University of Central Lancashire
UCLan

Central Lancashire Online Knowledge (CLoK)

Title	Familiarity effects in the construction of facial-composite images using modern software systems.
Type	Article
URL	https://clok.uclan.ac.uk/id/eprint/4279/
DOI	https://doi.org/10.1080/00140139.2011.623328
Date	2011
Citation	Frowd, Charlie, Skelton, Faye Collette, Butt, Neelam, Hassan, Amal, Fields, Fields and Hancock, Peter J.B. (2011) Familiarity effects in the construction of facial-composite images using modern software systems. Ergonomics, 54 (12). pp. 1147-1158. ISSN 0014-0139
Creators	Frowd, Charlie, Skelton, Faye Collette, Butt, Neelam, Hassan, Amal, Fields, Fields and Hancock, Peter J.B.

It is advisable to refer to the publisher's version if you intend to cite from the work.
https://doi.org/10.1080/00140139.2011.623328

For information about Research at UCLan please go to http://www.uclan.ac.uk/research/

All outputs in CLoK are protected by Intellectual Property Rights law, including Copyright law. Copyright, IPR and Moral Rights for the works on this site are retained by the individual authors and/or other copyright owners. Terms and conditions for use of this material are defined in the http://clok.uclan.ac.uk/policies/

Familiarity effects in the construction of facial-composite images using modern software systems

Charlie D. Frowd[1]*, Faye C. Skelton[1], Neelam Butt[1], Amal Hassan[1] and Stephen Fields[2]

[1]University of Central Lancashire

[2]HMP Edinburgh

Relevance

The work is of relevance to practitioners who construct facial composites with witnesses to and victims of crime, as well as for software designers to help them improve the effectiveness of their composite systems.

*Corresponding author: Charlie Frowd, School of Psychology, University of Central Lancashire, Preston PR1 2HE, UK. Email: cfrowd@uclan.ac.uk. Phone: +44 1772 893439.

Abstract

We investigate the effect of target familiarity on the construction of facial composites, as used by law-enforcement to locate criminal suspects. Two popular software construction methods were investigated. Participants were shown a target face that was either familiar or unfamiliar to them and constructed a composite of it from memory using a typical 'feature' system, involving selection of individual facial-features, or one of the newer 'holistic' types, involving repeated selection and breeding from arrays of whole faces. The study found that composites constructed of a familiar face were named more successfully than composites of an unfamiliar face; also, naming of composites of internal and external features were equivalent for unfamiliar targets, but internal features were better named than external features for familiar targets. These findings applied to both systems, although benefit emerged for the holistic type due to more accurate construction of internal features and evidence for a whole-face advantage.

Keywords: facial composite; eyewitness; evolutionary algorithms; facial feature; memory

Familiarity effects in the construction of facial composite images using software systems

1 Introduction

We are extremely good at recognising familiar faces (e.g. Bruce, 1986; Bruce, Henderson, Newman & Burton, 2001; Kemp, Towell & Pike, 1997). Face recognition occurs automatically and accurately for well-known people, and from different viewpoints, lighting conditions and physical environments – a bank, park, supermarket, etc. Even in less than ideal situations, such as a grainy wedding photograph or a poor quality CCTV image, recognition accuracy and confidence levels are very high for familiar persons (Burton, Wilson, Cowan & Bruce, 1999); we can even recognise a familiar face accurately when heavily pixelated (Lander, Bruce & Hill, 2001). The robust nature of familiar face recognition is in contrast to the recognition of unfamiliar faces: when a face has been seen a few times, or just once, recognition judgements are less accurate (e.g. Young, Hay & Ellis, 1985a) and errors are often made by confusing faces of similar appearance (Cohen & Nodine, 1978; Davies, Shepherd & Ellis, 1977). Unfamiliar face recognition is slower (e.g. Bruce, 1986) and confidence judgements are lower in comparison to the familiar face (Burton et al., 1999).

Different regions of the face also have varying importance for recognition depending on familiarity. Ellis, Shepherd and Davies (1979) compared photographs of *internal facial features* of celebrity faces—the region comprising of brows, eyes, nose and mouth—with photographs of *external facial features*—hair, face outline and ears. These regions were compared with a control condition comprising photographs of complete faces. They found that, irrespective of familiarity, intact faces were named the best, familiar faces were recognised better by their internal than external features but, for unfamiliar items, internal and external features were equally identifiable. The work also found evidence for a whole-

face advantage: seeing the complete face led to more successful recognition than combining recognition scores from internal features and external features seen separately (for evidence of this, see footnote [1]). Young, Hay, McWeeny, Flude and Ellis (1985b) found faster reaction times for matching internal relative to external features of familiar faces, thus supporting an internal features advantage for identities that have been seen many times. These basic effects have been replicated many times (Campbell et al., 1999; Clutterbuck & Johnston, 2002; de Haan & Hay, 1986; Endo et al., 1984; Ge et al., 2008; Want, Pascalis, Coleman & Blades, 2003; Young, 1984).

When a crime is committed, eyewitnesses may be asked to construct a facial composite of the perpetrator. This is a visual representation of the face and is traditionally made by eyewitnesses describing a criminal's appearance and selecting individual features—hair, face shape, eyes, nose, mouth, etc. Most composites nowadays are constructed in this way using software programs, but some police forces use forensic artists who follow a similar process with pencils or crayons. Either way, the resulting image is published in the media with the aim that someone will name it to the police and provide a further line of enquiry. As such, a face is normally constructed using unfamiliar face perception, as most witnesses only see a criminal on one occasion, but to be named, the composite must be recognised by somebody familiar with the person depicted. Thus the construction and recognition of facial composites involve both types of face processing.

Based on the recognition biases described above, and using such human interfaces to access facial memory, the implication is that the internal features should be constructed as accurately as the external features (since the face is unfamiliar to the person constructing it).

[1] In Ellis et al. (1979), photographs of familiar complete faces were correctly named (M) at 80%, internal features at 50% and external features at 30%. The probabilities of *not* correctly naming the face—calculated as (1 – M/100)—are 0.2, 0.5 and 0.7, respectively. Using the product rule, the probability of not correctly naming both internal and external features seen separately is 0.5 * 0.7, or 0.35. Converting this probability (p) into percent correct—(1 – p)*100—gives 65%, but correct naming was much higher when seeing the complete face (80%). This analysis provides evidence for a whole-face (holistic) recognition advantage.

The lack of an advantage for the internal features here is likely to be problematic since recognition will be attempted later by someone familiar with that identity—a person who will rely mainly on this region of the face. It is entirely possible that this change of processing focus is one of the reasons why composites tend to be named infrequently (e.g. Ellis, Davies & Shepherd, 1978; Frowd et al., 2005a, 2007b; Koehn & Fisher, 1997). This observation also suggests that methods which can help witnesses to construct a better set of internal features are likely to promote a more identifiable image, and so be of greater value to law enforcement.

A second and well-known issue with these traditional 'feature' systems is that witnesses are required to describe and select individual features (Davies, Shepherd & Ellis, 1978). These are unnatural and difficult tasks, and together is perhaps the reason why target familiarity effects have not been observed (Davies et al., 2000), or are very weak (Frowd et al., 2007b) when a person constructs a face from memory. Familiarity effects can be relevant practically to face construction as well, since sometimes witnesses do see offenders on more than one occasion, in some prolonged criminal cases of deception for example (but do not know the correct identity of person, hence the need to construct an image of the face). It is in this situation that face recognition tends to be very good but, as mentioned above, face construction using the 'feature' approach is little better than if the target face were unfamiliar. In general, faces are seen as complete entities, and are recognised best as such (e.g. Davies & Christie, 1982; Shapiro & Penrod, 1986; Tanaka & Farah, 1993; Tanaka & Sengco, 1997). The basic feature-by-feature approach to face construction would therefore appear to be theoretically flawed.

There is a glimmer of hope, though, since 'holistic' systems are emerging that are based on the selection of complete faces rather than on facial parts (EvoFIT: Frowd et al., 2004; EigenFIT/EFIT-V: Gibson et al., 2003; ID: Tredoux et al., 2006). The basic procedure

with the holistic types involves repeated selection and breeding from arrays of complete faces. A Genetic Algorithm is used to provide further choices for selection, based on items previously selected, to allow a composite to be 'evolved'. At least for the EvoFIT system, there is evidence that the approach does produce more recognisable images than those from traditional feature methods (Frowd et al., 2006b, 2007c, 2010). While the holistic approach in general is assumed to engage a witness in the process of recognition, which tends to be automatic and fairly stable over time (e.g. Shepherd, 1983), the feature approach involves the more difficult task of face recall, and such information is forgotten more rapidly (e.g. Davies, 1983; Ellis et al., 1980). As such, being based on whole-face processing, one would expect holistic systems to show strong target familiarity effects. Were this to be demonstrated, but not for feature systems, this would provide evidence that recent systems behave more like the natural processes involved in face recognition: essentially, being a better interface to memory.

In the current study, participants were recruited to construct composites of familiar and unfamiliar targets using feature and holistic software systems. The resulting images were later evaluated by asking further participants to identify them. Complete composites were expected to be identified the best, internal to be better named than external features for composites of familiar targets, but internal and external composites of unfamiliar targets to be equally identifiable; also, that composites from the holistic system to be better identified overall than those from the feature system. We also tested the hypothesis that complete composites would show a whole-face advantage relative to combined naming of composites of internal and external features seen separately (e.g. Ellis et al., 1979, for photographs of faces; also see footnote [1]). The two stages required to carry out the study are presented below: (2.1) composite construction and (2.2) composite evaluation.

2 Stage 1: Composite construction

2.1 Design

The study required target familiarity to be manipulated at the construction stage: half of the participants would make a composite of a target they were familiar with, with the other half constructing an unfamiliar target. Note that in the subsequent evaluation stage, to be carried out when the composites had been constructed, all of those participants were required to have high target familiarity, to allow them to name the composites. These requirements were achieved using photographs of UK international level footballers. In the UK, there is a fairly strong gender split between people who are familiar with such identities, football fans, who are mostly male, and non-football fans, mostly female: this design thus allowed fairly easy recruitment of participants throughout.

It is perhaps worth mentioning that gender effects tend to be fairly small in face recognition studies (e.g. Ellis, Shepherd & Davies, 1980; Shapiro & Penrod, 1986); they have not been observed in the composite literature (e.g. Frowd et al., 2005b). Therefore, gender of participants is unlikely to impact upon the quality of the composites, or their naming, using this design. Further, it is also unlikely that footballer targets would adversely affect results (e.g. by involving participants who are unfamiliar or very familiar with footballers). This is based on evidence that similar performance has been measured for one system (EvoFIT) under similar testing conditions using different target identities—snooker players, footballers, staff working in a retail store, and university staff (Frowd et al., 2008b, 2010, under revision, in press): in all of these studies, mean composite naming ranged from 20 to 25% correct.

PRO-fit software was chosen for the current work. It is a feature system used by UK police, and is typical of composite systems of this type (Frowd et al., 2007e). It has been found to perform equivalently to the other UK feature system, E-FIT (Frowd et al., 2005a, 2005b). PRO-fit contains a large repertoire of individual features that have been cut from

photographs of faces of a given race, age and gender. In use, an eyewitness works with a police operative (or other such suitably-trained person) to select, size and position these features to build a likeness of the face. There is an artwork package available to remove or add hair, and to enhance the likeness by the addition of shading, lines, wrinkles, etc. Unique to PRO-fit is an additional software tool, PRO-warp, to enable the shape of features to be changed on demand.

EvoFIT was selected as the holistic system. A recent review of it, along with a general discussion of feature systems, may be found in Frowd, Bruce and Hancock (2009). EvoFIT was designed to produce identifiable images when deployed several days after a face had been seen, as typical in police investigations. Witnesses first select a suitable hairstyle and are then presented with screens of 18 faces. They first choose from facial 'shapes', whereby the shapes and position of facial features change, and then facial 'textures', the greyscale colourings of the features and the overall skin tone. The shape and texture that has the best likeness is identified and all selections are bred together using a Genetic Algorithm, to mix facial characteristics, and produce more faces for selection. When repeated a couple of times, the faces become more like each other, and more like the face in the witness's memory; ultimately, the item with the best likeness is saved as the composite.

Using EvoFIT in this way produces faces that are named 11% on average when the target delay is two days; it is normally 5% or less under the same conditions for a feature system (e.g. Frowd et al., 2007c). Two recent developments have improved the performance of EvoFIT and are worth noting here. The first applied a Gaussian (or 'blur') filter to the external features of faces (Frowd et al., 2008b, 2010). This type of image filtering has the aim of encouraging witnesses to base selections on internal features, the region proposed to be important for composite naming. The second development is a set of software tools (scales) that allow an evolved face to be enhanced, should it be noticeably inaccurate. It is

possible, for example, to change its apparent age, weight, masculinity, and other holistic properties (Frowd et al., 2006a). In a recent study (Frowd et al., 2010), composites constructed using this improved EvoFIT (from a two day memory of an unfamiliar face) had mean naming of 25% correct, and using a feature system, 5%.

Composites from feature systems also tend to be named fairly well, at around 20% correct, when the target delay is fairly short, from a few minutes to a few hours in duration (Brace et al., 2000; Bruce et al., 2002; Davies et al., 2000; Frowd et al., 2004, 2005b, 2007c, 2008a). However, as mentioned above, performance is normally very poor with this method of face production when the target delay is several days long (Frowd et al., 2005a, 2005c, 2007c, 2007e, 2010). In the current work, to avoid low naming levels for faces produced from the feature system, we chose a short delay: participants looked at a target and then constructed the face within a single experimental session.

Other aspects of the design mirrored real-world construction as far as possible in the laboratory. Our "witnesses" were given a version of the cognitive interview (CI) to help them recall details of the face (see Wells, Memon & Penrod, 2007, for a recent review of the CI). We involved components or *mnemonics* of this interview that are typical of face construction with real witnesses in the UK (Frowd et al., 2005b). In brief, these included *context reinstatement*, where participants were asked to think back to when the photograph was seen and to form a mental picture of the face; *free recall*, to describe the face in an unhindered format; and *cued recall*, to elicit further details. The study made use of both artwork enhancement and PRO-warp for all composites, including images produced by EvoFIT, to limit differences by system (as far as possible) and to promote optimal likenesses. The Experimenter was trained 'in house' and practiced extensively (and roughly equally) on both systems: she produced about 30 composites in total prior to starting the current study.

Each participant constructed a single composite and so the design for this stage was between-subjects for target familiarity (familiar / unfamiliar) and system (EvoFIT / PRO-fit).

2.1.1 Participants

Thirty-two students from the University of Central Lancashire were recruited on a voluntary basis. There were 14 males and 18 females, aged 18 to 26 years (M = 21.5 years, SD = 1.8 years). Half of these were chosen as they claimed to be a football fan: the other half as they expressed no interest in the sport.

2.1.2 Materials

Eight target faces of UK international-level footballers were located via an Internet search. Each image depicted a front view of the face in a near-neutral expression, without glasses, and with little or no stubble. They were of Joe Cole, Peter Crouch, Frank Lampard, Gary Neville, Alan Smith, Ole Gunnar Solskjaer and John Terry. Photographs were printed in colour on A4 plain paper, one per page, using a good quality printer to a size of 9cm (wide) x 13cm (high). Targets were reproduced four times for the construction phase (and put into four separate envelopes); each picture was given a randomized identifying code. The stimuli were prepared by the first author to allow the Experimenter, the third author, to be blind to them until all composites had been constructed.

2.1.3 Procedure

Opportunity sampling was used to recruit 16 participants who were familiar with football and 16 who were not. They were tested individually throughout and informed at the start that they would be constructing a composite with a system used by the UK police. Participants were randomly assigned, with equal sampling, to construct a face with either

PRO-fit or EvoFIT. The eight targets were constructed four times across the experiment in a 2 (target familiarity) x 2 (construction method) design to produce a total of 32 composites. The Experimenter worked with participants throughout, presenting stimuli, conducting the interview and operating the face-construction software. Participants worked at their own pace.

The following procedure allowed the Experimenter to present target photographs to participants but remain blind to their identity (to avoid her unwittingly influencing the quality of the composites, which is conceivable had she seen the targets). She presented participants with one of four envelopes containing the relevant set of targets – for the familiar or unfamiliar group, and for construction with EvoFIT or PRO-fit – and then turned her back on the participant, explaining that she must not see any of the targets, as for 'real' witnesses. She asked that a single photograph be removed. For the target unfamiliar group, if the face was familiar, they were told to return the photograph to the envelope and select another; this was repeated until the first unknown face was selected (this occurred on three occasions: the first photograph was known but the second one was not). Participants were given 60 seconds to memorize the face; they then reported the target code before placing it in a second 'used' envelope (i.e. non-replacement sampling was used). The familiar group did the opposite, looking at the first face that was familiar, and memorizing it for the same length of time (this time, all participants recognised the first face removed from the given envelope).

Participants were informed that a cognitive interview would be administered next, to recall details of the target face, and a composite would be constructed thereafter using either EvoFIT or PRO-fit (depending on the relevant assignment); also that interview procedure used would follow that of real witnesses.

An overview of the cognitive interview was provided. Participants were told that, in a few moments, they would be asked to both think back to when the target was seen and try to

visualize the face. They should then describe it in a free-recall format, providing as many details as possible in their own time. While this was being done, the Experimenter would not interrupt, but would take notes. When participants indicated that they were ready to start, a description of the face was collected using this procedure. Afterwards, it was mentioned that the description given for each feature would be repeated and participants should try to remember more information. To do this, the description from the first feature on the sheet was read aloud – face shape – and further recall was prompted. This was repeated for the remaining features in the following order: hair, brows, eyes, nose, mouth and ears.

For those assigned to construct using PRO-fit, a short summary of the system was provided. It was also explained that, as individual features of the face were cut from pictures, only a likeness of their target was possible, but an editing package could be used to improve the appearance of features by adding shading, facial hair, wrinkles etc; also, a further tool could change the shape of any selected feature. It was mentioned that such enhancements are normally done towards the end of the session once facial features had been selected.

The Experimenter explained that the first stage was to enter their description into PRO-fit, to locate about two dozen examples per feature. This would result in an 'initial' composite, a face with features that matched the description, that would be used as a basis to build the face. The description was entered and the initial composite presented. A short demonstration was then provided of how features could be selected and manipulated. Participants were given the choice of which feature to work on initially, which was usually hair. Examples were presented and participants asked to select the closest match. Once a feature had been selected, it was sized and positioned, and brightness and contrast levels adjusted, as necessary. The procedure was repeated for the remaining features. Participants were asked whether artistic enhancement was necessary and, if it was, the artwork program was started and appropriate changes carried out. Lastly, it was explained that small changes

could be made to the shapes of features using PRO-warp: parts of any feature could be stretched or squashed. If participants required this, PRO-warp was started and changes applied. The final face was saved to disk as the composite.

For participants assigned to EvoFIT, an overview of the basic selection and breeding procedure was provided. They were also told that the first stage was to select hair, then facial shapes and textures. Afterwards, the shape and texture closest to the target would be chosen, and faces bred together to produce more examples for selection. This selection and breeding procedure would be repeated twice more to allow the faces to increasingly resemble the target. EvoFIT was then started and, similar to the procedure for PRO-fit, participants chose an appropriate hairstyle from the set of alternatives. It was mentioned that the external parts of faces would appear 'blurred', to enable focus on the central part that is important for later recognition of their composite by another person. The first screen of 18 (random) shapes were presented, with blurred externals, and participants told that the faces changed by feature shape and feature placement. Four such screens were presented in total and participants were asked to select the six closest matches to their target. They were then presented with (random) textures and told that the (greyscale) colouring of features now changed. As for shape, participants similarly inspected four texture screens and selected six faces. They went on to look at 'combinations' of selected shape and texture over two screens (6 shapes x 6 textures = 36 faces = 2 screens of 18) to locate the overall 'best' likeness at this stage. All selected items were bred together, to combine characteristics, and the above selection and breeding procedure was repeated twice more; each time, the 'best' likeness was identified.

Participants were told that the blurring would now be removed, and their 'best' likeness enhanced using 'holistic' tools (scales). These tools would change the perceived age, weight and other overall properties of the face. Blurring was then removed and holistic tools started. The scales were presented in the following order: age, weight, attractiveness,

extroversion, health, honesty, masculinity and threatening. Additional scales, to adjust the colouring of individual features, were offered, and these applied as required. The resulting image was transferred to PRO-fit and the same procedure for artistically enhancing the face was used, including PRO-warp. The resultant face was saved to disk as the composite.

Composites took approximately an hour to construct.

2.2 Stage 2: Composite evaluation

2.2.1 Design

A different set of participants named the 32 composites constructed in the first stage of the experiment; they inspected complete composites, or those depicting either internal or external features. This resulted in a mixed-factorial design with system (EvoFIT / PRO-fit) and target familiarity (familiar / unfamiliar) as within-subjects factors, and composite-type (complete face / internal features / external features) as between-subjects. It was expected that (1) complete and part-face composites would be named better for familiar than unfamiliar targets, (2) composite-type and target familiarity would interact: the internal features would be better named than the external features when the target was familiar to the person constructing the face, but equivalent when unfamiliar, (3) complete composites would be named the best irrespective of familiarity and (4) system would interact with target familiarity: EvoFITs would be named better than PRO-fits for both complete and internal features; external features would be named equivalently as hair was selected in a similar way for both systems (see Procedure section, above, Stage 1); also, familiarity effects would be observed for EvoFIT but not for PRO-fit.

The number of participants required was chosen to permit small effect sizes to be observed. This was based on a G*Power analysis (Faul, Erdfelder, Lang & Buchner, 2007) with an effect size, $f = 0.2$; alpha, $\alpha = .05$; power, $1 - \beta = .8$; an equal number of participants

in the composite-type between-subjects factor; and a fairly-low correlation (between repeated measures), r = .35, which is typical for composites (Frowd et al., 2005a).

2.2.2 Materials

Two additional sets of composites were prepared in Adobe Photoshop, one each for internal and external features for the 32 composites. The internal features were produced using the Elliptical and Rectangular Marquee Tools to extract the central region of the face; the remaining image was taken as the external features. No changes were made to the original, complete composite. Examples can be seen in Figure 1. Composites were printed in greyscale using a good quality printer on A4 paper to approximately 9cm (wide) x 13cm (high). A set of eight target photographs taken from the construction stage were also used.

Figure 1 about here.

2.2.3 Participants

The participants who named the composites comprised of 43 male and 17 female volunteers, aged 19 to 30 years (M = 22.5 years, SD = 2.8 years). All professed to regularly watch international-level football in the UK, and were based in the Preston and Manchester areas, UK. None had participated in the face-construction stage.

2.2.4 Procedure

Participants were tested individually. They were told that they would be evaluating a set of facial composites of UK international level footballers by attempting to name them. It was mentioned that there were repeated identities in the set. Participants were randomly assigned, with equal sampling, to one of three booklets containing internal features, external

features or complete faces. Participants worked through their booklet at their own pace and offered names as requested. They were not encouraged to guess. After the composites had been seen, participants were similarly shown the target photographs and named those. The order of presentation of composites and targets was randomised for each person.

2.2.5 Results

All target photographs were correctly named by all participants and therefore familiarity with the target identities was very high. Each composite was correctly named by at least one person in either the whole or one of the part-face conditions. The best named composite from PRO-fit was of Ole Gunnar Solskjær, with a mean naming at 61.1%, and the best EvoFIT was of Joe Cole, at 94.4%. As can be seen in Table 1, complete composites were named overall by far the best, followed by composites of internal features and then external features. Composites of familiar targets were named much better than composites of unfamiliar targets; similarly, EvoFITs were named much better than PRO-fits.

Table 1 about here

The participant correct naming (summed accuracy) scores were analysed using a 2 (System: EvoFIT, PRO-fit) x 2 (Familiarity of target face at construction: unfamiliar, familiar) x 3 (Composite type: complete, internal features, external features) Mixed Factorial Analysis of Variance (ANOVA). This was significant for system, $F(1,51) = 81.0$, $p < .001$, $\eta_p^2 = .61$, as EvoFITs were named better than PRO-fits overall, and for familiarity, $F(1,51) = 78.3$, $p < .001$, $\eta_p^2 = .61$, as composites of familiar targets were named better than composites of unfamiliar targets. Composite-type was also significant, $F(2,51) = 114.8$, $p < .001$, $\eta_p^2 = .82$, and simple contrast of the ANOVA confirmed that complete composites were named

overall better than composites of internal features ($p < .001$), which in turn were better named than composites of external features ($p < .001$). There was, however, some evidence that these main effects were not consistent since the two-way interactions were significant, as presented below, although the three-way interaction was not, $F(2,51) = 3.1$, $p = .06$, $\eta_p^2 = .11$.

A simple-main effects analysis was conducted on these composite data to explore the two-way interactions. Firstly, familiarity x composite-type was significant, $F(2,51) = 9.7$, $p < .001$, $\eta_p^2 = .28$, as internal features were superior to external features when the target was familiar at construction ($p < .001$), but equivalent when unfamiliar ($p = .61$). Note, though, the consistent effect of these two factors: (a) complete composites were named significantly better than either of the part-face conditions irrespective of familiarity, and (b) the familiarity effect (better naming for composite of a familiar than an unfamiliar target) extends to all composite-types (complete, internal and external features). Secondly, system x composite-type was significant, $F(2,51) = 19.9$, $p < .001$, $\eta_p^2 = .44$, since EvoFITs were named better than PRO-fits for complete and internal feature composites ($p < .001$), but not for external features ($p = .80$); also, while internal features were superior to external features for EvoFIT ($p < .001$), these part-face conditions were equivalent for PRO-fit ($p = .21$). Thirdly, system x familiarity was significant, $F(1,51) = 5.9$, $p = .018$, $\eta_p^2 = .10$, due to floor effects. (This interaction is not of great interest here, indicating the presence of low values, as tends to occur in composite naming tasks—e.g. Frowd et al., 2008a.)

An ANOVA of the correct naming scores for the composites – a by-items analysis – revealed the same pattern of significant and non-significant main effects and interactions, except that the system x familiarity interaction was now not significant, $F(1,7) = 3.8$, $p = .10$,

therefore no longer detecting the previously observed floor effects in the data. See also footnote [2].

Next, we tested the hypothesis that complete composites would be named more successfully than combined naming of internal and external features seen separately, as was found by Ellis et al. (1979) using photographs of faces. We carried out separate analyses for PRO-fit and for EvoFIT systems using (by-item) correct naming scores averaged across conditions for complete, internal and external features—means shown in data rows three and four of Table 1. For PRO-fit, using the same procedure as described in footnote [1], mean probability of incorrect naming was 0.84 for internal and 0.87 for external composite-features. This yields a combined probability of 0.74, or 26.4% correct naming. This figure is slightly less than 29.2% for naming of complete PRO-fits, but not significant so using a one-tailed paired-samples t-test, $t(9) = 0.9$, $p = .202$. Repeating the analysis for EvoFIT, combined correct naming for internal and external features is 42.2%, which is significantly less than mean naming of complete faces, at 53.8%, $t(9) = 2.1$, $p = .031$. These results thus indicate that a whole-face advantage applies to EvoFIT but not to PRO-fit composites.

Only three incorrect names were given in total. This indicates that the level of guessing was very low overall, and that participants provided names that were almost exclusively correct. Due to low values, no inferential statistics were conducted on these data.

2.2.5.1 Correlational analysis.

Partial correlations were carried out on the correct naming data between complete and part-face conditions, to provide an indication of the relative importance of internal and

[2] We collected additional naming data from participants who did not inspect composites of repeated identities—as, repeating such items may elevate naming levels and potentially promote different results. To do this, an additional 60 football fans (people different to those in Section 2.2.3) named eight composites from one of the eight conditions (familiarity x system x composite-type), selected randomly with equal sampling. Overall naming was 22.3% correct for the additional data without repeats, which is obviously only slightly less than the above, 26.5%, with repeated-items. We also found the same pattern of effects by system and familiarity.

external features to naming of complete images. For this, mean scores were used from all 32 composites in each condition only, as smaller groupings (by system or by familiarity) are likely to be unreliable (as there then would be too few items). The correlation was medium sized, $r(29) = .54$, and significant between internal and complete with external as the partial ($p = .002$), but non-significant with internal as the partial ($p = .54$). This suggests that internal features are more important than external features to naming of complete composites.

2.3 Discussion

Facial composites are often an important part of a criminal investigation. A better understanding of how humans construct these images is not only theoretically interesting, but it could potentially help to improve their effectiveness for detecting those that commit crime, as recent research has demonstrated (e.g. Frowd et al., 2006a, 2006b, 2007d, 2008a, 2008b, 2010, under revision, in press). In the current work, we explored the influence of target familiarity at construction (familiar / unfamiliar), face-production system (PRO-fit / EvoFIT) and facial region (complete / internal / external features) on composite naming. It was found that while complete faces were the most identifiable representation, composites of internal features (the region including the brows, eyes, nose and mouth) were named better than those of external features (hair, face shape, ears) when the target was familiar to the person constructing the face, but these part-face regions were equivalent when the target was unfamiliar. It was also found that EvoFITs were better named than PRO-fits for the complete and the internal features region, and that naming was better for EvoFITs from their internal than external features, but the same for PRO-fits.

The current work would appear to be the first to observe a target familiarity effect for face construction from memory. When the target was familiar to the constructor, composites were named significantly better than when the target was unfamiliar. This familiarity effect

was consistent across whole and part-face preparations, as expected, as well as for EvoFIT and, surprisingly, for PRO-fit. The overall (by-items) effect size was by far the strongest for internal features, with a very large Cohen's d of 1.95, but was also large for complete, d = 0.84, and external features, d = 1.0. This suggests that while a better memory (a familiar face) assists in the construction of the external features, it is the internal features that benefit the most – a result that was supported by the correlational analysis. Such a suggestion is entirely sensible: there is an accumulating face perception literature which suggests that learning a face involves developing a robust and stable representation for the internal part of the face (e.g. Bonner, Burton & Bruce, 2003; Clutterbuck & Johnston, 2005; O'Donnell & Bruce, 2001). More generally, as mentioned earlier, it is apparent that such a representation yields a substantial benefit for recognising a face (e.g. Bruce, 1986; Ellis et al., 1979). With composite construction, the result suggests that being familiar with a person's face is valuable to the tasks required to the construct the face: in this case, to the main tasks of selecting either individual facial-features, for PRO-fit, or complete faces, for EvoFIT.

The familiarity x composite-type interaction revealed significantly better naming for internal than external features for composites of familiar targets, and with a very large effect size, d = 1.96, but equivalent naming between the internal and external features from unfamiliar targets. This pattern of effects was also expected and suggests that composites were constructed more accurately in terms of their internal features when the target was familiar, an *internal features advantage*, but equally when unfamiliar. This result is the same as that found with photographs of faces (e.g. Campbell et al., 1999; Ellis et al., 1979; Endo et al., 1984). The data also revealed that, irrespective of target familiarity, an intact face was named significantly better than either of the part-face preparations. While such a result may appear obvious – i.e. better performance as more information is shown in the intact condition – other composite evaluating methods have reported otherwise. Frowd et al. (2007b) asked

participants to match composites to target photographs, and found equivalent matching scores between complete and external features regions, with both of these types being matched significantly better than for internal features. In their case, the result is likely to have occurred due to the task used (matching) and the people who evaluated the composites: essentially, their stimuli were largely unfamiliar, and so participants were likely to have focussed on external features. In our case, a naming task was used, engaging participants in familiar face processes to provide an indication of the amount of *identifying* information therein. For the external features, this information was either less useful for naming than the internal features (construction of a familiar target) or the same (unfamiliar). Overall, the result here is the same as that found for non-composites, or photographs of faces, with superiority for the complete face (e.g. Ellis et al., 1979; Endo et al., 1984).

Significant differences were observed by composite system. EvoFITs were better named than PRO-fits for intact composites and those of internal features, and both with very large effect sizes ($d = 1.55$ and 1.62, respectively). One of the aims of the EvoFIT designers was to produce a system with a good interface to human memory (Frowd et al., 2004). They did this by presenting arrays of whole faces for users to repeatedly select, with the system presenting alternatives based on previous selections. This basic approach does produce a somewhat more identifiable composite than that produced by selecting individual features (e.g. Frowd et al., 2007c). Recent developments to EvoFIT have increased naming levels further; these include external-features blurring and holistic face transformations (Frowd et al., 2010), procedures that generally promote a more recognisable set of internal features.

The system x composite-type interaction did, as predicted, reveal that external features were equally identifiable between the holistic (EvoFIT) and feature (PRO-fit) systems. The external features comprise the hair, face shape, ears and neck. Of these, hair is arguably the most important for face recognition (e.g. Ellis et al., 1980; Wright & Sladden, 2003). It is

also the feature that is selected at the start of the construction process for both types of system: both systems have about 200 alternatives available for a given age range. As such, the process used to select the hair is equivalent and therefore the production of similar quality external features (particularly hair) is reasonable.

We did not predict a familiarity effect for PRO-fit: naming by composite-type was not expected to increase when targets were familiar compared to unfamiliar, unlike EvoFIT. This was based mainly on Davies et al. (2000). In their study, a familiarity effect was found for a feature system (E-FIT), but only when both the target was familiar to the person constructing the face and a photograph was available for reference; a familiarity effect was absent when constructing from memory. There are similarities between our study and theirs: the same number of targets, a short target delay and naming using a within-subjects design. The difference may simply be one of experimental power: their work recruited 12 participants in the naming phase, which is likely to detect a small-to-medium effect sizes (based on a similar G*Power analysis as above: a 2x2 repeated-measures design and with 12 participants, d = 0.41). Ours involved 60, in order to detect small *d*s, which turns out to be the size of the familiarity effect for complete PRO-fits here, d = 0.31. Overall, face construction from memory by the selection of individual features appears to be only slightly sensitive to target familiarity, even in fairly favourable conditions (a short target delay). It also underlines the ineffectiveness of the general approach for accessing memory. The notion fits well with Frowd et al. (2007b) who reported weak familiarity effects using a feature system tested similarly.

The study chose a short interval between participants inspecting a target photograph and face construction. In police work, construction very occasionally occurs on the same day in which a crime occurred, but it is normally one, two or more days later. A short delay was chosen here to promote fairly good naming levels for the feature system (Frowd et al., 2004,

2005b, 2007b, 2008a), thus creating favourable conditions to observe a target familiarity effect. The approach was successful, with a (small) effect found for the complete face. When the delay is much longer, however, likenesses from feature systems tend to be very poor (e.g. Frowd et al., 2005a). While unfamiliar face recognition is known to remain fairly stable for a week or more (e.g. Shepherd, 1983), the ability to recall information about faces decays rapidly (e.g. Davies, 1983; Ellis et al., 1980). The impact is quite marked for building a traditional composite: we simply cannot access features from memory, or place them on the face in order to produce a recognisable representation; and, these tasks become increasingly more difficult with time. Given our promising results from the holistic EvoFIT, however, familiarity effects would be expected under more forensically-relevant intervals.

The work was able to explore another theoretical issue, one concerning whole-face processing. Research has established that seeing a photograph of a complete face promotes better naming than combining naming of photographs of internal features and external features seen separately (e.g. Ellis et al., 1979, footnote [1]). Using mean correct naming scores for complete, internal and external features, we were able to demonstrate that a whole-face benefit extends to one type of composite technology, EvoFIT, a result not previously reported in the literature. In the light of superiority of EvoFIT over PRO-fit, this perhaps highlights the important role played by an intact image for recognition of EvoFITs over and above recognition of constituent regions, and with a large effect size (Cohen's $d = 0.75$). This test does provide a new method for assessing system effectiveness: composites that demonstrate a whole-face benefit are likely to advantageous relative to those that do not. System designers could usefully make improvements, in particular to feature types, with this metric in mind.

We were also able to show that the internal region of the face is the most important when recognising a composite, but also that the external features do contribute to some

extent. Thus, software system developers could usefully work on techniques to promote better construction of the internal features, as has been done for EvoFIT using external feature blurring and holistic tools (Frowd et al., 2010). Indeed, more recent research (Frowd et al., in press) has found that masking external features entirely during face construction promotes even more identifiable images. Such an approach could also lead to improvements to traditional feature and sketch systems—and, initial work confirms such an advantage. That said, sensible techniques for improving interfaces to face memory have not always been successful (e.g. Frowd et al., 2005a, 2007c, 2008b), but this is perhaps partly due to the difficult nature of the problem: witnesses are required to produce recognisable faces from their memory. However, designers should be aware that substantial benefits to performance do not necessary require changes to the system itself. For example, both feature and holistic composites are rendered considerably more identifiable by asking constructors to focus on whole-face properties of a face after having described it in detail (e.g. Frowd et al., 2005c, 2008a, under revision). In addition, a person's ability to correctly name an identity improves substantially by seeing a (finished) composite being progressively caricatured, by exaggerating distinctive shape information in the face (Frowd et al., 2007d).

Ultimately, such efforts promote better identification of internal relative to external features, thereby increasing the effectiveness of the whole-face representation (unlike what was observed here for PRO-fit, with internal and external features named equivalently overall). It is apparent that EvoFIT can outperform a typical UK feature system, not only when the delay to construction is fairly long, as indicated in other work (e.g. Frowd et al., 2007c, 2008b, 2010, under revision, in press), but also when very short. It will be interesting to see whether such an advantage extends to other types of holistic system. In some criminal investigations, however, witnesses are familiar with an offender's face (although do not know the correct name of the offender). Our work suggests that, in such a circumstance, face

construction using a feature system is only likely to be slightly better than for construction of a face seen only once; much better performance, however, would be expected from the holistic EvoFIT (and potentially from other holistic systems).

In summary, we were interested in how target familiarity and composite system would interact with each other and how these factors would impact on the recognition of different regions of the face. The work suggested that, for a familiar target, internal features were constructed very much better than external features, but there was little difference between these regions when the target was unfamiliar. Complete composites were of best quality overall. The data nicely fit the pattern found using photographs rather than composites of faces. Also, EvoFIT emerged as a better interface to our memory, by promoting more identifiable images for the complete and internal features region, and by producing an overall image that was recognised better than its constituent parts combined (unlike the feature composites); the external features were equivalent across systems presumably due to similarities in the method used to select hair. The work highlights the importance of developing systems that produce composites with identifiable internal features.

2.4 References

Bonner, L., Burton, A.M., & Bruce, V. (2003). Getting to know you: how to learn new faces. *Visual Cognition*, 10, 527-536.

Brace, N., Pike, G., & Kemp, R. (2000). Investigating E-FIT using famous faces. In A. Czerederecka, T. Jaskiewicz-Obydzinska & J. Wojcikiewicz (Eds.) *Forensic Psychology and Law* (pp. 272-276). Krakow: Institute of Forensic Research Publishers.

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

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

Bruce, V., Ness, H., Hancock, P.J.B, Newman, C., & Rarity, J. (2002). Four heads are better than one. Combining face composites yields improvements in face likeness. *Journal of Applied Psychology*, 87, 894-902.

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

Clutterbuck, R. & Johnston, R.A. (2002). Exploring Levels of Face Familiarity by Using an Indirect Face-Matching Measure. *Perception*, 31, 985-994.

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

Cohen, M.E., & Nodine, C.F. (1978). Memory processes in facial recognition and recall. *Bulletin of the Psychonomic Society*, 12, 317-319.

Davies, G.M. (1983). Forensic face recall: the role of visual and verbal information. In S.M.A. Lloyd-Bostock and B.R. Clifford (Eds.). *Evaluating witness evidence* (pp. 103-123). Chichester: Wiley.

Davies, G.M., & Christie, D. (1982). Face recall: an examination of some factors limiting composite production accuracy. *Journal of Applied Psychology*, 67, 103-109.

Davies, G.M., Shepherd, J., & Ellis, H.D. (1977). Similarity effects in face recognition. *American Journal of Psychology*, 92, 507-523.

Davies, G.M., Shepherd, J., & Ellis, H.D. (1978). Remembering faces: acknowledging our limitations, *Journal of Forensic Science*, 18, 19-24.

De Haan, E.H.F., & Hay, D. (1986). The matching of famous and unknown faces, given either the internal or external features:Astudy on patients with unilateral brain lesions. In H.D. Ellis, M.A. Jeeves, F. Newcombe, & A. Young (Eds), *Aspects of face processing* (pp. 302–309). Dordrecht: Martinus Nijhoff.

Ellis, H.D., Davies, G.M., & Shepherd, J.W. (1978). A critical examination of the photofit system for recalling faces, *Ergonomics*, 21, 297-307.

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

Ellis, H.D., Shepherd, J.W., & Davies, G.M. (1980). The deterioration of verbal descriptions of faces over different delay intervals. *Journal of Police Science and Administration*, 8, 101-106.

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, 23-34.

Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioural, and biomedical Sciences. *Behavior Research Methods*, 39, 175-191.

Frowd, C.D., Bruce, V., Gannon, C., Robinson, M., Tredoux, C., Park., J., McIntyre, A., & Hancock, P.J.B. (2007a). Evolving the face of a criminal: how to search a face space more effectively. In A. Stoica, T. Arslan, D.Howard, T. Kim and A. El-Rayis (Eds.) 2007 ECSIS Symposium on Bio-inspired, Learning, and Intelligent Systems for Security, (pp. 3-10). NJ: CPS. (Edinburgh).

Frowd, C.D., Bruce, V., & Hancock, P.J.B. (2009). Evolving facial composite systems. *Forensic Update*, 98, 25-32.

Frowd, C.D., Bruce, V., McIntyre, A. & Hancock, P.J.B. (2007b). The relative importance of external and internal features of facial composites. *British Journal of Psychology*, 98, 61-77.

Frowd, C.D., Bruce, V., McIntyre, A., Ross, D., Fields, S., Plenderleith, Y., & Hancock, P.J.B. (2006a). Implementing holistic dimensions for a facial composite system. *Journal of Multimedia*, 1, 42-51.

Frowd, C.D., Bruce, V., Ness, H., Bowie, L., Thomson-Bogner, C., Paterson, J., McIntyre, A., & Hancock, P.J.B. (2007c). Parallel approaches to composite production. *Ergonomics*, 50, 562-585.

Frowd, C.D., Bruce, V., Plenderleith, Y., & Hancock, P.J.B. (2006b). Improving target identification using pairs of composite faces constructed by the same person. *IEE Conference on Crime and Security* (pp. 386-395). London:IET.

Frowd, C.D., Bruce, V., Ross, D., McIntyre, A., & Hancock, P.J.B. (2007d). An application of caricature: how to improve the recognition of facial composites. *Visual Cognition*, 15, 1-31.

Frowd, C.D., Bruce, V., Smith, A., & Hancock, P.J.B. (2008a). Improving the quality of facial composites using a holistic cognitive interview. *Journal of Experimental Psychology: Applied*, 14, 276-287.

Frowd, C.D., Carson, D., Ness, H., McQuiston, D., Richardson, J., Baldwin, H., & Hancock, P.J.B. (2005a). Contemporary Composite Techniques: the impact of a forensically-relevant target delay. *Legal & Criminological Psychology*, 10, 63-81.

Frowd, C.D., Carson, D., Ness, H., Richardson, J., Morrison, L., McLanaghan, S., & Hancock, P.J.B. (2005b). A forensically valid comparison of facial composite systems. *Psychology, Crime & Law*, 11, 33-52.

Frowd, C.D., Hancock, P.J.B., & Carson, D. (2004). EvoFIT: A holistic, evolutionary facial imaging technique for creating composites. *ACM Transactions on Applied Psychology (TAP)*, 1, 1-21.

Frowd, C.D., McQuiston-Surrett, D., Anandaciva, S., Ireland, C.E., & Hancock, P.J.B. (2007e). An evaluation of US systems for facial composite production. *Ergonomics*, 50, 1987–1998.

Frowd, C.D., McQuiston-Surrett, D., Kirkland, I., & Hancock, P.J.B. (2005c). The process of facial composite production. In A. Czerederecka, T. Jaskiewicz-Obydzinska, R. Roesch & J. Wojcikiewicz (Eds.) *Forensic Psychology and Law* (pp. 140-152). Krakow: Institute of Forensic Research Publishers.

Frowd, C.D., Nelson, L., Skelton F.C., Noyce, R., Heard, P., Henry, J., Morgan, D., Fields, S., McIntyre, A., & Hancock, P.J.B. (Under revision). Interviewing techniques for Darwinian facial composite systems. *Applied Cognitive Psychology*.

Frowd, C.D., Park., J., McIntyre, A., Bruce, V., Pitchford, M., Fields, S., Kenirons, M. & Hancock, P.J.B. (2008b). Effecting an improvement to the fitness function. How to evolve a more identifiable face. In A. Stoica, T. Arslan, D. Howard, T. Higuchi, and

A. El-Rayis (Eds.) 2*008 ECSIS Symposium on Bio-inspired, Learning, and Intelligent Systems for Security*, (pp. 3-10). NJ: CPS. (Edinburgh).

Frowd, C.D., Pitchford, M., Bruce, V., Jackson, S., Hepton, G., Greenall, M., McIntyre, A., & Hancock, P.J.B. (2010). The psychology of face construction: giving evolution a helping hand. *Applied Cognitive Psychology*. DOI: 10.1002/acp.1662.

Frowd, C.D., Skelton F., Atherton, C., Pitchford, M., Hepton, G., Holden, L., McIntyre, A., & Hancock, P.J.B. (in press). Recovering faces from memory: the distracting influence of external facial features. *Journal of Experimental Psychology: Applied*.

Ge, L., Anzures, G., Wang, Z., Kelly, D.J., Pascalis, O., Quinn, P.C., Slater, A.M., Yang, Z., & Lee, K. (2008). An inner face advantage in children's recognition of familiar peers. *Journal of Experimental Child Psychology*, 101, 124–136.

Gibson., S.J., Solomon, C.J., & Pallares-Bejarano, A. (2003). Synthesis of photographic quality facial composites using evolutionary algorithms. In R. Harvey and J.A. Bangham (Eds.) *Proceedings of the British Machine Vision Conference* (pp. 221-230).

Kemp, R., Towell, N., & Pike, G. (1997). When seeing should not be believing: photographs, credit cards and fraud. *Applied Cognitive Psychology*, 21, 297-307.

Koehn, C.E., & Fisher R.P. (1997). Constructing facial composites with the Mac-a-Mug Pro system. *Psychology, Crime & Law*, 3, 215-224.

Lander, K., Bruce, V., & Hill, H. (2001). Evaluating the effectiveness of pixelation and blurring on masking the identity of familiar faces. *Applied Cognitive Psychology*, 15, 101-116.

McQuiston-Surrett, D., Topp, L. D., & Malpass, R. S. (2006). Use of facial composite systems in U.S. law enforcement agencies. *Psychology, Crime and Law*, 12, 505–517.

O'Donnell, C., & Bruce, V. (2001). Familiarisation with faces selectively enhances sensitivity to changes made to the eyes. *Perception*, 30, 755-764.

Shapiro, P. N., & Penrod, S.D. (1986). Meta-analysis of facial identification rates. *Psychological Bulletin*, 100, 139-156.

Shepherd, J.W. (1983). Identification after long delays. In S.M.A. Lloyd-Bostock, and B.R. Clifford (eds.), *Evaluating witness evidence* (pp. 173–187). Chichester: Wiley.

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

Tanaka, J.W., & Sengco, J.A. (1997). Features and their configuration in face recognition. *Memory & Cognition*, 25, 583-592.

Tredoux, C.G., Nunez, D.T., Oxtoby, O., & Prag, B. (2006). An evaluation of ID: an eigenface based construction system. *South African Computer Journal*, 37, 1-9.

Want, S.C., Pascalis, O., Coleman, M., & Blades, M. (2003). Recognizing people from the inner or outer parts of their faces: Developmental data concerning 'unfamiliar' faces. *British Journal of Developmental Psychology*, 21, 125-135.

Wells, G., Memon, A., & Penrod, S.D. (2007). Eyewitness evidence: improving its probative value. *Psychological sciences in the public interest*, 7, 45-75.

Wright, D.B., & Sladden, B. (2003). An own gender bias and the importance of hair in face recognition. *Acta Psychologica*, 114, 101-114.

Young, A.W. (1984). Right cerebral hemisphere superiority for recognizing the internal and external features of famous faces. *British Journal of Psychology*, 75, 161-169.

Young, A.W., Hay, D.C., & Ellis, A.W. (1985a). The faces that launched a thousand slips: everyday difficulties in recognising people. *The British Journal of Psychology*, 76, 495-523.

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

List of figure and table titles

Table 1. Percentage correct naming by composite-type for target familiarity and composite system. Values in parentheses are (by-item) standard deviations.

Figure 1. Example composites of the footballer, Ole Gunnar Solskjaer, here a familiar target to the two people who constructed his face. On the top row are composites from EvoFIT, bottom from PRO-fit. Complete composites are on the left, internal features are in the centre and external features to the right.

List of figures and tables

Table 1.

	Composite type			
	Complete	Internal	External	Mean
Target				
Unfamiliar	35.1 (21.3)	12.5 (13.4)	= 9.7 (7.7)	19.1 (18.8)
Familiar	47.9 (26.6)	37.2 (20.9)	16.7 (10.9)	33.9 (24.0)
System				
PRO-fit	29.2 (17.4)	16.0 (15.2)	= 12.8 (11.8)	19.3 (16.3)
EvoFIT	53.8 (25.0)	33.7 (23.3)	= 13.5 (8.1)	33.7 (26.0)
Mean	41.5 (24.6)	24.8 (21.3)	13.2 (10.0)	26.5 (22.7)

Note. All differences between adjacent conditions (by column and by row) for target, system and composite-type are significant ($p < .05$ or less) except for the three contrasts indicated by an equals symbol (=) which do not differ significantly. This includes main effects (column and row means) and the two significant interactions (the 12 means in the centre of the table) between composite-type and target, and between composite-type and system.

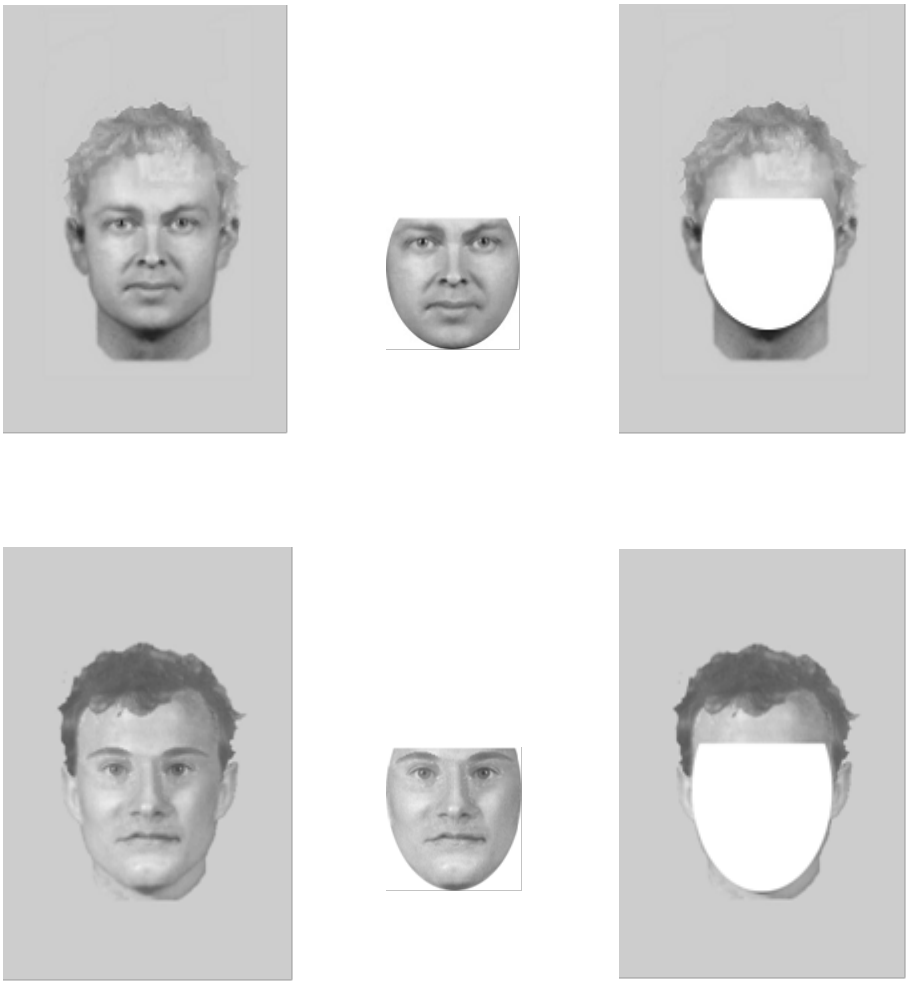

Figure 1.

Acknowledgement

The authors would like to thank Farah Hanif, Chanda Tanvir and Mahwish Kiran for assisting with data collection in the project.