Article:

Young, Andrew William orcid.org/0000-0002-1202-6297 (2018) Faces, people and the brain:the 45th Sir Frederic Bartlett lecture. The Quarterly Journal of Experimental Psychology. pp. 1-26. ISSN: 1747-0226

https://doi.org/10.1177/1747021817740275

Author's final copy of manuscript accepted for publication in *Quarterly Journal of Experimental Psychology*.
Date of acceptance: 9th October 2017.

Faces, people and the brain: The 45th Sir Frederic Bartlett Lecture

Andrew W. Young
Department of Psychology
University of York
York YO10 5DD
England

Correspondence: Andy Young, Psychology, University of York, Heslington, York YO10 5DD, England.
email: andy.young@york.ac.uk
ORCID: 0000-0002-1202-6297

Keywords: face perception, person perception, face recognition, facial expressions, first impressions.

Acknowledgements

It is an honour and a privilege to be invited to give this lecture, not only for me personally but for its recognition of the field of face perception as a whole. During my lifetime this has become a well-established field of inquiry, linked to widespread interest and a substantial body of evidence. My own part in this has piggy-backed on the work of colleagues and students too numerous to mention here, but all deserving grateful acknowledgement of their contributions. Hopefully, they know who they are. None the less, particular thanks are due to Mike Burton, Tim Andrews, Stefan Schweinberger and Tony Atkinson for help in preparing the lecture and this paper.

Abstract

The fact that the face is a source of diverse social signals allows us to use face and person perception as a model system for asking important psychological questions about how our brains are organised. A key issue concerns whether we rely primarily on some form of generic representation of the common physical source of these social signals (the face) to interpret them, or instead create multiple representations by assigning different aspects of the task to different specialist components. Variants of the specialist components hypothesis have formed the dominant theoretical perspective on face perception for more than three decades, but despite this dominance of formally and informally expressed theories the underlying principles and extent of any division of labour remain uncertain. Here, I discuss three important sources of constraint. First, the evolved structure of the brain. Second, the need to optimise responses to different everyday tasks. Third, the statistical structure of faces in the perceiver's environment. I show how these constraints interact to determine the underlying functional organisation of face and person perception.

Nowadays, many papers on face perception begin by noting the range of social signals conveyed by the face. From looking at someone's face we can infer something about their age, gender, and ethnic background, their moods and feelings, and we form impressions of their state of health, attractiveness and even their personality. This is what makes faces so fascinating. No other stimuli carry such a wide range of meanings for us, though voices (Belin, Fecteau, & Bédard, 2004; Campanella & Belin, 2007) and bodies (de Gelder, 2016; Yovel & O'Toole, 2016) may well come close. Although many studies still focus on face recognition, the task of recognising whether or not we know someone is embedded in this much wider context of interpersonal perception and communication.

This means that we can also exploit the fact that the face is a source of diverse social signals and use face perception as a model system for asking important psychological questions about how our brains are organised. Given the range of inferences we can make from looking at someone's face, a key issue concerns whether we rely primarily on some form of generic representation of the common physical source of these social signals (the face) to interpret them, or instead create multiple representations by assigning different aspects of the task to different specialist components.

The theoretical paper by Bruce and Young (1986) was among the first to see the matter like this. Bruce and Young came down firmly on the side of the specialist components hypothesis, and proposed a rather baroque organisation that was mainly intended as a first stab at describing the theoretical landscape but has since proved surprisingly resilient (Schweinberger & Burton, 2011).

----- FIGURE 1 ABOUT HERE -----

The Bruce and Young (1986) model is shown in Figure 1. Broadly speaking, it proposed that following some form of initial perceptual encoding different aspects of facial signals such as identity and expression are analysed in parallel. It further emphasised the difference between processing the identities of familiar faces (recognised via 'face recognition units') and unfamiliar faces ('directed visual processing'). It also suggested some form of hierarchical organisation of components within each functional pathway.

We don't need to concern ourselves here with all of the details of Bruce and Young (1986), but it is worth reiterating its relationship to contemporary models of word and object recognition. First, ideas about naming mechanisms (Oldfield, 1966; Ratcliff and Newcombe, 1982) were represented in the different components of the pathway responsible for familiar face recognition. Second, Morton's (1979) revised logogen model inspired the idea of face recognition units (Hay & Young, 1982). Third, dual-route models of reading (Coltheart, 1981; Marshall & Newcombe, 1973) were reflected in the difference between familiar faces (cf. words) and unfamiliar faces (treated as like pronounceable nonwords in some respects), with the dual-route aspect leading to the difference between the 'face recognition unit' and 'directed visual processing' pathways.

Discussion of the strengths and weaknesses of Bruce and Young (1986) can be found elsewhere (Schweinberger & Burton, 2011; Young & Bruce, 2011). All that needs to be said here is that different variants of the specialist components hypothesis have formed the dominant theoretical perspective across the three decades since Bruce and Young (1986). What I want to address here, however, is not the detailed question of which model variant is the best fit to what we now know, but rather the deeper and I think more profound question of why our brains should be organised in this way. What is it that drives the underlying functional separation? This is a key issue, yet the question is seldom raised. When Calder and Young (2005) approached the question more than a decade ago in reviewing studies of the relation between recognition of facial identity and facial expression, the underlying principles and extent of any division of labour remained uncertain. Now I hope we can do a bit better.

There have been some attempts at relatively high-level explanations of reasons underlying functional specialisation. Fodor's (1983) views on modularity have been widely discussed, and studies in cognitive neuropsychology were strongly influenced by and supportive of assumptions of modularity or discrete functional pathways (Caramazza, 1984, 1986; Coltheart, 2017; Ellis & Young, 1988). One idea that Vicki Bruce and I found particularly interesting (though for some reason we didn't cite it in Bruce & Young, 1986) was Marr's (1982) suggestion that "any large computation should be split up into a collection of small, nearly independent, specialized subprocesses" (‘Vision’, 1982, p.325). Unless this is done, Marr pointed out that a change in one part of the system could have unintended consequences throughout.

My aim here is to see whether it is possible to offer something that can give a more detailed understanding than these overarching principles. I will discuss examples from the field of face and person perception that reflect three important sources of constraint, and show how and why each may exert its influence. First, the evolved structure of the brain. Second, the need to optimise responses to different everyday tasks. Third, the statistical structure of faces in the perceiver's environment. I realise that this tripartite division has some rough edges and overlapping parts, but I don't think this matters much for my primary purpose, which is to consider these sources of constraint in turn whilst showing how they may interact to determine the underlying functional organisation of the face perception system.

The evolved structure of the brain

Constraints that result from the evolved structure of the brain are perhaps the first that will spring to mind. Certainly they seem to be the most widely discussed and debated. We know that there is a long evolutionary background to human social abilities (Dunbar & Shultz, 2007), making it possible that there is some form of evolved neural substrate for face and person perception (Kanwisher, 2000). Indeed co-evolution between the structure of our faces and brains is clearly suggested by Waller, Cray and Burrows' (2008) finding of high consistency across individuals in the anatomy of facial muscles essential to producing what are often considered to be universal facial expressions, implying that these may be subject to a selection pressure for effective nonverbal communication. Similarly, it has also been suggested that the structure of human faces has evolved to signal individual identity (Sheehan & Nachman, 2014). Consistent with such ideas, recent studies have shown significant heritability and genetic specificity of face recognition ability (Shakeshaft & Plomin, 2015; Verhallen et al., 2014; Wilmer et al., 2010; Zhu et al., 2010).

In line with earlier and less precise behavioural methods (e.g. Rizzolatti, Umilta & Berlucchi, 1971; Young & Bion, 1980; Young & Ellis, 1976), functional brain imaging has revealed highly consistent differential regional responses to faces in the adult brain. The key technique was introduced by Kanwisher, McDermott and Chun (1997), who used fMRI to identify regions in individual participants' brains that were more highly responsive to faces than to many other visual stimuli. This 'functional localiser' technique has been used in many subsequent studies (Kanwisher, 2017). Even though the method is usually applied at the individual participant level, it consistently identifies regions (shown in the upper panel of Figure 2) in the lateral fusiform gyrus (often called the fusiform face area, or FFA), in inferior occipital gyri (the occipital face area, or OFA), and in the posterior part of the superior temporal sulcus (STS/pSTS) (Andrews, Baseler, Jenkins, Burton, & Young, 2016; Andrews, Davies-Thompson, Kingstone, & Young, 2010; Baseler, Harris, Young, & Andrews, 2014; Haxby, Hoffman, & Gobbini, 2000; Kanwisher et al., 1997; Kanwisher, 2017).

----- FIGURE 2 ABOUT HERE -----

There has been discussion as to whether these localisable regions themselves represent anatomically discrete functional components of the face perception system or instead correspond to the regions of peak activation in a more distributed system (Haxby, Gobbini, Furey, Ishai, Schouten & Pietrini, 2001), but either view accepts that there is a remarkably consistent organisation, and this has also been noted in other studies (Hasson, Nir, Levy, Fuhrmann, & Malach, 2004; Polk, Park, Smith, & Park, 2007).

An influential theoretical paper by Haxby, Hoffman and Gobbini (2000) suggested that OFA, FFA and pSTS form a core system involved in the visual analysis of faces, as shown in the upper panel of Figure 2. Whilst it undoubtedly involves both feedforward and feedback projections, Haxby et al. (2000) proposed that this core system has distinct functional pathways for dealing with properties of faces such as expression that can change from moment to moment (an OFA to pSTS pathway) and properties such as identity that are relatively invariant (an OFA to FFA pathway). This core system then acts in concert with other brain regions to form an extended system that allows further processing of facial information and integration with other sources.

As Figure 2 also shows, and Haxby et al. (2000) acknowledged, their model offers a synthesis of the more functional questions that concerned Bruce and Young (1986) with evidence concerning neural organisation. To emphasise the strong family resemblance to Haxby et al. (2000), the Bruce and Young (1986) model has been rotated through 90 degrees in Figure 2 (see Calder & Young, 2005). Importantly, however, the Haxby et al. model also shows how the organisation of face perception may dovetail with more general principles underlying brain organisation, such as dorsal and ventral streams (see Bruce & Young, 2012). Striking findings that fit Haxby et al.'s model have been made in cognitive neuroscience studies (Baseler et al., 2012; Hoffman & Haxby, 2000; Pitcher, 2014), though some questions have been raised too (e.g. Atkinson & Adolphs, 2011; Bernstein & Yovel, 2015; Duchaine & Yovel, 2015; Pitcher, Duchaine, & Walsh, 2014; Rossion, 2008).

None the less, generic evolved constraints can't offer the full picture, since we know that brain organisation is to some extent under the influence of learning and adaptation to the environment. This is evident in brain imaging studies demonstrating that although the core face-responsive brain regions are present from an early age, they clearly undergo protracted development (Golarai et al.,2007; Deen et al., 2017). We know too that although even newborn infants can show responses to face-like stimuli (Johnson, Dziurawiec, Ellis, & Morton, 1991), these innate abilities work together with the prevalence of faces in the visual world of young infants (Fausey, Jayaraman, & Smith, 2016; Jayaraman, Fausey, & Smith, 2017) to create conditions that promote effective learning (Johnson, Senju, & Tomalski, 2015; Morton & Johnson, 1991). Indeed, the role of learning and development is clearly evident in many infant studies, including the interesting work on perceptual narrowing showing that infants become less adept at remembering pictures of monkey faces as they become more experienced with human faces (Kelly et al., 2007; Pascalis, de Haan, & Nelson, 2002; Pascalis et al.,

2005) in a way that forms an interesting potential parallel to perceptual narrowing in language development (Pascalis, Dole, & Loevenbruck, 2017; Vihman, 2017). Then of course as adults we remain susceptible to effects of previous experience, as shown for example by other-race effects (Rossion & Michel, 2011; Yan, Andrews, & Young, 2016; Yan, Andrews, Jenkins, & Young, 2016; Yan, Young, & Andrews, 2017a, 2017b) whose origin can again be traced back to experience in early years (e.g. Kelly et al., 2007; Lee, Quinn, & Pascalis, 2017; Tham, Bremner, & Hay, 2015, 2017).

Whilst the evolutionary background clearly offers an important source of constraint, then, we need to understand how it interacts with other factors. The demands of everyday life offer important examples.

The need to optimise responses to different everyday tasks

I will use recognition of face identity and facial expression to illustrate how everyday tasks can shape the optimal organisation of perceptual mechanisms. Identity and expression offer a good example to discuss because they were so clearly flagged as distinct by Bruce and Young (1986) and Haxby et al. (2000).

Bruce and Young's (1986) starting point was that a core requirement of face recognition is to be able to recognise the faces of people we know across different expressions, whereas a core requirement of facial expression recognition is to be able to recognise expressions across different identities. So at some level there must be some degree of separation between mechanisms involved in recognising identity and expression. Otherwise we would find ourselves susceptible to errors that would severely affect our everyday lives, such as failing to recognise a familiar face with an unusual expression. This doesn't happen; even highly unusual expressions such as those created by the Thatcher illusion have little impact on our ability to recognise familiar face identities (Psalta, Young, Thompson, & Andrews, 2014a, 2014b; Thompson 1980; Thompson, Anstis, Rhodes, Jeffrey, & Valentine, 2009).

Some findings suggest that this separation between identity and expression arises at relatively early stages of perception, because there are clear differences in the susceptibility of identity and expression recognition to different stimulus transforms. For example in behavioural studies contrast negation of photographs makes it very hard to recognise face identity, but has only a limited effect on recognising facial expressions (Bruce & Young, 1998; White, 2001). Parallel findings have been noted in fMRI, where responses from the FFA are sensitive to contrast polarity changes whilst responses from pSTS are relatively insensitive (Harris, Young, & Andrews, 2014a).

Such data imply differences in the types of information that are used to analyse identity and expression. More specifically, they have been taken to suggest that surface texture patterns that are disrupted by contrast negation are critical to recognising face identity and that feature shapes are important to interpreting facial expression. These differences, though, are clearly relative rather than absolute, since manipulations that minimise variation in surface texture or differences in feature shapes can both affect facial expression recognition, showing that both types of information can contribute (Sormaz, Young, & Andrews, 2016b).

A useful exercise is to look more closely at the everyday demands of face identity and expression recognition. I already mentioned that a core requirement is to be able to recognise familiar faces across different expressions and to recognise facial expressions across different identities, but there are other important differences too. Consider the complexity of the task. Most of us can recognise hundreds (perhaps thousands) of familiar faces. In contrast, although we don't know how many distinct facial expressions exist, and a complicating factor is that some expressions seem to be universal whilst others are more culture-specific, most theories put the number of recognisable expressions well below a hundred. However, as well as recognising this relatively small number of expressions we need to interpret their intensities; seeing whether someone is a bit frightened or very frightened can be a critical difference. Moreover, it is important to note that blends of different expressions can be meaningful (Du, Tao, & Martinez, 2014). Comparable requirements mostly don't exist for face identity recognition, where differences between images of the same face are identity-irrelevant (a point whose implications will be looked at in more detail later).

The implications for our behaviour are different, too. Recognising someone's identity allows you to access previously stored semantic and episodic information that facilitates appropriate interaction. In contrast, emotional expressions modulate ongoing priorities (Oatley & Johnson-Laird, 1987, 2014); if someone looks afraid or angry, you immediately interrupt what you were doing and search for a reason. Finally, the temporal demands of identity recognition are relatively low, because once you have recognised someone their identity doesn't change during a social encounter. In contrast, expression recognition has high temporal demands (Young & Bruce, 2011); someone's mood can change in an instant, and such changes need constantly to be monitored.

How do these pervasive differences in everyday task demands influence the functional organisation we have noted? I will approach this question by looking at the implications of neuropsychological findings. In line with burgeoning interest in the 1980s in using neuropsychological evidence to test and refine cognitive models (Ellis & Young, 1988; Shallice, 1988), Bruce and Young (1986) placed considerable emphasis on the consequences of brain injury for recognising face identity and expression. From the limited evidence available to them, they concluded that there was something akin to a neuropsychological double dissociation between impairments affecting face identity and facial expression. This turned out in part to be a misrepresentation, but the reasons why it misrepresented things have proved important.

Let's start by considering neuropsychological impairments of face identity recognition. The most widely-documented of these is prosopagnosia, with case descriptions dating back to the Nineteenth Century (for examples, see Della Sala & Young, 2003; Ellis & Florence, 1990; Young & van de Wal, 1996). The loss of ability to recognise familiar faces that is a key defining symptom of acquired prosopagnosia due to brain injury has a number of consistent characteristics (Meadows, 1974; Hécaen 1981; Young, 2011). It is severe, such that even close family members may not be recognised. It is pervasive, with nearly all faces being affected and no previously recognised categories spared (for example, there are no reported cases of patients who can still recognise politicians but can't recognise television personalities they knew before their brain injury). It is specific to the visual modality, with familiar people still being recognised from non-facial cues such as their voices (Liu et al., 2016) or names (Young, Hellawell, & de Haan, 1988). Finally, it is selective to face recognition, in the sense that other aspects of face perception (including recognition of expression) may be less severely compromised (Bruyer et al.,1983; Riddoch, Johnston, Bracewell, Boutsen, & Humphreys, 2008).

In sum, cases of acquired prosopagnosia show that face recognition is to some extent dissociable from other aspects of face perception and from recognition based on the person’s voice or name. They form an interesting contrast with cases involving loss of memory for people, in which a more central semantic deficit leads to failure to recognise familiar people from their face, voice, or name (de Haan, Young, & Newcombe, 1991; Ellis, Young, & Critchley,1989; Hanley, Young, & Pearson, 1989). In general, the brain lesions associated with prosopagnosia are located in relatively ventral and posterior occipito-temporal cortex whereas loss of memory for people follows more anterior temporal lobe damage (Barton & Corrow, 2016; Gainotti, 2014). In fMRI, image-invariant familiar face adaptation also involves relatively anterior regions (Weibert et al., 2016).

Based on evidence available at the time, Bruce and Young (1986) tended to think of neuropsychological impairments of facial expression recognition as having a complementary pattern to acquired prosopagnosia. That is expression recognition impairments were expected to be face-specific (with recognition of non-facial expressive cues being less affected), to affect the recognition of all facial expressions, and to have little effect on ability to recognise face identity. These complementary patterns would, of course, create what Shallice (1988) called a strong double dissociation between impairments of face identity and facial expression recognition.

Some later studies using careful techniques initially tended to support these views (Parry, Young, Saul, & Moss, 1991; Young, Newcombe, de Haan, Small, & Hay, 1993; Young et al., 1995), but with the benefit of hindsight they suffered two limitations. First, they only used overall scores for facial expression recognition, and didn't disaggregate the different emotions used in tasks. Second, they didn't test expressions of emotion outside the facial domain.

The first limitation was dramatically exposed in a study by Adolphs, Tranel, Damasio and Damasio (1994). Their participant, case SM, had suffered bilateral calcification of the amygdala due to Urbach-Wiethe disease. Rather than measuring SM's overall ability to recognise facial expressions, however, Adolphs et al. (1994) used an approach which showed that not all emotions were equally severely affected. Instead, SM's recognition of facial expressions of fear was particularly poor.

At the time, my colleagues and I were investigating a case that also involved selective bilateral amygdala damage, but with a different aetiology resulting from surgery for relief of otherwise intractable epilepsy (Young et al., 1995). We had found that our participant, DR, showed relatively well-preserved recognition of familiar face identity but experienced problems in perceiving direction of gaze and recognising facial expressions. What we had not done was to look separately at her recognition of different emotions.

----- FIGURE 3 ABOUT HERE -----

To produce a sensitive test of facial expression recognition we developed an 'emotion hexagon' procedure, as shown in Figure 3 (Calder et al.,1996; Young, Perrett, Calder, Sprengelmeyer, & Ekman, 2002). This took the images of prototype expressions of six basic emotions (happiness, surprise, fear, sadness, disgust, and anger) for one of the models from the Ekman and Friesen (1976) series of facial expressions, located these expressions around the perimeter of a hexagon in which each emotion was placed next to those it was most likely to be confused with, and then created 30 morphed expression images that traversed this perimeter. By presenting these 30 images in random order and asking which of the 6 emotions each was most like, we could chart DR's recognition of emotion across different levels of task difficulty (as reflected in the degree to which the test image was morphed away from the nearest prototype expression). The results (see Figure 3) revealed that DR showed normal recognition of happiness and sadness and particularly poor recognition of fear and anger (Calder et al., 1996). A comparable test of face identity recognition based on a morphed hexagon of familiar faces showed that DR had no difficulty with this, demonstrating normal recognition of familiar identity (Calder et al., 1996).

The findings of Adolphs et al. (1994) and Calder et al. (1996) fitted neatly together to show clearly that facial expression recognition impairments can be to some extent category-specific, affecting the recognition of some emotions more than others. What these studies didn't do, though, was to test recognition of emotions from sources other than facial expression. When we tested DR's recognition of auditorily expressed emotions (Scott et al., 1997), we again found particularly poor recognition of fear and anger; exactly the same emotions as were poorly recognised from facial expressions. Multi-modal deficits of emotion recognition affecting recognition of fear were also found in another case of bilateral amygdala damage reported by Sprengelmeyer et al. (1999). In fact, deficits in the experience of the emotion of fear itself were also evident in bilateral amygdala cases that we studied (Broks et al., 1998; Sprengelmeyer et al., 1999). Similarly, although an early report suggested intact recognition of emotion from speech prosody (Adolphs & Tranel, 1999), later follow-ups of SM have shown impaired recognition of emotion from music (Gosselin, Peretz, Johnsen, & Adolphs, 2007) and highly atypical experience of fear (Feinstein, Adolphs, Damasio, & Tranel, 2011).

The effect of bilateral amygdala damage on fear can be contrasted with other neuropsychological cases that show selective problems with disgust following damage to the insula and putamen (Calder, Keane, Manes, Antoun, & Young, 2000) or problems encompassing all emotions in frontal variant frontotemporal dementia (Keane, Calder, Hodges, & Young, 2002; Van den Stock et al., 2015). These deficits typically encompass not just face and voice perception, but interpretation of body cues as well. Experience of emotion can also be affected.

Converging evidence from fMRI has also emphasised the contribution of different brain regions to different emotions (Calder, Lawrence, & Young, 2001; Morris et al.,1996; Phillips et al., 1997), but for now I want to focus on the implications of the fact that the patterns of neuropsychological impairment of emotion recognition are so unlike those we were seeking. Using the analogy with prosopagnosia, we expected to find problems affecting recognition of all facial expressions but leaving the recognition of emotion from voices or bodies relatively intact. Instead, what has gradually come to light is that emotion recognition impairments are intrinsically multimodal but at the same time they can affect some emotions more than others. We have to ask why this should be the case.

An appealing hypothesis is again that the patterns of breakdown reflect an underlying organisation that is strongly influenced by everdyay task demands (Young & Bruce, 2011). Consider identity recognition. This will often involve a unimodal input (the person may be seen but not heard), but identity does not change from moment to moment during a social encounter (once you have recognised someone, you don't need to keep doing it). Hence a unimodal system can offer a good solution. For emotion recognition, however, things are quite different. We need constantly to monitor rapidly changing signals (someone's mood can change in an instant), these signals can be of high priority (for example, reflecting different types of threat), and the cues to emotion are often simultaneously expressed across multiple channels (face, voice, and body). Hence it is adaptive that emotion recognition seems to be to some extent organised around emotion categories and to be able to make use of multi-modal inputs (Carroll & Young, 2005; Young, 2016). In these respects, it differs markedly from familiar face recognition in ways that are consistent with the differing demands of these everyday tasks.

Much the same point is evident from behavioural repetition effects. Whilst recognising a face leads to a long-lasting facilitation for recognising the same face again (Bruce & Valentine, 1985; Ellis, Young, Flude, & Hay, 1987), recognising the face's expression confers no long-term advantage to recognising that expression at a later date (Ellis, Young, & Flude, 1990). This otherwise puzzling pattern can again be seen to reflect the fact that it is useful to retain some degree of facilitation for recognising the identities of recently encountered faces, in case they are encountered again. However, it would be counterproductive to expect that someone will always have the same expression, as this would interfere with detecting potentially important changes in mood; hence the absence of long-term priming of expression recognition. Again, the characteristics of human face perception seem to fit the environmental demands (see also Taubert, Alais, & Burr, 2016).

This point is underlined if we look more closely at the multi-modal properties of emotion recognition and compare them to speech perception as another example in which concurrent cues to a rapidly changing signal are often available from our voices (speech sounds) and faces (the movements needed to produce each sound). Although we are used to a 'common sense' way of thinking about speech perception entirely in terms of decoding the acoustic signal, and of course we know that purely acoustic analyses can support speech perception when we listen to a radio or talk to someone on the telephone, there is none the less substantial evidence that seeing someone's facial movements can make an important contribution.

The most well-known example is the McGurk illusion (McGurk & MacDonald, 1976), in which a video showing the face of a person saying one phoneme (for example, "ga") is combined with a different phoneme (for example, "ba") on the soundtrack. Remarkably, the heard phoneme can then correspond neither to the auditory nor the visual part of the video, but is usually a fusion of the two (heard as "da" in the example we used). The illusion shows that in hearing what someone says we make use of the correspondence between movements of their lips (and tongue) and the speech sounds.

An important clue to why this happens comes from a classic study of speech perception in noise reported by Miller and Nicely (1955), who noted a substantial improvement when the speaker's face was visible. In considering the cause of this effect, they noted that ‘The place of articulation, which was hardest to hear in our tests, is the easiest of features to see on a talker's lips. The other features are hard to see, but easy to hear’ (Miller & Nicely, 1955, p. 352). Considered more generally, it seems that because speech signals involve rapid temporal changes that have to be decoded as they occur, integrating complementary information from face and voice offers an optimal way of dealing with these temporal constraints. Studies of infants suggest that sensitivity to these audio-visual correspondences begins early in life (Kuhl & Meltzoff, 1982; Patterson & Werker, 2003).

Functional brain imaging studies offer an important contribution here by identifying brain regions that are involved in lipreading. Calvert and her colleagues (Calvert, Campbell, & Brammer, 2000; Calvert Hansen, Iversen, &

Brammer, 2001) introduced the stringent criterion that an unequivocally multi-modal region will show a supra-additive response to audio-visual stimuli. That is, a multi-modal brain region will show more activation to an audiovisual stimulus than the sum of its responses to purely auditory stimuli or to purely visual stimuli. A region that has repeatedly been found to meet Calvert's supra-additive criterion in audio-visual integration studies using talking faces is located in the vicinity of the left posterior superior temporal sulcus and perhaps left superior temporal gyrus (Calvert, 2001). The importance of left pSTS to audio-visual integration has been confirmed by demonstrating that TMS to this region disrupts the McGurk effect (Beauchamp, Nath, & Pasalar, 2010).

Even so, posterior left STS should not be overinterpreted as the only region involved in audiovisual integration of speech; it clearly forms part of a larger network that is apparent in studies that have used different criteria (Hall, Fussell, & Summerfield, 2005; Macaluso, George, Dolan, Spence, & Driver, 2004; Wright, Pelphrey, Allison, McKeown, & McCarthy, 2003). This network includes other regions along the superior temporal sulcus and superior temporal gyrus that include classical auditory areas. The point of principal interest here, however, is that the region of left pSTS that shows a supra-additive response to audiovisual speech is close to or likely part of Haxby et al.'s (2000) core system that they identify as heavily involved in perceiving changeable aspects of faces.

With this background of key facts about audio-visual integration in speech perception in mind, we need to look again at the possibility of audio-visual integration in emotion recognition. The first thing that needs to be said is that although our moods can change from moment to moment, the speed and complexity of such changes is unlikely to approach the demands of speech perception. However, there is behavioural as well as neuropsychological evidence consistent with a multi-modal contribution to emotion recognition (de Gelder, Stienen, & Van den Stock, 2013; Lewis, Lefevre, & Young, 2016; Schirmer & Adolphs, 2017). For example, de Gelder and Vroomen (2000) asked participants to classify morphed images of faces from a happy to sad continuum as happy or sad, but with each face presented either on its own or accompanied by a semantically neutral sentence read with a happy or sad tone of voice. Participants were instructed to ignore the voice, but they were unable to do this; compared to the 'no voice' baseline the proportion of sad classifications was increased by the sad voice and decreased by the happy voice. A further experiment by de Gelder and Vroomen (2000) showed that the interaction between faces and voices is bidirectional, with facial emotion interfering with vocal emotion and vice versa.

Such findings are not exactly the same as the McGurk effect; the result can be interpreted as a bias rather than a novel emotion percept. In this respect, neuroimaging studies again offer important evidence. Hagan et al. (2009) used purely nonverbal audio-visual stimuli (Ekman faces and nonverbal sounds) in MEG to demonstrate a supra-additive response to audiovisual emotion from right STG/STS; this had a clear posterior focus but also included much of the right STS. Because of the excellent temporal resolution of MEG, this activation could be seen within 200 ms of stimulus onset, as shown in Figure 4. This supra-additive response to audio-visual emotion is therefore centred on a region in the right hemisphere that is opposite the left hemisphere region that responds supra-additively to audio-visual speech, and its early onset (within 200 ms) implies that it reflects an involuntary integrative mechanism.

----- FIGURE 4 ABOUT HERE -----

A follow-up MEG study by Hagan et al. (2013) used a combination of faces with neutral words spoken with emotional prosody to again reveal a fast supra-additive response to audio-visual emotional stimuli from right posterior STG/STS, showing that this region responds to audio-visual emotion even in a context involving spoken words. It seems that, like speech perception, emotion recognition is to some extent a multi-modal phenomenon, and that whereas the left posterior STS/STG region is implicated in audio-visual integration of speech (Calvert, 2001) an equivalent region on the right side of the brain is implicated in audio-visual integration of emotion. As for lipreading, we should note that this is likely to be only a part of a more extensive network for multi-modal analysis (Park et al., 2010). None the less, it seems to form a critical part of this network and this right posterior STG/STS region is again close to or overlapping with Haxby et al.'s (2000) posterior STS region hypothesised to be involved in analysing changeable aspects of faces. However, although forming part of Haxby et al.'s 'core visual system' for faces, posterior STS also seems to be a key component of a system involved in biological motion perception (Hein & Knight, 2008; Pelphrey, Morris, Michelich, Allison, & McCarthy, 2005), predictive coding (Johnston et al., 2017) and multi-modal integration as well as purely visually-driven responses to faces (Calder and Young, 2005; Peelen, Atkinson, & Vuilleumier, 2010).

Although the details of the various studies I have mentioned may seem complex, the key point I want to emphasise here is the way that these different findings fit together from an overarching perspective. Like speech perception, emotion recognition has temporal characteristics that make it useful to take advantage of any cross-modal cue complementarity, and as for speech perception we find that a brain region that shows an involvement in multi-modal responses more generally plays an important role. The everyday task demands and the evolved structure of the brain work in concert to determine the functional organisation.

The statistical structure of faces in the perceiver's environment

I mentioned already that studies of infant learning show perceptual narrowing toward human faces as the baby acquires more experience with these. Interestingly, the extent of this perceptual narrowing can be reduced if non-human faces are more strongly present in the environment. Both phenomena show how the faces that are seen can shape perceptual organisation (Kelly et al., 2007; Pascalis et al., 2002, 2005). Whilst it seems likely that this ability is never entirely lost, it may well be reduced in adulthood in much the same way that adults experience difficulty with many things (such as learning a new language) that would have been easier earlier in life. We know, for example, from numerous studies of other-race effects that such influences can be difficult to modify in adulthood (Rossion & Michel, 2011). None the less, perceptual adaptation effects also show the possibility of short-term recalibration within the boundaries of what our visual systems have already experienced as well-known (Hsu & Young, 2004; Webster, Kaping, Mizokami, & Duhamel, 2004). Such studies show something of the ability of our perceptual systems to adapt to the 'diet' of the faces in the current environment (Rhodes, 2017; Webster & MacLeod, 2011), though the precise relation between short-term and long-term changes remains uncertain.

However, the point I want to emphasise now is different from these more general observations of plasticity and its limits. It is becoming clear that the image properties of faces themselves can influence how optimally to extract different types of information. This has profound implications.

Let's return to the question of the relation between recognising face identity and facial expression. We already saw that there are abundant differences in the demands of these everyday tasks that have a substantial impact on how best to achieve them. However we also noted that there are potential differences in the types of information that our visual systems use to analyse identity and expression (Harris et al., 2014a; Sormaz et al., 2016b), and I want to explore these a bit more carefully.

We can think of a face photograph, or any image of a face that falls on the retina of our eyes, as involving two distinct properties (Bruce & Young, 1998, 2012; Sutherland, Rhodes, & Young, 2017a). First, it represents 2D facial shape; the positions and shapes of features (eyebrows, nose, mouth, chin, etc.) as they are located in the image. Second, it represents the facial surface; the brightness and colour of features, skin and hair, including shading cues to 3D facial shape from ambient lighting. Computer image manipulation methods often begin by establishing the location of a set of landmark fiducial points that define the image's 2D shape and then use these fiducials to subdivide the image into a large number of smaller regions that can represent surface colour and brightness values (Tiddeman, Burt, & Perrett, 2001; see Sutherland et al., 2017a).

Using similar methods Calder, Burton, Miller, Young and Akamatsu (2001) investigated the image statistics that underlie representations of facial identity and expression through a principal component analysis (PCA) of the shape and surface properties of images from the Ekman and Friesen (1976) series of 'Pictures of facial affect'. In this context image shapes are defined through the 2D fiducial locations in each photograph, and surface properties can then be compared by reshaping all of the photographs to a common (averaged) set of fiducial positions. PCA is a data reduction technique that finds a set of principal components (PCs) that best accounts for the observed variations in shape and surface properties across the different images.

Calder et al. (2001) analysed the Ekman and Friesen (1976) photographs because these are a well-validated set used in many previous studies and they include expressions of different emotions and different models (identities). Having used PCA to find PCs of shape and surface variation, Calder et al. used linear discriminant analysis (LDA) to try to decode the expression or the identity of the faces from the photographs in the set. They found that different combinations of PCs can be used successfully to decode expression or identity, but with two important caveats. First, multiple PCs are always needed. Second, whilst some PCs are mainly useful for decoding expression and some are mainly useful for decoding identity, other PCs are useful for both identity and expression.

Before discussing the implications of this, we need to note that it seems unlikely that the brain itself uses PCA as a perceptual mechanism. Its purpose here is simply that the combination of PCA and LDA offers a convenient way to explore any underlying image statistics the brain might be able to exploit. Being able to do this is important in its own right.

From this perspective, Calder et al.'s (2001) data are clear. A simple way to think of the findings is that they show that expression and identity can be represented through combinations of PCs, but not in a fully exclusive manner. Whilst there are differences between the visual properties underlying identity and expression, there is also substantial covariation.

These findings concerning image properties may help explain data that are inconsistent with a complete segregation between facial identity and expression processing. Many studies have sought to explore the limits of the separation between identity and expression. They have shown that some cross-talk between facial identity and expression does seem to occur in certain circumstances. For familiar faces, characteristic expressions can slightly benefit recognition (Kaufmann & Schweinberger, 2004). For unfamiliar faces interference in the Garner paradigm reveals that changes in face identity can influence judgements about facial expression (Schweinberger & Soukup, 1998; Schweinberger, Burton, & Kelly, 1999; Atkinson, Tipples, Burt,

& Young, 2005). In a parallel with findings based on Garner interfence, adaptation paradigms show that adaptation to expression is influenced by a change in identity (Ellamil, Susskind, & Anderson, 2008; Fox & Barton, 2007; Pell & Richards, 2013; Rhodes et al., 2015) whereas adaptation to identity is less affected by a change in expression (Fox, Oruc & Barton, 2008). In our previously mentioned case study of the effect of bilateral amygdala lesions (Calder et al., 1996; Young et al., 1995), participant DR experienced difficulties in unfamiliar face identity matching tasks when the faces' identities were discrepant with their expressions; for example, in deciding that two photographs with different expressions showed the same face (Young, Hellawell, van de Wal, & Johnson,1996). Similarly, interactions between identity and expression have also been noted in an early component of ERPs from neurologically normal participants (Fisher, Towler, & Eimer, 2016). The underlying covariation between some aspects of facial identity and expression revealed by PCA helps us to understand why these interactions may occur.

Of course, the potential role of facial movements also needs to be considered. Although a particular strength of the Ekman and Friesen (1976) stimuli is that they are grounded in a careful analysis of the muscle movements involved in producing different expressions, as photographs they don't contain any information that might derive from the timing of the movements themselves. We need to keep in mind that the fact that facial expressions shown in photographs can be recognisable does not mean that movement is unimportant. On the contrary, there is evidence that the timing of facial movements is carefully balanced between the needs of the sender and the intended recipient, even for a facial signal as apparently simple as raising the corners of the mouth in a smile (Leonard, Voeller, & Kuldau, 1991). For expressions that are too subtle to be easily seen in static displays, too, a role for patterns of movement has been found; movement can draw attention to small but critical changes (Ambadar, Schooler, & Cohn, 2005). However, the good recognition of photographs of normal intensity basic emotions such as those used by Ekman and Friesen (1976) shows, for these emotions at least, either that the apex of the set of muscle contractions forms a recognisable configuration of the facial features or that we are very skilled at estimating the implied motion (Martinez, 2003). Likewise, studies I have been involved with have not found much in the way of differences between moving and static expressions of basic emotions (Johnston, Mayes, Hughes, & Young, 2013; Harris, Young, & Andrews, 2014b). So, despite its intuitive appeal, we need also to be careful not to overstate the role of movement in facial expression recognition (Young, 2016).

Further evidence of the usefulness of image statistics can be found if we consider the perceptual similarity between different facial expressions. Some expressions, for example surprise and fear, look more similar than others, such as surprise and disgust. This perceptual similarity was of course the underlying principle used to create the emotion hexagon shown in Figure 3, but it can be expanded by considering the similarity between every possible pairing of expressions of basic emotions. Importantly, image statistics concerning the shape and surface properties of the images can be used to model perceptual similarity ratings for facial expressions (Sormaz et al., 2016a); the expressions people judge as being more similar have more similar shape and surface properties. Interestingly, it also turns out that neural responses recorded with fMRI from Haxby et al.'s (2000) core regions of OFA and pSTS track the perceptual similarity of expressions (Sormaz, Watson, Smith, Young, & Andrews, 2016a), with more similar patterns of activation in these regions to the more similar expressions. These core regions represent Haxby et al.'s (2000) pathway for analysing changeable aspects of faces, but in Sormaz et al.'s (2016a) study their responses were measured with purely static stimuli.

Understanding image statistics, then, offers a valuable perspective on how we recognise facial identity and expression, and exemplifies a distinct type of constraint on perceptual mechanisms. But I think that similar approaches have wide applicability, and I will make what at first seems like a digression to demonstrate the point in studies of facial first impressions.

Our impressions of other people from their appearance have long attracted interest, and were of course central to ideas now usually dismissed as pseudoscientific, such as the overstated claims of the physiognomists (see Bruce & Young, 1998, 2012; Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015; Todorov, 2017). Nowadays, few would claim that someone's character and other traits can validly be inferred from their facial appearance, though it remains possible that there is a 'kernel of truth' that leads to slightly better than chance performance for some attributions (Berry, 1990; Kramer & Ward, 2010; Penton-Voak, Pound, Little, & Perrett, 2006).

Although they are of questionable validity, however, we form surprisingly consistent subjective impressions of traits such as friendliness or intelligence from images of faces. So how do we do this? The question is of some practical as well as theoretical interest because of the remarkable proliferation of online relationships and interactions in which a posted photo may form a primary source of information about someone you haven't actually met; for example, Meetic Group (a parent company owning online dating agencies) claims to have had more than six billion unique visitors to its websites and to have introduced six million European couples (Meetic Group, 2017).

To try to capture the full range of potential cues my colleagues and I have adopted a data-driven approach to understanding first impressions of faces (Oldmeadow, Sutherland, & Young, 2013; Santos and Young, 2005, 2008, 2011; South Palomares & Young, 2017; Sutherland et al., 2013, 2015a; Sutherland, Oldmeadow, & Young, 2016; Sutherland, Young, Mootz, & Oldmeadow, 2015b). We began by collecting 1,000 everyday face images from the internet. These were all non-famous Caucasian adults (to avoid other-race effects in initial studies) and the photos were as varied as possible, with no constraints on ages (other than being of adult appearance), expressions, poses, accessories (facial hair, piercings, glasses), or lighting and image quality. We then had each face image rated on several characteristics; its apparent trustworthiness, intelligence, attractiveness, and so on. These characteristics were subjectively, not objectively defined, in the sense that we didn't actually know how trustworthy or intelligent each person might really be, but were instead interested in how trustworthy or intelligent they *looked* in that photo.

The first thing we found was that inter-rater agreement was good, even with these unstandardised images. Participants agree with each other as to which images look trustworthy or whatever, in line with findings using more controlled images (Oosterhof & Todorov, 2008; Todorov et al., 2015). Moreover, this agreement seems to reflect the fact that participants make use of consistent cues. By averaging images rated high or low on a given trait we can create prototypes that seem to represent its essential characteristics, as shown in Figure 5 (from Sutherland et al., 2013). This averaging can only 'work' well if the individual images contain consistent cues or combinations of cues; inconsistent or idiosyncratic cues will largely be averaged out (Sutherland et al., 2017a). By morphing between the high and low prototypes we could create graded levels of each trait (see Figure 5), which again demonstrates that key cues have been successfully captured.

----- FIGURE 5 ABOUT HERE -----

Being able to visualise the cues that create different impressions represents a step forward, but we were then able to build on pioneering work by Oosterhof and Todorov (2008) to capture the underlying structure of these subjective impressions. Oosterhof and Todorov (2008) had applied PCA to ratings encompassing several different traits and found that these could be modelled as falling along two underlying dimensions that could be approximated by perceived trustworthiness and perceived dominance. From a factor analysis of ratings of 13 traits across our 1,000 highly varied face photographs we identified a three-factor solution (Sutherland et al., 2013).

The first factor identified by Sutherland et al. (2013) corresponds to approachability; it is similar to the trustworthiness factor found by Oosterhof and Todorov (2008). The second factor we labelled youthful attractiveness. The third factor involves perceived dominance; it is comparable to Oosterhof and Todorov's (2008) second factor.

Sutherland et al.'s (2013) three-factor model is thus an expanded version of Oosterhof and Todorov's (2008) two-factor scheme. In Oosterhof and Todorov's (2008) study, perceived attractiveness resulted from a combination of their two factors, whereas in Sutherland et al. (2013) it emerged as a factor in its own right. The reason is almost certainly because Sutherland et al. (2013) used faces from a wider range of ages than Oosterhof and Todorov (2008), underlining the importance of the nature of the stimulus sample to data-driven approaches (Todorov et al., 2015; Todorov, 2017).

Sutherland et al.'s (2013) factors are visualised in Figure 5 by creating factor loadings for each image and then averaging the 20 highest and 20 lowest images for each factor. Like the averages for specific traits shown in Figure 5, these averaged images representing each factor show that each factor involves multiple interacting cues. This has important implications for understanding facial first impressions. For example, the averaged face-like images representing high levels of approachability depict smiling individuals, whereas the images representing low levels contain more neutral or even slightly hostile expressions, which is consistent with previous research (Hess, Adams, & Kleck, 2004; Said, Sebe, & Todorov, 2009). However, the face-like averages also show clearly that smiling does not in itself constitute an exclusive cue to approachability. Instead, high versus low levels of all three factors involve differences in smiling; what is important is possibly the type of smile, and certainly the way smiling is combined with other cues such as skin tone, age, and face shape. To understand first impressions from faces, then, we will need to understand how different cues are interpreted in combination with each other (Santos & Young, 2011; South Palomares & Young, 2017; Todorov, 2017) instead of investigating each cue in isolation. The same point had been made in seminal studies by Secord (1958) and is of course in line with evidence of the importance of holistic processing in many aspects of face perception (Abbas & Duchaine, 2008; Calder, Young, Keane, & Dean, 2000; Chen, Ren, Young, & Liu, 2017; Rossion, 2013; Sormaz, Andrews, & Young, 2013; Tanaka & Farah, 1993; Todorov, Loehr, & Oosterhof, 2010; Young, Hellawell, & Hay, 1987).

This point about the importance of combinations of cues became very clear when we sought to model subjective impressions from objective physical attributes (Vernon, Sutherland, Young, & Hartley, 2014). Using the same 1,000 everyday images, Vernon et al. (2014) created loadings for each image on Sutherland et al.'s (2013) three underlying factors from a factor analysis of 16 traits rated by human observers. They then positioned 179 fiducial points onto each face and used these together with colour and brightness values to objectively define 65 physical ‘attributes’ such as feature positions, feature shapes, skin and lip colour. These 65 physical attributes were then used as inputs to a linear network that was trained to fit the three Sutherland et al.

(2013) factors (its outputs) to a subset of the 1,000 images, and the network's performance was evaluated through its ability to predict human ratings for untrained images (Vernon et al., 2014). This procedure was repeated a number of times, to arrive at estimates of the network's ability to predict the approachability, youthful attractiveness and dominance of every one of the 1,000 face images.

----- FIGURE 6 ABOUT HERE -----

Correlations between the factor scores from human ratings and scores predicted from a linear network that had been trained on entirely different images are shown in Figure 6. Impressively, the linear network was able to explain 58% of the variance in the evaluations of these highly varied face images by human observers. Moreover, in this case the performance of networks capable of finding non-linear relationships showed no overall improvement.

When Vernon et al. (2014) looked at which physical attributes are critical to different types of impression, they found that most of the 65 attributes were significantly correlated with more than one factor. This reinforces the implications from Sutherland et al.'s (2013) findings that it is how different features are combined with each other that is critical; the same feature can mean different things in different combinations.

The implications of these simulations are that our first impressions of faces can largely be driven by making use of covariation among available cues. Analysing the image statistics again gives us a clearer picture of the brain's task (see also Dotsch, Hassin, & Todorov, 2016). Although this represents significant progress in understanding the mechanisms involved, we still have a relatively poor understanding of the underlying factors that make these impressions so pervasive in our everyday lives. At one extreme, it is tempting to postulate an evolutionary background. For example, evaluations of trustworthiness may involve the perceived intention of a conspecific to help or harm you, dominance involves capability to carry out these intentions, and attractiveness may be a signal of genetic fitness (Oosterhof & Todorov, 2008; Sutherland et al., 2013). But a powerful alternative is that the basis lies not so much in our primate ancestry as in the overgeneralisation of contingencies present in our everyday environments (Secord, 1958; Zebrowitz, 2017) and from social stereotypes which may themselves reflect weak environmental contingencies (Sutherland et al., 2015). Of course, these are not mutually exclusive possibilities, but one promising way to tease apart their contributions may again be through neuropsychological studies. For example, Sprengelmeyer et al. (2016) found that atypical first impressions of faces by patients with Huntington's disease were correlated with impairments affecting recognition of facial expressions; a finding that is very much in line with Secord's (1958) and Zebrowitz's (2017) ideas.

Another finding that fits in well with Zebrowitz's (2017) general approach in terms of overgeneralisation from perceived cues concerns the importance of within-person image variability. It is very tempting to treat differences in perceived trustworthiness, attractiveness or dominance as properties of the *face* being viewed, but often they are just as much properties of a specific *photograph*. This was emphasised by Jenkins, White, Van Montfort and Burton (2011), who showed that differences in the rated attractiveness of different photographs of the same person's face can be as large as the differences between faces of different individuals. Sutherland, Young and Rhodes (2017b) found that this is equally true for perceived trustworthiness and dominance (see Figure 7) and demonstrated how differences in viewpoint and expression make interacting contributions to these impressions.

----- FIGURE 7 ABOUT HERE -----

This point about the variability between different views of the same face also has important implications for understanding how we recognise face identity. Although image variability can create very different first impressions of a face, it is largely irrelevant to determining the face's identity; we want to be able to recognise people we know across many different views. A powerful insight into the role of image variability in recognising face identity is offered by a sorting task devised by Jenkins et al. (2011). Participants were given a set of 40 everyday photographs of faces like those shown in Figure 8 and asked to sort these into piles of photographs of the same person.

----- FIGURE 8 ABOUT HERE -----

In fact there are only two faces in Figure 8, 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 to16 different piles (identities). In other words, with varied images of unfamiliar faces, participants always thought there more identities than were actually present in the set of 40 photos, and often substantially more.

This finding 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. Instead, Jenkins et al.'s (2011) data show that participants are more likely to see photos of unfamiliar 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). Similar difficulties occur in other perceptual matching tasks, such as comparing someone's face to their passport photograph, where performance with unfamiliar faces can be surprisingly error-prone (Bruce et al., 1999; Hancock, Bruce, & Burton, 2001; Kemp, Towell, & Pike, 1997).

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 can 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 8 shows, within-person variability can be substantial.

As was already noted, Jenkins et al.'s (2011) sorting task is much easier with familiar faces, with participants then rarely putting photos of the two different individuals into the same pile (less than 1% of trials). Indeed most of us can recognise familiar faces without difficulty across changes in pose and expression, and in very variable lighting. Yet, in stark contrast, we have seen that equivalent image changes create significant problems in recognising the faces of people we don't know well. Understanding how view-invariant recognition of familiar faces is achieved, and why recognition of unfamiliar faces across equivalent image changes is relatively poor, are therefore key theoretical tasks (Burton, 2013; Burton, Jenkins, Hancock, & White, 2005; Burton, Jenkins, & Schweinberger, 2011; Bindemann & Johnston, 2017; Davies & Young, 2017; Young & Burton, 2017).

A recent study by Burton, Kramer, Ritchie and Jenkins (2016) marks a substantial advance. Burton et al. analysed how images of the same face vary by applying PCA to everyday photographs of the same person. As we already noted, PCA is a statistical technique that can reduce the dimensionality of photographs to a relatively small number of principal components (PCs). This is a widely used technique in the computer science literature, but it is mostly applied to images of different faces. That is, researchers usually run PCA across images of many different faces to find the PCs of faces in general. Often, too, the images used in PCA are photographed under standard conditions, to eliminate changes in pose, expression, lighting and so on.

Burton et al.'s (2016) application of PCA to everyday images of the same face represents a different approach. Instead of asking what image properties can distinguish between all faces (by finding PCs across different faces), they asked what properties characterise the highly variable images that correspond to a specific face identity. What Burton et al. found was that different faces have different PCs. Put simply, the ways in which one person's face varies from image to image is different from the way in which someone else's face will vary. The characteristics that vary or remain relatively consistent across images differ between one person and another.

These differences are comprehensively demonstrated in Burton et al.'s (2016) paper, and Kramer, Jenkins and Burton (2017a) have made available the software tools used in the analyses. The implication of Burton et al.'s (2016) data is that we have to learn separately the relatively variant and invariant characteristics of each of the faces we know, and this immediately explains why unfamiliar face identity can be problematic. Variability across images is to some extent identity-specific. We need to learn which characteristics of a particular face are relatively consistent and which are variable. For unfamiliar faces our brains can't readily interpret whether image differences are identity-relevant or not.

In this light we need to think carefully about the widely used concept of face expertise. It is often said that that we spend so much time looking at faces that we are all 'face experts' (Carey, 1992; Diamond & Carey, 1986). Certainly, phenomena such as the other-race effect show that adult perceptual mechanisms have become tuned to dealing with particular types of faces (see Rossion & Michel, 2011). However, the idea of expertise is usually applied in the case of face identity, and is often invoked in studies involving recognition of unfamiliar faces, but Burton et al.'s (2016) findings show that the concept of expertise seems to apply more closely to our ability to recognise familiar rather than unfamiliar faces. In effect, we have become particularly expert at recognising each of the faces we know (Young & Burton, 2017; see also Johnston, Overell, Kaufman, Robinson, & Young, 2016). In contrast, unfamiliar face identity can be problematic because some of the relevant image statistics are unknown, which can often make unfamiliar face learning surprisingly image-dependent (Bruce, 1982; Longmore, Liu. & Young, 2008; Longmore et al., 2017). This sets limits on our expertise with unfamiliar face identity.

This is not to deny that there are many things we can reliably judge from unfamiliar faces, as the studies of facial expression and first impressions have shown. These may also reflect some form of expertise, but it seems to involve characteristics that reflect relatively generic visual properties. The generic properties also include important social categories of gender, race and apparent age (Bruce & Young, 2012). For these, the idea of expertise may again have some merit, but the evidence is mixed. Although we are often thought to have become so expert at perceiving such categories in faces that they will be seen automatically (e.g. Martin, et al., 2015), they actually show only a somewhat limited form of automaticity if stringent criteria are applied (Yan, Young, & Andrews, 2017c); see Palermo and Rhodes (2007) for a review of automaticity and face perception.

There has also been debate about how perceived social categories of gender and ethnicity relate to face identity. Bruce & Young (1986) maintained that there must be some separation between the coding of gender and identity because we can easily classify the gender of unfamiliar faces, and Bruce, Ellis, Gibling and Young (1987) showed that having a gender-stereotypical appearance affected decisions about a familiar face's sex but had no effect on recognising its identity. Bruce and Young (1986) therefore drew a distinction between identity-specific semantic codes based on recognising a familiar face and the visually-derived semantic codes (such as gender or race) created by its appearance. However, other widely discussed models such as Haxby et al. (2000) elide this distinction by shifting the focus onto the fact that characteristics such as gender, race, and identity represent relatively invariant facial attributes. Consistent with this type of account, some data support the idea of a more integral representation of gender and identity (Goshen-Gottstein & Ganel, 2000; Rossion 2002; Zhao & Hayward, 2013).

Recently, we approached these questions in a novel way, by asking how social categories of gender and race might be learnt (Kramer, Young, Day, & Burton, 2017c). A natural intuition is that learning the difference between male and female faces must involve a great deal of practice, perhaps driven by evolved mechanisms for sexual selection and assisted by more salient cues from body shape, voice, and in many cultures clothing. Similarly, one might think that learning about race is driven by social mechanisms concerning group membership, and such factors feature prominently in some theories of other-race effects in face recognition (Hugenberg, Miller, & Claypool. 2007; Sporer, Trinkl, & Guberova, 2007). However, and rather to our surprise, we discovered that such explicit learning of gender or race isn't necessary; these social categories can instead be an emergent property of learning to recognise a small number of familiar faces from multiple different images of each person (Kramer et al., 2017c).

Kramer et al. (2017c) used a combination of PCA and LDA to identify highly varied images of different face identities; these trained faces can then be considered as 'familiar' to the model. The trained model showed properties analogous to human face recognition in that it performed well at ‘recognising’ new (untrained) images of familiar faces and performed less well at establishing whether images of unfamiliar (i.e. untrained) faces were of the same person. Remarkably, following training for recognising identity, the first dimension from the LDA separated male from female faces with over 95% correct accuracy for untrained images of both familiar and completely novel unfamiliar faces. Moreover, high levels of performance at separating face images by gender were found even with only a small number of trained identities (>90% correct after training with 5 men and 5 women). Comparable findings were found for classifying faces by race, which emerged as the second dimension from the LDA (Kramer et al., 2017c).

In sum, incidental learning of gender and race can be a natural consequence of learning a small number of familiar face identities. It seems that the cues needed to distinguish the gender or race of any face covary with those needed to recognise familiar identities. This is interesting because learning to recognise a small number of familiar individuals closely approximates the task facing human infants in many societies. In contrast, Kramer et al. (2017c) found that a model trained only to classify gender offered no benefit to classifying identity.

The findings I have summarised make a powerful case that our brains can often make good use of image statistics, but we need to be careful not to overstate this. If we return to issues concerning plasticity, it is clear that as we grow up we become particularly tuned to some image properties at the expense of others (e.g. Rossion & Michel, 2011) and that whilst this confers advantages in everyday life, it incurs a cost in more unusual circumstances. For example Kramer, Jenkins, Young and Burton (2017b) found that adults can only use image statistics when faces are presented in familiar formats.

----- FIGURE 9 ABOUT HERE -----

Kramer et al. (2017b) asked participants to watch episodes of TV soap operas they had never seen before in different formats that involved normal colour videos, upside-down colour videos, normal greyscale videos, and contrast-reversed greyscale videos. They were then tested for their ability to recognise new (unseen) images of the characters from the soap operas; this is a strong test of face familiarisation. Recognition was tested in the same format or the opposite format to that in which participants had watched the TV show.

Results from Kramer et al.'s (2017b) study are summarised in Figure 9. In terms of general image statistics there is no real difference between normal and inverted videos, or between normal greyscale and contrast-reversed videos; the ranges and types of variability are equivalent in each case. Yet only faces that were both learnt and tested in a normal format showed effective learning. In other words, participants were unable to use information that was present in the unusual formats. The mere presence of multiple, variable views in the seen videos was not in itself useful; this variability could only be exploited to learn characteristics of the faces if they were presented in a format that would fit with participants' previous experience. We know, for example, that upright faces are much more commonly seen than inverted faces from early infancy (Sugden & Moulson, 2017).

Overview

The First Bartlett Lecture was given by Carolus Oldfield in 1966, on 'Things, words and the brain'. Oldfield (1966) discussed the cognitive and neural processes involved in retrieving words in speech, using object naming as a particularly useful paradigm and bringing to bear on it evidence from both experiments (e.g. Oldfield & Wingfield, 1965) and neuropsychological phenomena (Newcombe, Oldfield, & Wingfield, 1965).

A number of us who worked in the relatively novel field of face recognition in the 1980s will recall that we tried to fill a theoretical vacuum identified by Hadyn Ellis (1975; see also Davies & Young, 2017) by taking Oldfield's (1966) agenda and transporting it from the study of words and things to the study of faces and people, whilst mixing in a few other ideas from contemporary cognitive models of word recognition (see Young & Bruce, 2011). So on the principle that imitation remains the sincerest form of flattery I decided to mimic Oldfield's title here, simply changing the topic from 'Things, words and the brain' to 'Faces, people and the brain'.

Although it proved a fruitful way to begin, if everything that the field of face recognition had accomplished was built on copying ready-made research agendas from elsewhere I doubt that it would still command so much interest. What has substantially enhanced the field's appeal is appreciation of its relation to broader questions about face and person perception. My main aim here has been to use the fact that the face is a source of diverse social signals to ask important questions about the underlying principles that determine how our brains are organised. We have looked at evidence relevant to understanding constraints that underpin the functional architecture of face and person perception resulting from the evolved structure of the brain, the need to optimise responses to different everyday tasks, and the statistical properties of facial and other signals in the environment. I have shown that these constraints are not mutually exclusive; instead, they interact to determine the optimal organisation. So, for example, differences between the recognition of identity and emotion reflect a combination of statistical properties of the input (where there is partial but not complete segregation), covariation and complementarity of cues across different modalities (such as face and voice), and whether the environment sets a premium on detecting change (emotion) or permits a working assumption of stability (identity).

Along the way I have also sought to bring out some more specific take-home messages about approaches to studying face and person perception. The first of these messages involves the usefulness of a combination of naturally varying ambient images and data-driven approaches in identifying key characteristics of our face perception abilities. This approach contrasts with the natural temptation to try to work with standardised images that eliminate as much variability as possible and home in straight away on specific cues. Although there is still a role for trying to tease apart the relative contributions of different cues (cf. Sutherland et al., 2017b), it needs to be tempered by the second message, which is that cue covariation and cue complementarity are common properties of many face and person perception tasks. This means that the quest for the specific 'diagnostic cue' that is exclusively needed to perceive a given characteristic will often prove futile. It is more important to understand how our brains exploit natural cue covariation than to search for any unique cues involved. This leads to my third message, which is that multi-attribute models can capture much of this covariation, but adult brains have reached a point where they can only use covariation that falls within a previously learnt range.

More generally, an emerging theme has been that the implications of the idea of creating stability from variation (Bruce, 1994, Burton, 2013) should be taken seriously. Image variability is not simply a nuisance, as is so often assumed, but neither is it everything that needs to be understood. Instead, we have seen that variability has to be used in different ways for different purposes. As is so often the case in psychology, we need to look at the interplay between different constraining factors to understand the way we do the things we do.

References

Abbas, Z-A., & Duchaine, B. (2008). The role of holistic processing in judgments of facial attractiveness. *Perception, 37*, 1187-1196.

Adolphs, R., & Tranel, D. (1999). Intact recognition of emotional prosody following amygdala damage. *Neuropsychologia, 37*, 1285-1292.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature, 372*, 669-672.

Ambadar, Z., Schooler, J. W., & Cohn, J. F. (2005). Deciphering the enigmatic face: the importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science, 16*, 404-410.

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.

Andrews, T. J., Baseler, H., Jenkins, R., Burton, A. M., & Young, A. W. (2016). Contributions of feature shapes and surface cues to the recognition and neural representation of facial identity. *Cortex, 83*, 280-291.

Andrews, T. J., Davies-Thompson, J., Kingstone, A., & Young, A. W. (2010). Internal and external features of the face are represented holistically in face-selective regions of visual cortex. *Journal of Neuroscience, 30*, 3544-3552.

Atkinson, A. P., & Adolphs, R. (2011). The neuropsychology of face perception: beyond simple dissociations and functional selectivity. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 366*, 1726-1738.

Atkinson, A. P., Tipples, J., Burt, D. M., & Young, A. W. (2005). Asymmetric interference between sex and emotion in face perception. *Perception and Psychophysics, 67*, 1199-1213.

Barton, J. J. S., & Corrow, S. L. (2016). Recognizing and identifying people: A neuropsychological review. *Cortex, 75*,132-150.

Baseler, H. A., Harris, R. J., Young, A. W., & Andrews, T. J. (2014). Neural responses to expression and gaze in the posterior superior temporal sulcus interact with facial identity. *Cerebral Cortex, 24*, 737-744.

Beauchamp, M. S., Nath, A. R., & Pasalar, S. (2010). fMRI-guided transcranial magnetic stimulation reveals that the superior temporal sulcus is a cortical locus of the McGurk effect. *Journal of Neuroscience, 30*, 2414-2417.

Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends in Cognitive Sciences, 8*, 129-135.

Bernstein, M., & Yovel, G. (2015). Two neural pathways of face processing: A critical evaluation of current models. *Neuroscience and Biobehavioral Reviews, 55*, 536-546.

Berry, D. S. (1990). Taking people at face value: evidence for the kernel of truth hypothesis. *Social Cognition, 8*, 343-361.

Bindemann, M,. & Johnston, R. A. (2017). Understanding how unfamiliar faces become familiar: introduction to a special issue on face learning. *Quarterly Journal of Experimental Psychology, 70*, 859-862.

Broks, P., Young, A. W., Maratos, E. J., Coffey, P. J., Calder, A. J., Isaac, C. L., et al. (1998). Face processing impairments after encephalitis: amygdala damage and recognition of fear. *Neuropsychologia, 36*, 59-70.

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, 47A*, 5-28.

Bruce, V., & Valentine, T. (1985). Identity priming in the recognition of familiar faces. *British Journal of Psychology, 76*, 363-383.

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

Bruce, V., & Young, A. (1998). *In the eye of the beholder: the science of face perception*. Oxford: Oxford University Press

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

Bruce, V., Ellis, H. D., Gibling, F., & Young, A. W. (1987). Parallel processing of the sex and familiarity of faces. *Canadian Journal of Psychology, 41*, 510-520.

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.

Bruyer, R., Laterre, C., Seron, X., Feyereisen, P., Strypstein, E., Pierrard, E., et al. (1983). A case of prosopagnosia with some preserved covert remembrance of familiar faces. *Brain and Cognition, 2*, 257-284.

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., 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.

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience, 6,* 641-651.

Calder, A. J., Burton, A. M., Miller, P., Young, A. W., & Akamatsu, S. (2001). A principal component analysis of facial expressions. *Vision Research, 41*, 1179-1208.

Calder, A. J., Keane, J., Manes, F., Antoun, N., & Young, A. W. (2000). Impaired recognition and experience of disgust following brain injury. *Nature Neuroscience, 3*, 1077-1078.

Calder, A. J., Lawrence, A. D., & Young, A. W. (2001). Neuropsychology of fear and loathing. *Nature Reviews Neuroscience, 2*, 352-363.

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configurational information in facial expression perception. *Journal of Experimental Psychology: Human Perception & Performance, 26*, 527-551.

Calder, A. J., Young, A. W., Rowland, D., Perrett, D. I., Hodges, J. R., & Etcoff, N. L. (1996). Facial emotion recognition after bilateral amygdala damage: differentially severe impairment of fear. *Cognitive Neuropsychology, 13*, 699-745.

Calvert, G. A. (2001). Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cerebral Cortex, 11*, 1110-1123.

Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology, 10*, 649-657.

Calvert, G. A., Hansen, P. C., Iversen, S. D., & Brammer, M. J. (2001). Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *NeuroImage, 14*, 427-438.

Campanella, S., & Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences, 11*, 535-543.

Caramazza, A. (1984). The logic of neuropsychological research and the problem of patient classification in aphasia. *Brain and Language, 21*, 9–20.

Caramazza, A. (1986). On drawing inferences about the structure of normal cognitive systems from the analysis of patterns of impaired performance. *Brain and Language, 5*, 41–66.

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

Carroll, N. C., & Young, A. W. (2005). Priming of emotion recognition. *Quarterly Journal of Experimental Psychology, 58A*, 1173-1197.

Chen, W., Ren, N., Young, A. W., & Liu, C. H. (2017). Interaction between social categories in the composite face paradigm. *Journal of Experimental Psychology: Learning, Memory and Cognition*, DOI:10.1037/xlm0000418.

Coltheart, M. (1981). Disorders of reading and their implications for models of normal reading. *Visible Language, 15*, 245-286.

Coltheart, M. (2017). The assumptions of cognitive neuropsychology: reflections on Caramazza (1984, 1986). *Cognitive Neuropsychology*.

Davies, G. M., & Young, A. W. (2017). Research on face recognition: the Aberdeen influence. *British Journal of Psychology*, DOI:10.1111/bjop.12243.

de Gelder, B. (2016). *Emotions and the body*. Oxford: Oxford University Press.

de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and eye. *Cognition and Emotion, 14*, 289-311.

de Gelder, B., Stienen, B. M. C., & Van den Stock, J. (2013). Emotions by ear and by eye. In P. Belin, S. Campanella, & T. Ethofer (Eds), *Integrating face and voice in person perception*. New York: Springer, 253-268.

de Haan, E. H. F., Young, A. W., & Newcombe, F. (1991). A dissociation between the sense of familiarity and access to semantic information concerning familiar people. *European Journal of Cognitive Psychology, 3*, 51-67.

Deen, B., Richardson, H., Dilks, D. D., Takahashi, A., Keil, B., Wald, L., et al. (2017). Organization of high-level visual cortex in human infants. *Nature Communications, 8*, 13995.

Della Sala, S., & Young, A. W. (2003). Quaglino's 1867 case of prosopagnosia. *Cortex, 39*, 533-540.

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

Dotsch, R., Hassin, R. R., & Todorov, A. (2016). Statistical learning shapes face evaluation. *Nature Human Behaviour, 1*, 0001.

Du, S., Tao, Y., & Martinez, A. M. (2014). Compound facial expressions of emotion. *Proceedings of the National Academy of Sciences, USA, 111*, E1454–E1462.

Duchaine, B., & Yovel, G. (2015). A revised neural framework for face processing. *Annual Review of Vision Science, 1*, 393-416.

Dunbar, R. I., & Shultz, S. (2007). Evolution in the social brain. *Science, 317*, 1344-1347.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, California: Consulting Psychologists Press.

Ellamil, M., Susskind, J. M., & Anderson, A. K. (2008). Examinations of identity invariance in facial expression adaptation. *Cognitive, Affective, and Behavioral Neuroscience, 8*, 273-281.

Ellis, A. W., & Young, A. W. (1988). *Human cognitive neuropsychology*. Hove: Psychology Press.

Ellis, A. W., Young, A. W., & Flude, B. M. (1990). Repetition priming and face processing: priming occurs within the system that responds to the identity of a face. *Quarterly Journal of Experimental Psychology, 42A*, 495-512.

Ellis, A. W., Young, A. W., Flude, B. M., & Hay, D. C. (1987). Repetition priming of face recognition. *Quarterly Journal of Experimental Psychology, 39A,* 193-210.

Ellis, A. W., Young, A. W., & Critchley, E. M. R. (1989). Loss of memory for people following temporal lobe damage. *Brain, 112*, 1469-1483.

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

Ellis, H. D., & Florence, M. (1990). Bodamer's (1947) paper on prosopagnosia. *Cognitive Neuropsychology, 7*, 81-105.

Fausey, C. M., Jayaraman, S., & Smith, L. B. (2016). From faces to hands: changing visual input in the first two years. *Cognition, 152*, 101-127.

Feinstein, J. S., Adolphs, R., Damasio, A., & Tranel, D. (2011). The human amygdala and the induction and experience of fear. *Current Biology, 21*, 34-38.

Fisher, K., Towler, J., & Eimer, M. (2016). Facial identity and facial expression are initially integrated at visual perceptual stages of face processing. *Neuropsychologia, 80*, 115-125.

Fodor, J. (1983). *The modularity of mind*. Cambridge, Massachusetts: MIT Press.

Fox, C. J., & Barton, J. J. S. (2007). What is adapted in face adaptation? The neural representations of expression in the human visual system. *Brain Research, 1127*, 80-89.

Fox, C. J., Oruc, I., & Barton, J. J. S. (2008). It doesn't matter how you feel. The facial identity aftereffect is invariant to changes in facial expression. *Journal of Vision, 8:11*, 1-13.

Gainotti, G. (2014). The neuropsychology of familiar person recognition from face and voice. *Psychologica Belgica, 54*, 298-309.

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

Golarai, G., Ghahremani, D. G., Whitfield-Gabrieli, S., Reiss, A., Eberhardt, J. L., Gabrieli, J. D. E., & Grill-Spector, K. (2007). Differential development of high-level visual cortex correlates with category-specific recognition memory. *Nature Neuroscience, 10*, 512-522.

Goshen-Gottstein, Y., & Ganel, T. (2000). Repetition priming for familiar and unfamiliar faces in a sex-judgment task: Evidence for a common route for the processing of sex and identity. *Journal of Experimental Psychology-Learning Memory And Cognition, 26*, 1198-1214.

Gosselin, N., Peretz, I., Johnsen, E., & Adolphs, R. (2007). Amygdala damage impairs emotion recognition from music. *Neuropsychologia, 45*, 236-244.

Hagan, C. C., Woods, W., Johnson, S., Calder, A. J., Green, G. G. R., & Young, A. W. (2009). MEG demonstrates a supra-additive response to facial and vocal emotion in the right superior temporal sulcus. *Proceedings of the National Academy of Sciences, USA, 106*, 20010-20015.

Hagan, C. C., Woods, W., Johnson, S., Green, G. G. R., & Young, A. W. (2013). Involvement of right STS in audio-visual integration for affective speech demonstrated using MEG. *PLoS One, 8*, e70648.

Hall, D. A., Fussell, C., & Summerfield, A. Q. (2005). Reading fluent speech from talking faces: typical brain networks and individual differences. *Journal of Cognitive Neuroscience, 17*, 939-953.

Hancock, P. J. B., Bruce, V., & Burton, A. M. (2001). Recognition of unfamiliar faces. *Trends in Cognitive Sciences, 4*, 330-337.

Hanley, J. R., Young, A. W., & Pearson, N. (1989). Defective recognition of familiar people. *Cognitive Neuropsychology, 6*, 179-210.

Harris, R. J., Young, A. W., & Andrews, T. J. (2014a). Brain regions involved in processing facial identity and expression are differentially selective for surface and edge information. *NeuroImage, 97*, 217-223.

Harris, R. J., Young, A. W., & Andrews, T.J. (2014b). Dynamic stimuli demonstrate a categorical representation of facial expression in the amygdala. *Neuropsychologia, 56*, 47-52.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science, 303*, 1634-1640.

Haxby, J. M., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science, 293*, 2425-2430.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*, 223-233.

Hay, D. C., & Young, A. W. (1982). The human face. In A. W. Ellis (Ed.), *Normality and pathology in cognitive functions,* (pp. 173–202). London: Academic Press.

Hécaen, H. (1981). The neuropsychology of face recognition. In G. Davies, H. Ellis, & J. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 39-54). London: Academic Press.

Hess, U., Adams, R. B. Jr., & Kleck, R. E. (2004). Facial appearance, gender, and emotion expression. *Emotion, 4*, 378–388.

Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience, 3*, 80-84.

Hsu, S.-M., & Young, A. W. (2004). Adaptation effects in facial expression recognition. *Visual Cognition, 11*, 871-899.

Hugenberg, K., Miller, J., & Claypool, H. M. (2007). Categorization and individuation in the cross-race recognition deficit: toward a solution to an insidious problem. *Journal of Experimental Social Psychology, 43*, 334–340.

Jayaraman, S., Fausey, C. M., & Smith, L. B. (2017). Why are faces denser in the visual experiences of younger than older infants? *Developmental Psychology, 53*, 38-49.

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

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40*, 1-19.

Johnson, M. H., Senju, A., & Tomalski, P. (2015). The two-process theory of face processing: Modifications based on two decades of data from infants and adults. *Neuroscience and Biobehavioral Reviews, 50,* 169-179.

Johnston, P., Mayes, A., Hughes, M., & Young, A. W. (2013). Brain networks subserving the evaluation of static and dynamic facial expressions. *Cortex, 49*, 2462-2472.

Johnston, P., Overell, A., Kaufman, J., Robinson, J, & Young, A. W. (2016). Expectations about person identity modulate the face-sensitive N170. *Cortex, 85*, 54-64.

Johnston, P., Robinson, J., Kokkinakis, A., Ridgeway, S., Simpson, M., Johnson, S., Kaufman, J., & Young, A. W. (2017). Temporal and spatial localization of prediction-error signals in the visual brain. *Biological Psychology, 125*, 45-57.

Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience, 3*, 759-763.

Kanwisher, N. (2017). The quest for the FFA and where it led. *Journal of Neuroscience, 37*, 1056-1061.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*, 4302-4311.

Kaufmann, J. M., & Schweinberger, S. R. (2004). Expression influences the recognition of familiar faces. *Perception, 33*, 399-408.

Keane, J., Calder, A. J., Hodges, J. R., & Young, A. W. (2002). Face and emotion processing in frontal variant frontotemporal dementia. *Neuropsychologia, 40*, 655-665.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L. Z., & Pascalis, O. (2007). The other-race effect develops during infancy - evidence of perceptual narrowing. *Psychological Science, 18*, 1084-1089.

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., Jenkins, R., & Burton, A. M. (2017a). InterFace: a software package for face image warping, averaging, and principal components analysis. *Behavior Research Methods*,

Kramer, R. S. S., Jenkins, R., Young, A. W., & Burton, A. M. (2017b). Natural variability is essential to learning new faces. *Visual Cognition*, DOI:10.1080/13506285.2016.1242522

Kramer, R. S. S., & Ward, R. (2010). Internal facial features are signals of personality and health. *Quarterly Journal of Experimental Psychology, 63,* 2273–2287.

Kramer, R. S. S., Young, A. W., Day, M. G., & Burton, A. M. (2017c). Robust social categorization emerges from learning the identities of very few faces. *Psychological Review, 124*, 115-129.

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science, 218*, 1138-1141.

Lee, K., Quinn, P. C., & Pascalis, O. (2017). Face race processing and racial bias in early development: a perceptual-social linkage. *Current Directions in Psychological Science, 26*, 256-262.

Leonard, C. M., Voeller, K. K. S., & Kuldau, J. M. (1991). When's a smile a smile? Or how to detect a message by digitizing the signal. *Psychological Science, 2,* 166-172.

Lewis, G. J., Lefevre, C. E., & Young, A. W. (2016). Functional architecture of visual emotion recognition ability: a latent variable approach. *Journal of Experimental Psychology: General, 145*, 589-602.

Liu, R. R., Pancaroglu, R., Hills, C. S., Duchaine, B., & Barton, J. J. S. (2016). Voice recognition in face-blind patients. *Cerebral Cortex, 26*, 1473-1487.

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

Longmore, C. A., Santos, I. M., Silva, C. F., Hall, A., Faloyin, D., & Little, E. (2017). Image dependency in the recognition of newly learnt faces. *Quarterly Journal of Experimental Psychology, 70*, 863-873.

Lundqvist, D., Flykt, A., & Ohman, A. (1998). *The Karolinska directed emotional faces (KDEF)*. Department of Clinical Neuroscience, Psychology section, Karolinska Institutet. ISBN 91-630-7164-9. Retrieved from http://www.emotionlab.se/resources/kdef

Macaluso, E., George, N., Dolan, R., Spence, C., & Driver, J. (2004). Spatial and temporal factors during processing of audiovisual speech: a PET study. *NeuroImage, 21*, 725-732.

Marr, D. (1982). *Vision*. San Francisco: Freeman.

Marshall, J. C., & Newcombe, F. (1973). Patterns of paralexia: a psycholinguistic approach. *Journal of Psycholinguistic Research, 2*, 175-199.

Martin, D., Swainson, R., Slessor, G., Hutchison, J., Marosi, D., & Cunningham, S. J. (2015). The simultaneous extraction of multiple social categories from unfamiliar faces. *Journal of Experimental Social Psychology, 60*, 51-58.

Martinez, A. M. (2003). Matching expression variant faces. *Vision Research, 43*, 1047-1060.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 746-748.

Meadows, J. C. (1974). The anatomical basis of prosopagnosia. *Journal of Neurology, Neurosurgery, and Psychiatry, 37*, 489-501.

Meetic Group. (2017). Meetic Group and key figures in Europe. Retrieved July 17th, 2017, from https://www.meetic-group.com/en/meetic-group-who-are-we/

Miller, G. A., & Niceley, P. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America, 27*, 338-352.

Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., et al. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature, 383*, 812-815.

Morton, J. (1979). Facilitation in word recognition: experiments causing change in the logogen model. In P. A. Kolers, M. Wrolstad, & H. Bouma (Eds.), *Processing of visible language*, (Vol. 1,pp. 259–268). New York: Plenum.

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: a two-process theory of infant face recognition. *Psychological Review, 98*, 164-181.

Newcombe, F., Oldfield, R. C., & Wingfield, A. (1965). Object-naming by dysphasic patients. *Nature, 207*, 1217-1218.

Oatley, K., & Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition & Emotion, 1*, 29-50.

Oatley, K., & Johnson-Laird, P. N. (2014). Cognitive approaches to emotions. *Trends in Cognitive Sciences, 18*, 131-140.

Oldfield, R. C. (1966). Things, words and the brain. *Quarterly Journal of Experimental Psychology, 18*, 340-352.

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology, 17*, 273-281.

Oldmeadow, J. A., Sutherland, C. A. M., & Young, A. W. (2013). Facial stereotype visualization through image averaging. *Social Psychological and Personality Science, 4*, 615-623.

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

Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia, 45*, 75-92.

Park, J., Gu, B., Kang, D., Shin, Y., Choi, C., Lee, J., & Kwon, J. S. (2010). Integration of cross-modal emotional information in the brain: An fMRI study. *Cortex, 46*, 161-169.

Parry, F. M., Young, A. W., Saul, J. S. M., & Moss, A. (1991). Dissociable face processing impairments after brain injury. *Journal of Clinical and Experimental Neuropsychology, 13*, 545-558.

Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science, 296*, 1321-1323.

Pascalis, O., Dole, M., & Loevenbroek, H. (2017). More evidence of the linkage between face processing and language processing. *British Journal of Psychology, 108*, 31-33.

Pascalis, O., Scott, L. S., Kelly, D. J., Shannon, R. W., Nicholson, E., Coleman, M., et al. (2005). Plasticity of face processing in infancy. *Proceedings of the National Academy of Sciences, USA, 102*, 5297-5300.

Patterson, M. L., & Werker, J. F. (2003). Two-month-old infants match phonetic information in lips and voice. *Developmental Science, 6*, 191-196.

Peelen, M. V., Atkinson, A. P., & Vuilleumier, P. (2010). Supramodal representation of perceived emotions in the human brain. *Journal of Neuroscience, 30*, 10127-10134.

Pell, P. J., & Richards, A. (2013). Cross-emotion facial expression aftereffects are identity-dependent. *Vision Research, 79*, 1-7.

Penton-Voak, I. S., Pound, N., Little, A. C., & Perrett, D. I. (2006). Personality judgments from natural and composite facial images: more evidence for a "kernel of truth" in social perception. *Social Cognition, 24*, 607-40.

Phillips, M. L., Young, A. W., Senior, C., Brammer, M., Andrew, C., Calder, A. J., et al. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature, 389*, 495-498.

Pitcher, D. (2014). Facial expression recognition takes longer in the posterior superior temporal sulcus than in the occipital face area. *Journal of Neuroscience, 34*, 9173-9177.

Pitcher, D., Duchaine, B., & Walsh, V. (2014). Combined TMS and fMRI reveal dissociable cortical pathways for dynamic and static face perception. *Current Biology, 24*, 2066-2070.

Polk, T. A., Park, J., Smith, M. R., & Park, D. C. (2007). Nature versus nurture in ventral visual cortex: a functional magnetic resonance imaging study of twins. *Journal of Neuroscience, 27*, 13921-13925.

Psalta, L., Young, A. W., Thompson, P., & Andrews, T. J. (2014a). The Thatcher illusion reveals orientation dependence in brain regions involved in processing facial expressions. *Psychological Science, 25*, 128-136.

Psalta, L., Young, A. W., Thompson, P., & Andrews, T. J. (2014b). Orientation-sensitivity to facial features explains the Thatcher illusion. *Journal of Vision,14*, 9, 1-10.

Ratcliff, G., & Newcombe, F. (1982). Object recognition: some deductions from the clinical evidence. In A. W. Ellis (Ed.), *Normality and pathology in cognitive functions,* (pp. 147-171). London: Academic Press.

Rhodes, G. (2017). Adaptive coding and face recognition. *Current Directions in Psychological Science, 26*, 218-224.

Rhodes, G., Pond, S., Burton, N., Kloth, N., Jeffery, L., et al. (2015). How distinct is the coding of face identity and expression? Evidence for some common dimensions in face space. *Cognition, 142*, 123-137.

Riddoch, M. J., Johnston, R. A., Bracewell, R. M., Boutsen, L., & Humphreys, G. W. (2008). Are faces special? A case of pure prosopagnosia. *Cognitive Neuropsychology, 25*, 3-26.

Rizzolatti, G., Umilta, C., & Berlucchi, G. (1971). Opposite superiorities of the right and left cerebral hemispheres in discriminative reaction time to physiognomical and alphabetical material. *Brain, 94*, 431-442.

Rossion, B. (2002). Is sex categorization from faces really parallel to face recognition? *Visual Cognition, 9*, 1003–1020.

Rossion, B. (2008). Constraining the cortical face network by neuroimaging studies of acquired prosopagnosia. *Neuroimage, 40*, 423-426.

Rossion, B. (2013). The composite face illusion: a whole window into our understanding of holistic face perception. *Visual Cognition, 21*, 139-253.

Rossion, B., & Michel, C. (2011). An experience-based holistic account of the other-race effect. In A. J. Calder, G. Rhodes, M. H. Johnson and J. V. Haxby (Eds.), *The Oxford handbook of face perception*. Oxford: Oxford University Press, 215-243.

Said, C. P., Sebe, N., & Todorov, A. (2009). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion, 9*, 260–264.

Santos, I. M., & Young, A. W. (2005). Exploring the perception of social characteristics in faces using the isolation effect. *Visual Cognition, 12*, 213-247.

Santos, I. M., & Young, A. W. (2008). Effects of inversion and negation on social inferences from faces. *Perception, 37*, 1061-1078.

Santos, I. M., & Young, A. W. (2011). Inferring social attributes from different face regions: evidence for holistic processing. *Quarterly Journal of Experimental Psychology, 64*, 751-766.

Schirmer, A., & Adolphs, R. (2017). Emotion perception from face, voice, and touch: comparisons and convergence. *Trends in Cognitive Sciences, 21*, 216-228.

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.

Schweinberger, S. R., & Soukup, G. R. (1998). Asymmetric relationships among perceptions of facial identity, emotion, and facial speech. *Journal of Experimental Psychology: Human Perception & Performance, 24*, 1748-1765.

Schweinberger, S. R., Burton, A. M., & Kelly, S. W. (1999). Asymmetric dependencies in perceiving identity and emotion: experiments with morphed faces. *Perception & Psychophysics, 61*, 1102-1115.

Scott, S. K., Young, A. W., Calder, A. J., Hellawell, D. J., Aggleton, J. P., & Johnson, M. (1997). Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Nature, 385*, 254-257.

Secord, P. F. (1958). Facial features and inference processes in interpersonal perception. In R. Taguiri & L. Petrullo (Eds.), *Person perception and interpersonal behavior*. Stanford, CA: Stanford University Press, 300-315.

Shakeshaft, N. G., & Plomin, R. (2015). Genetic specificity of face recognition. *Proceedings of the National Academy of Sciences, USA, 112*, 12887-12892.

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge: Cambridge University Press.

Sheehan, M. J., & Nachman, M. W. (2014). Morphological and population genomic evidence that human faces have evolved to signal individual identity. *Nature Communications, 5*, 4800.

Sormaz, M., Andrews, T. J., & Young, A. W. (2013). Contrast negation and the importance of the eye region for holistic representations of facial identity. *Journal of Experimental Psychology: Human Perception and Performance, 39*, 1667-1677.

Sormaz, M., Watson, D. M., Smith, W. A. P., Young, A. W., & Andrews, T. J. (2016a). Modelling the perceptual similarity of facial expressions from image statistics and neural responses. *NeuroImage, 129*, 64-71.

Sormaz, M., Young, A. W., & Andrews, T. J. (2016b). Contributions of feature shapes and surface cues to the perception of facial expressions. *Vision Research, 127*, 1-10.

South Palomares, J. K., & Young, A. W. (2017). Facial first impressions of partner preference traits: trustworthiness, status, and attractiveness. *Social Psychological and Personality Science,* DOI:10.1177/1948550617732388.

Sporer, S. L., Trinkl, B., & Guberova, E. (2007). Matching faces: differences in processing speed of out-group faces by different ethnic groups. *Journal of Cross-Cultural Psychology, 38*, 398–412.

Sprengelmeyer, R., Young, A. W, Baldas, E-M., Ratheiser, I., Sutherland, C. A. M., Müller, H-P., et al. (2016). The neuropsychology of first impressions: evidence from Huntington's disease. *Cortex, 85,* 100-115.

Sprengelmeyer, R., Young, A. W., Schroeder, U., Grossenbacher, P. G., Federlein, J., Büttner, T., et al. (1999). Knowing no fear. *Proceedings of the Royal Society, London, B: Biological Sciences, 266*, 2451-2456.

Sugden, N. A,. & Moulson, M. C. (2017). Hey baby, what's "up"? One- and 3-month-olds experience faces primarily upright but non-upright faces offer the best views. *Quarterly Journal of Experimental Psychology, 70*, 959-969.

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

Sutherland, C. A. M., Oldmeadow, J. A., & Young, A. W. (2016). Integrating social and facial models of person perception: converging and diverging dimensions. *Cognition, 157*, 257-267.

Sutherland, C. A. M., Rhodes, G., & Young, A. W. (2017a). Facial image manipulation: a tool for investigating social perception. *Social Psychological and Personality Science*, DOI:10.1177/1948550617697176.

Sutherland, C. A. M., Rowley, L. E., Amaoku, U. T., Daguzan, E., Kidd-Rossiter, K. A., Maceviciute, U, & Young, A. W. (2015a). Personality judgements from everyday images of faces. *Frontiers in Psychology, 6*, 1616, 1-11.

Sutherland, C. A. M., Young, A. W., & Rhodes, G. (2017b). Facial first impressions from another angle: how social judgments are influenced by changeable and invariant facial properties. *British Journal of Psychology, 108*, 397-415.

Sutherland, C. A. M., Young, A. W., Mootz, C. A., & Oldmeadow, J. A. (2015b). Face gender and stereotypicality influence facial trait evaluation: Counter-stereotypical female faces are negatively evaluated. *British Journal of Psychoiogy, 106*, 186-208.

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

Taubert, J., Alais, D., & Burr, D. (2016). Different coding strategies for the perception of stable and changeable facial attributes. *Scientific Reports, 6*, 32239.

Tham, D. S. Y., Bremner, J. G., & Hay, D. C. (2015). In infancy the timing of emergence of the other-race effect is dependent on face gender. *Infant Behavior and Development, 40*, 131-138.

Tham, D. S. Y., Bremner, J. G., & Hay, D. C. (2017). The other-race effect in children from a multiracial population: a cross-cultural comparison. *Journal of Experimental Child Psychology, 155*, 128-137.

Thompson, P. (1980). Margaret Thatcher - a new illusion. *Perception, 9*, 483-484.

Thompson, P., Anstis, S., Rhodes, G., Jeffery, L., & Valentine, T. (2009). Thompson's 1980 paper. *Perception, 38*, 921-932.

Tiddeman, B., Burt, M., & Perrett, D. I. (2001). Prototyping and transforming facial textures for perception research. *Computer Graphics and Applications, IEEE, 21*, 42–50.

Todorov, A. (2017). *Face value: the irresistible influence of first impressions*. Princeton, NJ: Princeton University Press.

Todorov, A., Loehr, V., & Oosterhof, N. N. (2010). The obligatory nature of holistic processing of faces in social judgments. *Perception, 39*, 514-532.

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology, 66*, 519–545.

Van den Stock, J., De Winter, F-L., de Gelder, B., Rangarajan, J. R., Cypers, G., et al. (2015). Impaired recognition of body expressions in the behavioral variant of frontotemporal dementia. *Neuropsychologia, 75*, 496–504.

Verhallen, R. J., Bosten, J. M., Goodbourn, P. T., Bargary, G., Lawrance-Owen, A. J., & Mollon, J. D. (2014). An online version of the Mooney Face Test: phenotypic and genetic associations. *Neuropsychologia, 63*, 19-25.

Vernon, R. J. W., Sutherland, C. A. M., Young, A. W., & Hartley, T. (2014). Modeling first impressions from highly variable facial images. *Proceedings of the National Academy of Sciences, USA, 111*, E3353-E3361.

Vihman, M. M. (2017). Learning words and learning sounds: advances in language development. *British Journal of Psychology, 108*, 1-27.

Waller, B. M., Cray, J. J., & Burrows, A. M. (2008). Selection for universal facial emotion. *Emotion, 8*, 435-439.

Webster, M. A., & MacLeod, D. I. A. (2011). Visual adaptation and face perception. *Philosophical Transactions of the Royal Society, London, B: Biological Sciences, 366*, 1702-1725.

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature, 428*, 557-561.

Weibert, K., Harris, R. J., Mitchell, A., Byrne, H., Young, A. W., & Andrews, T. J. (2016). An image-invariant neural response to familiar faces in the human medial temporal lobe. *Cortex, 84*, 34-42.

White, M. (2001). Effect of photographic negation on matching the expressions and identities of faces. *Perception, 30*, 969–981.

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., et al. (2010). Human face recognition ability is specific and highly heritable. *Proceedings of the National Academy of Sciences, USA, 107*, 5238-5241.

Wright, T. M., Pelphrey, K. A., Allison, T., McKeown, M. J., & McCarthy, G. (2003). Polysensory interactions along lateral temporal regions evoked by audiovisual speech. *Cerebral Cortex, 13*, 1034-1043.

Yan, X., Andrews, T. J., & Young, A. W. (2016). Cultural similarities and differences in perceiving and recognizing facial expressions of basic emotions. *Journal of Experimental Psychology: Human Perception and Performance, 42*, 423-440.

Yan, X., Andrews, T. J., Jenkins, R., & Young, A. W. (2016). Cross-cultural differences and similarities underlying other-race effects for facial identity and expression. *Quarterly Journal of Experimental Psychology, 69*, 1247-1254.

Yan, X., Young, A. W., & Andrews, T. J. (2017a). Differences in holistic processing do not explain cultural differences in the recognition of facial expression. *Quarterly Journal of Experimental Psychology*, *70*, 2445-2459.

Yan, X., Young, A. W., & Andrews, T. J. (2017b). Cross-cultural similarities and differences in the perception and recognition of facial expressions. In M. Bindemann and A. M. Megreya (Eds.), *Face processing: systems, disorders and cultural differences*. New York: Nova Science Publishers, 305-324.

Yan, X., Young, A. W., & Andrews, T. J. (2017c). The automaticity of face perception is influenced by familiarity. *Attention, Perception, & Psychophysics*, *79*, 2202-2211.

Young, A. W. (2011). Disorders of face perception. In A. J. Calder, G. Rhodes, M. H. Johnson and J. V. Haxby (Eds.), *The Oxford handbook of face perception*. Oxford: Oxford University Press, 77-91.

Young, A.W. (2016). *Facial expression recognition*. Hove, East Sussex: Psychology Press.

Young, A. W., & Bion, P. J. (1980). Absence of any developmental trend in right hemisphere superiority for face recognition. *Cortex, 16*, 213-221.

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

Young, A. W., & Burton, A. M. (2017). Recognizing faces. *Current Directions in Psychological Science, 26*, 212-217.

Young, A. W., & Ellis, H. D. (1976). An experimental investigation of developmental differences in ability to recognise faces presented to the left and right cerebral hemispheres. *Neuropsychologia, 14*, 495-498.

Young, A. W., & van de Wal, C. (1996). Charcot's case of impaired imagery. In C. Code, C-W. Wallesch, Y. Joanette and A. R. Lecours (Eds.), Classic cases in neuropsychology. Hove, East Sussex: Erlbaum (UK) Taylor & Francis, 31-44.

Young, A. W., Aggleton, J. P., Hellawell, D. J., Johnson, M., Broks, P., & Hanley, J. R. (1995). Face processing impairments after amygdalotomy. *Brain, 118*, 15-24.

Young, A. W., Hellawell, D., & de Haan, E. H. F. (1988). Cross-domain semantic priming in normal subjects and a prosopagnosic patient. *Quarterly Journal of Experimental Psychology, 40A*, 561-580.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*, 747-759.

Young, A. W., Hellawell, D. J., van de Wal, C., & Johnson, M. (1996). Facial expression processing after amygdalotomy. *Neuropsychologia, 34*, 31-39.

Young, A. W., Newcombe, F., de Haan, E. H. F., Small, M., & Hay, D. C. (1993). Face perception after brain injury: selective impairments affecting identity and expression. *Brain, 116*, 941-959.

Young, A. W., Perrett, D. I., Calder, A. J., Sprengelmeyer, R., & Ekman, P. (2002). *Facial expressions of emotion: stimuli and tests (FEEST)*. Bury St. Edmunds: Thames Valley Test Company.

Yovel, G., & O'Toole, A. J. (2016). Recognizing people in motion. *Trends in Cognitive Sciences, 20*, 383-395.

Zebrowitz, L. (2017). First impressions from faces. *Current Directions in Psychological Science, 26*, 237-242.

Zhao, M., & Hayward, W. G. (2013). Integrative processing of invariant aspects of faces: effect of gender and race processing on identity analysis. *Journal of Vision, 13(1)*, 1–18.

Zhu, Q., Song, Y., Hu, S., Li, X., Tian, M., Zhen, Z., et al. (2010). Heritability of the specific cognitive ability of face perception. *Current Biology, 20*, 137-142.

FIGURES

Figure 1: Functional model of face perception suggested by Bruce & Young (1986).

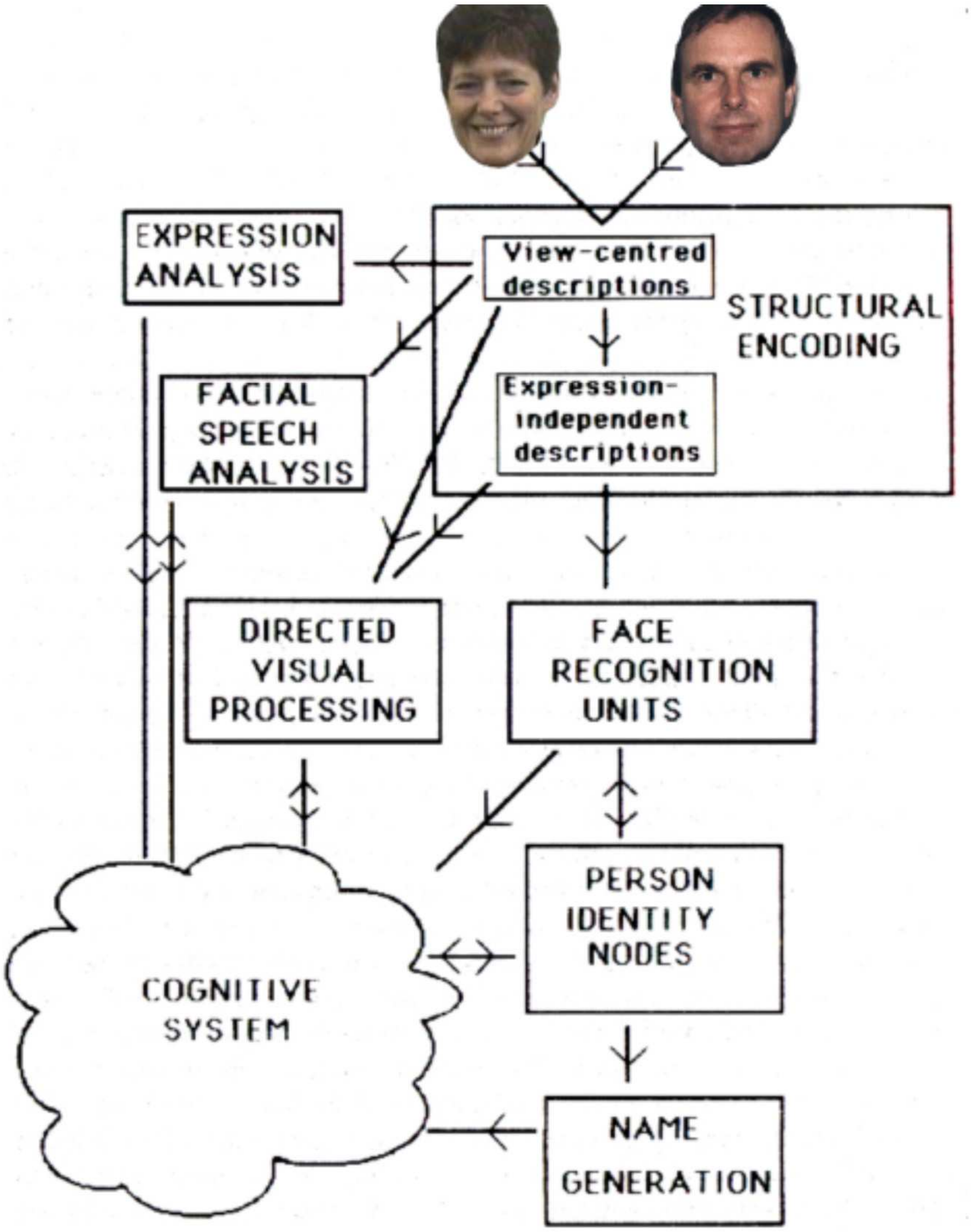


Figure 2: The upper panel shows the location of face-responsive regions from fMRI: the occipital face area (OFA), fusiform face area (FFA), and posterior superior temporal sulcus (STS) alongside Haxby et al.'s (2000) proposals concerning functional organisation of the neural network involved in face perception. The lower panel redraws the Bruce and Young (1986) model to bring out its similarities to Haxby et al.

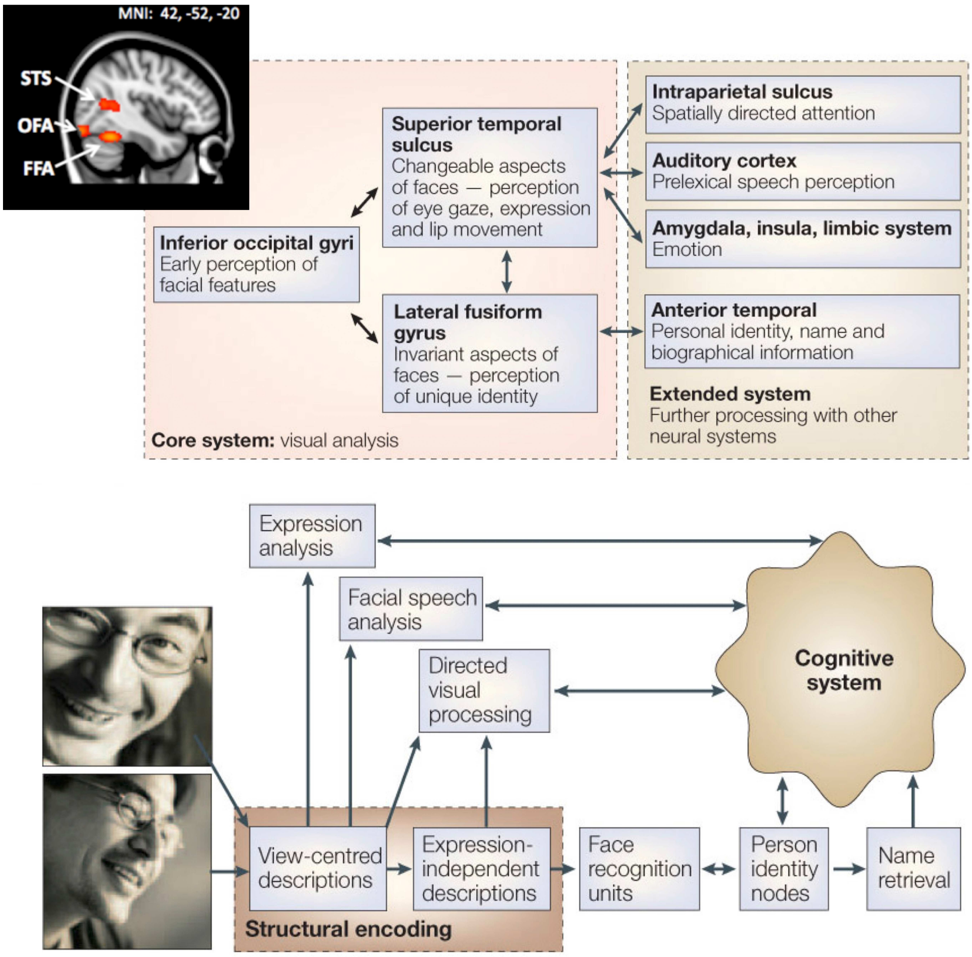


Figure 3: Upper panel: schematic arrangement of emotion hexagon stimuli from Calder et al. (1996) and Young et al. (2002). Each emotion is placed next to that it is most often confused with. The images are then morphed between happiness and surprise (top row), surprise and fear (second row), fear and sadness (third row), sadness and disgust (fourth row), disgust and anger (fifth row), and anger and happiness (bottom row). The lower panel shows performance by participant DR (Calder et al., 1996). The 30 images from the emotion hexagon are set out along the horizontal axis, with the proportion of correct recognition by DR (solid lines) and control participants (dashed lines) represented along the vertical axis. The percentages show DR's ability to recognise each emotion expressed as a percentage of control performance. (Adapted from Calder et al., 1996, and reproduced with permission from Taylor & Francis).

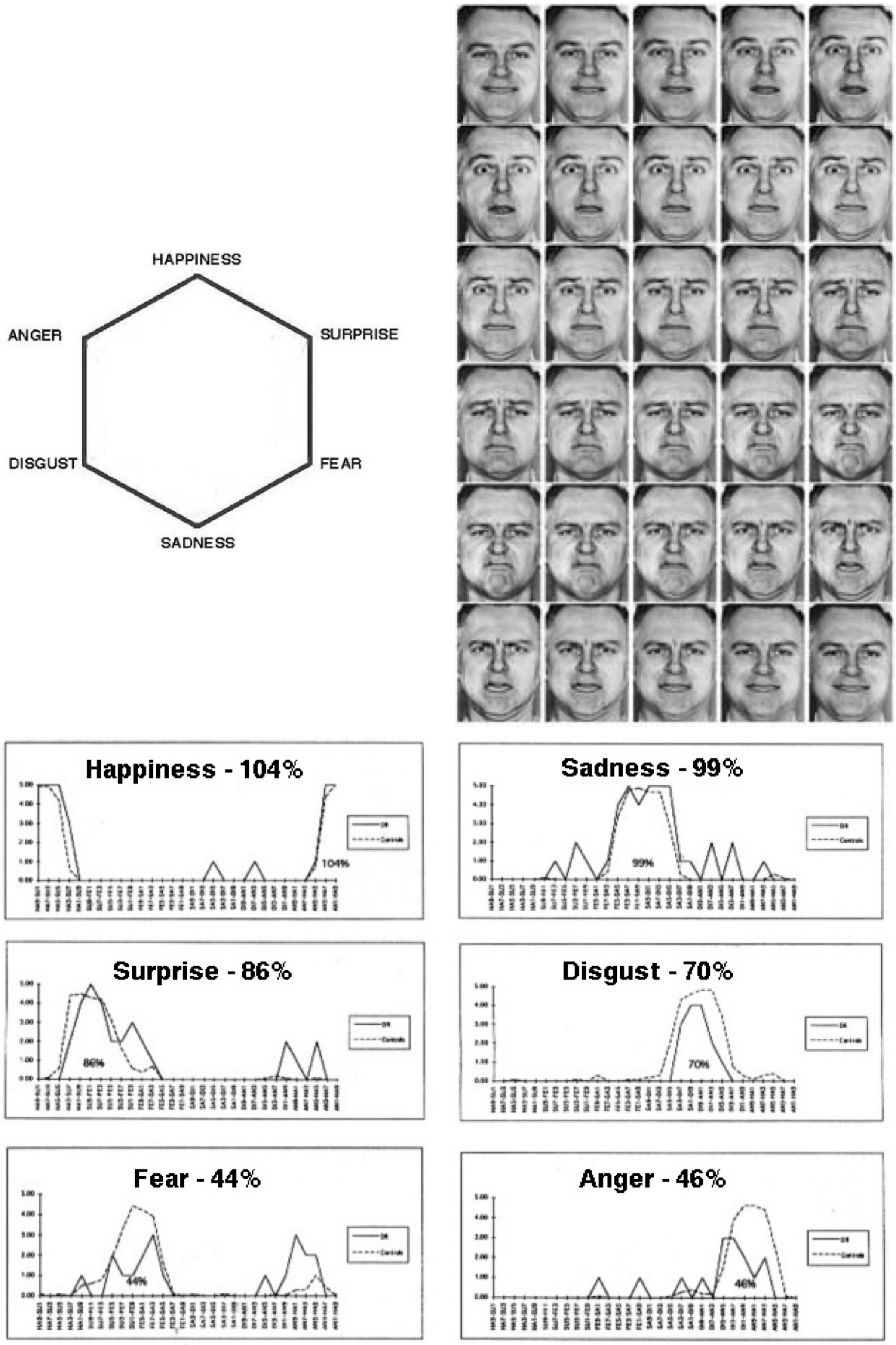

HAPPINESS
SURPRISE
FEAR
SADNESS
DISGUST
ANGER
Happiness - 104%
104%
Sadness - 99%
99%
Surprise - 86%
86%
Disgust - 70%
70%
Fear - 44%
44%
Anger - 46%
46%

Figure 4: Results from Hagan et al.'s (2009) study. Stimuli were fearful or neutral Ekman faces and nonverbal sounds presented in auditory (A), visual (V), or audio-visual (AV) conditions. The analysis identifies a region in right STG/STS (crosshairs MNI 60, -46, 18) showing a multi-modal response to audio-visual fear that meets a stringent criterion of supra-additivity (AV>A+V) based on a broadband (3-80Hz) increase in power. This is plotted across a 500ms moving window at 50 ms intervals. Note that the response is present at early latencies and fades by 200 ms.

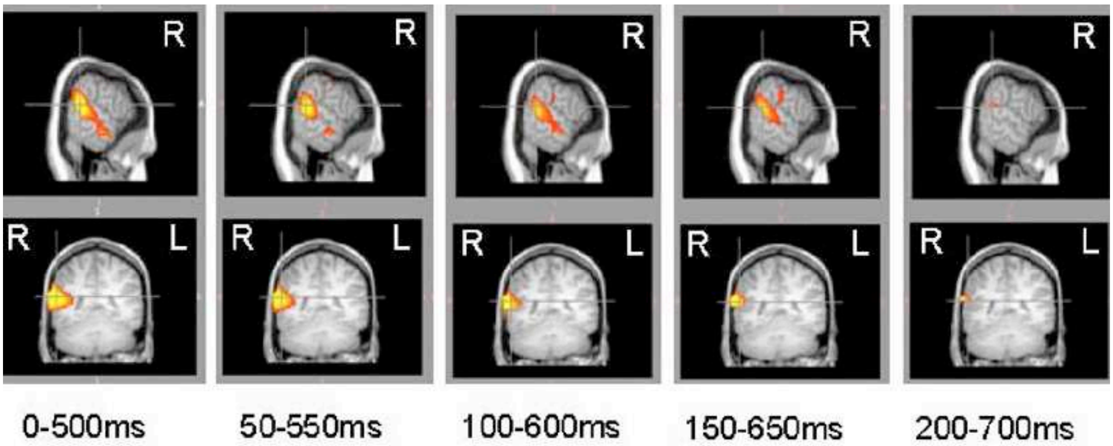
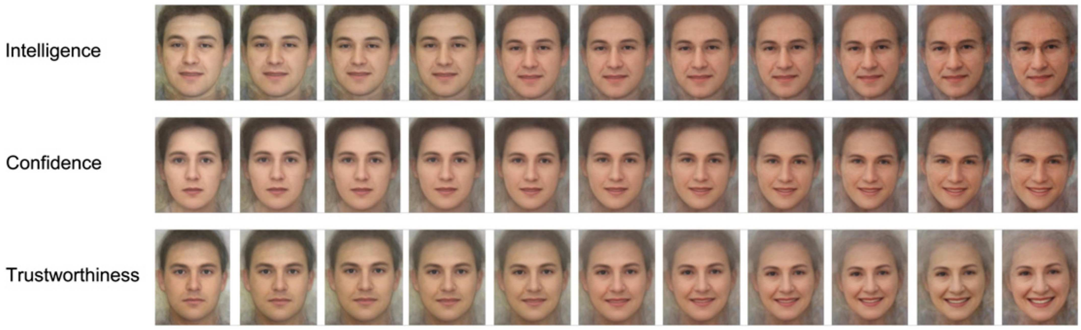


Figure 5: The upper panel shows everyday face images similar to those used by Sutherland et al. (2013). The central panel shows representative continua (intelligence, confidence, and trustworthiness) created by morphing between averages of 20 images rated as low (on the left) and 20 images rated as high (on the right) for each trait. The lower panel shows averages of the 20 images loading highest or lowest on factors of approachability, youthful-attractiveness, and dominance. (Central and lower panels reprinted from Sutherland et al., 2013, part of Figure 1, p. 109 and Figure 2A, p. 113, with permission from Elsevier).


Intelligence
Confidence
Trustworthiness

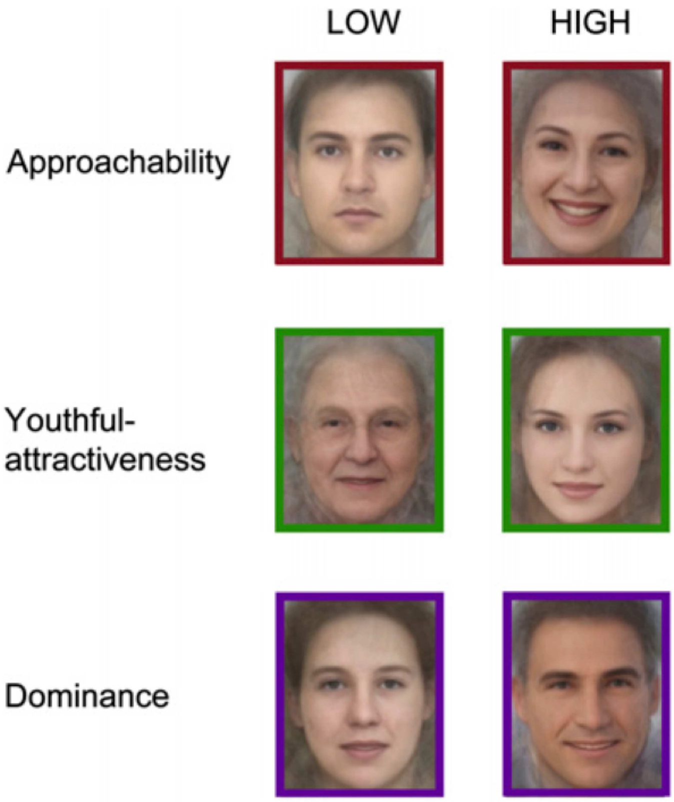

LOW
HIGH
Approachability
Youthful-attractiveness
Dominance

Figure 6: Scatterplots from Vernon et al.'s (2014) study showing the correlations between experimentally derived factor scores (from human raters) and the corresponding predictions for untrained images, derived from a linear neural network. Each data point (n = 1,000, for all axes) represents the observed and predicted ratings for a distinct face image. Overall correlations are 0.90 for approachability (scatterplot A), 0.70 for youthful attractiveness (scatterplot B), and 0.67 for dominance (scatterplot C).

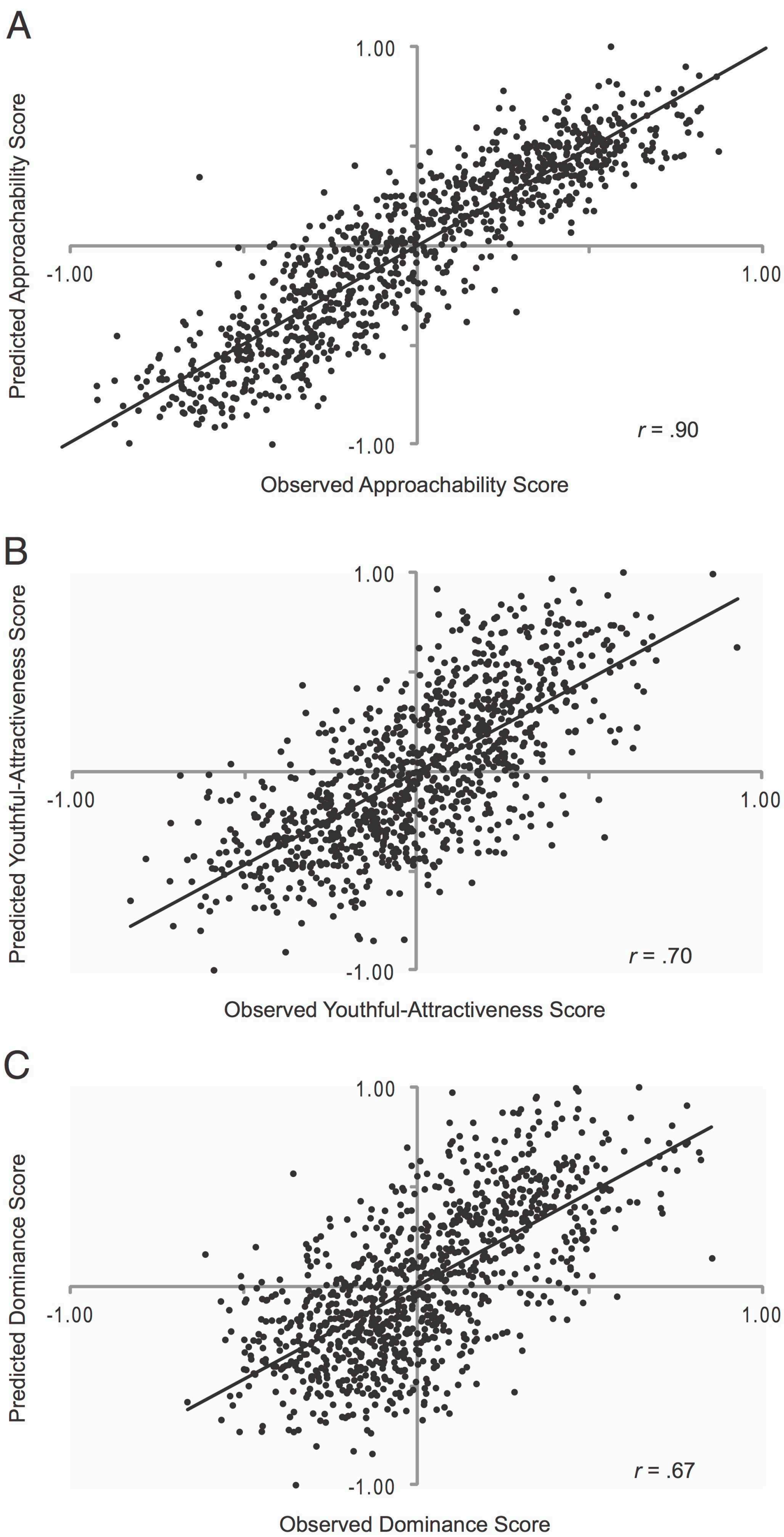

A
1.00
-1.00
1.00
-1.00
Predicted Approachability Score
Observed Approachability Score
r = .90
B
1.00
-1.00
1.00
-1.00
Predicted Youthful-Attractiveness Score
Observed Youthful-Attractiveness Score
r = .70
C
1.00
-1.00
1.00
-1.00
Predicted Dominance Score
Observed Dominance Score
r = .67

Figure 7: Data from Sutherland et al. (2017b). Mean trustworthiness (top), dominance (middle), and attractiveness (bottom) ratings of face photographs from the KDEF set (Lundqvist et al., 1998) of images varying in terms of identity, expression, and viewpoint. Results are plotted separately for female faces (left) and male faces (right). Each column represents a single identity, and each point represents a single photograph, with the overall mean rating for each identity shown as a darker point. The horizontal axis represents the between-person variability (the face identities, ranked by their overall mean trustworthiness, dominance, or attractiveness). The vertical axis represents the within-person variability (the different photographs of each person). Within-person variability (the differences between different images of the same face) is typically as large as between-person variability (the differences between different faces).

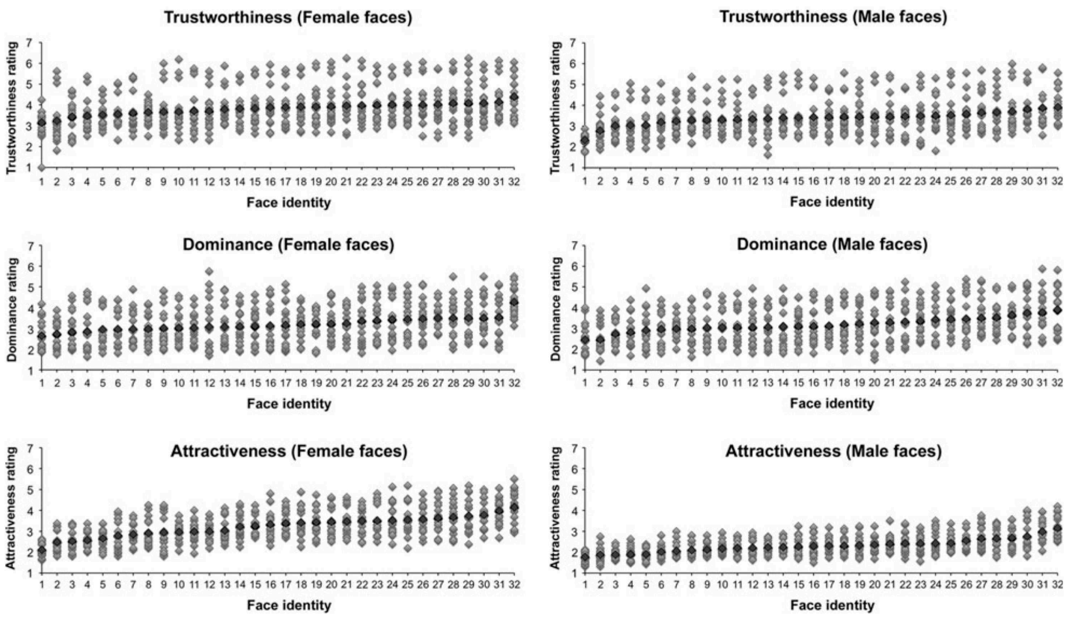


Figure 8: An example of the task used by Jenkins et al. (2011). Participants are asked to sort the 40 images into the different face identities. Most people only arrive at the correct solution if they already know the faces. For unfamiliar faces, participants tend to mistake differences between the images for differences in identity, leading them to overestimate the number of faces in the display. (Reprinted from Jenkins et al., 2011, Figure 2, p. 316, with permission from Elsevier).

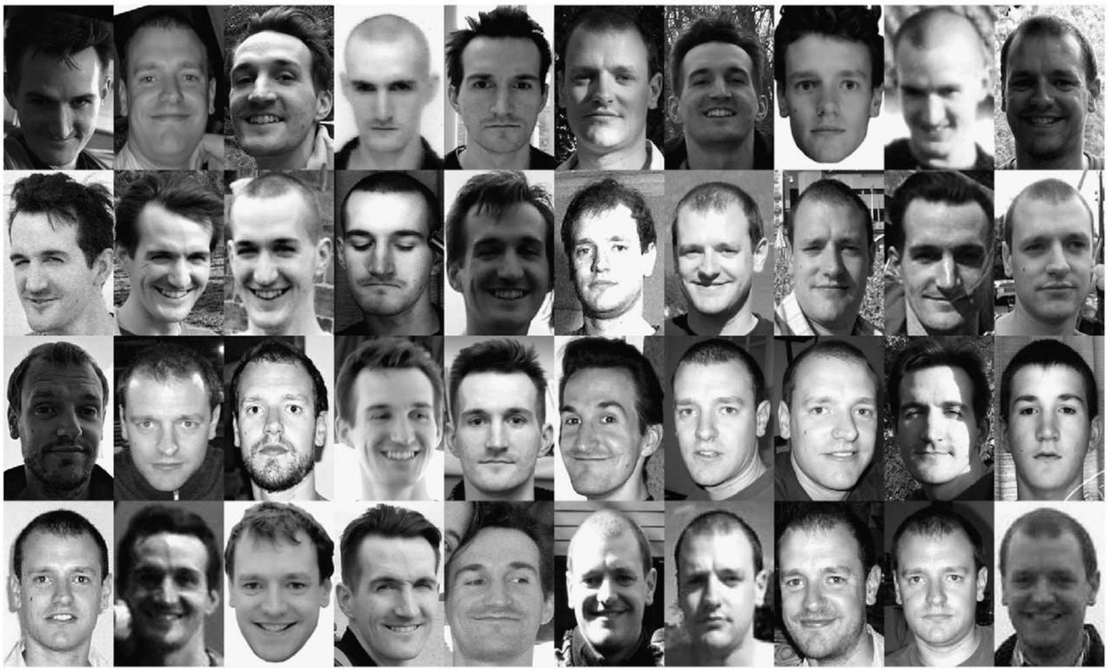


Figure 9: Kramer et al.'s (2017b) study. The upper panel shows sensitivity indices (d') for recognition of faces learnt from upright and inverted video, when tested with upright and inverted photographs. The lower panel shows sensitivity for recognition of faces learnt from positive and negative contrast greyscale videos, when tested with positive and negative contrast greyscale photographs. Errors bars represent 95% confidence intervals. An example test image in each format is shown alongside the relevant data. (Reproduced with permission from Taylor & Francis).