*In press: JEP:HPP, October 2015*

**A familiarity disadvantage for remembering specific images of faces**

Regine G.M. Armann, Rob Jenkins & A. Mike Burton

Department of Psychology, University of York, UK

**Acknowledgements**

The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n.323262, and from the Economic and Social Research Council, UK [ES/J022950/1].

**Correspondence to:**

A. Mike Burton

Department of Psychology

The University of York

Heslington, York

YO10 5DD, UK

[mike.burton@york.ac.uk](mailto:mike.burton@york.ac.uk)

Running head: Memory for specific images of faces

**Abstract**

Familiar faces are remembered better than unfamiliar faces. Furthermore, it is much easier to match images of familiar than unfamiliar faces. These findings could be accounted for by quantitative differences in the ease with which faces are encoded. However, it has been argued that there are also some qualitative differences in familiar and unfamiliar face processing. Unfamiliar faces are held to rely on superficial, pictorial representations, whereas familiar faces invoke more abstract representations. Here we present two studies which show, for one task, an *advantage* for unfamiliar faces. In recognition memory, viewers are better able to remember the *specific picture* they saw earlier, if it depicts an unfamiliar face. For familiar faces, viewers are more accurate in remembering that they have seen ‘this person’ than ‘this picture’. This rare advantage for unfamiliar faces supports the notion that familiarity brings about some representational changes, and further emphasises the idea that theoretical accounts of face processing should incorporate familiarity.

**Introduction**

A long history of experimental research demonstrates that people are better at remembering familiar faces than unfamiliar faces. For example, in recognition memory tests, participants are more accurate and faster to recognise familiar faces (e.g., Yarmey, 1971; Ellis, Shepherd, & Davies, 1979; Klatzky & Forrest, 1984). Furthermore, changes in expression, lighting, and viewpoint generally do not affect people’s ability to remember familiar faces, but are very damaging to unfamiliar face memory (e.g., Hill & Bruce, 1996; O’Toole, Edelman, & Bülthoff, 1998; Patterson & Baddeley, 1977). The fallibility of memory for unfamiliar faces is a major contributor to the difficulty experienced by eye-witnesses when attempting to describe or pick-out someone involved in an earlier event (see Linsday et al, 2011, for a review).

This advantage for familiar faces is also observed in matching tasks, where viewers have to decide whether simultaneously-presented images show the same or different people. This is a surprisingly difficult task for unfamiliar viewers, even when images are shown in good lighting and similar pose (Bruce et al, 1999; Burton, White & McNeill, 2010; Megreya & Burton, 2006). However, the task is very easy for familiar viewers, who are highly accurate when matching different images of the same person, even when images are poor quality (Bruce, Henderson, Newman, & Burton, 2001; Burton, Wilson, Cowan & Bruce, 1999; Megreya & Burton, 2007). Furthermore, these effects generalise beyond photo-to-photo matching. When asked to match a live person to an image (as in photo-ID checking), people are highly error-prone for unfamiliar faces (Davis & Valentine, 2009; Kemp, Towell, & Pike, 1997; Megreya & Burton, 2008) and this difficulty extends to people such as passport officers, whose daily job involves face matching (White, Kemp, Jenkins, Matheson & Burton, 2014). The ability to match two images of a face is, in fact, closely tied to the viewer’s familiarity. In a series of experiments, Clutterbuck & Johnston (2002, 2004, 2005) show a graded effect, such that viewers become more accurate with increasing familiarity.

These performance differences between familiar and unfamiliar faces extend to a wide range of tasks in memory and perception, for a review see Johnston & Edmonds (2009). There is also evidence for some differential neural processing of familiar and unfamiliar faces (see Natu & O’Toole, 2011 for a review). For example, in fMR-adaptation, a neuronal population is first adapted by repeated presentation of the same stimulus, and then presented with a variation of the stimulus to establish whether neurons are sensitive to the varied property, and thus recover from adaption. Such adaptation occurs across viewpoints for familiar but not unfamiliar faces (Ewbank & Andrews, 2008). Some fMRI studies also report that face-selective regions show higher responses to familiar than unfamiliar faces, and that differences between responses to familiar and unfamiliar faces in rFFA correlate with individual differences in matching performance (Weibert & Andrews, 2015). In ERP studies, the face-sensitive N250r component is greater for familiar than unfamiliar faces (Schweinberger, Pickering, Jentzsch, Burton & Kaufmann, 2002; Wiese & Schweinberger, 2008). Further physiological differences between the two classes of stimulus have been found in skin conductance responses where familiar faces evoke increased electrodermal activity compared to unfamiliar faces (Ellis, Quayle & Young, 1999; Ellis, Young & Koenken, 1993).

Many laboratory studies use face images which are highly controlled, both for gross properties such as pose and hairstyle and for lower-level image properties such as lighting. However, there has recently been a growing interest in ‘ambient images’, i.e. face images showing the large range of variability over which they are recognised day to day – such as those which might be returned by an internet search on a particular celebrity’s name (Jenkins, White, van Montfort & Burton, 2011; Sutherland et al, 2013; Todorov & Porter, 2014). Over this real-world range, variability between different pictures of the same person can sometimes be as large as, or even larger than, the variability between different people (Adini, Moses & Ullman 1997). Such natural variation makes it very difficult to ‘tell faces together’ for unfamiliar viewers (i.e. to see two images as the same person), whereas familiar viewers remain highly accurate at recognising and matching faces (Jenkins et al, 2011).

These differences have led researchers to ask whether familiar and unfamiliar faces might be, to a certain extent, represented *qualitatively* differently. We have argued that the representations of unfamiliar faces are more closely tied to surface, pictorial level information, whereas recognition of familiar faces relies to a greater extent on abstractive representations (Burton & Jenkins, 2011; Hancock, Bruce & Burton, 2000; Megreya & Burton, 2006). A range of experiments, some mentioned above, have shown that unfamiliar face recognition tends to be tied to specific images, with recognition performance falling as a function of changes between study and test. Even extended learning of unfamiliar faces through repeated exposure to different views of the same face has been shown to remain highly image-specific (e.g., Liu, Bhuiyan, Ward & Siu, 2009; Longmore, Liu & Young, 2008), and not to generalise well to novel views or other variations of the test images. In contrast to this reliance on pictorial information, perception of a familiar face appears to be associated with mandatory recognition of the *person* depicted – for example, information known about that person appears to be automatically accessed, giving rise to priming at later tests (Ellis, Young & Flude, 1990; McNeill & Burton, 2002).

If *qualitative* differences exist in the processing of familiar and unfamiliar faces, then this is important theoretically, as many contemporary models of face recognition do not incorporate this difference. The early influential model by Bruce and Young (1986) incorporated both *pictorial*  and *structural* codes, describing superficial picture-level and abstractive information respectively. In this model pictorial codes apply to all images of faces, familiar or unfamiliar. However, structural codes exist for familiar faces, and must be derived for unfamiliar faces over the course of learning. Later models either exclude analysis of unfamiliar faces (Burton, Bruce, & Hancock, 1999), or do not acknowledge a distinction between unfamiliar and familiar faces at all (e.g., Haxby, Hoffman, & Gobbini, 2000). Contemporary empirical work also tends to ignore the familiar/unfamiliar distinction. Indeed, recognition of individual face images is often conflated with recognition of a person, despite the early research highlighting these differences (see above). For some recent examples see Gold, Barker, Barr, et al., 2014, on feature- versus whole-face recognition; Sessa & Dalmaso, 2015, and Towler, Kelly, & Eimer, 2015 on visual working memory; Croydon, Pimperton, Ewing, Duchaine, & Pellicano, 2014, on a new test for face recognition in children; and Weigelt, Koldewyn, Dilks, Balas, McKone, & Kanwisher, 2014, on the development of perceptual discrimination, memory and domain-specificity for faces versus other classes of stimuli).

Is it necessary to acknowledge familiar/unfamiliar differences in theories of face processing? We have argued that treating ‘faces’ as an undifferentiated stimulus class may have contributed to the fact that progress in has been greatest in understanding the signals common across all faces (e.g. perception of sex, emotion or social attention), while the processes underlying the recognition of identity remain poorly understood (Burton, 2013). In particular, our understanding of face learning remains weak. Historically, learning studies have tended to emphasise improved recognition following greater exposure, either in duration or number of encounters (e.g. Ellis, 1981; Read, Vokey & Hammersly, 1990; Shapriro & Penrod, 1986). An explanation of familiarity effects based on quantitative, monotonic improvements leading to a generic familiarity advantage appears to provide an intuitive and straightforward explanation for many phenomena. However, an alternative explanation is offered by Bruce (1994): perhaps familiarisation arises by generalizing across superficial differences between encounters with a particular face, essentially stripping away those aspects of any encounter which are not relevant to identity. Since the final objective of an abstracted representation of Harrison Ford (say) is to recognize him across new encounters, superficial aspects of previous encounters may be unnecessary, or even harmful, to successful identification. In contrast, representations of unfamiliar faces generalize poorly because it is difficult to separate the key information from irrelevant detail over a small number of encounters.

Bruce’s (1994) proposal, which she names ‘stability from variation’, has received little empirical attention. Nevertheless, it raises a very interesting question. If superficial details of encounters with familiar faces are eliminated from one’s representations of these people, do these details remain available for retrieval when required? Is information which is present for representations of unfamiliar faces actually *lost* for familiars? If this were true, it would support the idea of a *qualitative* difference between familiar and unfamiliar faces, because the nature of the representation would change over learning. Since familiar faces have an advantage for almost all tasks, performance differences can be, and usually are, explained by quantitative differences in ease of processing. However, if tasks exist in which unfamiliar faces show an *advantage*, then it is harder to support a purely quantitative account of familiarity differences.

In this paper we test the hypothesis that unfamiliar face processing is image-bound, by contrast to abstractive familiar face processing. To do this, we make a novel prediction: memory for *specific images* of an unfamiliar face (that the observer has never seen before) will be better than memory for *specific images* of a familiar face, (i.e. a well-known celebrity). Following Bruce and Young (1986), this hypothesis is based on the idea that pictorial information dominates one’s representation for an unfamiliar face, making it (comparably) easy to recognise when a specific image has been seen before. However, for a familiar person, *personal information* will come into play, automatically, and as we hypothesize, at the expense of pictorial information. So, for example, participants might find themselves thinking ‘Yes, I know I saw President Obama before, but I’m not sure if I saw that particular photo of him’.

In previous recognition memory experiments for faces, the recognition phase often confounds recognition of photo and person. In very many of these experiments, the same photo is used at study and test. In experiments where this is not the case (such as those cited at the beginning of this section), the task at test is to identify whether the same person has been seen in the earlier phase. We know of no experiments in which participants are asked to recognise the same *image* as seen in the first phase, responding ‘unseen’ to images that they have not seen before, regardless of whether they have seen the particular person at study. The experiments below introduce this manipulation. To recap, when asked to recognise a *person*, we predict the standard advantage for familiar faces. However, when asked to recognise an *image*, we predict an advantage for unfamiliar faces.

**Experiment 1**

In this experiment we presented observers with a memory task in which they were shown a set of faces during an initial phase, T1, and instructed to remember the people shown. In a subsequent recognition phase, T2, participants were shown a further set of faces. Half the participants were asked, for each image, *“did you see this person before?”* and a second group was asked *“did you see this exact picture before?”*

**Method**

*Stimuli*

Stimuli comprised images of 99 familiar and 99 unfamiliar faces. Familiar faces were UK celebrities, known to our British participants, whereas unfamiliar faces were Australian celebrities, unknown to the British participants. The images were taken from a database developed for bilateral UK/Australian research and were all sourced from an internet search. Images showed roughly frontal pose, and none had occlusions (e.g. from hands, hats etc). They were cropped to show a rectangle around the head, but otherwise were ‘ambient images’, of the kind typically recognisable by familiar viewers. Copyright restrictions prevent us from reproducing most images here, but readers can easily get an impression of our stimulus selection by using celebrities’ names as Google Image search terms. A pilot ratings exercise (31 British observers) prior to the experiment confirmed the familiar/unfamiliar status of the people shown.

*Participants*

Forty-four participants (19 female, mean age 23.6, SD=8.1) were recruited from the university community. All reported having grown up in the UK, as necessary for the familiarity manipulation. Participants received course credit or a small payment.

*Design and Procedure*

In an encoding phase T1, participants were presented with 132 pictures comprising 66 familiar UK celebrities and 66 unfamiliar Australian celebrities on a computer screen. Faces were shown one at a time, with familiar and unfamiliar faces being randomly inter-mixed. Participants were instructed to *“look at every person”* and *“remember all these people”*. Study of the images was self-paced and participants skipped to the next image by pressing the space bar.

A subsequent test phase, T2, followed immediately. Participants saw 198 images: 99 familiar and 99 unfamiliar faces, inter-mixed in random sequence. For each set (familiar/unfamiliar), 33 of the images were exactly the same as in T1, 33 images contained the same people but were different pictures, and 33 showed celebrities who had not been shown at T1. Between participants, images were counter-balanced across conditions.

Participants were randomly assigned into one of two groups, each with a specific task to perform at test. One group were asked, for each test item, *“did you see this person before?”*, whereas the other group was asked “*did you see this exact picture before?”* Participants answered by pressing keys on the computer keyboard, with yes- and no-keys counterbalanced across participants.

**Results**

**Figure 1**. Mean accuracy across task groups (person/picture task) and test items (unseen/seen same/seen different) for familiar and unfamiliar faces. Error bars are SE.

Accuracy data are shown in Figure 1. To anticipate the analysis, this pattern shows the following. First, there are no differences between conditions in correctly rejecting ‘unseen’ items (leftmost panel of fig 1). Second, for test items which repeat the same pictures as seen in phase 1, we find the traditional advantage for familiar over unfamiliar faces (middle panel of fig 1). Third, for test items comprising novel pictures of people seen in phase 1 (rightmost panel of fig 1) there is an advantage for familiar faces when participants are asked to judge whether they have previously seen the same person. However, this reverses, showing an advantage for *unfamiliar* faces when participants are asked whether they have previously seen the same image.

Minor violations of the parametric normality assumption were well within the range over which ANOVA is robust (Schmider, Ziegler, Danay, Beyer & Bühner, 2010), and so this analysis strategy was adopted. Mean accuracy data were entered into a x 3 (*Trial type*: unseen/seen-same/seen-different) x 2 (*Task group*: person/picture) x 2 (*Familiarity*: familiar/unfamiliar faces) mixed-design ANOVA, with *Task group* as a between-subject factor. There was no main effect of group (F(1, 42) < 1) but a main effect of *trial type* (F(2, 84) = 37.92, p < .001, ηρ2 = .47) and a main effect of *Familiarity* (F(1, 42) = 46.76, p < .001, ηρ2 = .53), with memory performance being higher overall for familiar than for unfamiliar faces. There was also a significant interaction between *Trial type* and *Task group* (F(2, 84) = 6.00, p < .005, ηρ2 = .12), between *Trial type* and *Familiarity* (F(2, 84) = 15.31, p < .001, ηρ2 = .27) and between *Task group* and *Familiarity* (F(1, 42) = 36.04, p < .001, ηρ2 = .46). Finally, we observed a 3-way interaction *Trial type x Task x Familiarity* (F(2, 84) = 27.09, p < .001, ηρ2 = .39).

We followed up the significant three way interaction with separate two way ANOVAs for each of the three trial types. For *Unseen* trials (fig 1, leftmost panel), there were no significant main effects or interaction. (Task: F < 1; Familiarity: F(1,42) = 2.15, ns, ηρ2 = 05; Task x Familiarity: F(1,42) = 1.54, ns, ηρ2 = 04). For *Seen Same Picture* trials (fig 1, middle panel) there was a main effect of Familiarity (F(1,42) = 38.53, p<0.001, ηρ2 = .48, but no main effect of Task, and no interaction (F(1,42) = 1.35, ns, ηρ2 = .03; and F<1, respectively). For *Seen Different Picture* trials (fig 1, rightmost panel) there was a main effect of Task (F(1,42) = 7.06, p<0.01, ηρ2 = .14), a main effect of Familiarity (F(1,42) = 22.8, p<0.001, ηρ2 = .35), and a Task x Familiarity Interaction (F(1,42) = 74.87, p<0.001, ηρ2 = .64).

Of key importance here, there were significant effects of familiarity in the *Seen Different Picture* trials, but these worked in opposite directions. Simple main effects analysis showed that the *person* task produced a significant advantage for familiar faces (F(1,42) = 90.14, p<0.001, ηρ2 = .68), whereas the *picture* task produced a significant advantage for the unfamiliar faces (F(1,42) = 7.52, p<0.01, ηρ2 = .15). Similarly, there were simple main effects of task, but these worked in opposite directions. For familiar faces, the person task (‘did you see this person?’) resulted in better performance (F(1,42) = 7.00, p<0.01, ηρ2 = .08) whereas for unfamiliar faces, the picture task (‘did you see this picture?’) produced better performance (F(1,42) = 49.68, p<0.001, ηρ2 = .37).

**Discussion**

These results seem to show a striking effect. First, and in common with all the previous literature, we find that people are better able to recognise that they have seen a familiar person than an unfamiliar person. However, if the memory task is image-specific, i.e. if people are asked to remember whether they have seen a particular image or not, then the effect reverses: there is a disadvantage for familiar faces. This pattern of results is consistent with the suggestion that the coding of familiar faces is less reliant on image-specific properties than the coding of unfamiliar faces. In short, we appear to rely on more abstract representations for familiar than unfamiliar faces, and this results in rather poor coding of pictorial information in images of the people we know.

An interesting further question is the locus of this effect. Does the loss of sensitivity to pictorial information for familiar faces occur at retrieval, T2 – i.e. does recalling having seen a familiar person somehow override other information? Alternatively, does encoding at T1 somehow fail to incorporate pictorial information? We know that familiarity signals are very fast – i.e. a decision can be taken reliably in under a second (Bruce, 1986). As our viewers have unlimited time to view each image at T1, it is plausible that they engage in different types of processing when viewing familiar people. For example, we know that access to semantic knowledge about familiar people is automatic (e.g. Burton, Kelly & Bruce, 1998) and so it may be that viewers tend to activate conceptual relations, at the expense of lower level aspects of the stimuli.

We should also note that Experiment 1 encourages viewers to code stimuli in an abstract manner. We specifically instructed participants at T1 to look at each *person* and to remember the *people*. It is therefore possible that we have encouraged an abstractive mode of processing through these instructions – and that such a mode of processing inevitably favours familiar people, and (where possible) plays down superficial stimulus properties. Previous experiments have demonstrated that instructions can affect some aspects of face processing, for example a viewer’s eye movements (e.g. Armann & Buelthoff, 2009; Schwarzer, Huber, & Dümmler, 2005). Therefore, in the next experiment, we change instructions at encoding, and this time emphasise that viewers should try to remember the specific images they are shown.

**Experiment 2**

This experiment was the same as Experiment 1, with the only difference that this time observers were instructed to *“remember the exact pictures”* during encoding.

**Method**

*Participants*

Forty-four new participants (31 female, mean age 21, SD=3.2) were recruited from the same university community as in Experiment 1. As before, all reported having grown up in the UK and they received course credit or a small payment for their participation.

*Stimuli, Design and Procedure*

The same stimuli were used as in Experiment 1. The design and procedure were identical to Experiment 1, except for the instructions at T1. In this experiment, viewers saw 66 familiar and 66 unfamiliar faces at T1, but were instructed to *“remember the exact pictures”*. Once again, the study phase was self-paced. At T2 participants saw 198 images: half familiar and half unfamiliar. For each of these sets of stimuli, one third of the people had not been seen at T1, one third had been seen in a different image, and one third had been seen in the same image.

*Results & Analysis*

**Figure 2**. Experiment 2: Mean accuracy across task groups (person/picture task) and test items (unseen/seen same/seen different) for familiar and unfamiliar faces. Error bars are SE.

Accuracy data are shown in Figure 2. To summarise the analysis, this pattern shows the following. First, when correctly rejecting ‘unseen’ items (leftmost panel of fig 2), viewers are more accurate when asked if they have seen this image, than seen this person. This is a change from Experiment 1 (see fig 1) and presumably reflects the instructions to remember images at T1. For test items which repeat the same pictures as seen in phase 1, we find the normal advantage for familiar over unfamiliar faces (middle panel of fig 1), just as in Experiment 1. Third, for test items comprising novel pictures of people seen at T1 (rightmost panel of fig 2) there is an advantage for familiar faces when participants are asked to judge whether they have previously seen the same person. However, just as in Experiment 1 this reverses, showing an advantage for unfamiliar faces when participants are asked whether they have previously seen the same image.

Mean accuracy data were entered into a x 3 (*Trial type*: unseen/seen-same/seen-different) x 2 (*Task group*: person/picture) x 2 (*Familiarity*: familiar/unfamiliar faces) mixed-design ANOVA, with *Task group* as a between-subject factor. There were main effects of *Trial type* (F(2, 84) = 67.89, p < .001, ηρ2 = .62), *Task group* (F(1, 42) = 21.49, p < .001, ηρ2 = .34) and *Familiarity* (F(1, 42) = 47.46, p < .001, ηρ2 = .53). Each of the two way interactions were also significant: *Trial type* x *Task group* (F(2, 84) = 5.07, p < .01, ηρ2 = .11); *Trial type* x *Familiarity* (F(2, 84) = 12.05, p < .001, ηρ2 = .22); *Task group* x *Familiarity* (F(1, 42) = 62.68, p < .001, ηρ2 = .60), and all these effects were qualified by a significant three way interaction. *Trial type x Task x Familiarity* (F(2, 84) = 51.51, p < .001, ηρ2 = .55).

We followed up the significant three way interaction with separate two way ANOVAs for each of the three trial types. For *Unseen* trials (fig 2, leftmost panel), there was a main effect of Task (F(1,42) = 15.47, p < 0.001, ηρ2 = .27) but no main effect of familiarity and no interaction (both Fs < 1). For *Seen Same Picture* trials (fig 2, middle panel) there was a main effect of Familiarity (F(1,42) = 37.56, p<0.001, ηρ2 = .47, but no main effect of Task, and no interaction (F(1,42) = 1.62, ns, ηρ2 = .04; and F<1, respectively). For *Seen Different Picture* trials (fig 2, rightmost panel) there was a main effect of Task (F(1,42) = 20.65, p<0.001, ηρ2 = .33), a main effect of Familiarity (F(1,42) = 17.29, p<0.001, ηρ2 = .29), and a Task x Familiarity Interaction (F(1,42) = 102.62, p<0.001, ηρ2 = .71).

As with Experiment 1, there were significant effects of familiarity in the *Seen Different Picture* trials, but these worked in opposite directions. Simple main effects analysis showed that the *person* task produced a significant advantage for familiar faces (F(1,42) = 102.07, p<0.001, ηρ2 = .71), whereas the *picture* task produced a significant advantage for the unfamiliar faces (F(1,42) = 17.84, p<0.01, ηρ2 = .30). Similarly, there were simple main effects of task, but these worked in opposite directions. For familiar faces, the person task (‘did you see this person?’) resulted in better performance (F(1,42) = 6.12, p<0.01, ηρ2 = .07) whereas for unfamiliar faces, the picture task (‘did you see this picture?’) produced better performance (F(1,42) = 94.65, p<0.001, ηρ2 = .53).

**Discussion**

These results are very similar to those found in Experiment 1. In the second experiment we directed all participants at T1 to remember specific images, and this produced a slightly different pattern to Experiment 1 for unseen test items. In Experiment 2, participants were generally more accurate to reject novel test images in the picture task than in the person task, presumably reflecting the fact that they are generally more attentive to pictorial information at T1, as instructed. For seen test items, the pattern of results is identical between Experiments 1 and 2. With same picture test items, the standard familiarity advantage is observed. However, with different picture test items, there is an advantage for unfamiliar faces in the picture task. It seems that the key difference between familiar and unfamiliar face recognition memory is unaffected by instruction: unfamiliar faces are encoded more pictorially and familiar faces more abstractly, regardless of instruction. This is rather a striking result, suggesting that these differences are outside viewers’ control.

We have emphasized a representational change account of these findings. However, is it possible that the results could be explained by a bias to say ‘yes’ to any familiar face? One possibility is that participants will have seen some of the same photos of familiar faces that we have used – after all, these were sourced from internet searches. If this were true, it might cause difficulty for viewers to make the correct response in some conditions (e.g. “*yes* – I have seen that photo before – but not in this experiment, so I should respond *no*”). This possibility is consistent with the lower performance for familiar faces in the *Seen Different Picture* condition. However, if this account were true, it seems likely that participants would also show familiarity effects in the *Unseen Person* condition – whereas in fact, they have no difficulty in rejecting highly familiar people as unseen within the experiment.

Along the same lines, one might ask whether the opposite bias, i.e., to say ‘no’ to unfamiliar faces might explain the results. Such a bias is consistent with performance in some conditions, i.e. high accuracy in trials that require observers to say ‘no’ to unfamiliar faces (*Unseen person* condition for both tasks and the *Seen Different Picture* condition for the Picture Task), as well as low accuracy in the *Seen Different Picture* condition for the Person Task (requiring a ‘yes’ response). However, in the *Seen Same Picture* condition, where the correct answer is ‘yes’ (*“I have seen this person/picture”*), even though accuracy for unfamiliar faces is lower than for familiar faces, performance around 70% across all experiments and conditions does not suggest that observers have a bias to reject unfamiliar faces in general. So, general biases for either familiar or unfamiliar faces do not seem sufficient to explain the full pattern of our results.

As another way of looking more closely at potential biases in our accuracy data that might explain the differences between familiar and unfamiliar faces, we calculated sensitivity (d-prime) and criterion c, as described below. It is important to note that the design of our experiment was not planned with a SDT approach in mind and does therefore not allow for calculation of d-prime and criterion *within* each condition (see below for details about how those measures were calculated). As a consequence, we cannot detect and compare changes in sensitivity or bias between conditions within each task of each experiment. What d-prime and criterion can tell us is (i) whether the same kind of face-stimulus is processed differently depending on observer familiarity with the person in the picture, and (ii) whether picture- and person-specific manipulations of experimental instructions (‘remember these *people*’ versus ‘remember these *pictures*’) and manipulations of the memory task (‘have you seen this *person?*’ versus ‘have you seen this *picture?*’) lead to differences in sensitivity and bias across experiments and tasks for familiar and unfamiliar faces.

*Sensitivity and bias*

After an identical encoding phase, the memory test phase in both Experiment 1 and 2 consists of three conditions where observers are asked to respond either ‘yes’ *(“I have seen this”)* or ‘no’ *(“I have not seen this”)* to an image they are shown on a screen. Depending however on the memory task *(“have you seen this picture?”* versus *“have you seen this person?”),* observers in the two different groups are asked to respond differently to the same stimuli in some of these conditions. To illustrate this, in the *Unseen Person* condition, observers are required to say ‘no’ to be correct, independent of whether they are in the ‘person’ or in the ‘picture’ group. Similarly, in the *Seen Same Picture* condition, all observers are required to say ‘yes’ to be correct, independent of whether they are judging the picture or the person. In the *Seen Different Picture* condition, however, observers in the two groups have to say ‘yes’ to the person but ‘no’ to the picture, respectively. To calculate d-prime and criterion c, therefore, in each experimental group, we averaged all available ‘hits’ (i.e., all correct ‘yes’ responses) and all available ‘false alarms’ (i.e., all incorrect ‘yes’ responses) for each observer across all conditions, irrespective of the actual number of observations of each measure in each group. Figure 3 (A) shows d-prime for familiar and unfamiliar faces in both groups, while Figure 3 (B) shows criterion c in the same format.

**A**

**B**

**Figure 3**. (A) D-prime and (B) criterion c for both experiments (Person task, Picture task) and familiar and unfamiliar faces in both groups (Person memory, Picture memory). Note that PERSON task refers to Experiment 1, while PICTURE task refers to Experiment 2. For details of how both measures were calculated, see text.

Data for each measure were entered into a 2 (*Task: person/picture*) x 2 (*Memory*: *person/picture*) x 2 (*Familiarity*: *familiar/unfamiliar* faces) mixed-design ANOVA, with *Task group* as a between-subject factor. (Note that *person task* refers to data from Experiment 1, while *picture task* refers to data from Experiment 2).

For d-prime data, there was no main effect of *Task* group (F(1, 84) < 1) and no main effect of *Memory* (F(1, 84) < 2), but a main effect of *Familiarity* (F(1, 84) = 78.04, p < .001, ηρ2 = .48) with d-prime being higher overall for familiar than unfamiliar faces. There also was an interaction between *Task* and *Memory* (F(1, 84) = 4.64, p < .01, ηρ2 = .05). In general, what we find is that observers show higher sensitivity when doing either task in either condition for familiar compared to unfamiliar faces. When we compare familiar and unfamiliar faces in each group directly, we see that the difference in sensitivity between familiar and unfamiliar faces is bigger in the *Person memory* compared to the *Picture memory* group (see Figure 3A) in both Tasks. Simple interactions show that while the difference is significant for the *person Memory* group in both *person* (F(1, 84) = 50.73, p < .001, ηρ2 = .38) and *picture* (F(1, 84) = 35.42, p < .001, ηρ2 = .30) Task , and for the *picture Memory* group in the *person* Task (F(1, 84) = 12.69, p < .001, ηρ2 = .13), it is not significant for the *picture Memory* group in the *picture* Task (F(1, 84) = 1.07, p > .05).

This pattern (smaller / no difference in sensitivity in the *picture* Task) directly relates to our performance data (compare to *Figures 1* and *2*). Computing sensitivity has required us to merge ‘yes’ responses across same and different picture conditions, and as a result, there is no longer an advantage evident for unfamiliar faces in the picture task. However, note that across the other conditions, there is a clear advantage for *familiar* faces (consistent with the behavioural data). Therefore, this pattern of results is consistent with the original analysis, in that viewers’ familiarity with the faces affects the pattern of responses differently in the picture and person tasks.

For the criterion data, there was no main effect of *Task* group (F(1, 84) < 1) but a main effect of *Memory* (F(1, 84) =16.67, p <.001, ηρ2 = .17) with higher criterion for *person Memory* and a main effect of *Familiarity* (F(1, 84) = 102.20, p < .001, ηρ2 = .55) with higher criterion for unfamiliar compared to familiar faces. What we see here is the following: There is a bias for observers to say ‘no’ (higher c) to unfamiliar faces, especially in the person task, but only a small bias to say ‘yes’ to familiar people for observers who are asked *“have you seen this picture?”*. When observers are asked, after the same encoding phase and for the same familiar faces, whether they have seen the *person* in the image, their criterion is near 0, indicating no bias at all. This speaks against a general confusion about familiar people and rather points to clear familiarity-driven differences in processing *pictures* versus *people* in the same images.

To summarize, we find differences in sensitivity and bias for familiar and unfamiliar faces and for different encoding and memory conditions. These differences are consistent with our proposal that familiar and unfamiliar faces are processed, to some extent, in different ways. Across the experiments, the same visual stimulus, a static image of a face, is processed in different ways as a function of observer familiarity. This analysis appears to eliminate an artefactual account of our findings, while being consistent with (but of course not proving) our representational account.

*Encoding Times*

One further source of potential artefacts in the explanation of these results lies in encoding times. The learning phase (T1) in both experiments was self-paced and observers were very clearly instructed to take as much time as they needed to remember the person or the picture. One explanation for lower performance on familiar faces in the picture condition could be that observers generally assume they are well able to remember familiar people and thus spend less time looking at these images. That strategy could lead to a drop in performance for the *picture* task because a bias to accept familiar faces as ‘seen’ would be detrimental. What we find however is that there is no difference between encoding times for familiar faces in the *person* (M=3784sec, SEM=411sec) and the *picture* (M=3657sec, SEM=334) condition and for unfamiliar faces in the *picture* condition (M=3553sec, SEM=334sec). Observers do indeed take more time to encode unfamiliar faces in the *“remember the people”* condition (M=4592sec, SEM=469sec), presumably in an attempt to gather more detailed information in view of their (conscious) lack of familiarity with the person they are trying to remember. We do not find shorter encoding times for familiar faces in general, and these encoding times do not explain why observers are worse when they take more time to encode (unfamiliar faces, *person* condition), or why they are worse for familiar faces (*picture* condition) after spending the same amount of time encoding than on unfamiliar faces.

**General Discussion**

Here we have shown, for the first time, a performance advantage of unfamiliar over familiar faces. Although we have demonstrated the standard familiarity advantage for a person memory test, unfamiliar faces are better recognised when the task is to remember a specific image. This finding is consistent with the notion that representations of unfamiliar faces are inherently more pictorial than those of familiar faces. The abstractive representations which allow familiar faces to be recognised over a very large range of visual conditions come at a small cost – pictorial detail appears to be lost from these representations, resulting in the pattern of results described here. Our results are therefore consistent with the idea that processing of familiar and unfamiliar faces involves some qualitative differences, and that increasing familiarity cannot simply be equated with increasing information leading to better general performance, as predicted by Bruce (1994).

Experiment 2 shows an interesting effect, in that the disadvantage for remembering pictorial details for familiar faces persists, even when participants are instructed explicitly to attend to this detail. Moreover, neither the disadvantage for pictorial information in familiar faces nor the disadvantage for person-related information in unfamiliar faces can be explained by differences in encoding times across both experiments. This suggests that the distinction we have observed is not a consequence of differences in attention – which might arise, for example, if viewers simply found familiar faces more interesting or unfamiliar faces more difficult to remember. Instead, the loss of pictorial information appears to be inherent in face processing.

We are aware of the fact that “familiarity” as it is used throughout this manuscript is an oversimplification – in the real world, faces have varying levels of familiarity, rather than being simply familiar and unfamiliar. Furthermore, since all familiar faces have, at some point, been unfamiliar, it is important to try to understand how representations change as one learns a new face. In fact, this problem has been rather resistant to investigation in the literature, and there is still rather little known about it. Early work showed that images of faces were better learned as exposure was increased (e.g. Shapiro & Penrod, 1986), but little is known about development of generalizable representations. More recently, it has been suggested that a key component of learning new faces is the level of variability to which one is exposed. As mentioned in the introduction, repeated exposure to the same face under different viewpoints does not seem to lead to a high degree of generalizability, as performance is highly affected when novel images are used at test (e.g., Longmore, Liu & Young, 2008; Liu, Bhuiyan, Ward & Siu, 2009). Yet showing participants widely-varying images of the type which occur naturally (rather than highly controlled, parametrically varied images typically used in the lab) appears to result in robust representations which can be used to recognise a good range of novel images (Andrews, Jenkins, Cursiter, & Burton, 2015; Murphy, Ipser, Gaigg & Cook, 2015). However, such studies are rather new, and there remains much to understand about natural face learning – despite the fact that we all have this experience many times through life.

The pattern we have observed here for face recognition also exists in many other areas of psychology. Expert levels of performance in many domains appear to be achieved by integration of conceptual level knowledge into task-relevant structures. However, this often comes at a cost, such that task-irrelevant details appear to be lost (for a review see Dror, 2011). For example, expert radiologists show an improved memory for clinically-relevant aspects of X-rays, as compared to less expert viewers. However, they are poorer than novices in remembering irrelevant aspects of X-rays (Myles-Worsley, Johnston & Simons, 1988). In a completely different domain, London taxi drivers are well known to be experts in knowing established routes in their city, but they are in fact poorer than novices at integrating novel layouts in their areas (Woollett & Macguire, 2010).

These examples highlight the fact that familiarity within a particular domain extends beyond development of sensory expertise. While we have emphasised the differences in visual processing between familiar and unfamiliar faces, there are of course many other differences between our representations of familiar and unfamiliar *people*. As we come to know somebody, we develop modality-independent representations (perhaps this person is a neighbour, a swimmer, a politician). These can be accessed through multiple routes, for example somebody’s name or voice, and form abstract representations of the type called *Person Identity Nodes (PINs)* by Bruce and Young (1986). Imaging studies illustrate that personally familiar faces activate not only the areas associated with visual face processing (occipital and fusiform face areas), but also areas associated with personal memory (including anterior and posterior cingulate). Interestingly, this ‘extended face processing system’ is rather insensitive to experimentally-learned unfamiliar faces with no associated semantics (Bobes, Castellanos, Quinones, Garcia & Valdes-Sosa, 2013).

The normally-observed advantage for familiar faces could arise through a combination of increasingly abstractive visual representations, and top-down support from amodal representations of a ‘person’. The relative contributions of these sources is an interesting problem which could, at least in principle, be studied by manipulating the relative exposure to visual and conceptual information about a new person. In the present paper we have added further support for the proposal that there are some qualitative changes in perception as the faces we encounter become familiar. Theoretical accounts of face recognition therefore need to incorporate familiarity, and acknowledge that faces are not a single stimulus class.

**References**

Adini, Y., Moses, Y., & Ullman, S. (1997). Face recognition: The problem of compensating for changes in illumination direction. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *19*(7), 721-732.

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

Armann, R., & Bülthoff, I. (2009). Gaze behavior in face comparison: the roles of sex, task, and symmetry. *Attention, Perception, & Psychophysics*, *71*(5), 1107-1126.

Bobes, M. A., Lage Castellanos, A., Quiñones, I., García, L., & Valdes-Sosa, M. (2013). Timing and Tuning for Familiarity of Cortical Responses to Faces. *PLoS ONE*, *8*(10), e76100.

Bruce, V. (1986). Influences of familiarity on the processing of faces. *Perception*, *15*(4), 387-397.

Bruce, V. (1994). Stability from variation: the case of face recognition. The M.D. Vernon Memorial Lecture. *Quarterly Journal of Experimental Psychology*, *47A*(1), 5–28.

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J. B., Burton, A. M., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied*, *5*(4), 339–360.

Bruce, V., Henderson, Z., Newman, C., & Burton, A. M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied, 7(3),* 207–218.

Bruce, V., & Young, A.W. (1986). Understanding face recognition*. British Journal of Psychology, 77(3)*, 305–327.

Burton, A. M. (2013). Why has research in face recognition progressed so slowly? The importance of variability. *Quarterly Journal of Experimental Psychology, 66(8),* 1467–1485.

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

Burton, A. M., & Jenkins, R. (2011). Unfamiliar face perception. In A. J. Calder, G. Rhodes, M. Johnson, & J. Haxby (Eds.), *Oxford handbook of face perception* (pp. 287–306). Oxford: Oxford University Press.

Burton, A. M., Kelly, S. W., & Bruce, V. (1998). The face-space typicality paradox: Understanding the face-space metaphor. *Quarterly Journal Of Experimental, 51A*(3), 475–483.

Burton, A. M., White, D., & McNeill, A. (2010). The Glasgow Face Matching Test. *Behavior Research Methods*, *42*(1), 286–291.

Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science*, *10*(3), 243–248.

Clutterbuck, R., & Johnston, R. A. (2002). Exploring levels of face familiarity by using an indirect face-matching measure. *Perception, 31(8),* 985–994.

Clutterbuck, R., & Johnston, R. A. (2004). Matching as an index of face familiarity. *Visual Cognition, 11(7)*, 857–869.

Clutterbuck, R., & Johnston, R. A. (2005). Demonstrating how unfamiliar faces become familiar using a face matching task. *European Journal of Cognitive Psychology, 17(1),* 97–116.

Croydon, A., Pimperton, H., Ewing, L., Duchaine, B. C., & Pellicano, E. (2014). The Cambridge Face Memory Test for Children (CFMT-C): A new tool for measuring face recognition skills in childhood. *Neuropsychologia*, *62*, 60-67.

Davis, J. P., & Valentine, T. (2009). CCTV on trial: Matching video images with the defendant in the dock. *Applied Cognitive Psychology*, *23*(4), 482–505.

Dror, I.E. (2011). The paradox of human expertise: Why experts can get it wrong. In N. Kapur (Ed.), *The Paradoxical Brain* (pp. 177-188). Cambridge, UK: Cambridge University Press.

Ellis, H. D. (1981). Theoretical aspects of face recognition. In Davies, G., Ellis, H. Shepherd J. (Eds.). *Perceiving and remembering* *faces.* San Diego, CA: Academic Press.

Ellis, H. D., Quayle, A. H., & Young, A. W. (1999). The emotional impact of faces (but not names): face specific changes in skin conductance responses to familiar and unfamiliar people. *Current Psychology*, *18*(1), 88-97.

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception, 8(4),* 431–439.

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. *The Quarterly Journal of Experimental Psychology Section A*, *42*(3), 495–512.

Ellis, H. D., Young, A. W., & Koenken, G. (1993). Covert face recognition without prosopagnosia. *Behavioural Neurology*, *6*(1), 27-32.

Ewbank, M. P., & Andrews, T. J. (2008). Differential sensitivity for viewpoint between familiar and unfamiliar faces in human visual cortex. *Neuroimage*, *40*(4), 1857-1870.

Gold, J. M., Barker, J. D., Barr, S., Bittner, J. L., Bratch, A., Bromfield, W. D, Good, R.A., Jones, M., Lee, D. & Srinath, A. (2014). The perception of a familiar face is no more than the sum of its parts. *Psychonomic Bulletin & Review*, *21*(6), 1465-1472.

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

Haxby, J., Hoffman, E., & Gobbini, M. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4(6),* 223–232.

Hill, H., & Bruce, V. (1996). The effects of lighting on the perception of facial surfaces. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 986.

Jenkins, R., White, D., van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. Cognition, 121(3), 313–323.

# Johnston, R. A., & Edmonds, A. J. (2009). Familiar and unfamiliar face recognition: A review. *Memory*, *17*(5), 577-596.

Kemp, R., Towell, N., & Pike, G. (1997). When seeing should not be believing: Photographs, credit cards and fraud. *Applied Cognitive Psychology, 11(3),* 211–222.

Klatzky, R. L., & Forrest, F. H. (1984). Recognizing familiar and unfamiliar faces. *Memory & Cognition, 12(1)*, 60–70.

Lindsay, R. C. L., Mansour, J. K., Bertrand, M. I., Kalmet, N., & Melsom, E. I. (2011). Face recognition in eyewitness memory. In A. J. Calder, G. Rhodes, M. Johnson, & J. Haxby (Eds.), *Oxford handbook of face perception* (pp. 307–328). Oxford: Oxford University Press.

Liu, C. H., Bhuiyan, A., Ward, J., & Sui, J. (2009). Transfer between pose and illumination training in face recognition. *Journal of Experimental Psychology: Human Perception and Performance, 35(4),* 939-947.

Longmore, C. A., Liu, C. H., & Young, A. W. (2008). Learning faces from photographs. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(1), 77.

McNeill, A., & Burton, A. M. (2002). The locus of semantic priming effects in person recognition. *Quarterly Journal Of Experimental Psychology Section A-Human Experimental Psychology*, *55*(4), 1141–1156.

Megreya, A. M., & Burton, A. M. (2006). Unfamiliar faces are not faces: Evidence from a matching task. *Memory & Cognition, 34(4),* 865–876.

Megreya, A. M., & Burton, A. M. (2007). Hits and false positives in face matching: A familiarity-based dissociation. *Perception & Psychophysics, 69(7)*, 1175–1184.

Megreya, A. M., & Burton, A. M. (2008). Matching faces to photographs: Poor performance in eyewitness memory (without the memory). *Journal of Experimental Psychology: Applied, 14(4),* 364–372.

Murphy, J., Ipser, A., Gaigg, S.B., Cook, R. (2015). Exemplar variance supports robust learning of facial identity. *Journal of Experimental Psychology: Human Perception and Performance, 41(3)*, 577-581.

Myles-Worsley, M., Johnston, W. A., & Simons, M. A. (1988). The influence of expertise on X-ray image processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 553.

Natu, V., & O’Toole, A. J. (2011). The neural processing of familiar and unfamiliar faces: A review and synopsis. *British Journal of Psychology*, *102*(4), 726–747.

O’Toole, A. J., Edelman, S., & Bülthoff, H. H. (1998). Stimulus-specific effects in face recognition over changes in viewpoint. *Vision Research, 38*, 2351–2363.

Patterson, K. E., & Baddeley, A. D. (1977). When face recognition fails. *Journal of Experimental Psychology: Human Learning and Memory, 3(4),* 406–417.

Read, J. D., Vokey, J. R & Hammersley, R. (1990). Changing photos of faces: Effects of exposure duration and photo similarity on recognition and the accuracy-confidence relationship. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16(5),* 870-882.

Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is It Really Robust?: Reinvestigating the robustness of ANOVA against violations of the normal distribution assumption. *Methodology*, *6*(4), 147–151.

Schwarzer, G., Huber, S., & Dümmler, T. (2005). Gaze behavior in analytical and holistic face processing. *Memory & Cognition*, *33*(2), 344-354.

Schweinberger, S. R., Pickering, E. C., Jentzsch, I., Burton, A. M., & Kaufmann, J. M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Cognitive Brain Research*, *14*(3), 398–409.

Sessa, P., & Dalmaso, M. (2015, in press). Race perception and gaze direction differently impair visual working memory for faces: An event-related potential study. *Social Neuroscience*.

Shaprio, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin, 100,* 139-256.

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*(1), 105–118.

Todorov, A., & Porter, J. M. (2014). Misleading First Impressions: Different for Different Facial Images of the Same Person. *Psychological Science*, *25*(7), 1404–1417.

Towler, J., Kelly, M., & Eimer, M. (2015, in press). The focus of spatial attention determines the number and precision of face representations in working memory. *Cerebral Cortex*.

Weibert, K., & Andrews, T. J. (2015). Activity in the right fusiform face area predicts the behavioural advantage for the perception of familiar faces. *Neuropsychologia*, 75, 588-596.

Wiese, H., & Schweinberger, S. R. (2008). Event-related potentials indicate different processes to mediate categorical and associative priming in person recognition. *Journal Of Experimental Psychology-Learning Memory And Cognition*, *34*(5), 1246–1263.

Weigelt, S., Koldewyn, K., Dilks, D. D., Balas, B., McKone, E., & Kanwisher, N. (2014). Domain‐specific development of face memory but not face perception. *Developmental science*, *17*(1), 47-58.

White, D., Kemp, R. I., Jenkins, R., Matheson, M., & Burton, A. M. (2014). Passport Officers’ Errors in Face Matching. *PLoS ONE*, *9*(8), e103510.

Woollett, K., & Maguire, E. A. (2010). The effect of navigational expertise on wayfinding in new environments. *Journal of environmental psychology*, *30*(4), 565-573.

# Yarmey, A. D. (1971). Recognition memory for familiar “public” faces: Effects of orientation and delay. *Psychonomic Science*, *24*(6), 286-288.