The importance of internal facial features in learning new faces

Running head: FACIAL FEATURES IN LEARNING NEW FACES

Christopher A. Longmore
University of Plymouth, UK

Chang Hong Liu
Bournemouth University, UK

Andrew W. Young
University of York, UK

Address of correspondence:

Chris Longmore
School of Psychology
Faculty of Health and Human Sciences
Plymouth University
Drake Circus
Plymouth
PL4 8AA

Tel: +44 (0)1752 584890
Fax: +44 (0)1752 584808

Email: chris.longmore@plymouth.ac.uk

Abstract

For familiar faces, the internal features (eyes, nose, and mouth) are known to be differentially salient for recognition compared to external features such as hairstyle. Two experiments are reported that investigate how this internal feature advantage accrues as a face becomes familiar. In Experiment 1, we tested the contribution of internal and external features to the ability to generalise from a single studied photograph to different views of the same face. A recognition advantage for the internal features over the external features was found after a change of viewpoint, whereas there was no internal feature advantage when the same image was used at study and test. In Experiment 2, we removed the most salient external feature (hairstyle) from studied photographs and looked at how this affected generalisation to a novel viewpoint. Removing the hair from images of the face assisted generalization to novel viewpoints, and this was especially the case when photographs showing more than one viewpoint were studied. The results suggest that the internal features play an important role in the generalisation between different images of an individual's face by enabling the viewer to detect the common identity-diagnostic elements across non-identical instances of the face.

Keywords: faces, learning, recognition, internal and external features, hairstyle

Author Notes

Christopher A. Longmore, Department of Psychology, University of Plymouth United Kingdom; Chang Hong Liu, Department of Psychology, Bournemouth University, United Kingdom; Andrew W. Young, Department of Psychology, University of York, United Kingdom

Correspondence concerning this article should be addressed to Christopher A. Longmore, School of Psychology, University of Plymouth, Plymouth, PL4 8AA, United Kingdom. Email: chris.longmore@plymouth.ac.uk

Introduction

The recognition of familiar faces differs substantially from that of unfamiliar faces (Johnston & Edmonds, 2009). Familiar faces are recognised well despite changes in viewpoint (Bruce, 1982), expression (Bruce, 1982; Clutterbuck & Johnston, 2002; Young, McWeeny, Hay, & Ellis, 1986) and image degradation such as is found in many CCTV recordings (Bruce, Henderson, Newman, & Burton, 2001; Burton, Wilson, Cowan, & Bruce, 1999). Unfamiliar face recognition is, in stark contrast, fragile. Changes in viewpoint (Bruce, 1982; Krouse, 1981; Longmore, Liu, & Young, 2008; Young et al., 1986), expression (Bruce, 1982), lighting direction (Braje, Kersten, Tarr, & Troje, 1998; Hill & Bruce, 1996; Longmore et al., 2008) and image size (Kolers, Duchnicky, & Sundstroem, 1985) have all been shown to disrupt unfamiliar face recognition, resulting in poor levels of performance in both recognition and perceptual matching tasks.

Understanding how this marked difference in the properties of recognition of familiar compared to unfamiliar faces arises is a key issue. One clue is that differences between familiar and unfamiliar face recognition have also been demonstrated in terms of the features that are used for recognition. By subdividing facial features into internal features (eyes, eyebrows, nose and mouth) and external features (comprising the hair, ears and face shape), Ellis, Shepherd and Davies (1979) found that familiar faces were better recognised from their internal features than they were from their external features, whereas unfamiliar faces demonstrated no benefit for viewing the internal features over the external features. This internal feature advantage for familiar faces has been replicated using a variety of methods and populations; for example in a perceptual matching task (Young, Hay, McWeeny, Flude, & Ellis, 1985), in children

aged as young as 5 years (Wilson, Blades, & Pascalis, 2007) and with Japanese faces and observers (Endo, Takahashi, & Maruyama, 1984).

Although it is clearly a robust phenomenon, the reason why there should be an internal feature advantage for familiar faces remains unclear, but at least two possibilities have been suggested. First, internal features convey important social information such as an individual's mood, and will therefore receive more attention over time than external features. Second, the internal features of a face are less likely to undergo significant change than the most salient external feature of the hairstyle (Shepherd, Davies, & Ellis, 1981), and therefore offer a more reliable cue. Using the hair as a primary diagnostic cue to identity would be a poor strategy. For example, in their study of the usefulness of photographic images on credit cards Kemp, Towell and Pike (1997) found that cashiers attempting to match the photograph on a credit card to the user of the card were less able to detect fraud in female shoppers than in male shoppers. Kemp et al. suggest that cashiers appeared to be influenced by whether the hairstyles of a female shopper and her accompanying photograph matched. Male shoppers tended to have less variable hair, leading cashiers to focus upon more valid internal features.

Whatever the underlying reason - and it is of course likely that both of the possibilities mentioned make a contribution - studies have shown that the internal feature advantage arises gradually with increased exposure to a face (Bonner, Burton, & Bruce, 2003; Clutterbuck & Johnston, 2002; 2004; 2005; Osborne & Stevenage, 2008) suggesting that it is not an all-or-nothing effect and develops progressively over

time as a face becomes more familiar. Here, we investigate this in more detail using techniques introduced by Longmore et al. (2008).

Longmore et al. (2008) were interested in the question of whether unfamiliar face recognition is particularly fragile across image transformations because the face has not been seen very often or because it has not been seen in many different views. To investigate this they pushed the issue to its limit by investigating how well recognition generalised to a different view from a face that had been learnt only from one or two photographs. Their study demonstrated two striking phenomena. First, when participants learnt to recognise a face from a single photograph the generalisation gradient for the fall-off in recognition performance across different unstudied views of the same face was as steep when the single photograph was over-learnt (by studying it repeatedly) as it was when the photograph was studied once only (Experiment 1). Studying the same photograph repeatedly thus led to better recognition of the same photo, but did not improve the robustness of this recognition to image transformations such as a change in viewpoint. Second, when participants were taught two different views (full-face and profile) of the same person's face, their generalisation to novel views of the face was no better than when only a single view had been learnt (Experiment 3). That is, participants seemed to learn each of these very different views as if it was an entirely separate instance, and proved unable to combine information across the views to form a more effective representation.

Here, we use both of these phenomena to probe the roles of internal and external features. In Experiment 1 participants learnt a set of faces from a single photograph of each that was studied once only or across multiple presentations, using the procedure

6

established by Longmore et al. (2008, Experiment 1). Like Longmore et al. we then tested recognition of the same faces across different views, but we also investigated whether the performance decrement created by a changed view was equivalent for recognition based on the internal or external features. Because Experiment 1 showed the potential value of learning internal features, Experiment 2 picked up the implications of this by removing the most salient external feature (the hairstyle) from the images and repeating Longmore et al.'s Experiment 3 study of the effect of learning more than one view of the face. For faces shown without hairstyles, we were able to demonstrate better ability to integrate information across different views.

Experiment 1: Generalisation from a single photograph

Experiment 1 examined whether faces learnt from a single view demonstrate any internal feature advantage in a subsequent recognition test. Participants were either trained to recognise a single photograph of a complete face to induce a level of familiarity, or received a single exposure to the facial image to leave it relatively unfamiliar. Previous work has suggested that learning faces from a single picture results predominantly in the learning of image properties (e.g. Bruce, 1982; Longmore et al., 2008). It may be however, that even in this highly image-dependent scenario, some internal feature advantage can arise. Thus, participants were tested for their recognition of studied faces from the internal and external features, and performance was examined with respect to recognition of the originally studied image and for generalisation to a different view.

Method

Design

Participants were randomly allocated to one of two groups. A single exposure group studied a single photograph of each face on one occasion only, and a multiple exposures group studied a single photograph of each face during a training session in which each photograph was each presented a number of times. After the study session with the face photographs, participants took part in a recognition memory test in which the studied target faces were mixed with unstudied distractor faces. They were asked to recognise the faces they had seen before from the same image as learnt (i.e. a change of 0°) and from a change involving a rotation in view of 31°. Recognition was also tested across three types of image: the whole face, the internal features only, and the external features only. The experiment therefore had a 2x2x3 design with training group, type of test image, and pose change as independent variables. The dependent variable was the number of faces correctly identified (hits) during the test phase.

Participants

Forty undergraduate students (35 female, 5 male) aged between 18 and 22 ($M = 19.35$, $SD = 1.06$) years took part in the experiment in return for course credit or payment. All participants were naïve as to the nature of the experiment, had normal or corrected-to-normal vision, and gave their full informed consent before completing the experiment.

Materials

Images of 20 faces (all Caucasian and male) from the PIE face database (Sim, Baker, & Bsat, 2002) were used. None of the images depicted an individual with facial hair or wearing glasses. The 20 identities were split into two groups of 10 faces each for counter-balancing purposes. Each face was used in two poses and three image types (whole face, internal features and external features), resulting in a total of 120 images. The images of the two poses were taken from directly in front of the model (0°) or to the right (31°). The original colour images were converted to greyscale and each image was manipulated to remove all irrelevant background information, leaving only the head visible with the background replaced with a homogenous mid-grey colour. Each image was resized so that it was 384 pixels high in order to normalize face height and the background was expanded to create a final image of 384x384 pixels, subtending a visual angle of 4.87° when viewed from a distance of 60cm. Two further variations of each image were made to show only the internal features (eyes, nose and mouth) or external features (hair, face shape and ears) by masking parts of the original image with an oval frame. Examples of the images used are shown in Figure 1.

[Insert Figure 1 here]

Apparatus

The faces were presented on a 17 inch LCD screen monitor, set to a resolution of 1280 x 1024 pixels and a colour depth of 32 bits per pixel using a custom written computer program created in the Microsoft Visual Basic programming language. Participants made their responses through the use of a standard mouse.

9

Procedure

Participants were randomly assigned to either the single or multiple exposure conditions. For half the participants in each group, the first set of faces was allocated as the target set and the second was allocated as distractors. For the other half, this order was reversed.

All participants completed a study phase and a test phase of the experiment. Those in the multiple exposures condition also received additional exposure to the faces during the study phase. The participant sat in front of the computer screen at a distance of approximately 60cms and was given written instructions before the experiment began.

Study phase

During the study phase, all of the images shown depicted the whole face. Participants saw ten faces for a duration of 5 seconds each with 0.5 seconds between each face. The faces were evenly distributed across the two poses (0° and 31°) so that five faces were seen in each pose. Each individual face was presented to the participant once and was accompanied by a first name, presented below the image of the face. These name/face pairings were randomly generated for each participant using a fixed set of ten first names (David, Robert etc.). During this phase, the participants were instructed to remember the name/face pairings as best they could.

The training task used in the study phase for the multiple exposures participant group was the same as that used in Longmore et al. (2008) and was divided into two parts. In the first part, the ten face photographs shown during the initial exposure phase were divided into two blocks containing five faces each. In the second part, all ten

face photographs were presented in a single block. The task was the same for both parts.

On each trial of the training task the participant was presented with a single whole face image and their task was to indicate the name of the individual. Name options were given in the form of on-screen buttons located below the image of the face and participants were required to make a mouse click on the name that they thought belonged to the face. After a response was made, immediate feedback that took the form (names given are examples) "Yes, this is David" (correct answer) or "No, this is Robert" (for an incorrect answer). In the event of the participant making an error, the faces that were correctly named were removed from the set and the remainder re-presented. This process was repeated until all faces in the block of trials had been correctly identified, upon which all the faces were re-entered into the set to begin the next block of training trials. To complete the training, the entire set of ten faces had to be identified without error three times. This naming task was used merely to ensure that the participants had successfully individuated the different facial images. Participants were not tested on their knowledge of the name/face pairs after the training phase.

Test phase (all participants)

The testing phase was divided into six blocks of 20 individually presented images. Each image depicted a whole face, the internal features of a face only or the external features only, for which the participant made a "yes/no" decision as to whether they had seen the presented face during familiarisation. All the images within a block were presented in the same combination of pose (0° or 31°) and image type (whole

image, internal features only or external features only). The order of the presentation of the resulting six types of test block was rotated across participants in a reduced Latin-square design.

For each block, ten of the individuals were those in the familiarisation set, whilst the other ten were taken from the distractor set. Faces were presented one at a time and two buttons (labelled "Yes" and "No") were used for responses. Participants were required to click on "Yes" if they thought they recognised the individual as a member of the training set and "No" if they did not.

Results

The number of hits (faces correctly recognised as being members of the familiarisation set) was used for analysis. An average percentage correct score was calculated for each participant based on the number of hits obtained during the test phase. These data are shown in Figure 2.

[Insert Figure 2 here]

The hit rates were entered into a mixed design 2x2x3 ANOVA with training condition (single or multiple presentation, between-subjects), pose change (0° or 31°, within-subjects) and test image type (whole face, internal features and external features, within-subjects) as independent variables and number of hits obtained in the recognition task as the dependent variable. The Huynh-Feldt correction for departures from sphericity was used throughout the analyses and effect sizes are calculated using

12

generalised eta-squared (Bakeman, 2005). The results of this analysis are summarised in Table 1.

[Insert Table 1 here]

Since there was a three-way interaction between pose, image type, and training type, the interaction effect was investigated using simple main effects analysis. When multiple exposures were given there was a significant effect of test image type; $F(1.83, 34.69) = 25.10$, $MSE = 2.23$, $\eta_G^2 = 0.199$, $p < .001$ and no interaction between test image type and pose ($F(1.38, 26.18) = 2.50$, $MSE = 1.94$, $\eta_G^2 = 0.012$, $p = .117$. Contrasts revealed that a whole face image was recognised significantly better than recognition of either the internal or external features alone; $F(1,19) = 73.17$, $MSE = 6.89$, $\eta_G^2 = 0.521$, $p < .001$. Interestingly an internal feature advantage was also found with images showing only the internal features recognised better than images showing only the external features; $F(1,19) = 6.19$, $MSE = 5.84$, $\eta_G^2 = 0.075$, $p = .022$. When a single exposure was given a significant interaction was found between pose and image type; $F(2,38) = 10.97$, $MSE = 1.22$, $\eta_G^2 = 0.045$, $p < .001$. This interaction was analysed by further simple main effects. When no pose change occurred, the whole face was recognised significantly better than the internal or external features ($F(1,38) = 47.83$, $MSE = 1.09$, $\eta_G^2 = 0.432$, $p < .001$) and no internal feature advantage over the external features was found ($F < 1$, ns). After a pose change however, there was no difference in the recognition of either of the three image types ($F < 1$, ns), a result in contrast to the multiple exposure condition. In fact, performance for all three image types was at chance ($t < 1$, ns for all three test image types) demonstrating that for faces seen only once participants had difficulty recognising them after a pose change.

13

Thus, when recognition is required after a pose change (rather than recognition of the same picture), an internal feature advantage emerges when the training image of the face has been extensively learnt, but not when only a single exposure to the face is given.

Discussion

The current experiment revealed that when presented with a single exposure there was no difference in participants' recognition of the studied faces from internal and external features. With multiple exposures however a different pattern emerged. Importantly, extensive learning of a single image of a whole face through training resulted in a clear benefit for recognition of the internal features over the external features and this benefit for the internal features persisted after a pose change. This suggests that the internal features contain more viewpoint-invariant information that may be used to recognise faces across transformations such as pose.

Experiment 2: Learning internal features of the face

Experiment 1 demonstrated that the internal features of a face that was thoroughly learnt from a single image could show an advantage in terms of recognition accuracy over the external features when recognising the face across a viewpoint change. Thus, the internal features of a face can be more useful for recognition across pose changes than the external features. Of the available external cues one in particular is especially salient – the hairstyle (Shepherd et al., 1981). Although a potentially easy cue to extract, overreliance on the hairstyle may be misleading in many circumstances (Kemp et al., 1997). We therefore sought to test whether removal of the hair from

learnt images might create better generalisation across different images by encouraging participants to use the internal features.

To examine this possibility, Experiment 2 investigated participants' ability to generalise across different views of faces learnt from internal features, using the same overall design as that of Longmore et al. (2008 experiments 3, 4 and 5) in which the recognition performance of participants was assessed for faces learnt from multiple exposures of a single view or from multiple exposures of two different views. Longmore et al. found that when faces were learnt from two different viewpoints (full-face and profile), recognition accuracy for an image showing a view intermediate between those the views learnt in the study phase was statistically no better than if just one of the views had been learnt. Hence it appeared that participants were unable to combine two very different images of the same person into a more robust (view-invariant) representation of the face. The faces used by Longmore et al. were whole face images; if encouraging participants to focus upon the internal features (by removing the hairstyles) is beneficial for the extraction of viewpoint invariant information, then recognition of a view that is intermediate between the two learnt views in the present Experiment 2 would be expected to be better than if a single viewpoint had been learnt.

Method

Design

As for Experiment 1, participants learnt images of faces in a study phase and their recognition of the studied faces was then tested. However, the study phase for all participants involved multiple presentations of the images that were being learnt - the

single presentation study phase used with one group of participants in Experiment 1 was not considered necessary, as it had produced little evidence of any generalisation to unstudied views. A 3x3 within-subjects factorial design was used for Experiment 2, with study view (full-face, profile or both) and test view (full-face, three-quarter or profile) as factors. The dependent variable was the number of faces correctly recognised during the test phase.

Participants

Sixteen undergraduate participants (14 females and 2 males) aged between 18 and 57 ($M = 21.69$, $SD = 9.15$) years took part in the experiment in return for course credit or payment. All participants had normal or corrected-to-normal vision and gave informed consent before beginning the experiment. None had participated in the previous experiment and were thus naïve as to the purpose of the study.

Materials

Images of 24 individuals from the PIE database (Sim et al., 2002) were used. The faces were arbitrarily split into two groups of 12 for counterbalancing reasons, and three images of each face were prepared – one from the full-face viewpoint (0°), one from a three-quarter viewpoint (31°) and one from a near-profile view (62°), resulting in a total of 72 images. Each image was prepared in the same way as the whole face images in Experiment 1 except that the hairstyle was removed from each of the faces, using a consistent template for each viewpoint to avoid the possibility of creating different shapes around the top of the head that might be used for recognition. Examples of the images are shown in Figure 3.

[Insert Figure 3 here]

Apparatus

The same apparatus as used in the previous experiment was employed.

Procedure

The experiment consisted of three phases, comprising initial exposure, training and test phases. Each participant was allocated one set of 12 faces to act as targets and the other as distractors. The allocation of the two sets was counter-balanced across participants. Within the training set, for each participant, four faces were randomly designated to be learnt from the full-face view only (yielding four images), four were randomly assigned to be learnt from the profile view only (yielding four images), and four faces were randomly assigned to be learnt from both views (yielding eight images). This resulted in 12 *individual faces* to be learnt by each participant from a total of 16 different *images*.

Study phase

Twelve images were used during the exposure phase. Eight of these 12 images consisted of the four faces to be learnt in full-face view and the four to be learnt in profile. Of the final four images, depicting individuals to be learnt in both views, two were randomly selected to be presented in full-face and the other two were presented in profile. Thus, only one view of each person was given during the exposure phase. As in Experiment 1, each face was presented with a name underneath it for 5 seconds with 0.5 seconds between each face and participants were instructed to remember the name/face pairs.

17

Training phase

In the training phase, all 16 images were used. For the first part of training, the 16 images were split into four blocks of four images each. Within each block, all faces were shown in *either* full-face or profile, with eight faces presented in each view. Participants learnt either the full-face views followed by profile views or vice-versa (counter-balanced across participants). As in the previous experiment, to progress through the blocks participants were required to name all four people in the block without making an error, on 3 separate runs.

The second phase of training presented the participant with all 16 images. During this phase, only the 12 possible names were given as selection choices. Participants were required to name all 16 images without making an error, on three separate occasions, to complete the training task.

Test phase

The test phase involved three blocks consisting of faces presented in full-face view, three-quarter view, or near-profile view. These blocks were counter-balanced across participants. Each block comprised 24 images, 12 of which were photographs of the faces learnt during the training phase and the other 12 faces were distractors. Within each block, all faces were presented from the same view. Participants were instructed to press "Yes" if they thought the person presented was a member of the set of 12 they originally learnt or "No" if they were not.

Results

The number of hits (faces correctly recognised as being members of the familiarisation set) was used for analysis purposes. Mean percentage correct scores were calculated from the number of hits for full-face only, profile only and both full-face and profile view learning conditions for the three test views. These data are shown in Figure 4.

[Insert Figure 4 here]

A 3x3 repeated-measures ANOVA was conducted with learning type (full-face only, profile only, or full-face and profile) and test view (full-face, three-quarter, or profile view) as independent variables and number of hits as the dependent variable. A summary of the results of this analysis is shown in Table 2.

[Insert Table 2 here]

The key finding from this analysis is the interaction indicating that performance on the three test views differed across the three learning conditions, and a simple main effects analysis was carried out to decompose this. From Figure 4, it is evident that recognition of the profile view was poor after the full-face view only was learnt, as was recognition of the full-face view after a profile view only was learnt, and this is a result which contrasts sharply with the high levels of accuracy obtained when the same view (and therefore image) was used for both learning and testing. Consequently, an expected effect of learning condition for both the full-face test view; $F(2,30) = 79.68$, $MSE = 0.51$, $\eta_G^2 = 0.728$, $p < .001$ and the profile test view; $F(2,30) = 81.83$, $MSE = 0.48$, $\eta_G^2 = 0.681$, $p < .001$ was found. In addition, an effect of

learning condition was also found at the three-quarter test view; $F(2,30) = 4.35$, $MSE= 0.52$, $\eta_G^2 = 0.081$, $p = .022$. This is particularly noteworthy as the three-quarter view was never studied in any of the different learning conditions. Contrasts revealed that the two-view learning condition led to higher levels of performance for three-quarter test views than did either the full-face only or profile view only learning conditions; $F(1,15) = 6.64$, $MSE = 3.05$, $\eta_G^2 = 0.200$, $p = .021$. In contrast, the full-face only and profile only learning conditions did not differ significantly from each other; $F(1,15) = 2.14$, $MSE = 1.05$, $\eta_G^2 = 0.042$, $p = .164$. The results indicate that learning two views of a face led to greater generalisation from the learnt images to the novel three-quarter view test image in contrast to learning a single view, a result in contrast to Longmore et al. (2008, Experiment 3, 4, 5).

Discussion

Experiment 2 examined how recognition of faces that were studied without a hairstyle generalised to novel viewpoints. It would appear that learning two views of the internal features of the face allows for more viewpoint-invariant identity-diagnostic information to be extracted that gives rise to better generalisation to novel views as demonstrated by the critical test involving recognition from the previously unseen three-quarter view. It was found that recognition of this previously unseen three-quarter test view was significantly higher after two views of the face had been learnt than if only a single view had been learnt. This indicates that participants were better able to extract viewpoint-invariant information after learning two views of a face in comparison to learning a single view, a result that is in direct contrast to that of Longmore et al. (2008). Furthermore, it was also found that performance on the three-quarter test view after two views of the face were learnt was not significantly

below that obtained for the learnt views themselves. Although ceiling effects were likely present in the data, participants were clearly demonstrating substantial levels of generalisation to a novel view of a face.

General discussion

The experiments presented in this paper examined which of the cues present in a face are most useful for generalisation to novel viewpoints. It has been previously reported that the internal features of already familiar faces (Ellis et al., 1979; Endo et al., 1984; Young et al., 1985) and faces learnt over an extended period from multiple images (Bonner et al., 2003; Clutterbuck & Johnston, 2002; 2004; 2005) are particularly useful for recognition. Experiment 1 extended this work and demonstrated that some degree of internal feature advantage can arise not only for familiar faces that have been seen in many different views, but even when the face is learnt from multiple exposures to a single image within a relatively short timeframe. Furthermore, this advantage was also present when the face was shown in a different viewpoint to the one studied, compared to the original studied image. However, this advantage was only observed when the face had been learned from repetitive exposures and not after the face had been seen only once (see also Liu & Chaudhuri, 1998). Previous work that has employed a short laboratory-based learning procedure has not shown such an advantage (eg. Ellis et al., 1979) and it would appear that the repetitive exposure procedure used in the multiple exposures condition of Experiment 1 was sufficient to evoke an internal feature advantage that persists across changes in pose. This result provides a hint that the reason the internal feature advantage arises in the first place is to aid with generalisation to novel views of a face. Despite only learning a limited number of images of each face in this and other (e.g. Longmore et al., 2008)

experiments, this process may mimic the real world processing of learning new faces. When encountering someone in the real world for the first time a limited number of instances of their face are available to facilitate later recognition. Over a period of time and with repeated encounters with the individual it may simply be the case that more instances of the face are stored, enabling the highly robust form of recognition seen with familiar faces.

Building on this observation, Experiment 2 showed how encouraging participants to rely on internal features by removing the salient external feature of the hairstyle promoted the integration of information across different study views of a face, leading to enhanced generalisation of recognition to a previously unstudied view. Previous work has demonstrated that presenting the same picture during learning and recognition can lead to a process more akin to image learning than face learning (e.g. Bruce, 1982). Here, though, we were able to demonstrate the beginnings of a more face-like, generalisable form of learning based primarily on internal features.

An important consideration is why the internal features should be more informative than the external features? The analyses offered by Bruce (1994) and by Burton (2013) offer a useful overarching perspective. We naturally tend to think of faces as being highly similar to each other (Galton, 1833), which makes us think of the problem of face recognition as being primarily one of *discriminating* different individual faces from each other. Intuitively appealing as this idea is, it seriously underestimates the huge range of differences between views of the same face at different times or in different photographs, and these differences make the problem of face recognition as much one of finding the underlying commonalities that allow one

to group together views that may superficially be very different from each other, rather than that of telling the views apart (Bruce, 1994; Burton, 2013; Jenkins, White, Van Montfort, & Burton, 2011). For example Jenkins et al. presented participants with an array of 40 photos made up of 20 images of two individuals. Participants who were unfamiliar with the two identities, and therefore had difficulty determining the commonality between different images of the same person, indicated that they thought that there were on average 7.5 identities in the array of images. In contrast, participants who were familiar with the identities correctly recognised that there were only two different people in the set of 40 images. This point of difficulty in determining commonality is clearly seen in our data and those of Longmore et al. (2008), where the errors participants make often involve failing to see that two different images are pictures of the same person, not failing to discriminate between individuals. In this search for perceptual commonalities that can allow different views to be grouped together as belonging to the same face, the internal features seem to play a critical role.

References

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods, 37*(3), 379-384. doi: 10.3758/BF03192707

Bonner, L., Burton, A. M., & Bruce, V. (2003). Getting to know you: How we learn new faces. *Visual Cognition, 10*(5), 527-536. doi: 10.1080/13506280244000168

Braje, W. L., Kersten, D., Tarr, M. J., & Troje, N. F. (1998). Illumination effects in face recognition. *Psychobiology, 26*(4), 371-380. WOS:000078391900008

Bruce, V. (1982). Changing faces - Visual and non-visual coding processes in face recognition. *British Journal of Psychology, 73,* 105-116. WOS:A1982NC88500012

Bruce, V. (1994). Stability from variation: the case of face recognition. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 47*(1), 5-28. WOS:A1994MY82900002

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. doi:10.1037/1076-898X.7.3.207

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. doi:10.1080/17470218.2013.800125

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. doi:10.1111/1467-9280.00144

Clutterbuck, R., & Johnston, R. A. (2002). Exploring levels of face familiarity by using an indirect face-matching measure. *Perception, 31*(8), 985-994. doi:10.1068/p3335

Clutterbuck, R., & Johnston, R. A. (2004). Matching as an index of face familiarity. *Visual Cognition, 11*(7), 857-869. doi:10.1080/13506280444000021

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. doi: 10.1080/09541440340000439

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology, 1*(1), 42-45.

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. doi: 10.1068/p080431

Endo, M., Takahashi, K., & Maruyama, K. (1984). Effects of observer's attitude on the familiarity of faces: Using the difference in cue value between central and peripheral facial elements as an index of familiarity. *Tohoku Psychologica Folia, 43*(1-4), 23-34.

Galton, F. (1833). *Inquiries into human faculty and its development*. London: Macmillan.

Hill, H., & Bruce, V. (1996). Effects of lighting on the perception of facial surfaces. *Journal of Experimental Psychology Human Perception and Performance, 22*(4), 986-1004. doi: 10.1037/0096-1523.22.4.986

Jenkins, R., White, D., Van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. *Cognition*, *121*(*3*), 313-323. doi: 10.1016/J.Cognition.2011.08.001

Johnston, R. A., & Edmonds, A. J. (2009). Familiar and unfamiliar face recognition: A review. *Memory, 17*(5), 577-596. doi: 10.1080/09658210902976969

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. WOS:A1997XG45700002

Kolers, P. A., Duchnicky, R. L., & Sundstroem, G. (1985). Size in the visual processing of faces and words. *Journal of Experimental Psychology Human Perception and Performance, 11*(6), 726-751. doi:10.1037//0096-1523.11.6.726

Krouse, F. L. (1981). Effects of pose, pose change, and delay on face recognition performance. *Journal of Applied Psychology, 66*(5), 651-654. doi:10.1037/0021-9010.66.5.651

Liu, C. H., & Chaudhuri, A. (1998). Are There Qualitative Differences Between Face Processing in Photographic Positive and Negative? *Perception, 27*(9), 1107-1122. doi:10.1068/p271107

Longmore, C. A., Liu, C. H., & Young, A. W. (2008). Learning faces from photographs. *Journal of Experimental Psychology-Human Perception and Performance, 34*(1), 77-100. doi: 10.1037/0096-1523.34.1.77

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology, 4*(2), 61-64.

Osborne, C. D., & Stevenage, S. V. (2008). Internal feature saliency as a marker of familiarity and configural processing. *Visual Cognition, 16*(1), 23-43. doi: 10.1080/13506280701238073

Shepherd, J., Davies, G., & Ellis, H. (1981). Studies of cue saliency. In G. Davies, H. Ellis & J. Sherherd (Eds.), *Perceiving and remembering faces* (pp. 105-133). London: Academic Press.

Sim, T., Baker, S., & Bsat, M. (2002). The CMU Pose, Illumination, and Expression (PIE) database. *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, May 2002*, 46-51. WOS:000176289900008

Wilson, R. R., Blades, M., & Pascalis, O. (2007). What do children look at in an adult face with which they are personally familiar? *British Journal of Developmental Psychology, 25*, 375-382. doi:10.1348/026151006X159112

Young, A. W., Hay, D. C., McWeeny, K. H., Flude, B. M., & Ellis, A. W. (1985). Matching familiar and unfamiliar faces on internal and external features. *Perception, 14*(6), 737-746. doi:10.1068/p140737

Young, A. W., McWeeny, K. H., Hay, D. C., & Ellis, A. W. (1986). Matching familiar and unfamiliar faces on identity and expression. *Psychological research, 48*(2), 63-68. doi: 10.1007/BF00309318