Authors' final version of manuscript accepted for publication in *Attention, Perception, & Psychophysics*. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article.

Date of acceptance: 11th June 2017.

THE AUTOMATICITY OF FACE PERCEPTION IS INFLUENCED BY FAMILIARITY

Xiaoqian Yan, Andrew W. Young and Timothy J. Andrews*

Department of Psychology, University of York, York YO10 5DD United Kingdom

Figures: 4

Word Count: 5570

Key words: face, race, identity, expression

Acknowledgements: XY was supported by an Overseas Research Scholarship (ORS) studentship from the University of York. We are grateful to Georgina Parratt, Kailie Cutler and Flossie Vice-Coles for their help with data collection in the early stages of this project. TJA is grateful to Isabel Gauthier a useful discussion on the question of expertise and automaticity in face perception. We also thank David White for helpful suggestions concerning the interpretation of the data.

ABSTRACT

In this study, we explore the automaticity of encoding for different facial characteristics and ask whether it is influenced by face familiarity. We used a matching task in which participants had to report whether the gender, identity, race or expression of two briefly presented faces was the same or different. The task was made challenging by allowing non-relevant dimensions to vary across trials. To test for automaticity, we compared performance on trials in which the task instruction was given at the beginning of the trial, with trials in which the task instruction was given at the end of the trial. As a strong criterion for automatic processing we reasoned that, if perception of a given characteristic (gender, race, identity or emotion) is fully automatic, the timing of the instruction should not influence performance. We compared automaticity for the perception of familiar and unfamiliar faces. Performance with unfamiliar faces was higher for all tasks when the instruction was given at the beginning of the trial. However, we found a significant interaction between instruction and task with familiar faces. Accuracy of gender and identity judgements to familiar faces was the same regardless of whether the instruction was given before or after the trial, suggesting automatic processing of these properties. In contrast, there was an effect of instruction for judgements of expression and race to familiar faces. These results show that familiarity enhances the automatic processing of some types of facial information more than others.

INTRODUCTION

The human face conveys a variety of different signals that are important for successful social interactions. The image of a face provides information about a person's identity, gender, race, emotional state and a range of other important attributes (Bruce & Young, 1986, 2012). The ease with which these properties appear to be detected and discriminated has led many researchers to suggest that face processing is automatic (Fiske & Neuberg, 1990; Freeman & Ambady, 2011; Öhman, 2002; Vuilleumier & Righart, 2011; for review see Palermo & Rhodes, 2007, 2010). However, other theories and findings suggest that extraction of social information from faces can also involve some degree of top-down control (Macrae & Bodenhausen, 2000; Santos & Young, 2005).

As is often the case in psychology, this debate may in part reflect different criteria for automaticity. Four main inter-related attributes have been suggested to characterise automatic from controlled processes (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977; Logan 1988; Moors, 2016; Gawronski & Creighton, 2013). First, automatic processes will be fast. Second, they may be to some degree non-conscious. Third, automatic processes are involuntary. Fourth, automatic processes require limited attentional resources.

In terms of these criteria, human faces are detected and categorized more quickly than many other non-face objects and even animal faces. Electrophysiological studies have found that faces can be categorized within 100ms of stimulus onset (Bentin et al., 1996; Liu et al., 2002; Oram & Perrett, 1992; Sugase et al., 1999), which is about 100ms earlier than for other objects (Pegna et al., 2004). Behavioural studies have also found that the detection of faces in natural scenes occurs earlier for faces than for animals (Rousselet et al., 2003) and that the threshold for detecting upright faces is lower compared to inverted or scrambled faces (Besson, et al., 2017; Purcell & Stewart, 1988). The detection and

discrimination of emotional faces has also been shown to occur within 100ms of stimulus onset (Eimer & Holmes, 2002; Pizzagalli et al., 1999; Pourtois et al., 2004). The ability to discriminate the identity of different individual faces takes longer, but some studies have sought to demonstrate that this can be achieved within 170ms (Heisz et al., 2006; Jacques & Rossion, 2006; Liu et al., 2002; Sugase et al., 1999; although see Schweinberger et al., 2002, and Ewbank et al., 2008).

A range of evidence also suggests that some information about faces can be processed in the absence of awareness. Behavioural studies in healthy participants have shown that faces that are not consciously perceived can nonetheless influence subsequent behavior. For example, the emotional expression from a briefly presented face that is not perceived can influence the subsequent perception of a neutral stimulus (Murphy & Zajonc, 1993; Rotteveel et al., 2001). Neuroimaging studies have also shown that responses to different facial expressions can be discriminated in the brain even when the participants are not conscious of seeing the faces (Pasley et al., 2004; Vuilleumier et al., 2002; Whalen et al., 1998; Williams et al., 2004; although see Phillips et al., 2004 and Pessoa et al., 2006). Consistent with these neuroimaging studies, studies of the blindsight patient GY have found above chance discrimination of facial expression when faces are presented in the blind visual field (de Gelder et al., 1999; Morris et al., 2001). Dissociations between awareness of recognition and behavioural or psychophysiological responses are also found in prosopagnosia, where differential responses to familiar compared to unfamiliar faces can be demonstrated despite the absence of conscious recognition (Bauer, 1984; Tranel & Damasio, 1985; de Haan, Young, & Newcombe, 1987; Young & Burton, 1999).

The extent to which faces are processed automatically can also be determined by the extent to which attention or task influences processing of the image. For example, evidence

for automatic processing of the identities of familiar faces has been shown in face-name interference tasks, in which the ability to categorize a name is affected by whether a distractor face is congruent or incongruent with the correct response (Lavie et al., 2003; Jenkins et al., 2003; Young et al., 1986). Other studies have shown that performance on a task involving one facial dimension (e.g. identity) can be influenced by changes in an irrelevant dimension (e.g. expression), suggesting automatic processing of the unattended dimension (Martin et al., 2015; Schweinberger & Soukup, 1998; Schweinberger, Burton, & Kelly, 1999). Neuroimaging studies also provide support for the mandatory processing of faces. For example, the response in the amygdala is not different to attended or unattended fearful faces (Vuilleumier et al., 2001). However, although these findings suggest that some facial properties are processed irrespective of the task, other studies have shown that neural responses to faces can be modulated by attention (Downing et al., 2001; Eimer, 2000; Holmes et al., 2003; O'Craven et al., 1999). The importance of attention is also demonstrated by the enhanced discrimination of faces that are attended (Palermo & Rhodes, 2002; Reinitz et al., 1994).

From this brief overview it is clear that the extent and limits of automaticity of face perception remain in some respects uncertain. A limitation of most previous studies of automaticity in face perception is that they have usually investigated only one dimension of the face. However, a range of evidence has suggested that different properties of the face may be processed independently (Duchaine & Yovel, 2015; Haxby et al., 2000; Young & Bruce, 2011). So, it remains unclear whether the level of automaticity varies across different facial dimensions. Another limitation in our understanding is that much of the key evidence is based on measures of brain imaging. Although these findings have helped our understanding, these measures are correlational in nature and it is not always clear how

patterns of brain response influence behavior. Finally, the majority of studies of automaticity have used unfamiliar faces (although see Gobbini et al., 2013; Jackson & Raymond, 2006; Tong & Nakayama, 1999; Visconti di Olleggio Castello & Gobbini, 2015), yet it is well-established that familiar faces are processed more effectively than unfamiliar faces (Hancock, Bruce & Burton, 2000; Burton et al., 2005; Jenkins & Burton, 2011). The question of whether the strong processing advantages that accrue to familiar faces extends beyond the perception of identity may therefore offer important insights into the nature and extent of automatic processing of different facial characteristics.

The aim of the current study, then, is to compare the automaticity of face perception across different characteristics and across unfamiliar and familiar faces. To achieve this, we introduce a new method that involves using a matching task in which participants report whether sequentially presented images of faces vary in gender, identity, race or facial expression. The task is made difficult because the faces are briefly presented, and because non-relevant dimensions can vary across trials (for example, faces to be matched for same or different gender may also vary in identity, race, or expression). To contrast automatic with controlled processing, we compare performance when the dimension on which the faces have to be matched is given at the beginning of each trial, allowing the possibility of more controlled processing, compared to when it is only given at the end of the trial, making performance therefore largely dependent on automatic processing of the preceding stimuli.

One of the main lessons of the extensive literature on automaticity is that it is not simply an 'all or none' phenomenon. To interpret the extent of automatic processing of different facial characteristics, we therefore use two different criteria for automaticity. Our first criterion is that a fully automatic process will not be affected by whether the task is

given at the beginning or at the end of the trial; that is, if a characteristic is processed automatically, you won't need to know in advance what to look for. Note that good performance when the task is only specified at the end of the trial will require that the participant can encode the target characteristic in both of the briefly presented faces, remember what was seen in the first face across the short inter-stimulus interval, and not be distracted from this memory by the second face in the sequence. It therefore represents a strong criterion that combines key properties of automaticity. Our second (weaker) criterion for automaticity is that performance when the task is given at the end of the trial will be above chance; this criterion accepts that some information that was automatically extracted from the first face may be lost across the brief interval in the matching trial, but stipulates that this loss should not reduce performance to chance level. Using these criteria, we investigated automatic processing of the gender, race, identity and facial expression of unfamiliar faces in Experiment 1, and then compared automatic processing of these characteristics between familiar and unfamiliar faces in Experiment 2.

EXPERIMENT 1

Experiment 1 investigated the automaticity of unfamiliar face perception using a matching task, in which the faces varied in gender (female or male), identity (same or different), race (Chinese or Caucasian) or emotional expression (happiness or surprise). The task instruction for what property to match (gender, identity, race, or emotion) in each trial was given to participants either at the beginning or the end of each trial.

Participants

Twenty-eight Caucasian participants from the University of York were recruited for this experiment (20 females; mean age, 20.1 years). All participants gave their written consent prior to the experiment. The University of York Department of Psychology Ethics Committee approved the study.

Images

Face stimuli were selected from two stimuli sets; the CAS - PEAL R1 face database (Gao, et al., 2008) posed by Chinese models, and the Karolinska Directed Emotional Faces (KDEF) (Lundqvist, Flykt, & Öhman, 1998) posed by Caucasian models. In each set, images of 40 different identities with each identity posing two facial expressions (one happiness and one surprise) were selected, for each gender (female and male) and each race (Caucasian and Chinese), resulting in a total of 320 images. These 320 images were then rated and screened by three raters for each of the principal categories of interest (race, gender, and emotion) to choose 20 images each with the highest consistency among raters, resulting in a total of 160 images.

All face stimuli were cropped around the outline of the face and resized to 300 x 240 pixels and converted to greyscale. When viewed from 57 cm away, each image extended to a visual angle of approximately 7.9 degrees high and 6.4 degrees wide. Figure 1 shows examples of stimuli used in Experiment 1.

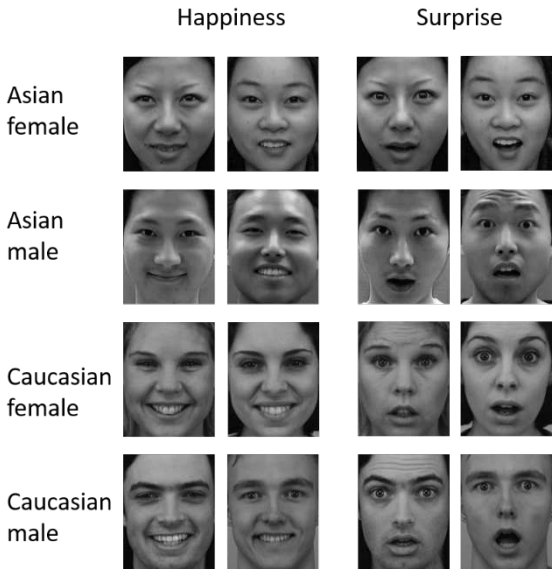

Figure 1 *Examples of stimuli used in Experiment 1, with two Asian and two Caucasian female and male models showing happiness and surprise expressions.*

Procedure

Participants viewed faces presented on a computer screen using PsychoPy (www. psychopy.org). Participants performed a same/different matching task (Figure 2). This involved judging whether the two face images were the same or different on one of four characteristics (gender, identity, race, emotion). A central fixation cross was presented throughout each trial. At the beginning of each trial, the fixation cross was presented on a grey background for 500ms, prior to the presentation of a face image for 100ms. This was followed by a phase-scrambled mask image for 900ms and then another face image for 100ms. Participants were instructed to respond as quickly and as accurately as possible after

the second image was presented. On half of the trials, the task instruction (i.e. the characteristic to be judged as same or different) was given at the beginning of the trial and on half the trials the task instruction was given at the end of the trial. Trials with different tasks (identity, race, gender, emotion) and different instruction timings (before, after) were randomly interleaved.

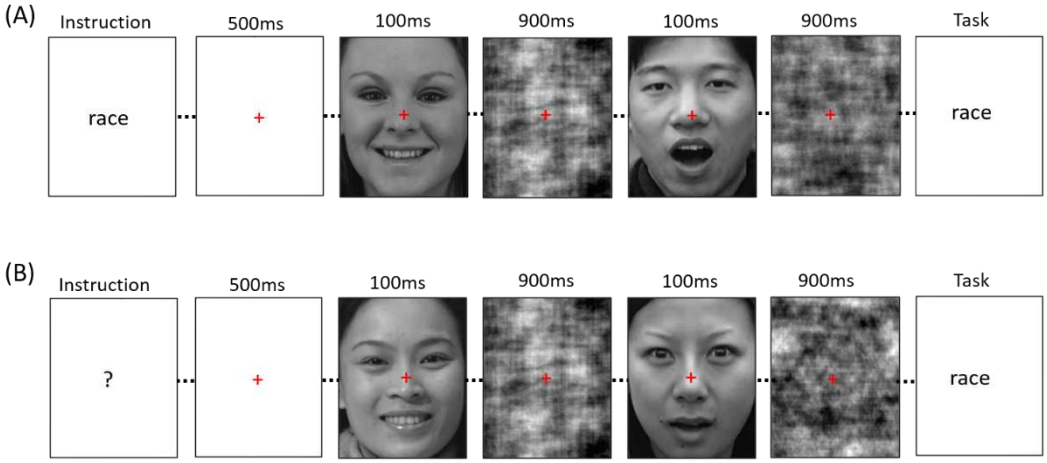

Figure 2 *Example trials for before task (A) and after task (B) instructions. The face pair in the before task instruction example shows a "Different race - Different gender - Different identity - Different emotion" pairing, and the face pair in the after task instruction example shows a "Same race - Same gender - Different identity - Different emotion" pairing. In both of these two example trials the participant is asked to make a judgment on race.*

For judgements of race, gender and emotion, it was possible to vary all non-relevant dimensions independently. For example, on a 'race' trial, there could be changes in the gender, emotion or identity of the face that were not relevant to the judgement of race. In contrast, judgements of identity are confounded by changes in gender and identity. For example, if two images differ in gender or race, it is possible to infer that the identity is also different. Thus, for judgements of identity, the only non-relevant dimension that could be varied was facial expression. We created 20 same trials and 20 different trials for judging each characteristic with the before and after task instruction timings, respectively. Within

these trials the non-judged characteristics were systematically varied to create all possible combinations. Hence (for example) a pair of face images to be judged as having the same expression might vary on gender, identity, or race, and a pair with different expressions might have the same gender, identity, or race. Ten additional trials were also included prior to the main experiment, to form a practice run. The whole experiment took approximately 30min. In line with standard practice for bounded data, % correct values were arcsine transformed before any statistical analysis was performed.

Results

Our principal interest is in the accuracy data. The mean accuracy for the different task variants is shown in Figure 3. A repeated-measures ANOVA was performed on the data with Instruction timing (before, after) and Task (gender, identity, race, emotion) as within-subjects factors. We found significant main effects of both Instruction ($F(1,27) = 106.8$, $p < .001$, partial $\eta^2 = 0.80$) and Task ($F(3,81) = 77.4$, $p < .001$, partial $\eta^2 = 0.74$). The effect of Instruction reflects higher accuracy when the task was given at the beginning of the trial compared to when it was given at the end of the trial. There was also a significant interaction of Instruction x Task ($F(3,81) = 4.0$, $p = .01$, partial $\eta^2 = 0.13$). This interaction reflects the fact that the effect of Instruction was greater for the race task, followed by emotion, identity, and then the gender task. However, for each task, performance was always greater when the instruction was given before the trial (Gender: $F(1,27) = 14.52$, $p < .001$; Identity: $F(1,27) = 25.80$, $p < .001$; Race: $F(1,27) = 60.47$, $p < .001$; Emotion: $F(1,27) = 32.11$, $p < .001$). Finally, we determined whether accuracy was above chance level (0.5) for each condition using a one-sample t-test. All conditions were significantly above chance (Gender-before: $t_{(27)} = 20.31$, $p < .001$, Gender-after: $t_{(27)} = 13.64$, $p < .001$; Identity-before:

$t_{(27)}$ = 11.17, p < .001, Identity-after: $t_{(27)}$ = 5.68, p < .001; Race-before: $t_{(27)}$ = 22.61, p < .001, Race-after: $t_{(27)}$ = 12.73, p < .001; Emotion-before: $t_{(27)}$ = 14.07, p < .001, Emotion-after: $t_{(27)}$ = 8.74, p < .001).

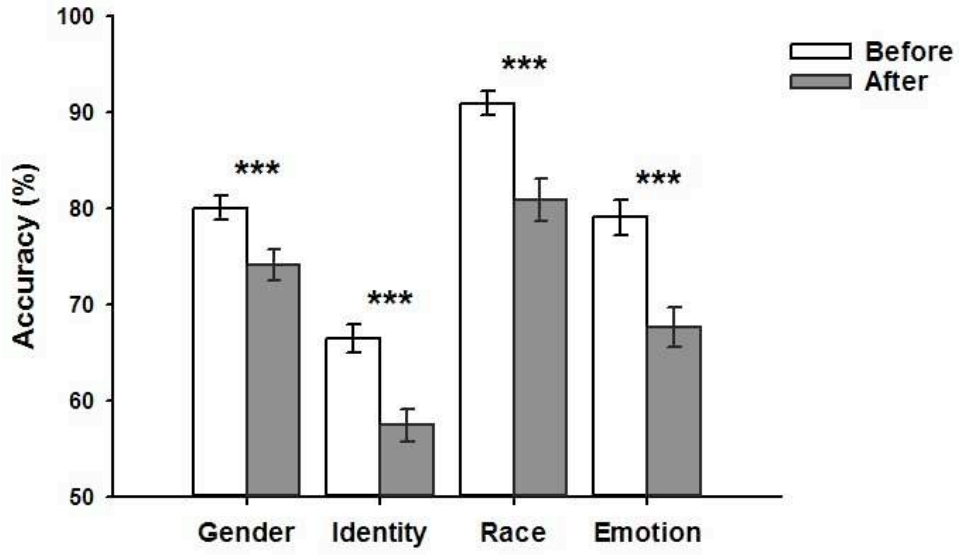

Figure 3 *Overall matching accuracies (with error bars) in different tasks for unfamiliar faces in Experiment 1. Asterisks denote higher recognition accuracies when the task instruction was given before (rather than after) the face images in the trial. ***: p < .001. Chance level performance = 50% correct.*

Discussion

In this experiment, we investigated the automaticity of unfamiliar face processing across gender, identity, race and emotional expression. Our strong criterion for automatic processing is that performance is not affected by whether the instruction is given at the beginning or at the end of trial. However, we always found a significant effect of instruction; when the instruction was given at the beginning of a trial, performance was significantly higher than when the instruction was given at the end of the trial. There was none the less an interaction between instruction and task, reflecting the fact that the effect of instruction

varied across the four tasks. Nevertheless, for all tasks there was a significant effect of instruction, ruling out evidence of automaticity on the strong criterion.

Next, we asked whether the data support the weaker criterion of automaticity, in which accuracy on trials in which the task instruction was given at the end were above chance. Using this weaker criterion, we found evidence of automaticity in the form of above chance performance for all tasks. Taken together, these results therefore show partial support for the automatic processing of unfamiliar faces, though we note that accuracy in the identity matching condition in particular was not far from chance level when the task instruction was only given at the end of the trial. In fact the generally poor performance of identity matching with these unfamiliar faces is consistent with other reports of limited ability to judge unfamiliar face identity (e.g., Hancock et al., 2000; Jenkins & Burton, 2011). In Experiment 2 we therefore contrasted automaticity across familiar and unfamiliar faces.

EXPERIMENT 2

The aim of this experiment was to determine the effect of familiarity on automaticity processing of different facial characteristics. This experiment used an equivalent matching task to Experiment 1, involving judgements of gender, identity, race, and emotion. However, we compared performance for familiar and unfamiliar face images. Again, we used two criteria for demonstrating automaticity. Our strong criterion would involve no difference in performance between trials in which the task instruction is given at the beginning or at the end of a trial. Our weaker criterion for automaticity is that performance on a task when the task instruction is given at the end of the trial should be above chance.

Participants

Twenty-one Caucasian participants from the University of the York were recruited (18 females; mean age, 21.45 years). One participant was removed from analyses because less than 80% (more than 8 identities out of 40) of famous face images were recognized. All participants gave their written consent prior to the experiment. The University of York Department of Psychology Ethics Committee approved the study.

Images

Images of familiar and unfamiliar white and black faces were obtained from the internet. There were 10 identities for each combination of race (white or black), gender (female or male), and familiarity (familiar or unfamiliar) dimensions. In addition, happy and neutral (instead of surprise, for ease of image selection online) face images were selected for each identity. This gave a total of 80 familiar faces and 80 unfamiliar faces. All face stimuli were cropped and resized to 300 x 240 pixels and converted to greyscale. When viewed from 57 cm away, each image extended to a visual angle of approximately 7.9 degrees high and 6.4 degrees wide.

Procedure

The sequential matching task from Experiment 1 was used, with the following differences: (1) Each face image was presented for 150ms, and the mask was presented for 850ms; (2) There were 10 same trials and 10 different trials for each dimension for each before/after task instruction; (3) Familiar and unfamiliar trials were randomly interleaved within each block.

We tested the familiarity of the familiar faces for each participant after the matching task was completed. Each participant was presented with a set of images of the familiar faces used in the experiment. None of these images were similar to those used in the main experiment. Participants were asked to write down the name or any relevant identifying information for each face. In this way we established that participants were able to recognize over 90% of the images of the celebrities used in the experiment. In the main experiment, trials that included a familiar face that was not recognized by a participant during this post-task screening test were not included in the analysis of that participant. Approximately 9% of trials were excluded on this basis.

Results

The mean accuracy of matching performance for familiar and unfamiliar faces is shown in Figure 4. A three-way repeated ANOVA was performed with Instruction timing (before, after), face Familiarity (familiar, unfamiliar), and Task (gender, identity, race, emotion) as within-subjects variables. There were significant main effects of Instruction ($F(1,19) = 46.54$, $p < .001$, partial $\eta^2 = 0.71$), Familiarity ($F(1,19) = 32.98$, $p < .001$, partial $\eta^2 = 0.63$), and Task ($F(3,57) = 41.0$, $p < .001$, partial $\eta^2 = 0.68$). However, there was also a significant three-way interaction of Instruction x Familiarity x Task ($F(3,57) = 3.83$, $p < .05$, partial $\eta^2 = 0.17$).

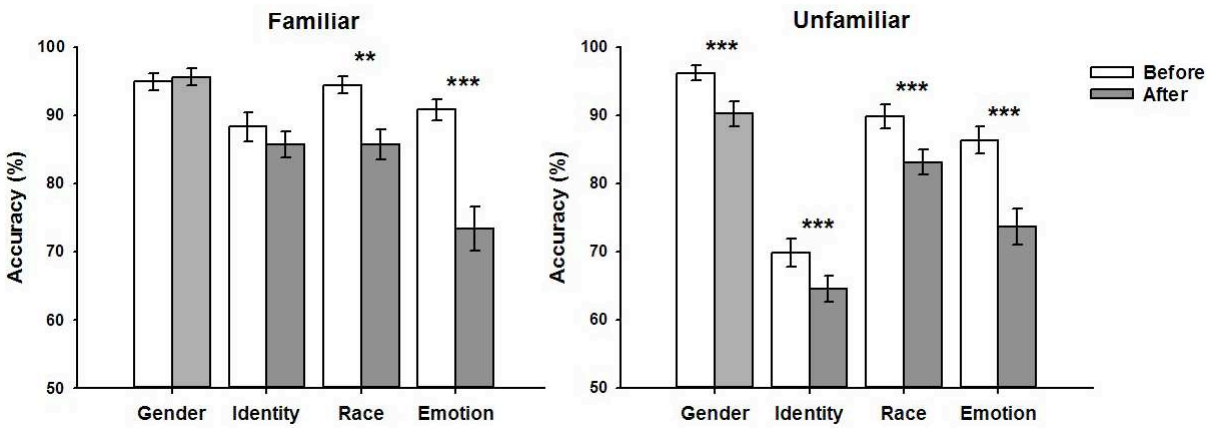

Figure 4 *Overall matching accuracies (with error bars) for familiar faces and unfamiliar faces in different tasks in Experiment 2. Asterisks denote higher recognition accuracies when the task instruction was given before (rather than after) the face images in the trial. **: p < .01, ***: p < .001. Chance level performance = 50% correct.*

To decompose the three-way interaction, separate two-way ANOVAs were conducted for familiar and for unfamiliar faces, with Instruction timing (before, after) and Task (gender, identity, race, emotion) as within-subjects factors. For familiar faces, there was a main effect of Instruction timing ($F(1,19) = 16.50$, $p < .001$, partial $\eta^2 = 0.47$) and Task ($F(3,57) = 14.11$, $p < .001$, partial $\eta^2 = 0.43$). There was also a significant interaction of Instruction x Task ($F(3,57) = 9.23$, $p < .001$, partial $\eta^2 = 0.33$). The interaction reflected the fact that there was no effect of instruction timing on some tasks but on other tasks there was better performance when the task instruction given at the beginning of the trial. For example, there was no difference in accuracy for both gender ($t(19) = 0.46$, $p > .1$) and identity ($t(19) = 1.14$, $p > .1$) when the instruction was given before or after the trial. In contrast, accuracy was higher when the instruction was given before the trial in judgements of race ($t(19) = 3.6$, $p < .01$) and expression ($t(19) = 6.81$, $p < .001$). None the less, one-sample t-tests showed that performance was above chance level (0.5) for all tasks when the instruction was given at the end of the trial (Gender-before: $t_{(19)} = 16.55$, $p < .001$, Gender-after: $t_{(19)} = 17.45$, $p < .001$; Identity-before: $t_{(19)} = 11.25$, $p < .001$, Identity-after: $t_{(19)} = 10.73$, $p < .001$; Race-before: $t_{(19)} = 16.32$, $p < .001$, Race-after: $t_{(19)} = 12.31$, $p < .001$; Emotion-before: $t_{(19)} = 13.82$, $p < .001$, Emotion-after: $t_{(19)} = 6.23$, $p < .001$).

For unfamiliar faces, there were main effects of Instruction timing ($F(1,19) = 30.97$, $p < .001$, partial $\eta^2 = 0.62$) and Task ($F(1,19) = 53.3$, $p < .001$, partial $\eta^2 = 0.74$). However, there

was no significant interaction between Instruction x Task ($F(3,57) = 1.65$, $p > .1$, partial $\eta^2 = 0.08$). Finally, one-sample t-tests showed that performance was above chance level (0.5) for all tasks when the instruction was given at the end of the trial (Gender-before: $t_{(19)} = 18.58$, $p < .001$, Gender-after: $t_{(19)} = 12.99$, $p < .001$; Identity-before: $t_{(19)} = 9.48$, $p < .001$, Identity-after: $t_{(19)} = 7.7$, $p < .001$; Race-before: $t_{(19)} = 13.40$, $p < .001$, Race-after: $t_{(19)} = 12.17$, $p < .001$; Emotion-before: $t_{(19)} = 11.91$, $p < .001$, Emotion-after: $t_{(19)} = 8.07$, $p < .001$).

Discussion

This experiment investigated automaticity for processing different characteristics of familiar and unfamiliar faces. Consistent with Experiment 1, we found that performance with unfamiliar faces was better when the instruction was given at the beginning of the trial compared to when it was given at the end of the trial. Although unfamiliar faces therefore again failed to meet our strong criterion of automaticity, performance on trials in which the task was given at the end of the trial were above chance, showing that they could satisfy the weak criterion for automaticity. The pattern of results for unfamiliar faces therefore replicates the main findings from Experiment 1.

In contrast, we found an interaction between the effect of instruction timing and task with familiar faces. This reflected the lack of any difference between whether the instruction was given before or after the trial for judgements of gender and identity. This satisfies our strong criterion for automaticity of gender and identity perception with familiar faces. However, we found that performance on judgements of the race or expression of familiar faces was higher when the task was given before the trial; for these characteristics only the weaker criterion of above-chance performance when the instruction was given at the end of the trial was met. These results show that the familiarity of the faces does

influence the automaticity with which they are processed, but that this influence is more pronounced for judgements of gender and identity.

GENERAL DISCUSSION

The aim of this study was to investigate whether face perception is automatic and to what extent this varies across different facial characteristics. Participants performed a matching task on two sequentially presented faces. Across trials, faces varied randomly in four dimensions: race, gender, identity, and emotion. At the end of each trial participants had to make a same/different judgement on one of the dimensions.

As a strong test for automaticity, we compared the responses to trials when the instruction for the trial was given at the beginning with trials in which the instruction was given at the end. If perception of a particular characteristic is fully automatic, we reasoned that there should be no difference in accuracy whether the task was given at the beginning or the end of the trial. In Experiment 1, however, we found higher accuracy with unfamiliar faces across all dimensions when the instruction was given at the beginning of the trial, showing a lack of full automaticity. Unfamiliar faces only met weaker criterion for automaticity of above chance performance when the instruction was given at the end of the trial.

In Experiment 2, we asked how familiarity influences automatic processing. Consistent with Experiment 1, accuracy with unfamiliar faces was higher across all dimensions when the instruction was given at the beginning of the trial; these findings with unfamiliar faces again did not meet our strong criterion for automatic processing. However, for familiar faces the accuracy of gender and identity judgements was not affected by the timing of the task instruction. None the less, race and expression judgements with familiar faces still only met our weaker criterion for automaticity of above chance performance in the delayed instruction condition. Taken together, these findings show that whilst the processing of

unfamiliar faces is not fully automatic, familiarity with a face increases automaticity for certain facial characteristics.

Our findings are relevant to current theories on the social categorization of faces. Categories such as race and gender can largely be determined from purely visual facial properties (Kramer, Young, Day & Burton, 2017) and some theories therefore suggest that their extraction may involve entirely bottom-up, pre-attentive, and automatic processes (Bargh, 1999; Brewer, 1988; Fiske & Neuberg, 1990; Freeman & Ambady, 2011). However, other theories maintain that there is also some top-down control of social categorization (Macrae & Bodenhausen, 2000). Our findings suggest that an interaction between bottom-up and top-down processes can explain the categorization of unfamiliar faces. We found that performance on all facial judgements with unfamiliar faces was above chance even when the characteristic that had to be matched was not given until the end of the trial. This demonstrates a clear contribution from bottom-up processing of different facial dimensions. Moreover, this bottom-up processing showed some limited form of automaticity because participants did not know to which aspect of the images they should attend until after both images were presented. However, we also found that performance with unfamiliar faces was always better when the task was given at the beginning of the trial, which provides a clear demonstration of the involvement of top-down control.

The majority of previous studies on automaticity have used unfamiliar faces. However, it is well-established that familiar faces are processed more effectively than unfamiliar faces (Burton et al., 2005; Hancock et al., 2000; Jenkins & Burton, 2011). In Experiment 2, we compared automatic processing of familiar and unfamiliar faces. Critically, we found an interaction between the timing of instruction, facial dimension and familiarity. This interaction was explained by the fact that judgements of gender and identity from

familiar faces were not different when the instruction was given at the beginning compared to the end of the trial. These findings meet our strong criterion for automaticity and are consistent with previous work that has shown fewer attentional resources are required to detect familiar compared to unfamiliar faces (Gobbini et al., 2013; Jackson & Raymond, 2006; Tong & Nakayama, 1999; Visconti di Olleggio Castello & Gobbini, 2015).

Interestingly, familiarity did not lead to strongly automatic processing of all facial dimensions. Accuracy on race and emotion trials was significantly lower when the task instruction was given at the end of the trial. It makes sense that familiarity helps with judgements of facial identity, as it is known that judgements of identity are much easier to make with familiar compared to unfamiliar faces (Hancock et al., 2000; Jenkins et al., 2011; Jenkins & Burton, 2011; Davies-Thompson et al., 2013; Young & Burton, 2017). Familiarity with a face also allows us to retrieve a range of semantic information associated with a person. The effect of familiarity on the automaticity of gender judgements may therefore suggest either that the perceptual representation of identity is linked to the perceptual representation of gender, or that the automatic recognition of familiar face identity allowed participants to infer the gender. In this respect it was surprising that familiarity did not influence the automaticity of race judgements. Although it may seem intuitively that a person's identity should be tightly linked to our representation of their race, previous studies have suggested that this is not always the case. For example, Phelps and colleagues (2000) showed that the difference in the amygdala response to black compared to white unfamiliar faces could be predicted by implicit measures of racial attitude, yet a similar effect was not evident when familiar black and white faces are evaluated. Likewise, behavioural studies show that familiar other-race faces do not show characteristics that are often associated with the classic other-race effect (McKone et al., 2007). This suggests that

when a face becomes familiar the neural representation of identity may become dissociated from the representation of race. We note, though, that despite the steps taken to establish that participants could recognise the familiar faces with unlimited presentation time, some errors were none the less made in the identity matching task, making it clear that the familiar faces were not always succesfully recognised with the brief masked presentations. It is therefore also possible that these failures in recognition might have influenced performance in the race classification task.

From some perspectives, it is perhaps not surprising that face familiarity did not confer any advantage in judgments of facial expression. Since the idea was put forward by Bruce and Young (1986), models of face processing have suggested that invariant properties of the face such as identity are processed independently of changeable properties such as expression (Andrews & Ewbank, 2004; Duchaine & Yovel, 2015; Haxby et al., 2000; Young & Bruce, 2011). Against this mainstream opinion, however, there have been reviews that have pointed to limitations in the available evidence (Calder & Young, 2005) and findings that point to some interaction between familiar face identity and expression (Kaufmann & Schweinberger, 2004; Martens, Leuthold & Schweinberger, 2010). Our findings of clear differences between the automaticity of judgements of identity and expression for familiar faces therefore have important implications as they imply substantial separation between the processing of identity and expression.

In conclusion, this study has developed a novel behavioral paradigm to investigate the automaticity of face perception. We found evidence for partial automaticity in the processing of unfamiliar faces. However, for all dimensions tested (gender, identity, race, emotion), there was evidence of significant top-down control. In contrast to unfamiliar faces, we found full automaticity for judgements of gender and identity to familiar face,

whereas judgements of race and emotion were again only partially automatic. These results demonstrate the importance of familiarity in the automaticity of face perception, but show that familiarity has differential effects on perceiving different facial characteristics.

REFERENCES

Andrews, T. J., & Ewbank, M. P. (2004). Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *NeuroImage, 23*(3), 905–913.

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer (Ed.), *Advances in social cognition: The automaticity of everyday life*, Vol. 10, (pp. 1–61). Mahwah, NJ: Lawrence Erlbaum.

Bauer, R. (1984). Automatic recognition of names and faces: A neuropsychological application of the guilty knowledge test. *Neuropsychologia, 22*, 457–469.

Bentin, S., & Carmel, D. (2002). Accounts for the N170 face-effect: A reply to Rossion, Curran, & Gauthier. *Cognition, 85*, 197−202.

Besson, G., Barragan-Jason, G., Thorpe, S. J., Fabre-Thorpe, M., Puma, S., Ceccaldi, M., & Barbeau, E. J. (2017). From face processing to face recognition: Comparing three different processing levels. *Cognition, 158*, 33−43.

Brewer, M.B. (1988). A dual-process model of impression formation. In Srull, T.K., Wyer, R.S., editors. *Advances in social cognition.* Vol. 1, pp. 1–36. Mahwah, NJ: Lawrence Erlbaum.

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

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

Burton, A. M., Jenkins, R., Hancock, P. J. B., & White, D. (2005). Robust representations for face recognition: the power of averages. *Cognitive Psychology, 51*(3), 256−284.

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience, 6*, 645−651.

Davies-Thompson, J., Newling, K., & Andrews, T. J. (2013). Image-invariant responses in face-selective regions do not explain the perceptual advantage for familiar face recognition. *Cerebral Cortex, 23*, 370−377.

De Gelder, B., Vroomen, J., Pourtois, G., & Weiskrantz, L. (1999). Nonconscious recognition of affect in the absence of striate cortex. *NeuroReport, 10*, 3759–3763.

De Haan, E. H. F., Young, A. W., & Newcombe, F. (1987). Face recognition without awareness. *Cognitive Neuropsychology, 4*, 385−415.

Downing, P., Liu, J., & Kanwisher, N. (2001). Testing cognitive models of visual attention with fMRI and MEG. *Neuropsychologia, 39*(12), 1329−1342.

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

Eimer, M. (2000). Attentional modulations of event-related brain potentials sensitive to faces. *Cognitive Neuropsychology, 17*, 103–116.

Eimer, M., & Holmes, A. (2002). An ERP study on the time course of emotional face processing. *NeuroReport, 13*, 427–431.

Ewbank, M. P., Smith, W. A. P., Hancock, E. R., & Andrews, T. J. (2008). The M170 reflects a viewpoint-depended representation for both familiar and unfamiliar faces. *Cerebral Cortex, 18*, 364−370.

Fisk, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology, 23*, 1−74.

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review, 118*(2), 247−279.

Gao, W., Cao, B., Shan, S. G., Zhou, D. L., Zhang, X. H., & Zhao, D. B. (2008). The CAS - PEAL large - scale Chinese face database and baseline evaluations. (2008). IEEE *Transactions on Systems, Man, and Cybernetics (Part A), 38*(1), 149−161.

Gawronski, B., & Creighton, L. A. (2013). Dual-process theories. In D. E. Carlston (Ed.), *The Oxford handbook of social cognition* (pp. 282−312). New York: Oxford University Press.

Gobbini, M. I., Gors, J. D., Halchenko, Y. O., Rogers, C., Guntupalli, J. S., Hughes, H., & Cipolli, C. (2013). Prioritized detection of personally familiar faces. *PLoS ONE, 8*(6), e66620.

Hancock, P. J. B., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in cognitive sciences, 4*(9), 330−337.

Haxby, J. V., & Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in cognitive sciences, 4*(6), 223−233.

Heisz, J. J., Watter, S., & Shedden, J. M. (2006). Automatic face identity encoding at the N170. *Vision Research, 46*(28), 4604−4614.

Holmes, A., Vuilleumier, P., & Eimer, M. (2003). The processing of emotional facial expression is gated by spatial attention: Evidence from event-related brain potentials. *Cognitive Brain Research, 16*, 174–184.

Jackson, M. C., & Raymond, J. E. (2006). The role of attention and familiarity in face identification. *Perception and Psychophysics, 68*(4), 543−557.

Jacques, C., & Rossion, B. (2006). The speed of individual face categorization. *Psychological Science, 17*, 485−492.

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

Jenkins, R., Lavie, N., & Driver, J. (2003). Ignoring famous faces: Category-specific dilution of distractor interference. *Perception & Psychophysics, 65*(2), 298−309.

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

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

Lavie, N., Ro, T., & Russell, C. (2003). The role of perceptual load in processing distractor faces. *Psychological Science, 14*(5), 510–515.

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: an MEG study. *Nature Neuroscience, 5*(9), 910−916.

Logan, G. D. (1988). Automaticity, resources, and memory: theoretical controversies and practical implications. *Human Factors, 30*, 583−598.

Lundqvist, D., Flykt, A., Öhman, A. (1998). *The Karolinska Directed Emotional Faces – KDEF* [CD ROM]. Stockholm: Karolinska Institutet, Department of Clinical Neuroscience, Psychology Section.

Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology, 51*, 93−120.

Martens, U., Leuthold, H., & Schweinberger, S. R. (2010). Parallel processing in face perception. *Journal of Experimental Psychology: Human Perception & Performance, 36,* 103−121.

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.

McKone, E., Brewer, J. L., MacPherson, S., Rhodes, G., & Hayward, W. G. (2007). Familiar other-race Faces show normal holistic processing and are robust to perceptual stress. *Perception, 36,* 224−248.

Moors, A. (2016). Automaticity: Componential, Causal, and mechanistic explanations. *Annual Review of Psychology, 67*, 263−287.

Morris, J. S., de Gelder, B., Weiskrantz, L., & Dolan, R. J. (2001). Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. *Brain, 124*, 1241–1252.

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality & Social Psychology, 64*(5), 723–739.

O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature, 401*, 584−587.

Öhman, A. (2002). Automaticity and the amygdala: Nonconscious responses to emotional faces. *Current Directions in Psychological Science, 11*, 62−66.

Oram, M. W., & Perrett, D. I. (1992). Time course of neural responses discriminating different views of the face and head. *Journal of Neurophysiology, 68*, 70−84.

Palermo, R., & Rhodes, G. (2002). The influence of divided attention on holistic face perception. *Cognition, 82*(3), 225–257

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

Pasley, B. N., Mayes, L. C., & Schultz, R. T. (2004). Subcortical discrimination of unperceived objects during binocular rivalry. *Neuron, 42*, 163−172.

Pegna, A. J., Khateb, A., Michel, C. M., & Landis, T. (2004). Visual recognition of faces, objects, and words using degraded stimuli: Where and when it occurs. *Human Brain Mapping, 22*, 300−311.

Pessoa, L., Japee, S., Sturman, D., & Ungerleider, L. G. (2006). Target visibility and visual awareness modulate amygdala responses to fearful faces. *Cerebral Cortex, 16*, 366–375.

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience, 12*, 729−738.

Phillips, M. L., Williams, L. M., Heining, M., Herba, C. M., Russell, T., Andrew, C., Bullmore, E. T., Brammer, M. J., Williams, S. C. R., Morgan, M. J., Young, A. W., & Gray, J. A. (2004). Differential neural responses to overt and covert presentations of facial expressions of fear and disgust. *NeuroImage, 21*, 1486−1498.

Pizzagalli, D. A., Regard, M., & Lehmann, D. (1999). Rapid emotional face processing in the human right and left brain hemispheres: An ERP study. *NeuroReport, 10*, 2691–2698.

Pourtois, G., Grandjean, D., Sander, D., & Vuilleumier, P. (2004). Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cerebral Cortex, 14*(6), 619–633.

Purcell, D. G., & Stewart, A. L. (1988). The face-detection effect: Configuration enhances detection. *Perception & Psychophysics, 43*, 355–366.

Reinitz, M. T., Morrissey, J., & Demb, J. (1994). Role of attention in face encoding. *Journal of Experimental Psychology: Learning Memory & Cognition, 20*, 161–168.

Rotteveel, M., de Groot, P., Geutskens, A., & Phaf, R. H. (2001). Stronger suboptimal than optimal affective priming? *Emotion, 1*(4), 348–364.

Rousselet, G. A., Macé, M. J.-M., & Fabre-Thorpe, M. (2003). Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes. *Journal of Vision, 3*, 440−455.

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

Schneider, W., & Shiffrin, R. M. (1977) Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*, 1-66.

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.

Schweinberger, S. R., Pickering, E. C., Burton, A. M., & Kaufmann, J. M. (2002). Human brain potential correlates of repetition priming in face and name recognition. *Neuropsychologia, 40*, 2057−2073.

Shiffrin, R. M., & Schneider, W. (1977) Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127−189.

Sugase, Y., Yamane, S., Ueno, S., & Yawano, K. (1990). Global and fine information coded by single neurons in the temporal visual cortex. *Nature, 400*, 869−872.

Tong, F., & Nakayama, K. (1999). Robust representations for faces: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1–20.

Tranel, D., & Damasio, A. R. (1985). Knowledge without awareness: An autonomic index of facial recognition by prosopagnosics. *Science, 228*, 1453–1454.

Visconti di Oleggio Castello, M., & Gobbini, M. I. (2015). Familiar Face Detection in 180ms. *PLoS ONE, 10*(8), e0136548.

Vuilleumier, P., Armony, J. L., Clarke, K., Husain, M., Driver, J., & Dolan, R. J. (2002). Neural responses to emotional faces with and without awareness: Event-related fMRI in a parietal patient with visual extinction and spatial neglect. *Neuropsychologia, 40*, 2156–2166.

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: An event-related fMRI study. *Neuron, 30*, 829−841.

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulates amygdala activity without explicit knowledge. *Journal of Neuroscience, 18*, 411–418.

Williams, M. A., Morris, A. P., McGlone, F., Abbott, D. F., & Mattingley, J. B. (2004). Amygdala responses to fearful and happy facial expressions under conditions of binocular suppression. *The Journal of Neuroscience, 24*(12), 2898–2904.

Young, A. W., & Bruce, V. (2011). Understanding person perception. *British Journal of Psychology, 102*(4), 959–974.

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

Young, A. W., Ellis, A. W., Flude, B. M., McWeeny, K. H., & Hay, D. C. (1986). Face–name interference. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 466–475.

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

Young, A. W., & Burton, A. M. (2017). Recognizing faces. *Current Directions in Psychological Science, in press*.