many more people in the population who could be the subject of similar mistaken identity decisions. Furthermore, passports are issued for 10 years, and the holder's appearance may substantially change in that time, making matching even more prone to error.

Identification errors rates are also more likely as a consequence of the typical positioning of CCTV cameras, often located above head height (Davies & Thasen, 2000; Thompson, Grattan, Rawding, & Buchholz, 2010). Differences in viewpoint, expression, size, and environmental lighting can reduce face matching and recognition accuracy (e.g., Bruce, 1982; Bruce *et al*., 1999; Bruce, Valentine, & Baddeley, 1987; Hill & Bruce, 1996; Hill, Schyns, & Akamatsu, 1997). A change in the distance from which images are taken, and in the focal length of a camera lens, can influence facial appearance (Harper & Latto, 2001).

From close-up, a face will appear 'long and narrow' with a greater distance between the eyes, whereas, from further away, it will appear 'rounder' with the eyes closer together.

Parallels between face recognition and face matching research

There is evidence that many of the variables that influence face recognition can be replicated in the absence of memory. Indeed, face recognition and eyewitness identification tasks correlate (e.g., Bindemann, Brown, Koyas, & Russ, 2012; Darling, Martin, Hellman, & Memon, 2009), and there are large consistent individual differences in face recognition and face matching ability (Bindemann *et al.,* 2012; Burton *et al*., 2010; Davis *et al*., 2013; Megreya & Burton, 2006; Russell, Duchaine, & Nakayama, 2009). In addition, familiar face recognition is primarily driven by the internal features of a face: - eyes, nose, mouth (Bonner & Burton, 2004; Ellis, Shepherd, & Davies, 1979; Young, Hay, McWeeny, Flude, & Ellis, 1985). In contrast, unfamiliar face processing is more influenced by external features: - hairstyle, face shape (Bruce *et al*., 1999, Experiment 4; Ellis *et al*., 1979; Young *et al*., 1985; although see Megreya & Bindemann, 2009). This partly explains why we still easily recognise close acquaintances following substantial changes to hairstyle, but minor changes to the appearance of someone less familiar may be far more disruptive.

To explore the role of internal and external features in simultaneous face matching, Bruce *et al.* (1999) presented high-quality images in three different conditions. In the *whole-face* condition no manipulation to images was conducted. In the *external feature* condition, a blank oval obscured the internal features; whereas, for the *internal feature* condition, the oval obscured the external features. Whole-face matching accuracy was highest. Accuracy was higher in the external feature condition than the internal feature condition. The authors suggest that participants may have primarily been employing a strategy based on matching hairstyle in the external feature condition, as this is the most salient feature. For this experiment, such a strategy would be successful as all images were taken on the same day – it is unlikely hairstyle would have significantly altered. However, it would be far less effective with images collected some time apart, or with the target in disguise. Interestingly, Megreya and Bindemann (2009) suggest that face matching performance may be influenced by cultural norms. In contrast to UK participants, those from Egypt demonstrated an advantage for matching own nationality unfamiliar faces using internal features. The authors suggest that such a strategy develops due to the common use of the headscarf hiding the external facial features in Middle Eastern countries.

Evidence from a study that recorded eye movements found that when comparing images taken at least a day apart, an internal feature processing style is advantageous particularly when viewing time is restricted (Fletcher, Butavicius, & Lee, 2008). Participants who naturally tended to focus more on internal features were more accurate at face matching. However, this effect was only found for images displayed for 2 seconds. Gaze direction had no influence on accuracy when images were depicted for 6 seconds, suggesting that the longer display time allowed for greater inspection. As would be expected given these results, disguises reduce face recognition and face matching accuracy (Davis & Valentine, 2009; Henderson *et al*., 2001), increasing the likelihood of both false positives and false negatives.

Another estimator variable that has been extensively replicated in eyewitness identification and face recognition studies is the cross-ethnicity effect. People tend to be better at recognising individuals from their own ethnicity than those from other ethnicities (see Chapter 7). It is likely that passport officers in particular will be required to judge the identity of people from many nationalities, and similar cross-ethnicity effects have been found in face matching tasks (Megreya, White, & Burton, 2011; Meissner, Susa, & Ross, 2013). Furthermore, in comparison to males, females demonstrate an advantage for matching same-gender faces (Megreya, Bindemann, & Havard, 2011), an effect also found in face recognition tasks (e.g., Wright & Sladden, 2003).

Another well-studied estimator variable that can influence face recognition is state anxiety. Participants stressed at the time of encoding a 'crime scene' are less likely to be able to identify the culprit later (e.g., Valentine & Mesout, 2009; for a meta-analysis see Deffenbacher, Bornstein, Penrod, & McGorty, 2004). Similar effects have been found in a face matching study (Attwood, Penton-Voak, Burton, & Munafo, 2013) in which acute state anxiety was induced in a laboratory by the administration of carbon dioxide. In this study, participants completed a face matching test twice – once while breathing normal air, and once while breathing carbon dioxide. State-induced anxiety reduced accuracy at matching two images of the same person, whilst having no effect on matching decisions to images depicting two different people. This replicates the negative effects of stress found in eyewitness research for target-present but not target-absent lineups (Deffenbacher *et al*., 2004).

This section has focussed on parallels between face recognition and face matching ability. However, there is at least one image manipulation technique that can improve face recognition ability, but does not appear to assist in face matching decisions. Caricaturing facial images enhances the idiosyncratic facial features that are most salient in an individual, making familiar faces more distinctive and thus recognisable (e.g., Brennan, 1985; Lee, Byatt & Rhodes, 2000; see Chapter 7 for an explanation of distinctiveness effects in facial recognition). The same technique reduces the likelihood that a foil will be selected from an array in an unfamiliar face matching deign, but it also reduces rates of correct target identifications (McIntyre, Hancock, Kittler, & Langton, 2013).

Theoretical models of face recognition and matching

Theories of face recognition distinguish between the processing of familiar and unfamiliar faces (Bruce & Young, 1986; Burton, Bruce, & Johnston, 1990; Hancock, Bruce, & Burton, 2000). These models propose that internal, structural, view-independent mental representations are employed in the recognition of familiar faces. In contrast, unfamiliar face recognition is primarily based on feature processing derived from surface information. As such, familiar faces are easily recognised across a wide range of viewing conditions, whereas unfamiliar face recognition or matching can be adversely affected by surprisingly minor image changes. Neurological support for this dissociation comes from some prosopagnosic patients who exhibit selective impairments to either familiar or unfamiliar faces (e.g., Malone, Morris, Kay, & Levin, 1982). Different patterns of brain activity were observed when non-clinical participants viewed familiar and unfamiliar faces (e.g., Leveroni, Seidenberg, Mayer *et al*., 2000; Rossion, Schlitz, & Crommelinck, 2003). There is also experimental evidence for this dissociation. When faces are familiar, even from relatively brief experimental exposure, face identification tasks display the *mirror effect*, in that there is a relationship between performance on individual stimuli in both target-present/match and target-absent/mismatch trials (Megreya & Burton, 2007). Faces that are easily recognised when present (hits), are easily rejected when not previously seen (correct rejections). This effect is common to many classes of stimuli (see Glanzer & Adams, 1985; Glanzer, Adams, Iverson, & Kim, 1993 for reviews). In contrast, the mirror effect is absent in tasks of unfamiliar face recognition (Deffenbacher, Johanson, Vetter, & O'Toole, 2000; Hancock, Burton, & Bruce, 1996; Vokey & Read, 1992) and matching (Megreya & Burton, 2007), in that there is no relationship between hit rates and correct rejection rates for the same stimuli. Perhaps counter-intuitively an unfamiliar face that, due to its distinctiveness, appears to be easy to recognise when it has been seen previously, is not necessarily equally easy to correctly reject if it has not been seen before. The face matching mirror effect is also absent when both unfamiliar and familiar faces are inverted (Megreya & Burton, 2007). Inversion is believed to disrupt the *holistic* or whole face processes that drive efficient face recognition. Megreya and Burton (2007) argue that inversion therefore bestows on a normally familiar face the processing status of an unfamiliar face.

Burton (2013; see also Burton, Jenkins, Hancock, & White, 2005; Burton & Jenkins, 2011; Burton, Jenkins, & Schweinberger, 2011; Hancock *et al*., 2000; Megreya & Burton, 2006) explains the dissociation between processing unfamiliar and familiar faces for identity as a consequence of unfamiliar faces being processed as 'simple visual patterns.' This does not suggest that unfamiliar faces have the same status as purely abstract visual information, because inverted unfamiliar faces are matched and recognised less well than upright unfamiliar faces (Megreya & Burton, 2006), which would not be the case with completely random patterns. Some salient identity information must be provided by an unfamiliar upright face. Nevertheless, Burton proposes that the primary reason that unfamiliar face matching is prone to failure is that humans effectively employ an unsophisticated image-matching strategy. The greater the variation between images being matched the greater the likelihood of error in such a task.

Burton (2013) draws on a recent study examining within-person variability in the appearance of photographic images by Jenkins, White, van Montfort, and Burton (2011) to illustrate this point. The authors entered the names of two different Dutch celebrities into Google. They collected the first 20 images of those celebrities that were of a pre-specified minimum quality, and asked participants to group the 40 images into 'identities' without specifying the actual number of people depicted. Dutch participants, familiar with the celebrities, almost always grouped the images correctly. However, UK participants, grouped images of the same unfamiliar celebrity into an average set of nine different identities. This study demonstrates that images depicting the same individual can possess characteristics that make them appear to belong to completely different people. In a later experiment, Jenkins *et al.* (2011) also found large within-person variability in likeness ratings to familiar face photographs. Some photos appear to capture an individual's identity better than others. Varying levels of image likeness may be why unfamiliar face recognition and matching ability is error prone. A simple pattern matching strategy using superficial surface cues may not provide enough information for an accurate identity match.

Face familiarisation may involve the development of a representation of the appearance of a familiar face 'averaged' across many images of their face (Burton *et al.*, 2005). An 'average' or 'prototype' forms the basic representation of each face in memory. As we encounter that person again, different perspectives of their face are added to that internal representation, until the face becomes highly familiar. To test this proposal, Burton *et al.* morphed 16 different photographs of famous faces to form a prototype of each celebrity. Participants were presented with a name and either a single photograph, or the morphed image of one of the celebrities. Name verification reaction times to the morphed images were consistently faster than to the single photographs. This supports the proposal that composite images are a better match than any single image to our internal representations. Interestingly, Jenkins and Burton (2008) found that the same technique improved the matching accuracy of an automatic face recognition system.

Practical aspects of face matching

Research has also been conducted with the aim of improving face matching performance using different types of training (e.g., Alenezi & Bindemann, 2013; Moore & Johnston, 2013; O'Toole, Phillips, Jiang, Ayyad, Penard, & Abdi, 2007; Towler, White, & Kemp, 2014; White, Kemp, Jenkins, & Burton, 2013b), by providing motivational rewards (Moore & Johnston, 2013), and by averaging responses using the wisdom-of-the-crowds technique (White, Burton, Kemp, & Jenkins, 2013a). Until recently, research into training techniques was disappointing (e.g., O'Toole *et al.*, 2007). For instance, Towler *et al.* (2014) demonstrated that a component of many training courses for detecting identity fraud – analysing the external shape of faces based on pre-determined templates - does not improve face matching. Some research however has demonstrated that success can be achieved (e.g., Alenezi & Bindemann, 2013; White *et al*., 2013b). It should be noted that face matching performance is improved by repeated exposure (Clutterbuck & Johnston, 2002; 2004; 2005; Megreya & Burton, 2006), although these effects tend to be stimulus-specific and do not necessarily generalise to the matching of novel faces. Indeed, matching extremely large sets of faces results in a gradual decrease in performance (Alenzi & Bindemann, 2013).

Alenzi and Bindemann (2013; Experiment 6) asked participants to make a series of 1,000 face matching decisions, with one face in profile view, the other facing the camera. Half the trials were matched. The authors found that participants increasingly made more errors in mismatch trials. Indeed, by the final set of trials, participants were responding at chance levels. In contrast, accuracy consistently increased in match trials. These results represent a response bias in that participants were increasingly likely to respond 'same' regardless of trial type, suggesting that fatigue may reduce the ability to tell two different but highly similar appearing people apart. This implies that officials, such as passport officers, tasked with the continual monitoring of facial identities may become less able to identify a holder of a fake photograph identification document – particularly as this is likely to be a rare occurrence anyway. In later experiments, Alenzi and Bindemann (2013) found that providing trial-by-trial performance feedback reduced this decline in mismatch trial performance. In particular, providing such feedback during the first few blocks of the experiment, helped to *maintain* performance when no feedback was provided. It should be emphasised that feedback did not *improve* face matching when compared with baseline accuracy established at the start of the experimental trials suggesting there is a limit to its positive effects.

Similarly, White *et al.* (2013b) found that providing trial-by-trial performance feedback on a preliminary face matching task involving equal numbers of match and mismatch stimuli, improved *some* participants' ability on subsequent face matching tasks with an entirely different stimulus set. No feedback was provided in the subsequent task.

Nevertheless, the performance of those with the highest aptitude at face matching, as measured using a pre-trial test, was not improved by this feedback. Only those with prior low face matching ability were assisted by this technique. They reached the same level of overall matching accuracy as the high ability group. Tests revealed that these positive effects were not a consequence of response bias, or an optimisation of response strategy, but were a consequence of greater sensitivity to both match and mismatch stimuli. The authors argue that some participants have a pre-conceived belief that face matching is easy, and the feedback drew their attention to the true difficulty of the task. They also suggest that the feedback may have encouraged the lower ability participants to focus on the stable internal features of the face, as these will normally provide more reliable cues to identity. Alenezi and Bindemann (2013) suggest that introducing face matching trial feedback on a daily basis to those tasked with matching faces in real-life situations (e.g., passport issuance officers) might reduce potential errors. The authors note that such procedures are applied in airport baggage screening to maintain the vigilance of operators (see Hofer & Schwaninger, 2005).

Offering motivational incentives can also improve face matching performance. Moore and Johnston (2013) demonstrated that inducements, of chocolate particularly, improved performance on mismatch trials when the ratio of match to mismatch trials was 16:16 (Experiment 1) and 30:2 (Experiment 2). The latter condition more closely replicates the expected rarity of mismatch trials in real life security settings although the authors acknowledge that the number of trials was far fewer than would be experienced by most security officials.

One recent proposal for improving decisions of face matching is to use crowd sourcing techniques, in which decisions of identity made by multiple untrained assessors are averaged and compared with the decisions made by the individual assessors (White, Burton, Kemp, & Jenkins, 2013a). Over a series of experiments, the authors found that employing this wisdom-of-the-crowds technique resulted in higher accuracy rates than the mean performance of the individual participants making up the ‘crowds’.

In summary, it is clear that positive effects may be derived from feedback training, incentives and crowd sourcing, but further research is required to ensure these effects are long lasting and practicable. Indeed, feedback may not assist those with the highest face matching ability, and it may not always be possible to employ a wisdom-of-the-crowds technique in security roles when the official is alone and time to process those under scrutiny is at a premium. Therefore it might be more appropriate (and easier) for occupational selectors to pre-screen those tasked with checking proof of photographic identity, to ensure they possess high levels of face recognition and matching ability in the first place.

Selection of naturally gifted face processors

There are large individual differences in face recognition and matching ability (Bindemann *et al.,* 2012; Burton *et al*., 2010; Davis *et al*., 2013; Megreya & Burton, 2006; Russell *et al*., 2009). For instance, in a study by Russell *et al*., four so called *super-recognisers* performed more than 2 standard deviations above the population mean on a standardised unfamiliar face recognition test (Cambridge Face Memory Test, CFMT). Anecdotes provided by the super-recognisers suggest that they are able, years later, to remember previously unfamiliar people encountered in a fleeting glimpse. Their scores on the CFMT test were the polar opposite of sufferers of *prosopagnosia* or face blindness. The prevalence of developmental prosopagnosia – a congenital inability to recognise faces in the population is approximately 1-2% (Kennerknecht, Grueter, Welling *et al*., 2006), and if face recognition ability is normally distributed, Russell *et al.* suggests that a similar prevalence of super-recognisers may exist – perhaps these individuals would be the most suitable for some security and policing roles?

For such a purpose in London, the Metropolitan Police Service (MPS) has recently instituted a unit of over 200 officers possessing superior face recognition ability from across the city's diverse geographical districts. As part of their occupational role these officers are provided with images of crime types matching their specialist knowledge, or in some cases, their locality within London. They have also been tasked with observing events with large crowds from CCTV operations rooms, in order to identify, track and direct other officers towards known offenders (see Taylor, 2013 for example).

In 2011, approximately 5,000 images were displayed on the MPS *Caught on Camera* website, and 18 of these officers identified over 600 suspects (Davis *et al*., 2013). One officer identified 190 suspects. The success of these officers dramatically contrast with the performance of the remaining 31,000 officers in the service, who are also encouraged to view the website - mostly to no effect. The 18 best-performing officers were examined by Davis *et al.* (in preparation; Davis *et al*., 2013), who employed a battery of face recognition and visual processing tests. Some, but not all of these officers performed far better than a control group of members of the general public (n = 104) on the combined scores from two unfamiliar face recognition tests, a familiar face recognition test, and an unfamiliar face matching test. One officer performed more than 3 standard deviations (SD) above the control mean, and a few others were more than 2 SD above the control mean. High performance on these tests is indicative of the ability to learn and remember facial images far more efficiently than those of average ability, as well as to be able to extrapolate identity from images taken from different viewpoints. The largest effect sizes were found with a familiar face recognition test based on 12-year-old poor-quality images of celebrities – designed to replicate an identification of a suspect not seen for many years from poor quality CCTV. However, the officers were no better than the controls at object recognition (flowers) or source monitoring, suggesting that their extreme abilities are limited to faces alone and may not generalise to other visual recognition tasks.

The utilisation of the skills of police ‘super-recognisers’ proved successful following a series of widespread riots in London in August 2011 as they were responsible for identifying approximately one-third of the 4000 riot suspects who were identified from over 200,000 hours of footage. In contrast, a face recognition software programme identified one suspect, probably a consequence of the generally poor quality of the riot images captured from above head height and often after dark (Associated Press, 2013). See Chapter 11 for a contrasting description of the extremely high performance of automatic face recognition algorithms when images are captured in ideal conditions.

Conclusions and Policy Recommendations

The consistent story from the research described in this chapter is that unless the viewer is familiar with those depicted, or they possess exceptional face processing abilities, errors in unfamiliar face matching may be common, and that these errors are similar to those made when memory is involved. Depending on the circumstances, both false positive and false negative errors are of equal concern, as can be illustrated by applying these findings to the security and police roles listed at the start of this chapter. For instance, a false negative error by a CCTV operator might mean they fail to realise the person under surveillance is the same person whose photograph they are simultaneously viewing on a ‘wanted’ person database. On the other hand, a false positive error may represent a scenario whereby an innocent suspect could be wrongly mistaken for the offender depicted in an image. The box describing the case of Jean Charles de Menezes illustrates the potentially disastrous consequences of a false positive error.

Some recent research suggests that providing performance feedback training or motivational incentives may improve face-matching performance. However, this research is in its infancy and feedback may be most effective with individuals possessing lower face recognition or matching ability. Bruce (2013) astutely asks "given the evident difficulties most of us have in matching or remembering images of faces, why do we continue to place so great an emphasis, in identity documents, and legal processes, on such a flimsy thing as an image of a face?" (p. 779). Perhaps until appropriate alternatives are developed, one of the most effective ways of reducing the potential of this error is to recruit individuals with exceptional face processing abilities for key security roles. The research on super-recognisers demonstrates that some individuals perform very much better in face recognition and matching tasks and this ability could be an occupational selection criterion for roles in border control, security checkpoints, CCTV operator control rooms, police officers with identification roles, and a multitude of other tasks where facial recognition ability is at a premium.

References

Alenzi, H.M., & Bindemann, M. (2013). The effect of feedback on face-matching accuracy. *Applied Cognitive Psychology*, *27*, 735-753. DOI: 10.1002/acp.2968

Associated Press (27 September 2013). UK cops hire 'super recognizers' to find criminals. *New York Post,* Retrieved 8th November 2013 from http://nypost.com/2013/09/27/uk-cops-hire-super-recognizers-to-find-criminals-in-a-crowd/

Attorney General's Reference, No.2 of 2002 (2003). EWCA Crim 2373.

Attwood, A.S., Penton-Voak, I.S., Burton, A.M., & Munafò, M.R. (2013). Acute anxiety impairs accuracy in identifying photographed faces. *Psychological Science*, 24, 1591-1594. DOI: 10.1177/0956797612474021

Barrett, D. (10 July 2013). One surveillance camera for every 11 people in Britain, says CCTV survey. *The Daily Telegraph*, Retrieved 5th November 2013 from http://www.telegraph.co.uk/technology/10172298/One-surveillance-camera-for-every-11-people-in-Britain-says-CCTV-survey.html

Bindemann, M., Attard, J., Leach, A., & Johnston, R.A. (2013). The effect of image pixelation on unfamiliar-face matching. *Applied Cognitive Psychology*. *27,* 707-717. DOI: 10.1002/acp.2970

Bindemann, M., Avetisyan, M., & Blackwell, K. (2010). Finding needles in haystacks: Identity mismatch frequency and facial identity verification. *Journal of Experimental Psychology: Applied*, *16*, 378-386. doi:10.1037/a0021893

Bindemann, M., Avetisyan, M., & Rakow, T (2012). Who can recognize unfamiliar faces? Individual differences and observer consistency in person identification. *Journal of Experimental Psychology: Applied, 18*, 277-291. doi:10.1037/a0029635

Bindemann, M., Brown, C., Koyas, T., & Russ, A. (2012). Individual differences in face identification postdict eyewitness accuracy. *Journal of Applied Research in Memory and Cognition, 1*, 96-103. doi:10.1016/j.jarmac.2012.02.001

Bonner, L., & Burton, A.M. (2004). 7-11-year-old children show an advantage for matching and recognizing the internal features of familiar faces: Evidence against a developmental shift. *Quarterly Journal of Experimental Psychology, 57A,* 1019 – 1029. doi:10.1080/02724980343000657

Borchard, E.M. (1932). *Convicting the Innocent: Sixty-five Actual Errors of Criminal Justice*. Garden City, New Jersey: Doubleday.

BBC News (16th August 2003). Tourist mistaken for missing doctor. *BBC News*, Retrieved 18th August 2003 from, http://news.bbc.co.uk/1/hi/uk_news/england/merseyside/3156819.stm

BBC News (1st November 2007). Police guilty over Menezes case. BBC News. Retrieved 20th December 2013 from http://news.bbc.co.uk/1/hi/uk/7069796.stm

Brennan, S.E. (1985). Caricature generator, the dynamic exaggeration of faces by computer. *Leonardo, 18*, 170–178. doi: 10.2307/1578048

Bruce, V. (1982). Changing faces: Visual and non-visual coding processes in face recognition. *British Journal of Psychology, 73,* 105 –116. doi:10.1111/j.2044-8295.1982.tb01795.x

Bruce, V. (2013). Comment. *Applied Cognitive Psychology*. *27,* 707-717. DOI: 10.1002/acp.2962

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

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. doi:10.1037/1076-898X.7.3.207

Bruce, V., Valentine, T., & Baddeley, A. (1987). The basis of the ¾ view advantage in face recognition. *Applied Cognitive Psychology, 1*, 110-120. doi:10.1002/acp.2350010204

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

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

Burton, A.M., Bruce, V., & Johnston, R.A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology, 81*, 361-380. doi:10.1111/j.2044-8295.1990.tb02367.x

Burton, A.M., & Jenkins, R. (2011). Unfamiliar face perception. In A.J. Calder, G. Rhodes, M.H. Johnson and J. Haxby (Eds.), *The Oxford Handbook of Face Perception* (pp. 287-306), Oxford University Press.

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

Burton, A.M., Jenkins, R., & Schweinberger, S.R. (2011). Mental representations of familiar faces. *British Journal of Psychology, 102,* 943-958. doi:10.1111/j.20448295.2011.02039.x

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

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. doi:10.1111/1467-9280.00144

Charman, S.D., & Wells, G.L. (2007). Eyewitness lineups: Is the appearance-change instruction a good idea? *Law and Human Behavior, 7*, 3–22. DOI: 10.1007/s10979-006-9006-3

Clutterbuck, R., & Johnston, R.A. (2002). Exploring levels of face familiarity by using an indirect face-matching measure. *Perception, 31,* 985-994. doi: 10.1068/p3335

Clutterbuck, R., & Johnston, R.A. (2004). Demonstrating the acquired familiarity of faces by using a gender-decision task. *Perception, 33,* 159 –168. DOI: 10.1068/p5115

Clutterbuck, R., & Johnston, R.A. (2005). Demonstrating how unfamiliar faces become familiar using a face matching task. *European Journal of Cognitive Psychology, 17,* 97 –116. doi:10.1080/09541440340000439

Cole, S.A., & Thompson, W.C. (2013). Forensic science and wrongful convictions. In C.R. Huff & M. Killias (Eds.), *Wrongful Convictions & Miscarriages of Justice* (pp. 111-135). Routledge: London.

Darling, S., Martin, D., Hellman, J. H., & Memon, A. (2009). Some witnesses are better than others. *Personality and Individual Differences, 47*, 369–373. doi: 10.1016/j.paid.2009.04.010

Davies, G., & Thasen, S. (2000). Closed-circuit television: How effective an identification aid? *British Journal of Psychology, 91,* 411-426. doi: 10.1348/000712600161907

Davis, J.P. (2007). The forensic identification of CCTV images of unfamiliar faces. *Unpublished thesis for the PhD Psychology at Goldsmiths, University of London*.

Davis, J.P., & Valentine, T. (2009). CCTV on trial: Matching video images with the defendant in the dock. *Applied Cognitive Psychology, 23*, 482-505. doi: 10.1002/acp.1490

Davis, J.P., Lander, K., & Jansari, A. (2013). I never forget a face! *The Psychologist, 26,* 726-729.

Davis, J.P., Valentine, T., & Wilkinson, C. (2012). Facial Image Comparison. In C. Wilkinson and C. Rynn (Eds.), *Craniofacial Identification*. (pp. 136-153). Cambridge University Press.

Deffenbacher, K.A., Bornstein, B.H., Penrod. S.D., & McGorty, E.K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior*, *28,* 687-706. doi: 10.1007/s10979-004-0565-x 10

Deffenbacher, K.A., Johanson, J., Vetter, T., & O'Toole, A.J. (2000). The face typicality-recognizability relationship: Encoding or retrieval locus? *Memory & Cognition, 28*, 1173-1182. doi: 10.1007/s10979-004-0565-x

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. doi: 10.1068/p080431

Fletcher, K.I., Butavicius, M.A., & Lee, M.D. (2008). Attention to internal face features in unfamiliar face matching. *British Journal of Psychology, 99*, 379–394. doi: 10.1348/000712607X235872

Fort, L. (11 July, 2013). Burglary case blunder as police break evidence rule. *Get Reading,* Retrieved August 13th, 2013, from http://www.getreading.co.uk/news/local-news/burglary-case-blunder-police-break-5066779

Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory and Cognition, 13*, 8-20. doi: 10.1037/0033-295X.100.3.546

Glanzer, M., Adams, J.K., & Kim, K. (1993). The regularities of recognition memory. *Psychological Review, 100*, 546-567. doi: 10.1037/0033-295X.100.3.546

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

Hancock, P.J.B., Burton, A.M., & Bruce, V. (1996). Face processing: human perception and principal components analysis. *Memory and Cognition, 24,* 26-40. DOI 10.3758/BF03197270

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

Heathrow Airport (2013). *Facts and Figures.* Retrieved 10th December 2013 from, http://www.heathrowairport.com/about-us/company-news-and-information/company-information/facts-and-figures

Henderson, Z., Bruce, V., & Burton, A.M. (2001). Matching the faces of robbers captured on video. *Applied Cognitive Psychology, 15,* 445-464. doi:10.1002/acp.718

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

Hill, H., Schyns, P.G., & Akamatsu, S. (1997). Information and viewpoint dependence in face recognition. *Cognition*, *62*, 201-222. doi:10.1016/s0010-0277(96)00785-8

Hofer, F., & Schwaninger, A. (2005). Using threat image projection data for assessing individual screener performance. *Safety and Security Engineering, 82*, 417-426.

Jenkins, R., & Burton A.M. (2008). 100% accuracy in automatic face recognition. *Science, 319,* 435. doi:10.1126/science.1149656

Jenkins, R., White, D., van Montfort, X., & Burton, A.M. (2011). Variability in photos of the same face. *Cognition, 121,* 313-323. DOI: 10.1016/j.cognition.2011.08.001

Kemp, R., Towell, N., & Pike, G. (1997). When seeing should not be believing: Photographs, credit cards and fraud. *Applied Cognitive Psychology, 11*, 211-222. doi:10.1002/(SICI)1099-0720(199706)11:3<211::AID-ACP430>3.0.CO;2-O

Kennerknecht, I., Grueter, T., Welling, B., Wentzek, S., Horst, J., Edwards, S., & Grueter, M. (2006). First report of prevalence of non-syndromic hereditary prosopagnosia (HPA). *American Journal of Medical Genetics, 140A*, 1617–1622. DOI: 10.1002/ajmg.a.31343

Lee, K., Byatt, G., & Rhodes, G. (2000). Caricature effects, distinctiveness, and identification: testing the face-space framework. *Psychological Science, 11*, 379–385. doi: 10.1111/1467-9280.00274

Lee, W.J., Wilkinson, C., Memon, A., & Houston, K.A. (2009). Matching unfamiliar faces from poor quality CCTV footage: an evaluation of the effect of training on facial identification ability. *Axis 1,* 19-28.

Leveroni, C.L., Seidenberg, M., Mayer, A.R., Mead, L.A., Binder, J.R., & Rao, S.M. (2000). Neural systems underlying the recognition of familiar and newly learned faces. *Journal of Neuroscience, 20*, 878-886. http://www.jneurosci.org/content/20/2/878.full

Liu, C.H., Chen, W., Han, H., & Shan, S. (2013). Effects of image preprocessing on face matching and recognition in human observers. *Applied Cognitive Psychology, 27,* 718-724. DOI: 10.1002/acp.2967

Loftus, G.R., & Harley, E.M. (2005). Why is it easier to identify someone close than far away? *Psychonomic Bulletin & Review, 12*, 43-65. DOI: 10.3758/BF03196348

Malone, D.R., Morris, H.H., Kay, M.C., & Levin, H.S. (1982). Prosopagnosia: a double dissociation between the recognition of familiar and unfamiliar faces. *Journal of Neurological Neurosurgical Psychiatry, 45*, 820-822. doi:10.1136/jnnp.45.9.820

Mcintyre, A.H., Hancock, P.J.B., Kittler, J., & Langton, S.R.H. (2013). Improving discrimination and face matching with caricature. *Applied Cognitive Psychology, 27*, 725-734. DOI: 10.1002/acp.2966

Megreya, A.M., & Bindemann, M. (2009). Revisiting the processing of internal and external features of unfamiliar faces: The headscarf effect. *Perception, 38,* 1831-1848. doi:10.1068/p6385

Megreya, A.M., Bindemann, M., & Havard, C. (2011). Sex differences in unfamiliar face identification: Evidence from matching tasks. *Acta Psychologica, 137*, 83-89. http://dx.doi.org/10.1016/j.actpsy.2011.03.003

Megreya, A.M., Bindemann, M., Havard, C., & Burton, A.M. (2012). Identity-lineup location influences target selection: Evidence from eye movements. *Journal of Police and Criminal Psychology, 27*, 167-178. doi:10.1007/s11896-011-9098-7

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

Megreya, A.M., & Burton, A.M. (2007). Hits and false positives in face matching: A familiarity-based dissociation. *Perception and Psychophysics, 69*, 1175-1184. DOI: 10.3758/BF03193954

Megreya, A.M., & Burton, A.M. (2008). Matching faces to photographs: Poor performance in eyewitness memory (without the memory). *Journal of Experimental Psychology: Applied, 14,* 364-372. doi:10.1037/a0013464

Megreya, A.M., Sandford, A., & Burton, A.M. (2013). Matching face images taken on the same day or months apart: the limitations of photo ID. *Applied Cognitive Psychology*. *27,* 700-706. DOI: 10.1002/acp.2965

Megreya, A.M., White, D., & Burton, A.M. (2011). The other-race effect does not rely on memory: evidence from a matching task. *Quarterly Journal of Experimental Psychology, 64, 1473-1483.* DOI: 10.1080/17470218.2011.575228

Meissner, C.A., Susa, K.J., & Ross, A.B. (2013). Can I see your passport please? Perceptual discrimination of own- and other-race faces. *Visual Cognition, 21*, 1287-1305. DOI:10.1080/13506285.2013.832451

Moore, R.M., & Johnston, R.A. (2013). Motivational incentives improve unfamiliar face matching accuracy. *Applied Cognitive Psychology, 27,* 754–760. DOI: 10.1002/acp.2964

Munsterberg, H. (1908/1925). *On the Witness Stand*: *Essays on Psychology and Crime*. Retrieved 1st August 2013, from http://psychclassics.yorku.ca/Munster/Witnessl

Murphy, T. (1999). The admissibility of CCTV evidence in criminal proceedings. *International Review of Law and Computers, 13,* 383-404. doi:10.1080/13600869955044

O'Toole, A.J., Phillips, P.J., Jiang, F., Ayyad, J., Penard, N., & Abdi, H. (2007). Face recognition algorithms surpass humans matching faces across changes in illumination. *IEEE: Transactions on Pattern Analysis and Machine Intelligence, 29,* 1642-1646. Doi:10.1109/TPAMI.2007.1107

Özbek, M., & Bindemann, M. (2011). Exploring the time course of face matching: Temporal constraints impair unfamiliar face identification under temporally unconstrained viewing. *Vision Research, 51,* 2145-2155 doi:10.1016/j.visres.2011.08.009

Papesh, M.H., & Goldinger, S.D. (2014). Infrequent identity mismatches are frequently undetected. *Attention, Perception, & Psychophysics, 76*, 1335-1349. DOI: 10.3758/s13414-014-0630-6

Police and Criminal Evidence Act (1984) *Codes of Practice (2011).* Retrieved 12th January 2013, from: http://www.homeoffice.gov.uk/publications/police/operational-policing/pace-codes/pace-code-d-2011

R v. Caldwell and Dixon (1993) CLR 862, UK

R v. Clare and Peach (1995) 2 Cr App R 333, UK

R v. Clarke (1995) 2 Cr App R 425, UK

R v. Dodson and Williams (1984) 79 Cr App R 220, UK

R v. Fowden and White (1982) Crim LR 588, UK

R v. Grimer (1982) Crim LR 674, UK

Rice, A., Phillips, P.J., & O'Toole, A. (2013). The role of the face and body in unfamiliar person identification. *Applied Cognitive Psychology, 27,* 761–768. DOI: 10.1002/acp.2969

Rossion, B., Schiltz, C., & Crommelinck, M. (2003). The functionally defined right occipital and fusiform "face areas" discriminate novel from visually familiar faces. *Neuroimage, 19*, 877-883. doi:10.1016/S1053-8119(03)00105-8

Russell, R., Duchaine, B., & Nakayama, K. (2009). Super-recognizers: People with extraordinary face recognition ability. *Psychonomic Bulletin & Review, 16*, 252-257. doi:10.3758/PBR.16.2.252

Smith v. The Queen (2001). HCA 50. Australia

Stainer, M.J., Scott-Brown, K.C., & Tatler, B. (2013). Expertise and spatial selection in the CCTV control room. *Frontiers in Human Neuroscience, 7,* 1-9. doi: 10.3389/fnhum.2013.00615

Strathie, A., McNeill, A., & White, D. (2012). In the dock: Chimeric image composites reduce identification accuracy. *Applied Cognitive Psychology, 26,* 140-148. Doi:10.1002/acp.1806

Strauss v. the Police (2013). SASC3, South Australia

Tanaka, J.W., & Farah, M.J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *46A*, 225-245. DOI: 10.1080/14640749308401045

Tanaka, J.W., & Sengco, J.A. (1997). Features and their configuration in face recognition. *Memory & Cognition*, *25*, 583-592. DOI: 10.3758/BF03211301

Taylor v. The Chief Constable of Cheshire (1987). 84 CR.APP.R. 191, UK.

Taylor, M. (24th August 2013). Police 'super recognisers' to keep watch over Notting Hill carnival. *The Guardian.* Retrieved 2nd September 2013, from http://www.theguardian.com/uk-news/2013/aug/23/police-super-recognisers-notting-hill-carnival

Thompson, B., Dunkelberger, N., & Vescio, S. (2013). Juror face matching accuracy: Do judicial instructions help? *Poster presented at the 25th Annual Convention of the Association for Psychological Science, Washington, D.C.,* May 2013.

Thompson, R., & Gerrard, G. (2011). Two million cameras in the UK. *CCTV Image, 42*, 10-12.

Thompson, W.B., Grattan, J., Rawding, J. & Buchholz, L. (2010). Identifying unfamiliar faces from overhead surveillance photos. *American Journal of Forensic Psychology, 28,* 71-78.

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.doi:10.1068/p7676

Troscianko, T., Holmes, A., Stillman, J., Mirmehdi, M., Wright, D., & Wilson, A. (2004). What happens next? The predictability of natural behaviour viewed through CCTV cameras. *Perception, 33*, 87–101.doi: 10.1068/p3402

Valentine, T., & Mesout, J. (2009). Eyewitness identification under stress in the London Dungeon. *Applied Cognitive Psychology, 23*, 151-161. doi:10.1002/acp.1463

Vokey, J.R., & Read, J.D. (1992). Familiarity, memorability, and the effect of typicality on the recognition of faces. *Memory & Cognition*, *20*, 291-302. DOI: 10.3758/BF03199666

White, D., Burton, A.M., Kemp, R.I., & Jenkins, R. (2013). Crowd effects in unfamiliar face matching. *Applied Cognitive Psychology, 27,* 769–777. DOI: 10.1002/acp.2971

White, D., Kemp, R.I., Jenkins, R., & Burton, A.M. (2014). Feedback training for facial image comparison. *Psychonomic Bulletin & Review, 21,* 100-106. doi: 10.3758/s13423-013-0475-3

Wright, D.B., & Sladden, B. (2003). An own gender bias and the importance of hair in face recognition. *Acta Psychologica, 114*, 101-114. DOI:10.1016/S0001-6918(03)00052-0

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. doi:10.1068/p140737

Figure 1: Hussain Osman (left) and Jean Charles de Menezes (centre), presented as a chimeric facial image (right). Note that this image has not been manipulated by standardising skin tone or stretching the images to align facial features as was suggested by the prosecution to have occurred in the Health and Safety court case reported in this chapter (kind permission Metropolitan Police Service, London)

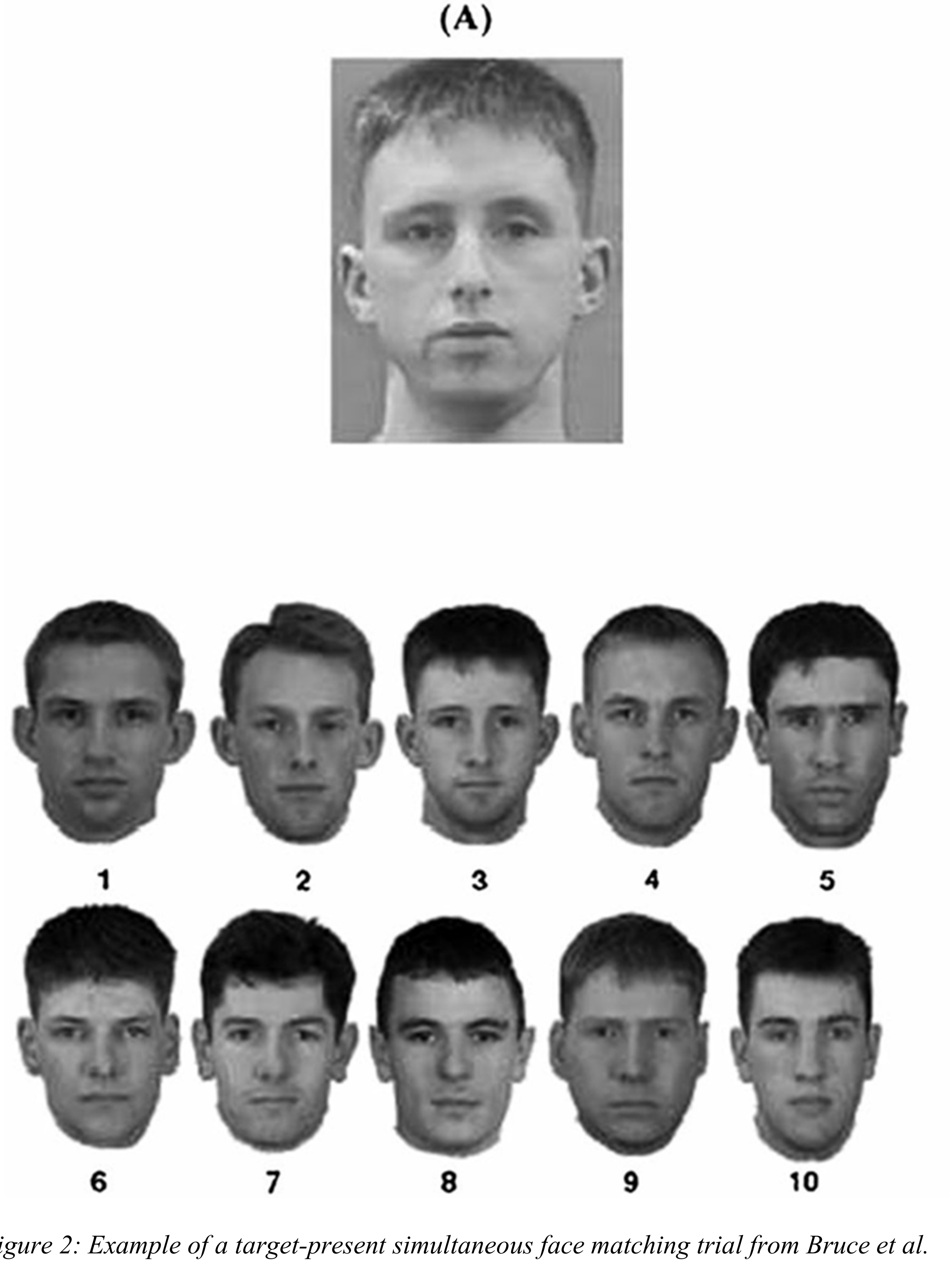

Figure 2: Example of a target-present simultaneous face matching trial from Bruce et al. (1999). The Target is number 3 in the array. Copyright © [1999] by the American Psychological Association. Reproduced with permission.