UNIVERSITY of York

Article:

Young, Andrew William orcid.org/0000-0002-1202-6297 and Burton, Anthony Michael orcid.org/0000-0002-2035-2084 (2017) Recognizing faces. Current Directions in Psychological Science. pp. 212-217. ISSN 1467-8721

https://doi.org/10.1177/0963721416688114

Recognizing faces

Andrew W. Young and A. Mike Burton

Department of Psychology, University of York,

Heslington, York YO10 5DD, England

Authors' version, as accepted for publication in *Current Directions in Psychological Science*.

Date of acceptance: December 13th 2016.

Acknowledgement

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n.323262. We are grateful to Dr. Nick Rule and three anonymous reviewers for helpful comments and suggestions.

ABSTRACT

The idea that most of us are good at recognising faces permeates everyday thinking and is widely used in the research literature. However, it is only a correct characterisation of familiar face recognition; the perception and recognition of unfamiliar faces can be surprisingly error-prone. We show how neglect of the important property of image variability has generated some misleading conclusions, and how studies that use and explore image variability can correct these and lead to substantial theoretical advances.

BACKGROUND

A widely held opinion is that most people are good at recognizing faces. Politicians call for citizens to carry photo ID, eyewitness evidence can seem very compelling in court, and researchers say that we are face experts (Carey, 1992). Although it has become well established that some people have problems with face recognition, these individuals are typically assumed to form a minority suffering from some form of pathological 'face blindness' (Behrmann & Avidan, 2005) in which our normally excellent face recognition abilities are somehow missing or switched off.

Our aim here is to review classic and recent findings that show the limitations of the above characterisation of human face recognition and replace it with something that is more securely grounded in evidence. This evidence has often been overlooked because it doesn't fit neatly into a set of assumptions that seem like "common sense". We will outline serious problems inherent in the common sense way of thinking and offer a more firmly established account and explanation of why we researchers have been misleading ourselves.

Modern research on face recognition can be traced back to the seminal review by Ellis (1975), which brought together a wide range of studies of normal adults, infants and children, and studies of the effects of brain injury to establish face recognition as a distinct field of psychological enquiry. We have used the same title of "Recognizing faces" to emphasize our indebtedness to Ellis (1975), but as we will explain, much has also changed across the intervening forty years.

FACE RECOGNITION AND PICTURE RECOGNITION

Ellis (1975) noted remarkably high levels of correct recognition of faces in standard recognition memory tasks. In such tasks, a participant studies a series of photographs of faces of people he or she has never seen before. We will call these *unfamiliar faces* to distinguish them from faces of already known individuals. These photographs of studied unfamiliar faces are then mixed with photographs of other unfamiliar faces and the participant is asked to pick out the faces seen in the study phase of the experiment. Ellis's (1975) review (and many subsequent studies) showed that high levels of performance can be found on such tasks even when the faces were only studied for a few seconds each, that performance is better for faces than many other types of visual stimuli, and better for upright than for inverted faces.

These laboratory findings contrasted markedly with miscarriages of justice in the 1970s in which witnesses were later found to have misidentified the perpetrators of various crimes. In some of these cases, witnesses were mistaken even when they were themselves certain that they were correct. The apparent paradox of the discrepancy between excellent laboratory performance (in recognition memory tasks) and fragile real-world performance (witness misidentifications) was resolved by Bruce (1982), who pointed out that standard recognition memory tasks asked participants to pick out face photographs that were identical to the ones they had studied. This procedure is as much one of picture recognition as of face recognition. A true face recognition test requires that participants can recognise the studied faces across different photographs. It turns out that for faces participants have only seen in a single photograph, generalisation of recognition to new images of the same faces is poor; performance falls off quickly as the difference between the studied and test photographs increases, even for face photographs that were very thoroughly learnt (Longmore, Liu, & Young, 2008). In consequence, the general applicability of laboratory face recognition performance was overestimated by relying on picture recognition tasks.

The difference between picture recognition and face recognition is easy to grasp, but it took researchers much longer to appreciate that our relatively poor performance at unfamiliar face recognition is as much a problem of perception as of memory. Try for yourself the face matching task shown in Figure 1. Even though the poses and expressions of the faces are much the same (and the pictures were taken on the same day so that there were no changes in hairstyle or other modifiable characteristics), most people find the task surprisingly tricky—with overall error rates around 30% (Bruce et al., 1999). Yet this is not a memory task; it is purely one of perceptual matching, and viewers can take as long as they like to compare the images.

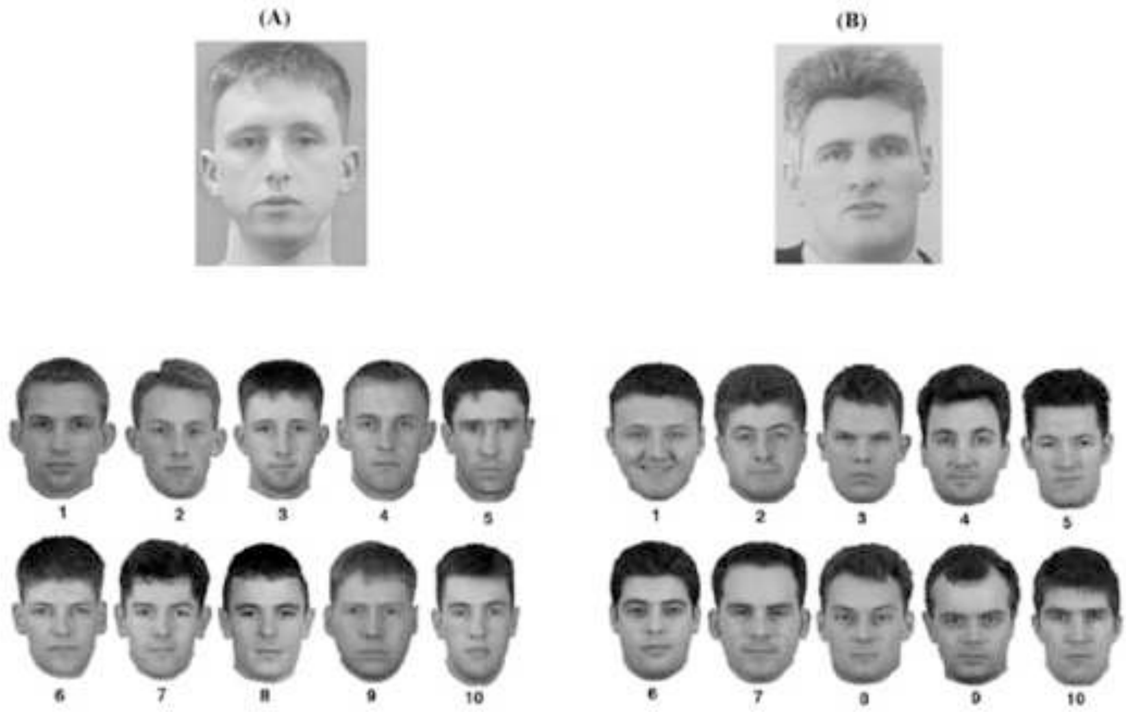

Figure 1: Two examples of the face matching task from Bruce et al. (1999). The lower arrays show photographs of 10 faces and the upper photographs were taken on the same day but with a different camera. Is the top person present in the lower array, or are they missing? Solutions given at the end of this paper.

It might be thought that the task shown in Figure 1 is unfair because it involves an artificially contrived problem of comparing one photograph to another but, in fact, the same problems arise for photo identification in everyday life. Kemp, Towell, and Pike (1997) conducted a study in a supermarket, using specially created identity (ID) cards with either a valid photograph (i.e., an actual photo of the person) or invalid photograph (i.e., a photo of another person). The supermarket's own cashiers were told to look out for the fake ID cards and told that a bonus payment would be made to the cashier with the best performance. Yet the cashiers challenged around 10% of people presenting a valid card and accepted 64% of the invalid cards if there was some similarity in appearance. Thus, even comparing a photo of a face to the person standing in front of you is not as easy as we usually take it to be.

It turns out that unfamiliar face matching is also a problem for professional groups that rely on this ability. For example, White et al. (2014) showed that working passport officers made about 10% errors when asked to verify passport photos of volunteers approaching an ID-check. Despite extensive training requirements, highly experienced officers performed, on average, no better than new recruits.

These average error rates actually conceal large individual variation, however. Most studies show very little effect of training on people's unfamiliar face matching skills but there appear to be substantial, and stable, differences between people's baseline abilities. These individual differences in face abilities are of considerable interest in themselves (Yovel, Wilmer, & Duchaine, 2014). In the present context they have led to an emphasis on selection, rather than training, in professional settings. Large organisations have the opportunity to select staff with particularly high face matching ability for face-related tasks (e.g. the London Metropolitan Police Force; see Robertson et al., 2016).

THE IMPORTANCE OF IMAGE VARIABILITY

The findings from unfamiliar face matching and photo ID tasks show that we have some kind of perceptual problem with recognising the identities of unfamiliar faces, or at least with photographs of unfamiliar faces. Though there are substantial individual differences, and a few ‘super-recognisers' do show unusually high ability (Russell, Duchaine, & Nakayama, 2009), for most people performance is both error-prone and surprisingly hard to improve.

A powerful insight into the nature of this problem derives from a sorting task devised by Jenkins, White, Montfort, and Burton (2011). Participants were given a set of 40 photographs of faces like those shown in Figure 2 and asked to sort these into piles of photographs of the same person.

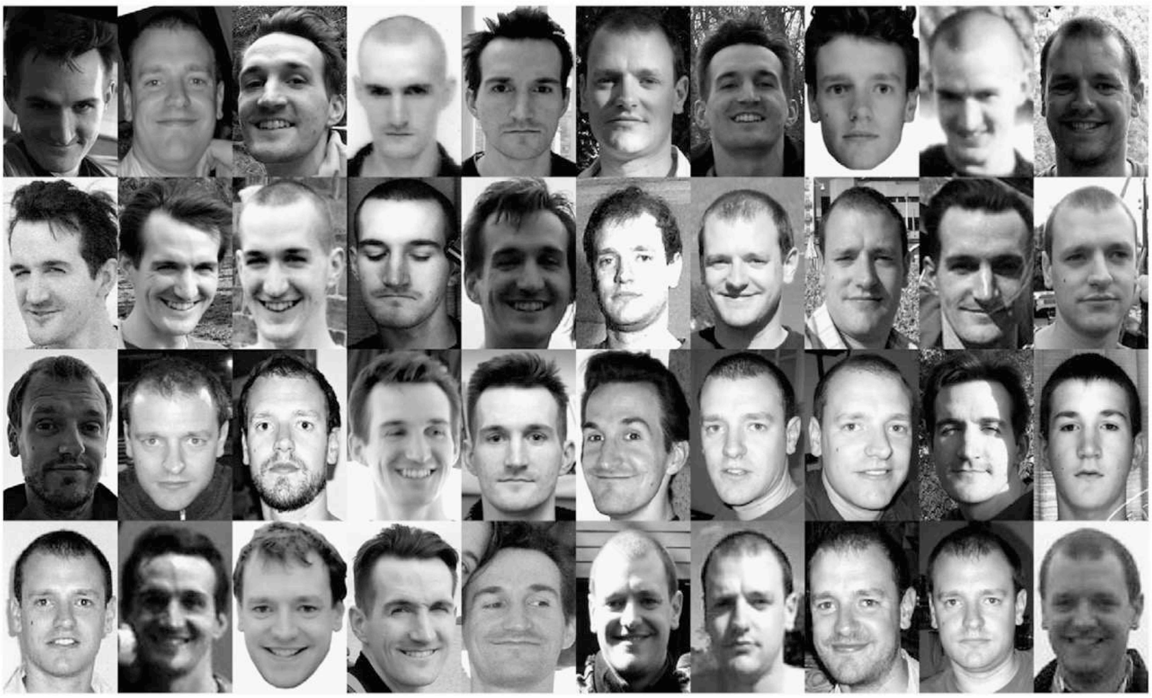

Figure 2: An example of the task used by Jenkins et al.

(2011). Can you sort the 40 images into the different face identities? Most people only arrive at the correct solution if they already know the faces. For unfamiliar faces, participants tend to mistake differences between the images for differences in identity, leading them to overestimate the number of faces in the display. With familiar faces, the task becomes easy. Reproduced from Jenkins et al. (2011). Solution given at the end of this paper.

In fact there are only two faces in Figure 2, and anyone who knows these people will experience little trouble in creating a fully correct solution of two piles of 20 images each (Jenkins et al., 2011). However, when the faces used by Jenkins et al. (2011) were unfamiliar to their participants, they created between 3–16 different piles (identities), with nine piles being the most common number. In other words, participants typically thought there were nine different individuals in the set of 40 photos when there were actually only two. In marked contrast, people rarely put photos of the two different individuals into the same pile (less than 1% of trials).

This finding is particularly interesting because it runs counter to a widely accepted intuition (which can be traced back at least as far as Galton,1883) that faces form a homogeneous class of visual stimuli and that, in consequence, people mainly struggle to tell similar faces apart. Rather, Jenkins et al.'s (2011) data show that participants are more likely to see photos of faces as more diverse than they actually are (thus, they create too many piles). The problem is as much one of seeing that very different images can represent the same unfamiliar face identity as of telling faces apart (Andrews, Jenkins, Cursiter, & Burton, 2015).

Jenkins et al. (2011) discussed their findings in terms of the idea of image variability. Photographs of faces differ in many ways that include pose, expression, lighting, camera, and lens characteristics. Importantly, real-life views of faces are also highly variable; this is true whether the faces are seen in person, in videos, or photographs. This variability might result from within-person variability (e.g., differences between different views of the same face) or between-person variability (e.g., differences between similar views of different faces). As Figure 2 clearly shows, within-person variability can be substantial. Researchers and government authorities therefore try to minimise its impact through the creation of sets of highly standardised images for laboratory tasks or passport photographs. This tactic is at best of limited use; the findings we have discussed show that image variability can easily confuse the visual system when people look at unfamiliar faces.

THE IMPORTANCE OF FACE FAMILIARITY

Although unfamiliar faces can present significant problems, these problems are remarkably negligible for familiar faces. This contrast between familiar and unfamiliar face recognition has been at the heart of cognitive models (Bruce & Young, 1986; Burton, Bruce & Hancock, 1999; Young & Burton, 1999;

Schweinberger & Burton, 2011; Young & Bruce, 2011). Image variability is generally not a problem for familiar faces, which can even be recognised despite severe image degradation (Burton, Jenkins, & Schweinberger, 2011).

Why should this be? In essence, the answer seems to be that we have seen familiar faces enough times to have learnt how to cope with their variability (Longmore et al., 2008), but can we arrive at something more precise? Bruce (1994) introduced the idea that ‘stability from variation’ might be useful for face recognition; multiple exposures to a familiar face allow our perceptual systems to separate transient within-person differences from the stable characteristics of that face. Developing this insight, Burton et al. (2005) offered a technique that can capture what is consistent across different images of the same person's face by using computer image manipulation techniques to average together multiple examples (see Figure 3). This ‘within-person’ average of someone's face turns out to be an excellent representation for automatic computer-based recognition (Jenkins & Burton, 2008) and has attractive properties for human perception too. Averaging leads to a robust representation that is not unduly influenced by the superficial differences that make simple image-comparison so difficult (e.g., lighting changes; Jenkins & Burton, 2011).

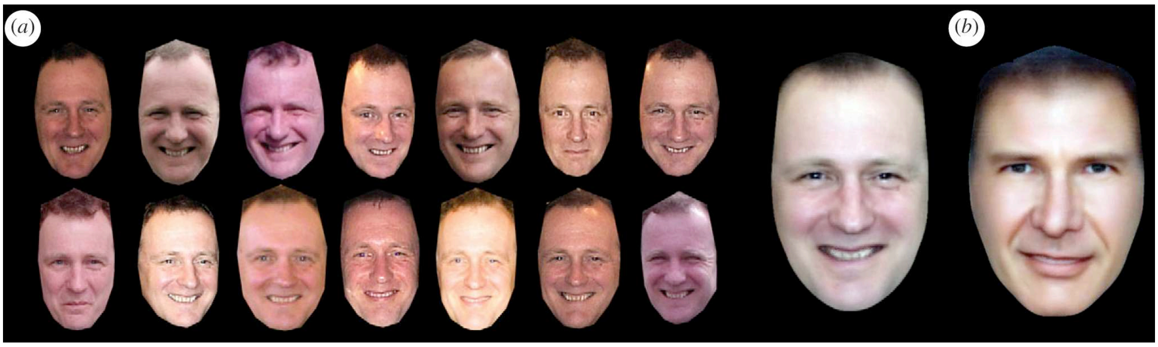

Figure 3: (a) Fourteen different photographs of one of the authors (AMB) and a computer-manipulated average of these (the larger image shown to the right of the 14 photos). Note how effective averaging is in removing the impact of identity-irrelevant differences in lighting, pose, and expression. (b) An average image of a face that should be familiar to many readers (Harrison Ford) created with an equivalent procedure. Images from Jenkins and Burton (2011).

IMAGE VARIABILITY AND FACE FAMILIARITY

Although averaging offers a powerful way to find the consistent cues that signal a face's identity, it seems curious that the visual system remains confused by variability between images of unfamiliar faces, given one's exposure to so many faces in everyday life. The reason turns out to be that the variability is to some extent *idiosyncratic* – that is, the ways in which face A varies across different images need not be the same as the ways in which face B varies across different images. Burton et al. (2016) analysed the statistical properties of face images, both within and between identities. Consistent with previous findings, the largest variations were common to all faces and corresponded to physical differences such as pose and lighting direction. Once this common variation was removed, the remaining differences among images of the same person were highly person-specific.

This idiosyncratic variability of faces explains how we can be an ‘expert’ for one face but not another (Burton et al., 2016; Kramer, Young, Day & Burton, in press). Learning how the face of Brad Pitt can vary – through seeing him in many settings – allows us to recognize new photos of him even in quite novel settings (e.g., if he is caught with an unusual expression or in poor lighting). This expertise may not generalize to another person’s face, however, because that person varies in different ways. For this reason, it is very easy to match two photos of a familiar person but very difficult to match photos of an unfamiliar person, whose range of variability is unknown. It also explains why attempts to train people to become better unfamiliar face matchers *in general* are likely to fail. We can train observers to recognize a particular new face very well, but this training will not generalise to recognizing a different face (see Dowsett et al., 2016).

IMAGE VARIABILITY AND THE CONCEPT OF EXPERTISE

Where do these findings leave us? A good place to begin considering their broader implications is to return to the idea that people are face experts, and the more general concept of visual expertise. This is usually put forward in terms of expertise in face recognition but, as we have explained, such a generic concept often does not hold. For face identity recognition, people mainly show something that can be considered expertise only for those faces they know well. One can be an expert at recognising Brad Pitt, Barack Obama or Scarlett Johansson, but this expertise is identity-specific and does not generalise to recognising the identities of unfamiliar faces. It is the ease with which people recognise familiar faces that seems to have misled them into overlooking their limitations with unfamiliar faces (Ritchie et al., 2015).

That said, there are other aspects of unfamiliar face perception where people do show expertise. For example, people can usually make a reasonably accurate estimate of a person's age or sex from his or her face, and they can interpret subtle differences in facial expressions and gaze direction (see Bruce & Young, 2012). What characterises such abilities is that, unlike face recognition, they involve cues that are highly consistent across many different faces (Kramer et al., in press).

From findings such as those we have presented, it is clear that considering image variability is critical to understanding face recognition. A key corollary of this is that we need to study *naturally occurring* images of faces, which Jenkins et al. (2011) have called *ambient images*. The traditional approach to face recognition has been to use highly standardized images in an attempt to minimize the impact of ‘nuisance’ variation from factors such as lighting, camera differences, etc. However, we have shown that this approach can obscure our understanding. By using the full range of ambient images of the type people recognize every day in their newspapers, televisions, and online, we can preserve the natural within-person variations that characterise our real world experience of faces. Far from being a nuisance, this variability is a necessity in allowing us to find consistent cues for recognising face identity (Bruce, 1994; Burton, 2013) and for other aspects of face perception (Bruce & Young, 2012; Sutherland et al., 2013; Vernon et al., 2014).

REFERENCES

Andrews, S., Jenkins, R., Cursiter, C., & Burton, A. M. (2015). Telling faces together: learning new faces through exposure to multiple instances. *Quarterly Journal of Experimental Psychology, 68*, 2041-2050.

Behrmann, M., & Avidan, G. (2005). Congenital prosopagnosia: face-blind from birth. *Trends in Cognitive Sciences, 9*, 180-187.

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

Bruce, V. (1994) Stability from variation: The case of face recognition. *Quarterly Journal of Experimental Psychology, 47*, 5-28.

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

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77*, 305-327.

Bruce, V., & Young, A. (2012). *Face perception*. Hove, East Sussex: Psychology Press.

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.

Burton, A. M., Bruce, V., & Hancock, P. J. B. (1999). From pixels to people: a model of familiar face recognition. *Cognitive Science, 23*, 1-31.

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.

Burton, A.M., Jenkins, R. & Schweinberger, S. R. (2011). Mental representations of familiar faces. *British Journal of Psychology, 102*, 943-958.

Burton, A. M., Kramer, R. S. S., Ritchie, K. L., & Jenkins, R. (2016). Identity from variation: Representations of faces derived from multiple instances. *Cognitive Science, 40*, 202–223.

Carey, S. (1992). Becoming a face expert. *Philosophical Transactions of the Royal Society, London, B: Biological Sciences, 335*, 95-103.

Dowsett, A. J., Sandford, A., & Burton, A. M. (2016). Face learning with multiple images leads to fast acquisition of familiarity for specific individuals. *Quarterly Journal of Experimental Psychology, 69*, 1-10.

Ellis, H. D. (1975). Recognizing faces. *British Journal of Psychology, 66*, 409-426.

Galton, F. (1883). *Inquiries into human faculty and its development*. London: Macmillan.

Jenkins, R., & Burton, A. M. (2008). 100% accuracy in automatic face recognition. *Science, 319*, 435.

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

Jenkins, R., White, D., Van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. *Cognition, 121*, 313–323.

Kemp, R., Towell, N., & Pike, G. (1997). When seeing should not be believing:

photographs, credit cards and fraud. *Applied Cognitive Psychology, 11*, 211-222.

Kramer, R. S. S., Young, A. W., Day, M. G., & Burton, A. M. (in press). Robust social categorization emerges from learning the identities of very few faces. *Psychological Review,*

Longmore, C. A., Liu, C. H., & Young, A. W. (2008). Learning faces from photographs. *Journal of Experimental Psychology: Human Perception & Performance, 34*, 77-100.

Ritchie, K. L., Smith, F. G., Jenkins, R., Bindemann, M., White, D., & Burton, A. M. (2015). Viewers base estimates of face matching accuracy on their own familiarity: Explaining the photo-ID paradox. *Cognition, 141*, 161-169.

Robertson, D. J., Noyes, E., Dowsett, A. J., Jenkins, R., & Burton, A. M. (2016). Face recognition by Metropolitan Police super-recognisers. *PLoS ONE, 11(2)*, e0150036.

Russell, R., Duchaine, B., & Nakayama, K (2009). Super-recognisers. People with extraordinary face recognition ability. *Psychonomic Bulletin & Review, 16*, 252-257.

Schweinberger, S. R., & Burton, A. M. (Eds.) (2011). Person perception 25 years after Bruce and Young (1986). *Special Issue of the British Journal of Psychology, 102*, 695-974.

Sutherland, C.A.M., Oldmeadow, J.A., Santos, I.M., Towler, J., Burt, D.M. and Young, A.W. (2013). Social inferences from faces: ambient images generate a three-dimensional model. *Cognition, 127*, 105-118.

Vernon, R. J. W., Sutherland, C. A. M., Young, A. W, & Hartley, T. (2014).

Modellng first impressions from highly variable facial images. *Proceedings of the National Academy of Sciences, USA, 111*, E3353-E3361.

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., & Bruce, V. (2011). Understanding person perception. *British Journal of Psychology, 102*, 959-974.

Young, A. W., & Burton, A. M. (1999). Simulating face recognition: implications for modelling cognition. *Cognitive Neuropsychology, 16*, 1-48.

Yovel, G., Wilmer, J. G., & Duchaine, B. (2014). What can individual differences reveal about face processing? *Frontiers in Human Neuroscience, 8*, 562, 1-9.

RECOMMENDED READING

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.

- A more detailed discussion of some of the key points made here.

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

- Shows that passport officers vary in their ability to perform unfamiliar face-matching tasks and that this variation is not related to experience or training.

Dowsett, A. J., Sandford, A., & Burton, A. M. (2016). Face learning with multiple images leads to fast acquisition of familiarity for specific individuals. *Quarterly Journal of Experimental Psychology, 69*, 1-10.

- Shows how image variability can enhance learning of specific faces.

Burton, A. M., Kramer, R. S. S., Ritchie, K. L., & Jenkins, R. (2016). Identity from variation: Representations of faces derived from multiple instances. *Cognitive Science, 40*, 202–223.

- Analyzes the variability across images of faces to demonstrate why each familiar face has to be individually learnt.

Bruce, V., & Young, A. (2012). *Face perception*. Hove, East Sussex: Psychology Press.

- A more extensive treatment that puts recognition into the context of a range of questions and issues concerning face perception.

Solutions to face matching problems

Figure 1. In Array A, the target image is face 3. The target is absent in Array B.

Figure 2. There are two identities, arranged as follows:

ABAAABABAB
AAAAABBBAB
BBBAAABBAA
BABAABBBBB