



The sibling familiarity effect: Is within-person facial variability shared across siblings?

Ailsa Strathie^{1*} , Naomi Hughes-White² and Sarah Laurence^{1,2}

¹Faculty of Arts and Social Sciences, School of Psychology and Counselling, The Open University, Milton Keynes, UK

²School of Psychology, Keele University, UK

Humans are experts at familiar face recognition, but poor at unfamiliar face recognition. Familiarity is created when a face is encountered across varied conditions, but the way in which a person's appearance varies is identity-specific, so familiarity with one identity does not benefit recognition of other individuals. However, the faces of biological siblings share structural similarities, so we explored whether the benefits of familiarity are shared across siblings. Results show that familiarity with one half of a sibling pair improves kin detection (experiment 1), and that unfamiliar face matching is more accurate when targets are the siblings of familiar versus unfamiliar individuals (experiment 2). PCA applied to facial images of celebrities and their siblings demonstrates that faces are generally better reconstructed in the principal components of a same-sex sibling than those of an unrelated individual. When we encounter the unfamiliar sibling of someone we already know, our pre-existing representation of their familiar relation may usefully inform processing of the unfamiliar face. This can benefit both kin detection and identity processing, but the benefits are constrained by the degree to which facial variability is shared.

Humans are experts at recognizing faces, but only when the faces belong to familiar individuals (see Young & Burton, 2017, 2018a). Familiar faces can be successfully identified across substantial changes in appearance (e.g., Davies & Milne, 1982; Noyes & Jenkins, 2019; Troje & Kersten, 1999) while even small alterations in pose or camera angle can thwart identification of unfamiliar faces (e.g., Bruce, 1982; Burton, White, & McNeill, 2010; Megreya & Burton, 2006, 2007). This contrast in performance may reflect underlying differences in the way in which familiar and unfamiliar faces are represented cognitively. The representation of unfamiliar faces is thought to be largely image-dependent (e.g., Bruce, 1982; Bruce & Young, 1986; Hancock, Bruce, & Burton, 2000), while familiar faces have been encountered across changes in lighting, angle, expression, age, hairstyle, etc., creating a robust representation that is largely independent of individual image properties (Andrews, Jenkins, Cursiter, & Burton, 2015; Burton, 2013; Burton, Kramer, Ritchie, & Jenkins, 2016; Dowsett, Sandford, & Burton, 2016; Jenkins, White, Van Montfort, & Burton, 2011; Ritchie & Burton, 2017).

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

*Correspondence should be addressed to Ailsa Strathie, School of Psychology and Counselling, The Open University, Milton Keynes, MK76AA, UK (email: ailsa.strathie@open.ac.uk).

As a consequence of these differences in the way familiar and unfamiliar faces are represented, while varied images of a familiar face can be effortlessly grouped by identity, multiple images of the same unfamiliar individual may appear to show several different people. This finding has been demonstrated in a number of studies (e.g., Andrews et al., 2015; Jenkins et al., 2011; Laurence & Mondloch, 2016; Zhou & Mondloch, 2016) using a card sorting task originated by Jenkins et al. (2011). The task requires participants to divide sets of 40 facial images (20×2 individuals) into their constituent identities. In the study by Jenkins et al., participants who were unfamiliar with the faces in the images divided them into a median of 7.5 piles, while participants familiar with the identities sorted the same image sets into a median of two piles. Participants made errors by separating each of the two identities into several different piles, but they rarely mixed the two identities. This suggests that the difference between two images of the same person (within-person variability) can be greater than the difference between images of two different people (between-person variability) and that people are experiencing difficulties in 'telling people together' rather than telling them apart.

Burton et al. (2016) propose that the way the face of each individual varies across encounters is idiosyncratic, and it is exposure to a face across varied conditions that fuels the transition from unfamiliar to familiar for that specific identity. This means that learning about the variation in one face will not improve recognition of novel faces. However, a newly encountered face may be similar in appearance to someone we already know, for example, a biological sibling (see Figure 1),¹ and little is known about the extent to which existing memory representations for familiar faces affect the way in which the faces of their previously unfamiliar relations are processed.

Support for the idea that face expertise is identity-specific and obtained through familiarity, rather than a more generic ability that applies to all facial stimuli, is offered by several complementary sources of evidence. First, there are large individual differences in unfamiliar face-matching performance (e.g., Burton et al., 2010; Fysh & Bindemann, 2017) that seem to persist after training. For example, passport officers who routinely verify identities as part of their job are no more accurate at unfamiliar face-matching tasks than student controls (White, Kemp, Jenkins, Matheson, & Burton, 2014). Second, research on face learning shows that repeated exposure to a specific face results in more accurate face matching for that identity (e.g., Bonner, Burton, & Bruce, 2003; Clutterbuck & Johnston, 2002, 2004) and that exposure to multiple different images of a face facilitates fast acquisition of familiarity for that individual identity (Andrews et al., 2015; Dowsett et al., 2016), but importantly, this performance benefit does not transfer to novel identities (Dowsett et al., 2016).

The findings from behavioural studies are supported by the results of principal component analysis (PCA). PCA is a method of data reduction that can be applied to a set of facial images in order to decompose them into a set of eigenfaces (e.g., Turk & Pentland, 1991). These eigenfaces can be used to represent the images from the training set in low-dimensional space. Once the PCA has been created with the training set, other facial images can be reconstructed within that space, based on the combined weightings of the eigenfaces. The error value produced from the reconstruction indicates how well a face can be represented using the eigenfaces. This approach has typically been used to identify the dimensions along which sets of different identities vary (e.g., between-person variation). However, Burton et al. (2016) applied PCA to many images of a single face, in

¹ The image shows Doug Pitt, brother of Brad Pitt



Figure 1. Image of the sibling of a well-known celebrity. See footnote

order to capture within-person variability. To do this, they collected multiple images of celebrities from online sources and used PCA to derive identity-specific components (or eigenfaces) for each celebrity. Novel images of (1) the same identity and (2) the other celebrity identities were projected into the components to create reconstructions. Reconstruction errors showed that novel images of the same identity were always better reconstructed in their own principal components than images of different identities, supporting the claim that variation is identity-specific.

Young and Burton (2018a, 2018b) acknowledge that experience results in some general improvements in unfamiliar face processing. For example, there is evidence that own-race and own-age biases in recognition are not observed in those who interact with people of other ages (e.g., Harrison & Hole, 2009) or races (see Meissner & Brigham, 2001). However, Young and Burton (2018a, 2018b) propose that it is experience with *familiar* faces and the way in which they vary that underlies these improvements in performance with unfamiliar faces. Specifically, exposure to the variability in a small number of familiar identities can influence the way we process unfamiliar faces, but this is constrained to aspects of appearance that are shared across the familiar and novel identities (e.g., faces of the same age or race) and cannot provide insight into the identity-specific variation that is most useful for individuation.

Current evidence offers clear support for the theory that variation in facial appearance is idiosyncratic, and that the benefits of familiarity cannot be transferred across *unrelated*

identities. However, facial appearance is partly determined by genetics, and as a result, the faces of biological relatives share structural similarities (e.g., Djordjevic, Zhurov, & Richmond, 2016; Tsagkrasoulis, Hysi, Spector, & Montana, 2017). For example, Djordjevic et al. (2016) studied morphological variation in twins and found that genetic factors explained more than 70% of the variation in facial size, nose, lip prominence, and distance between the eyes. Research on the recognition of evolutionary signals in faces has established that people can use these visual cues to detect kinship in high-quality photographs of strangers with reasonable accuracy (e.g., DeBruine et al., 2009; Kaminski, Dridi, Graff, & Gentaz, 2009; Maloney & Martello, 2006), and there is a close relationship between the similarity ratings attached to image pairs and judgements of kinship. There is also some evidence that kin detection is affected by familiarity. Participants shown a target face and asked to identify a relative from a choice of two (where the correct option is always one of the two) are more accurate when the target is familiar than unfamiliar (Hancock, 2020; Hancock & Bulloch, 2008). This suggests that due to their structural similarities, related faces may vary in similar ways to each other, raising the possibility that some of the benefits of familiarity may transfer across *related* identities.

The observation that related individuals share some visible facial resemblance may seem unremarkable in the light of everyday experience, but the effect of kin relationships on face identification has been largely overlooked. The cognitive processes that underpin kin detection are not yet clearly defined, and kin recognition is not incorporated into cognitive models of face processing (e.g., Bruce & Young, 1986; Haxby, Hoffman, & Gobbini, 2000; Young & Bruce, 2011). However, there are clear parallels between the task of kin detection and face matching, and an understanding of the cognition that informs kinship decisions also has implications for our understanding of face processing more broadly.

While facial variation may be idiosyncratic, given that facial appearance is partly heritable, we theorize that structural similarities in the facial appearance of related individuals will result in some shared variability. If so, familiarity with one identity (e.g., Brad Pitt) will provide an advantage for processing the face of their unfamiliar sibling (e.g., Doug Pitt). Here we present three studies designed to empirically assess this theory. In experiment 1, we use a behavioural kin detection study to investigate whether familiarity and kin detection interact. Next, we conduct a behavioural face-matching study (experiment 2) and a PCA to investigate whether similarities between siblings give rise to shared dimensions of within-person variability.

Experiment 1

In this experiment, we investigated whether familiarity improves kin detection. Participants were asked to decide if face pairs showed related or unrelated individuals. The target identities were celebrities, but only half of the target identities were likely to be familiar to people in the UK. The siblings and foil images were always unfamiliar. We predicted that kin detection would be more accurate for familiar target identities than unfamiliar target identities. This prediction was based on the evidence that (1) familiar face matching is more accurate than unfamiliar face matching, (2) exposure to a face across different conditions creates familiarity by providing an understanding of the way in which that individual face varies, and (3) facial appearance is partially genetically determined, so some aspects of facial variability will be shared by related individuals.

Method

Participants

The experiment was run online, and participants accessed it via web browser. Participants were recruited using adverts placed on the [The Open University Staff] intranet and the [HERC] Wordpress site. Consent was provided at the beginning of the experiment, and participants were able to withdraw during the experiment by closing the browser window. Participants who did not complete all portions of the study were treated as withdrawals, and their data were removed. Forty-five participants completed the experiment. Four participants who experienced technical difficulties during the experiment, and four participants who reported familiarity with four or more of the ‘unfamiliar’ identities, were excluded from analysis. This left data from a total of thirty-seven participants. Thirty of the participants reported their gender as female, six as male, and one as other. Participants were aged between 21 and 79 [$M = 44.5$, $SD = 15.5$].

Design

The experiment had a 2×2 repeated measures design. The first independent variable was target familiarity with two levels, familiar or unfamiliar. The second independent variable was trial type with two levels, related or unrelated. The dependent variable was the accuracy of the related/unrelated decisions which was measured using sensitivity (d') and criterion (C) values.

Stimuli

A total of eighty celebrity identities (40 familiar, 40 unfamiliar) were used as targets in this study, and images of the targets and their siblings were gathered from those freely available online. Following previous research (e.g., Burton et al., 2016; Jenkins et al., 2011; Ritchie & Burton, 2017), uncontrolled images were used as stimuli in order to capture the natural variation in facial appearance that occurs across different encounters. For each of the targets, two image pairs were created: a ‘related’ pair and an ‘unrelated’ pair.

To find identities for the ‘familiar’ related pairs, a list of celebrities that were (1) known to have siblings and (2) likely to be familiar to the UK public was compiled. We included celebrities of a range of different ages. Google image searches were conducted on each of the siblings of these target identities, and targets whose siblings had two good quality, publicly available images were selected for inclusion. This produced a set of 40 ‘familiar sibling’ identities. Good quality images of each of the target celebrities were also gathered. Images of the 40 celebrity identities were paired with their siblings to create 40 ‘related’ pairs for the ‘familiar’ stimuli set.

To control for the possibility that celebrities are distinctive in some way (either facially or because of image quality), the ‘unfamiliar’ identities were also celebrities, but those who were unlikely to be familiar to the UK public. A list of Spanish-speaking celebrities who (1) are known to have siblings and (2) work primarily in the Spanish language, so are unlikely to be known in the UK, was compiled. Categorization was on language, not nationality, so celebrities who work in both Spanish and English, and are therefore likely to be familiar to the UK public (e.g., Penelope Cruz), were classified as ‘familiar’. Google image searches were conducted using the names of each of the siblings of the unfamiliar target identities and those with two good quality, publicly available images were selected for inclusion. This produced a set of 40 ‘unfamiliar sibling’ identities. Good quality images

of each of the celebrities to whom the siblings were related were also gathered. The 40 celebrities were paired with their siblings to create 40 ‘related’ pairs for the ‘unfamiliar’ stimuli set.

As thirty of the sibling pairings (15 familiar, 15 unfamiliar) were opposite sex (brother/sister), the unrelated pairings were always created by pairing each target with an image of an unrelated person judged to be of similar appearance to the *sibling* rather than the target. To create the unrelated pairings, two images of each of 120 Spanish-speaking celebrities (not part of the target stimuli set) were collected to create a pool of 240 filler images. The fillers were drawn from people who work in the same fields as the target identities (e.g., actors, pop stars, presenters, politicians) and included people of a similar age range. To ensure the images in the ‘unrelated’ pairings shared superficial similarities, the eighty sibling identities (40 familiar, 40 unfamiliar) were randomly ordered, and two researchers matched each in turn to the image they judged to be the most similar filler image from the pool based on age, sex, and facial appearance. The filler images were then paired with the celebrity targets to create the ‘unrelated’ image pairings. No filler image was used more than once.

All images were cropped to show only the face and were resized to a height of 300 px. The face pairs were created by placing the two images side by side with a gap of 50 px between them. In total, there were 160 image pairs, and these were divided into two counterbalancing sets so that the participants saw each target only once (i.e., in either a related or unrelated pairing).

Procedure

The experiment was administered online using Qualtrics software (<http://www.qualtrics.com>). On each trial, participants were shown an image pair positioned in the centre of the screen and were asked to decide whether the images showed two related individuals or two unrelated individuals. Responses were entered via radio buttons on screen. To ensure they understood the task, participants completed two practice trials and received feedback about the correct responses before they commenced the main experiment.

In the main experiment, each participant saw eighty image pairs (20 per condition) presented in random order. The experiment was self-paced, and the image pair stayed on screen until a response was selected. No feedback was provided in the main experiment. Next, participants were shown a list of the names of the familiar and unfamiliar identities used in the experiment and were asked to indicate those they knew. If a participant indicated they were familiar with one of the sibling identities or with one of the Spanish ‘unfamiliar’ identities, the response for that pair was removed from the analysis. A single response was removed for twelve participants, while two participants had two responses removed. Participants who were familiar with four or more of either the ‘unfamiliar’ celebrities or sibling identities were excluded from the study. As the participants were of a wide age range, we did not expect that every item would be familiar for every individual and we did not exclude items in the ‘familiar’ group if participants did not choose them from the list of names.

Results

Table 1 summarizes mean percentage accuracy broken down by familiarity and trial type.

Signal detection theory (SDT; Green & Swets, 1966; Macmillan & Creelman, 2005) was applied to responses across related and unrelated conditions to form a measure of

Table 1. Mean percentage accuracy with standard deviations for Experiment 1

Familiarity	Trial type	Mean	SD
Familiar	Related	62	12
	Unrelated	63	15
Unfamiliar	Related	46	11
	Unrelated	67	14

sensitivity (d') and criterion (C) for each participant, in each condition. Sensitivity and criterion were compared across familiarity in two separate paired t -tests using JASP (JASP Team, 2020). JASP (Version 0.14) [Computer software].

For sensitivity, there was a significant effect of familiarity [$t(36) = 3.08, p = .004, d = 0.51$] with higher sensitivity (d') in the familiar condition ($M = 0.73, SD = 0.59$) than in the unfamiliar condition ($M = 0.37, SD = 0.42$). There was also a significant effect of familiarity on criterion (C), [$t(36) = -4.83, p < .001, d = -0.79$] with evidence of a more conservative response bias (e.g., increased tendency to respond 'unrelated') in the unfamiliar condition ($M = 0.28, SD = 0.3$) than in the familiar condition ($M = 0.03, SD = 0.34$).

To explore this further, two separate, one sample, z -tests were conducted to compare criterion against a hypothetical score of zero (i.e., no bias). Relative to zero, no response bias was observed in either the familiar condition [$z = .19, p = .85$], or the unfamiliar condition [$z = 1.7, p = .09$].

Discussion

The significant effect of sibling familiarity on sensitivity shows that people are more accurate at deciding whether two people are siblings when one of the targets is familiar versus unfamiliar. This is in line with our prediction and offers support for the theory that some of the identity-specific variation that creates a robust representation of a familiar face may be shared by close biological relatives. There was also a significant effect of sibling familiarity on criterion, with evidence of a more conservative bias in the unfamiliar condition than the familiar condition. The more balanced pattern of responses in the familiar condition is accompanied by higher sensitivity, so the shift in criterion seems to be driven by an increased ability to identify that two images show related people, rather than a general bias towards responding 'related'.

We had predicted that accuracy would be higher with familiar than unfamiliar faces, but not that there would be a difference in response bias between the two conditions. However, the more conservative responding with unfamiliar than familiar faces is paired with lower discrimination. Accuracy with the unfamiliar sibling pairs was only 46% on related trials in our study which suggests the difference is driven by performance with related pairings (accuracy for unrelated pairings is similar across familiar and unfamiliar conditions). This is consistent with research by Jenkins et al. (2011) who found that, as a result of variation, images of the same unfamiliar person can appear to show different people. In addition, research on face matching has shown familiarity is beneficial for identity matching tasks, but only on same trials (e.g., Ritchie & Burton, 2017). In the current study, mean d' was 0.37 with unfamiliar pairs and 0.73 for familiar pairs, which is lower than the d' for siblings in the study by DeBruine et al., 2009 (Mean $d' = 1.19$).

However, DeBruine et al. used high-quality, full-face images, captured by the same camera on the same day. In contrast, the images in our experiment were deliberately unconstrained in order to capture natural within-person variability. This suggests that higher levels of variation in our image set made it more difficult to discern the relevant kinship cues in the absence of an existing representation of the face.

Experiment 2

Experiment 1 demonstrated that it is easier to decide that two people are related when one of them is familiar. We theorize that shared variability underpins the advantage conferred by familiarity on this task. However, in experiment 1 the familiar individual was always present in the pairing, so we cannot discount the possibility that our findings reflect a more general advantage for perceptual matching tasks which include familiar faces, rather than evidence of shared variability. If some facial variability is shared by related individuals, then an advantage in processing the unfamiliar sibling faces would be predicted even in the absence of their familiar relations. To test this, in experiment 2 we used an unfamiliar face-matching task in which half of the targets were the siblings of familiar individuals. We predicted that accuracy would be higher for the siblings of familiar people than for those of unfamiliar individuals.

Method

Participants

The experiment was run online, and participants accessed the study via web browser. Participants were recruited from adverts placed on the [Open University Staff] intranet, the [HERC] Wordpress site, and the [Keele University Psychology] Facebook Group. Participants gave consent to participate at the beginning of the experiment and were able to withdraw at any time by closing the browser window. Participants who did not complete all portions of the study were treated as withdrawals and their data were removed. Forty-three participants completed the experiment. Two participants who reported familiarity with four or more of the sibling identities, and one who reported familiarity with four of the identities in the ‘unfamiliar’ stimuli set, were excluded from the study. This left data from a total of forty participants. Of these, 33 participants reported their gender as female, seven as male. Participants were aged between 20 and 62, [$M = 40.2$, $SD = 12.4$].

Design

The experiment had a 2×2 repeated measures design. The first independent variable was Sibling Familiarity with two levels, Familiar or Unfamiliar. The second independent variable was Trial Type with two levels, Same or Different. The dependent variable was the accuracy of the same/different face-matching decisions which was measured using sensitivity (d') and criterion (C) values.

Stimuli

The eighty sibling identities (40 siblings of familiar identities, 40 siblings of unfamiliar identities) from experiment 1 were used as targets in experiment 2. For each of the targets,

two image pairs were created: a 'same' pair and a 'different' pair. The 'same' identity pairs consisted of two different images of the same sibling identity (e.g., two different images of Doug Pitt). To create the 'different' identity pairs, the image on the right of the pairing was replaced with an image of a different person that had been matched to the appearance of the target during experiment 1.

All images were cropped to show only the face and were resized to a height of 300 px. The face pairs were created by placing the two images side by side with a gap of 50 px between them. In total, there were 160 image pairs, and these were divided into two counterbalancing sets so that each participant saw each target only once (i.e., in either a same or different pairing).

Procedure

The experiment was administered online using Qualtrics. On each trial, participants were shown an image pair positioned in the centre of the screen and were asked to decide whether the pair showed two images of the same individual or images of two different individuals. They were not told that some of the targets were siblings of familiar individuals. Responses were entered via radio buttons on screen. Participants completed two practice trials and received feedback about the correct responses before they commenced the main experiment. There was no feedback in the main experiment.

Each participant saw eighty image pairs (20 per condition) presented in random order. The experiment was self-paced, and the image pair stayed on screen until a response was selected. At the end of the experiment, participants were shown a list of the identities used in the experiment and were asked to indicate those they knew. If a participant indicated they were familiar with any of the sibling identities (e.g., Doug Pitt) or the celebrities who work in the Spanish language, the response for that pair was removed from the analysis. Nine participants had a single response removed. Data from participants who were familiar with four or more unfamiliar celebrities or sibling identities were excluded from the experiment. As in experiment 1, we did not expect every target to be familiar for every individual, so we did not exclude items in the 'familiar' group if participants did not indicate they knew them on the list of names. This meant that not every item was familiar for every participant. To check that participants were naïve to the purpose of the experiment, they were asked what they thought the study was measuring before debriefing. None of the participants identified that some of the targets were siblings of familiar individuals.

Results

Table 2 summarizes mean percentage accuracy broken down by sibling familiarity and trial type.

Signal detection theory (SDT) (Green & Swets, 1966; Macmillan & Creelman, 2005) was applied to responses across same and different conditions to form a measure of sensitivity (d') and criterion (C) for each participant in each condition. Sensitivity and criterion were compared across familiarity in two separate paired t -tests using JASP (JASP Team, 2020). JASP (Version 0.14) [Computer software].

For sensitivity, there was a significant effect of familiarity [$t(39) = 5.28, p < .001, d = 0.83$] with higher sensitivity (d') in the familiar condition ($M = 2.06, SD = 0.67$) than in the unfamiliar condition ($M = 1.47, SD = 0.67$). There was also a significant effect of familiarity on criterion (C), [$t(39) = -4.06, p < .001, d = -0.64$] with evidence of a more

Table 2. Mean percentage accuracy with standard deviations for Experiment 2

Familiarity	Trial type	Mean	SD
Familiar	Same	82	14
	Different	82	12
Unfamiliar	Same	69	18
	Different	79	13

conservative response bias (e.g., increased tendency to respond ‘different’) in the unfamiliar condition ($M = 0.17$, $SD = 0.47$) relative to the familiar condition ($M = -0.03$, $SD = 0.46$).

To explore this further, two separate, one sample z -tests were conducted to compare criterion against a hypothetical score of zero (i.e., no bias). Relative to zero, no response bias was observed in the familiar condition [$z = -0.18$, $p = .86$] or the unfamiliar condition [$z = 1.07$, $p = .28$].

Discussion

As predicted, face-matching accuracy as measured by sensitivity is higher when the targets are the siblings of familiar individuals than when they are the siblings of unfamiliar people, even though the participants were not explicitly aware of the relationship. However, there is a difference in criterion, with more conservative responding to unfamiliar face pairs than to familiar face pairs. This suggests that participants are more likely to respond ‘different’ with unfamiliar faces than with familiar faces. The difference seems to be driven by responses to the ‘same’ pairs as accuracy for ‘different’ pairs is similar across familiar and unfamiliar conditions. This pattern of results is consistent with experiment 1 and with previous research on face matching which found familiarity was beneficial only on same trials (e.g., Ritchie & Burton, 2017). It also fits with evidence that errors in unfamiliar face matching are driven by problems in ‘telling people together’, rather than telling them apart (Andrews et al., 2015; Jenkins et al., 2011).

Importantly, the results of experiment 2 demonstrate that the sibling familiarity advantage observed in experiment 1 is not restricted to tasks where the familiar sibling is present. People are more accurate at face-matching tasks when the targets are related to familiar people than when they are not. This lends further support to the theory that although learning about the way in which an individual face varies does not provide a global benefit on face-matching tasks, the representation created for that face may offer an advantage when matching the face of someone whose face is structurally similar, such as a close relative.

Principal component analysis

In addition to the behavioural studies reported in experiments 1 and 2, we used PCA to examine whether within-person variability is shared between siblings. To conduct a PCA, a specific grid is placed on anatomical landmarks in each facial image so that the images can be morphed to a standard shape average. The manipulation allows separation of eigenfaces for shape and texture (information after shape information has been excluded,

e.g., lighting). After creating this low-dimensional face space, novel images can be reconstructed using the components from the PCA. Mean square error (MSE) offers a measure of the goodness of encoding between the original image and the reconstruction. It is also possible to generate an 'average' face based on the images in the training set by calculating the mean x y coordinates for each facial landmark (shape) and the mean intensity value at each pixel (texture) across images (Jenkins & Burton, 2011).

Burton et al. (2016) used PCA to provide evidence that within-person variation is idiosyncratic. However, if people share common dimensions of variability with their siblings, then a set of dimensions that characterize the variability associated with a particular face (e.g., Brad Pitt) should code the face of their sibling (Doug Pitt) better than the faces of other unrelated identities. If, however, dimensions of variability are not shared between siblings then the variability associated with Brad Pitt should represent related and unrelated identities equally well.

We theorize that the information captured by the PCA of the celebrity will be more representative of their sibling's shape and texture components compared to those of other people. Therefore, we predict that identities reconstructed in the principal components of a sibling will generate an MSE that is smaller than the average reconstruction error of the other same-sex identities.

Images

Following the procedure described by Burton et al. (2016), images of 10 celebrities and their same-sex siblings (e.g., Brad Pitt and his brother Doug Pitt) were collected via a Google Image search in order to capture the natural variation in appearance that is encountered during everyday exposure to faces (see Burton, 2013; Burton et al., 2016). Five of the image pairs showed male sibling pairs and five showed female sibling pairs. Fifty-five images were required for each celebrity, while 25 were required for each sibling. To perform a PCA, the image set must meet specific criteria (Kramer, Jenkins, & Burton, 2017). The name of the celebrity/sibling was entered into the search box, and the first images that satisfied the following criteria were selected: (1) no part of the face should be obscured; (2) the individual should be mainly front facing to assist in landmark placement; (3) the image should be in colour; (4) the individual should not be lying down to ensure head angle is relatively upright. Each image was scaled to 380 pixels wide \times 570 pixels high and represented in RGB colour space using a bitmap image format.

Method

Following Burton et al. (2016), we used PCA to represent the dimensions of variability associated with a particular identity. Face shape was derived by adding a standard grid of 82 xy coordinates to anatomical landmarks for each image. To standardize this process, a Matlab add-on, which allows points to be allocated semi-automatically after five points are manually selected (see Kramer, Young, & Burton, 2018; Kramer, Young, Day, & Burton, 2017 for details), was employed. The specific grid was then manually altered to ensure points were accurately placed in accordance with the anatomical landmarks described in Kramer, Young, et al. (2017).

We used Interface software (Kramer, Young, et al., 2017) to carry out a person-specific PCA using 30 images of each of the 10 celebrities. This produced a training set for each celebrity identity. Next, we carried out the following reconstructions within the person-specific training set components of each of the 10 celebrity identities: (1) 25 novel images

of the same celebrity identity; (2) 25 images of the celebrity’s sibling; (3) 25 images of each of the four other celebrity siblings.

Results & discussion

MSE for (1) novel images of the same celebrity identity; (2) images of the celebrity’s sibling; and (3) images of the other celebrity siblings are shown in Figures 2 (shape) and Figure 3 (texture). In the ‘same identity’ condition (1), novel images of each celebrity are reconstructed in their own principal components (e.g., novel images of Brad Pitt reconstructed in Brad Pitt’s PCs). For the ‘related identity’ condition (2), the celebrity’s sibling (e.g., Doug Pitt) was reconstructed in the training set derived from the celebrity (e.g., Brad Pitt’s PCs). For unrelated identities (3), the unrelated siblings were reconstructed in each celebrity’s principal components. For example, Daniel Baldwin, Casey Affleck, Liam Hemsworth, and Kevin Dillon were reconstructed in the principal components of Brad Pitt. One-way ANOVAs were conducted separately for texture and shape using JASP.

Texture

For texture, there was a significant effect of image type for both females ($F(2, 372) = 62.08, p < .001, \eta^2_p = .25$) and males ($F(2,372) = 61.88, p < .001, \eta^2_p = .25$). *Post-*

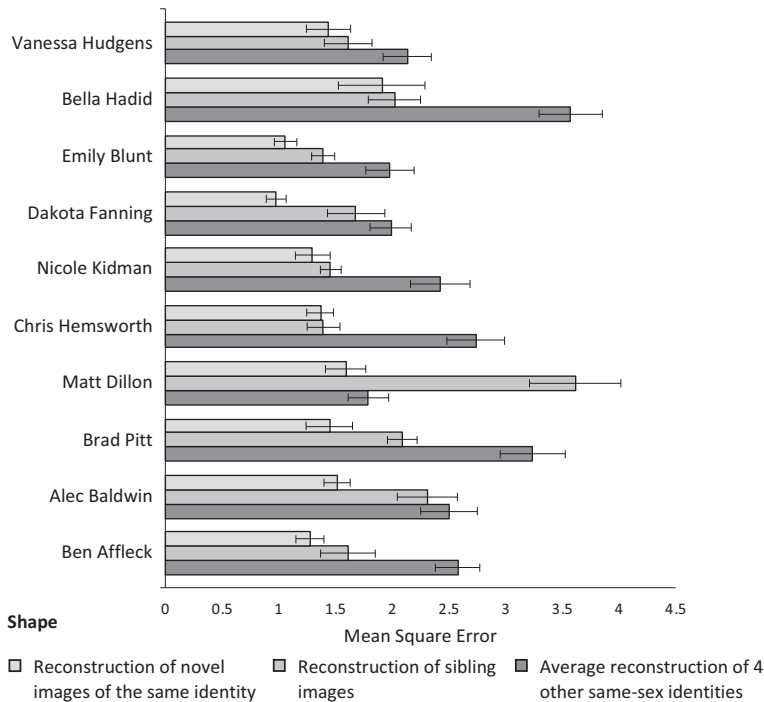


Figure 2. Shape reconstruction errors (MSE) for each identity with novel images reconstructed in their own PCs, images of the siblings constructed in the celebrities’ PCs, and average reconstruction errors of the four other same-sex identities reconstructed in the celebrities’ PCs. Error bars show standard error.

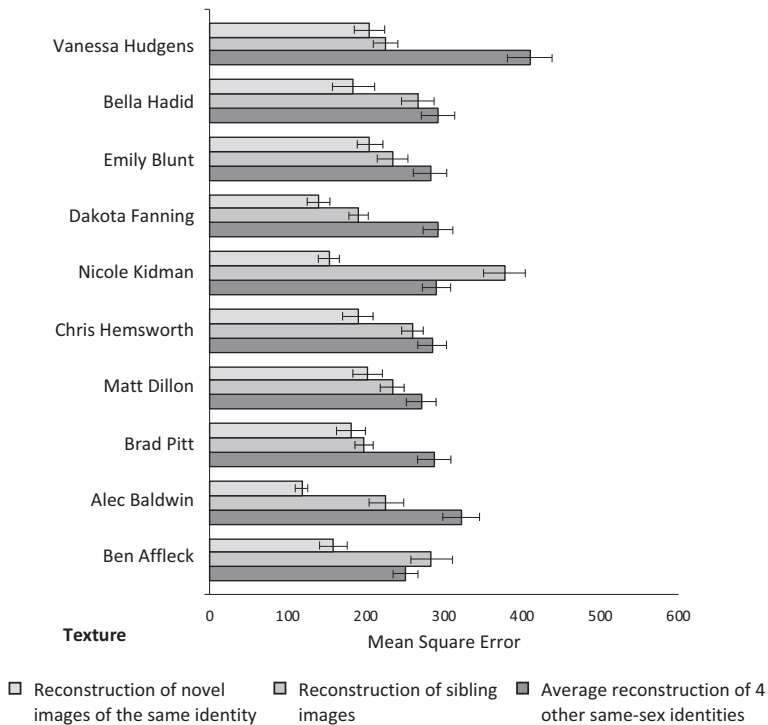


Figure 3. Texture reconstruction errors (MSE) for each identity with novel images of the identity reconstructed in their own PCs, images of the siblings constructed in the celebrities' PCs, and average reconstruction errors of the four other same-sex identities reconstructed in the celebrities' PCs. Error bars show standard error.

boc tests with a Tukey correction revealed that for both females and males, reconstruction errors were smaller when novel images of a celebrity were reconstructed using their own components compared with reconstructing their sibling in those components (both $p < .001$). This replicates the findings of Burton et al. (2016) and is consistent with the theory that faces vary idiosyncratically.

Post-boc tests also revealed that the reconstruction errors for both males and females were smaller for the siblings (2) than for the unrelated identities (3) ($p < .001$), lending support to the idea that while faces may vary idiosyncratically, the variance captured by the PCA is more representative of the siblings than the unrelated identities.

Shape

The same pattern was observed for shape. There was a significant effect of image type for both males and females (females $F(2,372) = 40.23$, $p < .001$, $\eta^2_p = .18$; males $F(2,372) = 36.74$, $p < .001$, $\eta^2_p = .17$). *Post-boc* tests once again revealed that reconstruction errors were smaller for (1) novel images of the same identity compared to (2) images of a sibling and also for (2) images of a sibling compared to (3) images of unrelated identities (all $p < .05$).

Visual inspection of Figure 2 reveals that numerically, smaller reconstruction errors for siblings relative to unrelated identities were not observed in three cases: Matt Dillon for

shape; Ben Affleck and Nicole Kidman for texture. Siblings vary in the extent to which they share genetic material, and as shape has been found to have a larger genetic component, it could be that the large MSE for shape observed with Kevin and Matt Dillon indicates that they share less common genetic material than the other siblings. Differences in texture are typically more determined by environment, and differences tend to increase as people age (Fasolt, Holzleitner, Lee, O'Shea, & DeBruine, 2019), so the relatively large MSEs for texture for the Kidman and Affleck siblings may suggest they have been differentially affected by environment in adulthood. The lack of a sibling effect in these three cases could also reflect larger reconstruction errors for that identity in general, perhaps due to the quality of the images we were able to obtain. We explore this further in a supplementary analysis section.

General discussion

The data presented here offer support for the theory that some facial variability is shared between related individuals, and that familiarity with one individual affects the way in which the face of their unfamiliar sibling is processed. In experiment 1, we found evidence that familiarity with one half of a sibling pair provided an advantage in a related/unrelated kin detection task. While previous research has established that people can detect kinship in unfamiliar faces, and that it may be modified by familiarity, experiment 1 also reveals that participants show a bias towards 'unrelated' responses with unfamiliar faces relative to familiar faces. This finding is consistent with evidence from face-matching tasks, where familiarity appears to be beneficial for the task of telling people together (Andrews et al., 2015; Ritchie & Burton, 2017).

Experiment 2 extends on these findings by offering evidence that familiarity with the face of a sibling is advantageous in a face-matching task even where the familiar sibling is not present and the relationship to the familiar individual is not cued. Participants in experiment 2 were more accurate at deciding two images showed the same individual when the target identities were siblings of familiar identities than when they were siblings of unfamiliar identities. The pattern of results is consistent with that observed in experiment 1 and with previous face-matching studies, in that the advantage offered by familiarity appeared to be restricted to the task of 'telling people together' (e.g., Andrews et al., 2015; Ritchie & Burton, 2017). We interpret these results as evidence that when the unfamiliar sibling of a highly familiar individual is encountered, some of the benefits of familiarity are transferred to the sibling. While two unrelated people may look similar in a single image pair, the similarities in appearance between biological siblings reflect genetic and environmental influences that are maintained across images and survive the variation introduced by changes in camera, lighting, angle, facial expression, etc. As a result, an understanding of the way in which one person varies across encounters may also offer a useful insight into the way in which their previously unfamiliar sibling varies in appearance.

The outcome of the principal component analysis supports the findings of Burton et al. (2016) in showing that novel images of the same identity are better reconstructed in their own components than in those of other individuals and extends this to show that same-sex sibling identities are generally better reconstructed in the components of their sibling than in those of unrelated individuals. This is consistent with the outcome of experiments 1 and 2 and lends support to the theory that structural similarities in the faces of siblings persist across variable images. Given that genetic similarity between fraternal siblings

varies (e.g., Fasolt et al., 2019) and as a result some sibling pairs look more alike than others, it is not unexpected that the same pattern was not observed for all identities. While the behavioural studies measured an average ‘sibling familiarity’ benefit across multiple trials, there was variation in accuracy between items, and the PCA suggests the effect is dependent on the degree of facial resemblance between sibling pairs. Thus, both sources of evidence may be telling the same story; we are better at processing faces belonging to the siblings of familiar identities, but only when the siblings look alike.

As texture is affected by environment, and similarities between relations on this dimension typically lessen with age (Fasolt et al., 2019), this may account for the mixed results we observed on this measure. A further issue is that while it is easy to source high-quality images of celebrities, there are relatively few publicly available images of their siblings. As such, we cannot discount the possibility that the variable quality of sibling images affected our findings on this measure.

Previous research has focussed on the importance of exposure to variation in creating familiarity with a face, and it has been argued that the benefits of learning about one individual’s facial variability do not transfer to other unrelated identities (e.g., Burton et al., 2016). Our findings support the view that variability is idiosyncratic but suggest that where two siblings are sufficiently similar in facial appearance, learning about the variability in the facial appearance of one sibling may confer benefits for processing the identity of the other. However, the MSE derived from reconstructing a face in the components of a sibling was always larger than the MSE derived from reconstructing a face in its own components. Similarly, although we observed higher matching accuracy with the siblings of familiar versus unfamiliar individuals, performance with the siblings of celebrities did not approach accuracy with familiar faces. Had we asked participants to match the faces of the celebrities, rather than the siblings, it is likely that they would have made few, if any, errors (e.g., Jenkins et al., 2011; Ritchie et al., 2015).

The current findings suggest that sibling familiarity can benefit face matching, but the effect depends on the extent to which facial variability is shared between the familiar and unfamiliar individuals. Exploring different kin relationships may offer additional insights into the aspects of facial variation that are most relevant for identity processing. Biological parents typically share more of their DNA with their biological children than is shared between siblings (Bettinger, 2017), so it seems reasonable to also predict a ‘parent familiarity’ advantage; however, differences in age and environment may constrain the effect of kinship on appearance in these relationships. Similarly, analysis of similarities between pairs of fraternal and monozygotic twins could inform our understanding of the relative importance of shape and texture in facial variation, and how they in turn influence identity processing. It is also worth considering whether a familiarity effect emerges with people who look similar to one another but are unrelated. It is clear that in a single image, two unrelated faces can look alike due to similarities in pose or superficial aspects of appearance, but superficial similarities are unlikely to be preserved across multiple images, so only ‘doppelgangers’ with structurally similar faces should generate the effect.

Young and Burton (2018a, 2018b) argue that we are experts at recognizing only the faces of familiar people and that the ability to categorize unfamiliar faces on other dimensions such as sex and race is derived from our experience of the way in which familiar faces vary. Research by Kramer, Young, et al. (2017) offers evidence to support their claim. Kramer et al. conducted LDA using multiple images of a small number of people in order to simulate the process by which we become familiar with individual faces. As predicted, after training, the system was good at classifying identity from novel images of the familiar people, and poor at classifying identity from novel images of

unfamiliar people. More surprisingly, race and gender classifications were not trained, but emerged incidentally from the system. Once established, the ability to classify the familiar faces by race and gender extended to unfamiliar stimuli.

The results of the current study extend upon the framework proposed by Young and Burton (2018a, 2018b). Our results show that kin detection is influenced by familiarity, and that familiarity with the sibling of an unfamiliar identity facilitates identity matching. Therefore, kin detection may likewise emerge as a property of prior experience with familiar faces. We theorize that knowledge of the ways in which siblings tend to resemble each other may be learned from familiar identities and applied to detect kinship in unfamiliar faces. However, where familiarity has created a robust representation of one of the siblings, knowledge both of how siblings are similar in general, and of the idiosyncratic variation in the familiar sibling, can usefully inform decisions about kinship and identity in the unfamiliar sibling.

Overall, our findings suggest that familiarity and kin recognition interact, with higher accuracy in ‘telling *siblings* together’ when one of the identities is familiar. In addition, a ‘sibling familiarity’ effect is evident in an unfamiliar face-matching task, with higher accuracy in ‘telling *people* together’ with the siblings of familiar identities. The results of the PCA support the behavioural findings by showing that the MSE for shape is generally smaller for identities reconstructed in the dimensions of a same-sex sibling than in the dimensions of an unrelated individual. Similarities in the facial appearance of siblings result in some shared variability, which means that familiarity with one sibling informs the way in which the unfamiliar sibling is processed.

The results are also consistent with Young and Burton’s (2017, 2018) theory that expertise in face recognition is limited to the faces of those with whom we are familiar. Familiarity is a product of exposure to an individual face across varied conditions, and this creates a robust representation of that individual’s within-person variation. The process of familiarization also results in incidental learning of the way in which facial appearance varies *across* the individuals with whom we are familiar, for example as a result of differences in sex or race, which is relevant to the processing of novel faces. We propose that exposure to within-person and within-family variation in facial appearance in the faces of familiar individuals also results in incidental learning of the ways in which siblings may look alike *in general*. This knowledge can inform decisions about kinship in unfamiliar faces, but familiarity with one half of the sibling pair offers an advantage. When we encounter the unfamiliar sibling of someone we already know, knowledge of the familiar sibling’s facial variation may usefully inform processing of the unfamiliar sibling’s face. This can benefit both kin detection and identity processing of the unfamiliar sibling, but the benefits are constrained by the degree to which facial variability is shared.

Acknowledgements

The authors would like to thank Peter Hancock for helpful discussions about sibling identification, and Mila Mileva for her valuable advice on PCA. Any errors are our own. Part of this work was supported by a Keele University summer studentship awarded to Sarah Laurence and Naomi Hughes-White.

Conflicts of interest

The authors report no conflicts of interest.

Author contribution

Ailsa Strathie: Conceptualization, Methodology, Investigation, Formal analysis, Project administration, Visualization, Writing – original draft, Writing – review & editing. Naomi Hughes-White: Writing – original draft, Writing – review & editing, Investigation, Formal analysis, Visualization. Sarah Laurence: Conceptualization, Methodology, Investigation, Funding acquisition, Formal analysis, Project administration, Visualization, Writing – original draft, Writing – review & editing.

Data availability statement

The data that support the findings of this study are openly available in OSF (<https://osf.io/>) at osf.io/2cwbm.

References

- Andrews, S., Jenkins, R., Cursiter, H., & Burton, A. M. (2015). Telling faces together: Learning new faces through exposure to multiple instances. *The Quarterly Journal of Experimental Psychology*, *68*(10), 2041–2050. <https://doi.org/10.1080/17470218.2014.1003949>
- Bettinger, B. (2017). August 2017 update to the shared cM Project. *The Genetic Genealogist*. Retrieved from <https://thegeneticgenealogist.com/2017/08/26/august-2017-update-to-the-shared-cm-project>
- Bonner, L., Burton, A. M., & Bruce, V. (2003). Getting to know you: How we learn new faces. *Visual Cognition*, *10*(5), 527–536. <https://doi.org/10.1080/13506280244000168>
- Bruce, V. (1982). Changing faces: Visual and non-visual coding processes in face recognition. *British Journal of Psychology*, *73*, 105–116. <https://doi.org/10.1111/j.2044-8295.1982.tb01795.x>
- Bruce, V., & Young, A. W. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305–327. <https://bpspsychub.onlinelibrary.wiley.com/doi/abs/10.1111/j.2044-8295.1986.tb02199.x>
- Burton, A. M. (2013). Why has research in face recognition progressed so slowly? The importance of variability. *Quarterly Journal of Experimental Psychology*, *66*(8), 1467–1485. <https://doi.org/10.1080/17470218.2013.800125>
- Burton, A. M., Kramer, R. S. S., Ritchie, K. L., & Jenkins, R. (2016). Identity from variation: Representations of faces derived from multiple instances. *Cognitive Science*, *40*, 202–223. <https://doi.org/10.1111/cogs.12231>
- Burton, A. M., White, D., & McNeill, A. (2010). The glasgow face matching test. *Behavior Research Methods*, *42*, 286–291. <https://doi.org/10.3758/BRM.42.1.286>
- Clutterbuck, R., & Johnston, R. A. (2002). Exploring levels of face familiarity by using an indirect face-matching measure. *Perception*, *31*, 985–994. <https://doi.org/10.1068/p3335>
- Clutterbuck, R., & Johnston, R. A. (2004). Matching as an index of face familiarity. *Visual Cognition*, *11*, 857–869. <https://doi.org/10.1080/13506280444000021>
- Davies, G., & Milne, A. (1982). Recognising faces in and out of context. *Current Psychological Research*, *2*(4), 235–246. <https://doi.org/10.1007/BF02684516>
- DeBruine, L. M., Smith, F. G., Jones, B. C., Roberts, S. C., Petrie, M., & Spector, T. D. (2009). Kin recognition signals in adult faces. *Vision Research*, *49*(1), 38–43. <https://doi.org/10.1016/j.visres.2008.09.025>
- Djordjevic, J., Zhurov, A. I., & Richmond, S., & Visigen Consortium. (2016). Genetic and environmental contributions to facial morphological variation: A 3D population-based twin study. *PLoS One*, *11*, e0162250. <https://doi.org/10.1371/journal.pone.0162250>
- Dowsett, A. J., Sandford, A., & Burton, A. M. (2016). Face learning with multiple images leads to fast acquisition of familiarity for specific individuals. *The Quarterly Journal of Experimental Psychology*, *69*(1), 1–10. <https://doi.org/10.1080/17470218.2015.1017513>

- Fasolt, V., Holzleitner, I. J., Lee, A. J., O'Shea, K. J., & DeBruine, L. M. (2019). Contribution of shape and surface reflectance information to kinship detection in 3D face images. *Journal of Vision, 19*, 9. <https://doi.org/10.1167/19.12.9>
- Fysh, M. C., & Bindemann, M. (2017). The kent face matching test. *British Journal of Psychology, 109*(2), 219–231. <https://doi.org/10.1111/bjop.12260>
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. John Wiley.
- Hancock, P. (2020). Familiar faces as islands of expertise. <https://doi.org/10.31234/osf.io/n2eug>
- Hancock, P. J. B., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences, 4*(9), 330–337. [https://doi.org/10.1016/S1364-6613\(00\)01519-9](https://doi.org/10.1016/S1364-6613(00)01519-9)
- Hancock, P. J. B., & Bulloch, A. L. (2008). Familiarity aids detection of family resemblance. *Perception ECVF Abstract, 37*, 77. <https://journals.sagepub.com/doi/pdf/10.1177/03010066080370s101>.
- Harrison, V., & Hole, G. J. (2009). Evidence for a contact-based explanation of the own-age bias in face recognition. *Psychonomic Bulletin & Review, 16*, 264–269. <https://doi.org/10.3758/PBR.16.2.264>
- 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. [https://doi.org/10.1016/S1364-6613\(00\)01482-0](https://doi.org/10.1016/S1364-6613(00)01482-0)
- JASP Team (2020). JASP (Version 0.14.1) [Computer software].
- Jenkins, R., & Burton, A. M. (2011). Stable face representations. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*, 1671–1683. <https://doi.org/10.1098/rstb.2010.0379>
- Jenkins, R., White, D., Van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. *Cognition, 121*(3), 313–323. <https://doi.org/10.1016/j.cognition.2011.08.001>
- Kaminski, G., Dridi, S., Graff, C., & Gentaz, E. (2009). Human ability to detect kinship in strangers' faces: Effects of the degree of relatedness. *Proceedings of the Royal Society B: Biological Sciences, 276*(1670), 3193–3200. <https://doi.org/10.1098/rspb.2009.0677>.
- Kramer, R. S. S., Jenkins, R., & Burton, A. M. (2017). InterFace: A software package for face image warping, averaging, and principal components analysis. *Behavior Research Methods, 49*, 2002–2011. <https://doi.org/10.3758/s13428-016-0837-7>
- Kramer, R. S. S., Young, A. W., & Burton, A. M. (2018). Understanding face familiarity. *Cognition, 172*, 46–58. <https://doi.org/10.1016/j.cognition.2017.12.005>
- 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*(2), 115–129. <https://doi.org/10.1037/rev0000048>
- Laurence, S., & Mondloch, C. J. (2016). That's my teacher! Children's ability to recognize personally familiar and unfamiliar faces improves with age. *Journal of Experimental Child Psychology, 143*, 123–138. <https://doi.org/10.1016/j.jecp.2015.09.030>
- Maloney, L. T., & Dal Martello, M. F. (2006). Kin recognition and the perceived facial similarity of children. *Journal of Vision, 6*(10), 1047–1056. <https://doi.org/10.1167/6.10.4>
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Lawrence Erlbaum Associates Publishers.
- Megreya, A. M., & Burton, A. M. (2006). Unfamiliar faces are not faces: Evidence from a matching task. *Memory & Cognition, 34*, 865–876. <https://doi.org/10.3758/BF03193433>
- Megreya, A. M., & Burton, A. M. (2007). Hits and false positives in face matching: A familiarity-based dissociation. *Perception and Psychophysics, 69*(7), 1175–1184. <https://doi.org/10.3758/BF03193954>
- Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, & Law, 7*, 3–35. <https://doi.org/10.1037/1076-8971.7.1.3>
- Noyes, E., & Jenkins, R. (2019). Deliberate disguise in face identification. *Journal of Experimental Psychology: Applied, 25*(2), 280–290. <https://doi.org/10.1037/xap0000213>
- Ritchie, K. L., & Burton, A. M. (2017). Learning faces from variability. *Quarterly Journal of Experimental Psychology, 70*, 897–905. <https://doi.org/10.1080/17470218.2015.1136656>

- Ritchie, K. L., Smith, F. G., Jenkins, R., Bindemann, M., White, D., & Burton, A. M. (2015). Viewers base estimates of face matching accuracy on their own familiarity: Explaining the photo-ID paradox. *Cognition*, *141*, 161–169. <https://doi.org/10.1016/j.cognition.2015.05.002>
- Troje, N. F., & Kersten, D. (1999). Viewpoint-dependent recognition of familiar faces. *Perception*, *28*, 483–487. <https://doi.org/10.1068/p2901>
- Tsagkraloulis, D., Hysi, P., Spector, T., & Montana, G. (2017). Heritability maps of human face morphology through large-scale automated three-dimensional phenotyping. *Scientific Reports*, *7*(45885), 1–18. <https://doi.org/10.1038/srep45885>
- Turk, M. A., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, *3*(1), 71–86. <https://doi.org/10.1162/jocn.1991.3.1.71>
- White, D., Kemp, R., Jenkins, R., Matheson, M., & Burton, A. M. (2014). Passport officers' errors in face matching. *PLoS One*, *9*(8), e103510. <https://doi.org/10.1371/journal.pone.0103510>
- Young, A. W., & Bruce, V. (2011). Understanding person perception. *British Journal of Psychology*, *102*(4), 959–974. <https://doi.org/10.1111/j.2044-8295.2011.02045.x>
- Young, A. W., & Burton, A. M. (2017). Recognizing faces. *Current Directions in Psychological Science*, *26*, 212–217. <https://doi.org/10.1177/0963721416688114>
- Young, A. W., & Burton, A. M. (2018a). Are we face experts? *Trends in Cognitive Sciences*, *22*(2), 100–110. <https://doi.org/10.1016/j.tics.2017.11.007>
- Young, A. W., & Burton, A. M. (2018b). The limits of expertise in face recognition: Response to Mackenzie Sunday and Isabel Gauthier: Face expertise for unfamiliar faces: A commentary on Young and Burton's "Are we Face Experts?". *Journal of Expertise*, *1*(2), 130–135.
- Zhou, X., & Mondloch, C. J. (2016). Recognizing "Bella Swan" and "Hermione Granger": No own-race advantage in recognizing photos of famous faces. *Perception*, *45*, 1426–1429. <https://doi.org/10.1111/bjop.12147>

Received 10 July 2020; revised version received 16 May 2021

Supporting Information

The following supporting information may be found in the online edition of the article:

Supinfo S1 Supplementary Analysis.

Figure S6 Chart showing reconstruction errors (MSE) for the 10 sibling identities reconstructed in the shape PCs of their celebrity siblings and the average reconstruction errors from reconstructing them in the PCs of the four other same-sex identities.

Figure S7 Chart showing reconstruction errors (MSE) for the 10 siblings reconstructed in the texture PCs of their celebrity siblings, and the average reconstruction error for reconstructing them in the PCs of the four other same-sex identities.

Figure S8 Chart showing errors (MSE) for reconstructions of the opposite-sex siblings' in the celebrities' shape PCs and the average reconstruction errors of the four other opposite-sex identities. The name on the Y-axis specifies the identity of the training identity.

Figure S1 Chart showing errors (MSE) for reconstructions of the opposite-sex siblings in the celebrities' texture PCs and the average reconstruction errors of the four other opposite-sex identities.