

*promoting access to White Rose research papers*



**Universities of Leeds, Sheffield and York**  
**<http://eprints.whiterose.ac.uk/>**

---

This is an author produced version of a paper published in **Science and Justice**.

White Rose Research Online URL for this paper:

<http://eprints.whiterose.ac.uk/78320/>

---

**Paper:**

Frowd, CD, Skelton, F, Hepton, G, Holden, L, Minahil, S, Pitchford, M, McIntyre, A, Brown, C and Hancock, PJB (2013) *Whole-face procedures for recovering facial images from memory*. *Science and Justice*, 53 (2). 89 - 97.

<http://dx.doi.org/10.1016/j.scijus.2012.12.004>

---

## Whole-face procedures for recovering facial images from memory

Charlie Frowd (1\*)

Faye Skelton (1)

Gemma Hepton (1)

Laura Holden (1)

Simra Minahil (1)

Melanie Pitchford (2)

Alex McIntyre (3)

Charity Brown (4)

Peter J.B. Hancock (3)

(1) School of Psychology, University of Central Lancashire, Preston PR1 2HE

(2) Department of Psychology, Lancaster University, LA1 4YF

(3) Psychology, School of Natural Sciences, University of Stirling, Stirling FK9 4LA

(3) Psychology, School of Natural Sciences, University of Stirling, Stirling FK9 4LA

(4) Institute of Psychological Sciences, University of Leeds, Leeds LS2 9JT

\*Corresponding author: Charlie Frowd, University of Winchester, Sparkford Road, Winchester, SO22 4NR, UK. [Charlie.Frowd@winchester.ac.uk](mailto:Charlie.Frowd@winchester.ac.uk). Phone: +44 1962 824943.

Running head: Optimising identification of facial memory

Keywords: facial composite, holistic face-processing, interview, witness, face recognition.

Paper accepted for publication in the journal Science and Justice.

NOTICE: this is the author's version of a work that was accepted for publication in Science and Justice. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published as: Frowd, C.D., Skelton, F., Hepton, G., Holden, L., Minahil, S., Pitchford, M., McIntyre, A., Brown, C., Hancock, P.J.B. (2013). Whole-face procedures for recovering facial images from memory. *Science and Justice*, 53, 89-97. DOI:10.1016/j.scijus.2012.12.004.

Research has indicated that traditional methods for accessing facial memories usually yield unidentifiable images. Recent research, however, has made important improvements in this area to the witness interview, method used for constructing the face and recognition of finished composites. Here, we investigated whether three of these improvements would produce even-more recognisable images when used in conjunction with each other. The techniques are holistic in nature: they involve processes which operate on an entire face. Forty participants first inspected an unfamiliar target face. Nominally 24 hours later, they were interviewed using a standard type of cognitive interview (CI) to recall the appearance of the target, or an enhanced 'holistic' interview where the CI was followed by procedures for focussing on the target's character. Participants then constructed a composite using EvoFIT, a recognition-type system that requires repeatedly selecting items from face arrays, with 'breeding', to 'evolve' a composite. They either saw faces in these arrays with blurred external features, or an enhanced method where these faces were presented with masked external features. Then, further participants attempted to name the composites, first by looking at the face front-on, the normal method, and then for a second time by looking at the face side-on, which research demonstrates facilitates recognition. All techniques improved correct naming on their own, but together promoted highly-recognisable composites with mean naming at 74% correct. The implication is that these techniques, if used together by practitioners should substantially increase the detection of suspects using this forensic method of person identification.

## 1.0 INTRODUCTION

It is common practice for police to invite eyewitnesses to help them solve crime. These witnesses and victims can be a valuable resource at the initial stage of an investigation when there is no apparent evidence available to locate an offender (e.g. from CCTV footage or DNA analysis). If a witness or victim has clearly seen either the offender's face or another person who is unknown but nevertheless may potentially have valuable evidence, he or she may be able to describe the appearance of the face and construct a facial composite. Composite images produced in this way can be circulated within a police force, or more generally in the media, with the aim that someone will name the face and thereby generate new lines of enquiry. Potential suspects can be identified or eliminated and ideally the offender brought to justice.

Facial composites have been in regular use for the principal purpose of detecting offenders for about 40 years (for recent reviews, see Frowd, Bruce & Hancock, 2008; Frowd, Bruce & Hancock, 2009; Frowd, 2012). There are three basic methods currently available to police to construct the face. First, there are feature-based computer systems which involve the selection of individual facial features (eyes, nose, mouth, etc.). Witnesses describe the face they have seen and this description is used to locate matching features within the composite system. They view these choices and are asked to select the best matches, which are then sized and positioned on the face with the aim of producing the best likeness. Software systems for constructing composites include E-FIT and PRO-fit in the UK and Europe, and FACES and Identikit 2000 in the US. Second, there are sketch artists who produce composites using pencils or crayons using a similar feature-by-feature method.

Research has indicated that feature-based methods produce fairly identifiable composites when the delay is fairly short between a person (observer) seeing a target face and constructing a composite (e.g. Brace, Pike, Kemp, 2000; Bruce, Ness, Hancock, Newman & Rarity, 2002; Davies, Van der Willik & Morrison, 2000; Frowd, Carson, Ness, Richardson, et al., 2005; Frowd, Bruce, Ness et al., 2007; Frowd, Skelton, Butt, Hassan, & Fields, 2011). When the retention interval is up to a few hours in duration, for example, composites are named with a mean accuracy of around 10 to 20%.

### 1. Improving the effectiveness of composites

Considerable research effort has been spent attempting to improve the performance of composites, and this has been met with success in three areas—

#### 1.1.1 Enhancing the witness interview

The first area of success relates to the interview. It is normal for witnesses (including victims) to initially undergo a cognitive interview (CI) to help them recall details of the target's (offender's) face; this allows appropriate facial features to be located within the composite system for witnesses to select from. The CI has been extensively developed (e.g. Wells, Memon & Penrod, 2007) for use in forensic and other applications and contains a set of memory-facilitating techniques or *mnemonics* to facilitate information recall. For face construction, mnemonics normally include a request to describe the face (a) in an uninterrupted (free-recall) format, (b) in as much detail as possible but (c) without guessing (for more details, see Frowd, Carson, Ness, McQuiston et al., 2005).

The first improvement involves administering an additional 'holistic' mnemonic to the CI (e.g. Frowd, Bruce, Smith & Hancock, 2008). This aims to improve witnesses' face recognition and consequently allow them to select facial features with greater accuracy. The technique is based on the established 'deep' over 'shallow' processing advantage for *encoding* a target: face recognition is facilitated when a target face is remembered by attributing overall (holistic) judgements (e.g. attractiveness, masculinity, distinctiveness) rather than by physical attribution such as size of mouth or length or nose (e.g. Shapiro & Penrod, 1986). The benefit of holistic over physical-feature attribution has similarly been found when carried out immediately prior to face recognition (e.g. Berman & Cutler, 1989).

With face construction, the holistic element of the interview requires witnesses to (i) think freely about the character of this face (silently to themselves for 1 minute) and (ii) make seven overall (holistic) judgements about it (e.g. intelligence, friendliness and distinctiveness). In Frowd, Bruce, Smith and Hancock (2008), composites constructed using the PRO-fit feature system after this 'holistic' cognitive interview (H-CI) were correctly named at 41.2% compared with 8.6% for composites constructed after the traditional CI.

### 1.1.2 Facilitating recognition (naming) of composites

The second area of success relates to the naming stage—that is, when another person attempts to recognise a composite, as would be carried out by a police officer or member of the public in a police appeal for information. The main problem is that composites constructed from memory contain error, with inaccuracies present in the composites' shape and placement of facial features: for this reason, composites do not evoke perfect recognition. Despite developing methods to reduce error, such as the H-CI, there are often marked differences between a person's appearance and a constructed likeness of him or her, the outcome of which is to limit recognition (e.g. Frowd, Skelton, Atherton et al., 2012). Two techniques have successfully facilitated recognition—

One technique is to artificially increase the level of distinctiveness by exaggerating distinctive features and configural relations (distances between features) through caricature. Positive caricature facilitates recognition of line drawings of faces (e.g. Benson & Perrett,

1994) and briefly-presented photographs of faces (e.g. Lee & Perrett, 1997). Observing a composite with a range of caricature levels, from negative through to positive, provides one frame (one level of exaggeration) that is a better probe to memory than the veridical composite, facilitating recognition. In Frowd, Bruce, Ross et al. (2007), multi-frame caricature improved correct naming of composites by about 50%.

The other technique emerged from a curious line of research involving transforming faces by linear stretch, an affine transformation. When applied to photographs, by doubling the height, Hole, George, Eaves and Rasek (2002) found that recognition accuracy of familiar (celebrity) faces did not decrease relative to unaltered images (although reaction times were slower in one of their four experiments). However, as described in Astbury (2011), doubling height or width of familiar-face (celebrity) composites substantially improved an observer's ability to correctly name the face. The technique is impractical for publishing witness composites in the media, as stretched images look silly for the serious application of identifying offenders. It was found, however, that the technique was also effective when participants were asked to look at the face sideways, to give the *appearance* of a thin face (there is also a smaller effect of perspective, discussed in section 3.2). Relative to looking at the face front-on, this 'perceptual' stretch about doubled correct naming, from 17% to 36%.

Hole et al. (2002) argued that, for photographs of faces, an observer's cognitive system may be normalizing a stretched face prior to recognition, to fit with how a face usually appears, or that the memory itself may be altered to accommodate the transformation, a 'deformable template' theory. More-recent research by the first author (unpublished) has found that the null effect of stretching extends to unfamiliar-face recognition, which argues for the former case: normalisation of an incoming face not deformation of memory. For composites, be it physical or perceptual stretch, the advantage appears to be based on reducing the perception of error: in the process of normalising the face for comparison with existing (un-stretched) memories, some inaccuracies in the composite are reduced. In particular, information in the horizontal plane is condensed with a side-on perspective, making it ambiguous: this ambiguity can be resolved in an observer's cognitive system in effect by 'filling in the details', facilitating recognition. In a related study, Frowd, Bruce, Gannon et al (2007) simplified information in a normal photographic-type composite, by converting it into a sketch-like image. For the same reason as being suggested here, this image transform facilitated naming of composites that were poorly recognised (i.e. images with inaccurate facial texture). The most likely candidates underlying the benefit of (perceptual) stretch are the individual features, whose likeness is known to positively correlate with composite naming (Frowd, Bruce, Smith and Hancock, 2008), and the region around eyes, in particular the eye spacing, known to be important for recognition (e.g. Diamond & Carey, 1996; Leder & Bruce, 2000; McIntyre, Frowd, Bruce & Hancock, 2010). So, inaccuracies present in both individual features and spacing in the eye region may be reduced with side-on viewing, augmenting recognition.

### 1.1.3 Problems with face construction

Unfortunately, in spite of these improvements, when the retention interval is upwards of 24 hours, the normal situation in police investigations, most people struggle to recall details of an offender's face and select individual features. It is under these circumstances that identifications arising from feature-based composites tend to be infrequent in police investigations (Frowd, Hancock et al., 2011), an observation supported by laboratory research (e.g. Frowd, Carson, Ness, McQuiston et al., 2005; Frowd, McQuiston-Surrett et al., 2007; Frowd et al., 2010; Koehn & Fisher, 1997; Kovera, Penrod, Pappas & Thill, 1997). In fact, many practitioners find it extremely hard to create a composite with witnesses that have poor face recall; also, under these circumstances, UK police guidelines advise against production of a feature-based composite (ACPO, 2009).

While the difficulty with face recall is by no means new (e.g. Davies, Shepherd & Ellis, 1978), there is an important theoretical issue: construction usually involves an unfamiliar face, a person that the witness has not knowingly seen previously, but recognition of the composite is carried out by a person familiar with the offender (Frowd, Bruce, McIntyre, & Hancock, 2007). It is known (e.g. Ellis, Shepherd & Davies, 1979) that while familiar-face recognition tends to involve internal facial features (the region including eyes, brows, nose and mouth), recognition of unfamiliar faces is biased strongly by external facial features (hair, face shape and ears). For face construction, the consequence is that the internal region tends not to be created accurately, which is arguably the reason why traditional composites are not often recognised correctly (e.g. Frowd, Bruce, McIntyre, & Hancock, 2007; Frowd & Hepton, 2009; Frowd, Skelton, Butt et al., 2011).

### 1.2 Holistic (recognition-based) composite systems

Alternative methods have been sought to access facial memory: the resulting recognition-based (or 'holistic') systems form the third area of success emerging from research. These systems do not have a database of individual facial features; instead, they create a face space (model) constructed from faces in their entirety (for a discussion on different types of face spaces, see Valentine, 1991). These models are built using the statistical technique, Principal Components Analysis (for a review, see Frowd, Hancock & Carson, 2004). The face space is holistic in nature, with each parameter representing an overall aspect of the face (e.g. Hancock, Bruce & Burton, 1997); one parameter might code for face width, for example, while others attend to age and complexion. When these holistic parameters are given random values, the result is realised as a synthetic yet plausible-looking face—it is, in fact, a point in face space. This can be repeated to produce other plausible-looking unfamiliar faces—other points in face space. The goal of the method is to search the space to locate a set of parameters which together represent the appearance of the offender's face, as seen by the witness. Estimating these parameter values accurately, based on feedback from a witness (user) during face construction, is the challenge.



The approach is generally based on the idea that faces are recognised as whole entities rather than by their isolated facial features (e.g. Davies & Christie, 1982; Tanaka & Farah, 1993). Also, while face recall decays quickly (e.g. Ellis, Shepherd & Davies, 1980), which is arguably the Achilles' heel of the feature systems (e.g. Frowd, Carson, Ness, McQuiston et al., 2005), face recognition remains largely intact after a forensically-relevant interval of several weeks (e.g. Shepherd, 1983). Accordingly, these systems principally use a face-recognition task: witnesses are asked to select the overall closest matches from an array of randomly-generated faces of appropriate age, gender and ethnicity. These choices are 'bred' together, usually using an Evolutionary Algorithm to combine parameter values for pairs of faces, to produce further choices for selection. When repeated a few times, a composite is 'evolved'—or, more specifically, the space is searched to locate a point in multidimensional space (a set of parameter values) that best represents the target.

There are three main recognition-based systems: ID (Tredoux, Nunez, Oxtoby & Prag, 2006), E-FIT-V (Gibson, Solomon, Maylin & Clark, 2009) and EvoFIT (Frowd, Bruce, & Hancock, 2008). While there does not appear to be much research published on the first two, E-FIT-V does appear to be somewhat successful, producing composites after a short target delay that are named about as accurately as feature-based composites (Valentine, Davis, Thorner, Solomon & Gibson, 2010).

### 1.3 EvoFIT

We have been researching and developing EvoFIT for about 15 years now (for a review, see Frowd, 2012). In the original design, face constructors selected external features to match those of the target, and these were presented in arrays of 18 faces from which users made repeated selections. Choices were based initially on facial shapes, then facial textures and finally on the overall appearance; selected items were 'bred' together, to combine characteristics, as described above, and the process repeated. There have been three milestones in development—

#### 1.3.1 Blurred external-features presentation

The first development facilitated construction of the internal features, the important region for composite naming (Frowd, Skelton et al., 2011). A Gaussian (blur) filter was applied to external features in the face arrays. The blurring level was such (8 cycles per face width) as to render recognition difficult if extended across the image (e.g. Thomas & Jordan, 2002); here, it was used to de-emphasise the external features and increase the saliency of the internal features. After the face had been evolved, blurring was disabled to allow the entire face to be seen clearly. Alongside this development, a set of 'holistic' tools were created for a user to enhance the overall likeness of an evolved face. These could be used to

change age, weight, masculinity and seven other overall properties of the face. In Frowd et al. (2010), composites constructed after a 2 day delay using blur and holistic tools were correctly named with a mean of 24.5% compared to 4.2% for composites where only holistic tools were used (i.e. with non-blurred external features presented in the face arrays).

### 1.3.2 Internal-features first construction

The second development considered in more detail the importance of the exterior region with face construction. It was expected that external features would, to some extent, provide a beneficial context in which to select internal features: some level of blur would be optimal for face selection. The idea is similar to the recognition benefit achieved when the background scene (context) does not differ greatly between a person seeing a target face and attempting to recognise it (e.g. Memon & Bruce, 1983). The outcome for composites was counterintuitive: higher levels of blur led to more-identifiable images, with an “infinite” level of blur emerging the best—filtering that revealed only the internal features. In Frowd, Skelton et al. (2012), correct naming approximately doubled from 22.7% for composites created using the previous external-features’ blur method to 45.6% with internals only. So, the mere presence of external features appears to be a distraction to the person constructing the face: best identification is achieved by presenting just internal features in the face arrays for selection (and adding external features after holistic-tool use).

A recent police-field trial using a version of EvoFIT including this development indicates very-good performance in the real world: 60% of EvoFIT composites constructed by witnesses and victims in a 12 month period directly led to identification of a suspect, and 29% of these identifications resulted in conviction (Frowd, Pitchford et al., 2012).

### 1.3.3 Holistic CI

The third development applied the H-CI to face construction using EvoFIT. In Frowd, Skelton et al. (2012), participants looked at a target video and, the following day, received either CI or HCI and then evolved the face using the blurred external-features procedure. Correct naming of composites increased from 24.1% made after a CI to 39.4% after an H-CI. We argued that the procedure helped a participant-witness’s ability to select faces from the arrays based on the overall appearance of the face, to create a representation that was more recognisable (since recognition is based on the perception of the entire face).

## 1.4 Current experiment

The aim of the current experiment was to explore the impact of combining the above techniques at face construction and naming. We did not know quite how recognisable composites would be were the H-CI to be used in conjunction with internals-only construction. Similarly, would performance increase further using techniques for facilitating recognition of finished composites? One would anticipate that effects would be additive; if this is the case, then augmenting face construction in this way is theoretically interesting as well as providing a worthwhile improvement for police practitioners. In the following experiment, we systematically investigated three factors that have previously been found to increase composite naming: interview type (CI / H-CI), construction type (external-features blur / internal-features only), and one of the procedures for facilitating recognition of a finished composite, angle of view at naming (front on / side on).

## 2.0 EXPERIMENT

Two stages were required to conduct the investigation, which are described below, composite face construction (Stage 1) and composite face naming (Stage 2).

### 2.1 Stage 1: Composite face construction

#### 2.1.1 Method

*Design.* The design used procedures that were similar to the construction and naming of composites in the real world. Participants were recruited to construct composites of targets with whom they were unfamiliar; later, participants familiar with the relevant identities were recruited to name the composites. There are a number of target categories and image-presentation formats that can satisfy this constraint on familiarity—some examples are listed in Table 2. A video format was chosen, principally due to the naturalistic presence of motion (cf. photographs). In addition, and somewhat arbitrarily (see discussion related to Table 2), we chose actors as targets from the Eastenders TV soap (see Materials), enabling us to recruit people who did not follow the programme for face construction and fans of the soap for naming. A nominal 24-hour delay was inserted between a person seeing the video and being interviewed to construct the composite.

To be confident that recognisable composites would be constructed after this forensically-relevant delay, EvoFIT was the method chosen to construct the composites (e.g. Frowd et al., 2010; Frowd, Skelton et al., 2012). Much is known about how this method accesses facial memory, allowing comparisons to be made between results of the current study and elsewhere (e.g. Table 2); we also argue (Discussion) that our results are likely to generalise to other recognition-type composite systems, and potentially to feature systems.

Two factors were manipulated at face construction: interview type and construction type. (A third factor, angle of view, was manipulated at composite naming—Stage 2.) For interview type, participants received either a cognitive interview (CI) or a holistic-cognitive

interview (H-CI): all underwent a CI, to elicit a description of the target face, but those assigned to the H-CI condition were then given procedures to focus on the holistic aspects of their target face. Part of the H-CI mnemonic involves character attribution using whole-face prompts, although the prompts themselves do not seem to be important for the H-CI to be effective (Frowd, Nelson et al., 2012); here, we chose prompts used in Frowd, Nelson et al. (2012)—see Procedure below. For construction type, the procedure used to create the face was the same, except that participants either (i) selected external features and saw them blurred when selecting from EvoFIT face arrays (blur construction), or (ii) were shown internal features (IF) in these face arrays and then chose external features at the end of constructing the face (IF construction).

*Materials.* Stimuli were ten video excerpts from the TV soap, Eastenders. Recordings were made over a one-month calendar period and these were edited into video clips that (i) had a duration of about 30 seconds, (ii) involved an interaction between two current, well-known characters on the programme and (iii) allowed us to freeze-frame (without motion blur) on a front-face view of the intended target (for 5 seconds). Five clips resulted in a different male actor and five in a different female actor. Videos were presented in colour, with sound.

A verbal-description sheet was used to record the description given by participants of their target face at interview. This is an A4 sheet typical of that used with real witnesses. It contains labels for each individual feature—hair, face shape, eyes, nose, mouth, ears and general observations—with space provided underneath for police personnel (or the experimenter here) to make written notes.

EvoFIT software version 1.5 was used to construct the composites.

*Participants.* An opportunity sample of 40 participants volunteered to construct the composites, recruited on the basis of being unfamiliar with actors in the Eastenders TV soap. Sampling was fairly wide, from Wakefield and Barnsley, UK. There were 15 males and 25 females and their age ranged from 25 to 55 ( $M = 36.3$ ,  $SD = 8.4$ ) years. Participants were assigned equally to the two between-subjects factors, to give four groups of 10 people.

*Procedure.* Participants were tested individually. They met with the experimenter on two occasions. In the first meeting, they were instructed to watch a video clip from the TV soap, Eastenders. Participants were told that the video contained an interaction between two characters and, at the end, would freeze-frame on one actor, the target for face construction on the following day. Participants were given headphones to listen to the dialog. After watching the video, they were asked whether the target was familiar. (No targets were reported as known, but if any had been, another video would have been selected randomly and presented.) Allocation of participants to video clips was random within the constraint that the 10 target videos were inspected in each of the two between-subjects factors of

interview (CI / H-CI) and construction (blur / IF). The experimenter remained blind to the identity of targets until all composites had been constructed.

The experimenter met with participants at the agreed time, between 20 and 28 hours after exposure to the target. Participants were told that the session would be in two parts. First, there would be an interview where they would be asked to recall the appearance of the target's face. Afterwards, a composite would be created using EvoFIT.

All participants received a cognitive interview (CI) to elicit a description of the target face. They were instructed that, in a few moments, they would be asked to think back to when the target face was seen and try to visualise it. Then, they would freely recall the face in their own words and in their own time, without guessing. Also, while participants were doing this, the experimenter would not interrupt but make written notes on a verbal-description sheet. A description of the face was obtained using this procedure.

Participants assigned to the CI then moved onto face construction (next paragraph), whereas those assigned to the H-CI condition received a holistic interview. These latter participants were told that they would now focus on the character of the target face, as this would help them to produce an identifiable composite. First, they were asked to think about the personality or character of this person's face, silently for 1 minute. After this timed period, participants were instructed that they would judge the face against a series of general characteristics; also, that these judgements should be made using the rating scale of 'low', 'medium' and 'high'. The first property, intelligence, was read aloud and participants duly provided a three-point rating. Six further properties were similarly repeated and a rating elicited from each in the following order: distinctive, pleasant, masculine, cold, caring and competent. The CI took about 10 minutes to administer and the holistic part, about another 5 minutes.

The session then moved onto face construction. The procedure for constructing the face with EvoFIT using the blur and IF method is fairly detailed and can be found in Experiment 3 of Frowd, Skelton et al. (2012). For the sake of brevity, an overview is provided here. Participants assigned to the blur condition were asked first to select external features that best matched external features on their target; these features were shown, with Gaussian blur applied, on the face arrays. Participants were asked to select faces from these arrays to match the target, for facial shape, for facial texture and for the overall appearance; selected items were 'bred' together and this selection procedure was repeated once more. Blur was then disabled, to clearly see the array of faces, and participants were asked to select the single item with the best likeness. This image was enhanced using the holistic tools, to adjust age, masculinity and other overall properties of the face. Participants were also offered use of a software tool to adjust the shape and placement of features. The resulting face was saved to disc as the composite. Participants assigned to IF construction did the same, except that they inspected arrays containing internal features only and, after using holistic tools, then selected external features. Composites took about an hour to

construct.

## 2.2 Stage 2: Composite face naming

In this part, the composites constructed in Stage 1 were evaluated by asking another group of participants, people who were familiar with the target identities, to name them.

### 2.2.1 Method

*Design.* Participants who were familiar with the target set (self-proclaimed Eastenders' fans) named a set of 10 composites produced by interview (CI / H-CI) and by face construction (blur / IF); they did this first by viewing the face front-on and then side-on. The design was therefore mixed-factorial, with interview and construction type as between-subjects factors and view type as a within-subjects factor. It was expected that correct naming would be higher for H-CI over CI, IF over blur, and side-on over front-on.

The number of participants required for naming was estimated by a G\*Power analysis (Faul, Erdfelder, Lang & Buchner, 2007). We assumed that effect sizes would be very large ( $w = 0.8$ ) for the two between-subjects factors, interview (Cohen's  $d = 2.0$ ) and construction ( $d = 1.3$ ), as indicated by Frowd, Nelson et al. (2012) and Frowd, Skelton et al. (2012), respectively; as the third factor, view of face at naming, was within subjects and has a large impact ( $d = 0.8$ ), the sample size was expected to be satisfied by the estimate from the between-subjects factors. Other model parameters were for power ( $1 - \beta = .8$ ), standard error probability ( $\alpha = .05$ ) and the planned  $X^2$ -type regression analysis ( $df = 2$ ). The analysis indicated that four participants per group would be sufficient, which did seem a rather small sample to establish reliability, and so we doubled this figure to a minimum of eight per group ( $N = 32$  in total)—to give a design that should have sufficient power to detect a large effect size ( $w = 0.55$ ).

*Materials.* Stimuli were the 40 composites constructed in Stage 1: 10 composites from each of the four conditions (CI and blur; CI and IF; H-CI and blur; H-CI and IF). See Figure 1 for example composites. A good-quality photograph for each of the 10 targets was located on the Internet. These photographs were presented to participants to name, after having inspected the composites, as a check that they were actually familiar with the relevant identities. All images were printed in greyscale on single sheets of A4 paper to dimensions of approximately 8cm (wide) x 10cm (high).



(a)

(b)

(c)

(d)

Figure 1. Example composites constructed in Stage 1 and presented to participants for naming in Stage 2. They are of Eastenders' character, Ian Beale, played by actor Adam Woodyatt. Each image was produced by a different person: (a) CI and blur, (b) CI and IF, (c) H-CI and blur and (d) H-CI and IF. For copyright reasons, we are unable to provide an image of the actor.

*Participants.* Thirty-two participants volunteered to name the composites. They were visitors to Blackburn Town Centre and were recruited on the basis of being familiar with characters from the TV soap, Eastenders. This was an opportunity sample comprising of 13 males and 19 females, with an age range from 20 to 41 ( $M = 25.5$ ,  $SD = 7.1$ ) years. They were assigned equally to the two between-subjects factors, eight per group.

*Procedure.* Participants were tested individually. They were informed that composites of actors in the Eastenders' soap would be shown for them to name. Participants were instructed to guess if unsure of the identity, but also that it was acceptable not to give a name—a "don't know" response. They were assigned, with equal sampling, to inspect composites constructed from one condition in Stage 1: by interview (CI / H-CI) or by construction (blur / IF). The assigned set of 10 composites were presented sequentially and participants provided a name (or not), as requested. Afterwards, it was mentioned that research had found that looking at a composite sideways, to give a different perspective, can help to recognise the face. Participants were invited to have a second attempt at naming by inspecting the composites in this way. Composites were presented again, sequentially and in the same order as before, and participants turned each one to the side themselves and provided a name (or not). Finally, the 10 target photographs were presented sequentially and participants asked to name those as well. Each person received a different random order of presentation for composites and target photographs. The task was self-paced and took about 10 minutes per person to complete.

### 2.2.2 Results

The target photographs were correctly named almost perfectly ( $M = 99.7\%$ ), with only one participant giving an incorrect (“don’t know”) response (in the CI / IF condition). Therefore, participants demonstrated excellent familiarity with the target identities and so, in theory anyway, should have been able to recognise all but one of the composites.

Composites were scored for accuracy: a coding of 1 was used when a correct name was given (for the soap character or actor name) and 0 otherwise. For the target photograph that was not correctly named, as mentioned above, we coded the response for the associated composite as ‘missing data’ so that it would be appropriately omitted from subsequent analyses. Accurate responses are summarised in Table 1—see *Note* for how these values were calculated. It can be seen that, even in the best condition (H-CI, IF construction, and side-on naming), correct naming was still much lower for composites ( $M = 73.8\%$ ) than for targets ( $M > 99\%$ ), but this is the usual situation as composites are error-prone stimuli and are rarely (if ever) identified perfectly. Composite naming substantially increased with each manipulation, as predicted: (i) from CI to H-CI, (ii) from blur to internal-features (IF) construction, and (iii) from front-on to side-on naming. Naming was fairly good ( $M = 23.8\%$ ) in the worst-predicted condition (CI, blur and front-on), but much higher ( $M = 73.8\%$ ) when all three techniques were used together (H-CI, IF and side-on).

Inferential analyses involved two between-subjects factors, interview type (CI / H-CI) and construction type (blur / IF), and a single within-subjects factor, view at naming (front-on / side-on). Generalized Estimating Equations (GEE) were employed, a type of regression model suitable for our dichotomous (0 and 1) responses with repeated measurements (e.g. Barnett et al., 2009). The model was full-factorial using a binary-logistic function; a covariance structure was chosen based on equal correlations for the within-subjects factor.

The GEE model revealed that all three fixed factors were significant: (i) interview [ $\chi^2(1) = 13.4, p < .001$ ], with H-CI superior to CI [ $B = 1.2, SE(B) = 0.3, Exp(B) = 3.4$ ]; (ii) construction [ $\chi^2(1) = 6.8, p = .009$ ], with IF superior to blur [ $B = 0.9, SE(B) = 0.3, Exp(B) = 2.4$ ]; and (iii) view [ $\chi^2(1) = 46.2, p < .001$ ], with side-on superior to front-on [ $B = 0.8, SE(B) = 0.2, Exp(B) = 2.4$ ]. For more details of the model, see Table 1. None of the two-way interactions were significant ( $\chi^2 < 1$ ). However, the three-way interaction was significant [ $\chi^2(1) = 4.3, p = .038$ ], which emerged with IF constructions for interview x view—in the table, refer to the four cells relating to IF. While the benefit of *view* was equivalent ( $p > .05$ ) under both CI and H-CI [factor increase in Odds Ratio relative to Constant model  $\Delta Exp(B)_c = 4.0$ ], the benefit of *interview* was reliably stronger ( $p < .05$ ) under side-on than front-on [ $\Delta Exp(B)_c = 5.6$ ]. In other words, the increase indicated by †<sup>1</sup> in the table is greater than that predicted by the sum of the three fixed effects.



Table 1. Accurate (correct) naming: an advantage of holistic (H-CI) over cognitive interview (CI), internal-features (IF) over blur construction, and side-on over front-on naming

View	CI		H-CI	
	Blur	IF	Blur	IF
Front-on	23.8 (19 / 80)	36.7† (29 / 79)	42.5 (34 / 80)	53.8† (43 / 80)
Side-on	37.5 (30 / 80)	45.6† <sup>1</sup> (36 / 79)	53.8 (43 / 80)	73.8† <sup>1</sup> (59 / 80)

Note. Values are correct-naming scores calculated by dividing responses shown in parentheses and expressed as a percentage. Parenthesised values are summed correct responses (numerator) and total (correct and incorrect) responses (denominator) for composites that participants correctly named the relevant target ( $N = 638$  out of 640). Model details for these data: Generalized Estimating Equations' Goodness of Fit,  $QIC = 844.4$ , and intercept [Beta (gradient) coefficient,  $B = -1.0$ ; standard error of  $B$ ,  $SE(B) = 0.3$ ; Odds ratio (effect size),  $Exp(B) = 0.36$ ]. †Significant three-way interaction involving these four cells. See text for more details.

The composite naming data were re-scored for inaccurate responses. Unlike naming of the target photos, “don’t know” as well as mistaken names were given for composites. An analysis of these data provides a measure of willingness to produce a name, or guess, and is analogous to the signal-detection measure of *criterion* (response bias). The analysis considered the number of mistaken names (coded as 1) relative to the total incorrect names (mistaken names + no names, both coded as 0)<sup>(1)</sup>; as for correct naming, it excluded the composite data point for which the relevant target had not been correctly named.

Incorrect names did not greatly vary between each cell of the design [ $42.6 < M(\%) < 52.4$ ]. When the GEE model was run in the same way as above, several standard errors [ $SE(B)$ ] were high relative to their beta coefficient ( $B$ )—for example,  $SE(B)$  was 0.6 [ $B = 0.2$ ] for construction x interview. This suggests that too many variables were being fitted to too few cases; for a dataset like ours that was fairly invariable by condition, this indicates an issue of complete separation (perfect prediction), and so a simpler model was built involving just the three fixed factors. The result [ $QIC: 487.5$ ; intercept:  $B = 0.1$ ,  $SE(B) = 0.3$ ,  $Exp(B) = 1.1$ ] was significant for view [ $\chi^2(1) = 4.3$ ,  $p = .038$ ], indicating that there were more mistaken

<sup>1</sup> It is perhaps worth mentioning that this measure of inaccuracy is calculated in this way to allow estimation of response bias. It is the proportion of mistaken to total inaccurate responses (mistaken + no-names) rather than to responses in total (correct + mistaken + no-names). Done in this way, it appears that incorrect names are very frequent, when this may not be the case: it depends upon the number of correct names. For example, in the condition that elicited most correct names (H-CI, IF and side-on naming), mistaken names amount to 13.8% of the responses in total (11 out of 80): it is 52.4% when relative to the total incorrect (11 out of 21).

names given for side-on ( $M = 50.3\%$ ) than front-on ( $M = 45.4\%$ ) [ $B = 0.11$ ,  $SE(B) = 0.05$ ,  $Exp(B) = 1.1$ ]. Interview and construction were not reliable fixed factors ( $X^2 < 1$ ).

### 3. DISCUSSION

Recovery of an identifiable face from memory can be of enormous value for identifying people who commit crime. Until quite recently, methods available to the police had not consistently produced identifiable faces. The feature systems in particular do not appear to provide a good interface to memory, especially when deployed after long intervals of time. In the current project, we investigated three techniques which have been found to improve the effectiveness of composites created using traditional feature systems or one of the more-recent recognition-based varieties, EvoFIT. Each technique relates to a different stage in composite usage: interview, system and naming. Participants viewed a video containing an unfamiliar target face and the following day recalled this face using either a cognitive interview (CI) or a holistic cognitive interview (H-CI), and then constructed a composite of it using EvoFIT by selecting from face arrays containing either blurred external features or just internal features (IF). Further participants who were familiar with the identities attempted to name the resulting composites in the normal way, by viewing the printed face front-on, and then by turning the page to view the face from the side.

The pattern of effects observed for correct composite naming replicated previous findings: there was an advantage of interview (H-CI was superior to CI), construction (IF was superior to blur), and view (face side-on was superior to front-on). Performance was worst when techniques were not used, CI with blur construction and front-on naming, and best when the three techniques were combined, H-CI with IF construction and side-on naming. There was also a significant three-way interaction: while the fixed factors gave rise to a reliable advantage throughout, for IF construction the benefit was greater for H-CI with side-on over front-on naming. In other words, the techniques were mainly additive, except for this additional advantage: a combined result that was greater than the sum of the main effects. The number of mistaken (incorrect) names given by participants was higher (albeit with a small effect size) when viewing the face side-on than front-on.

#### 3.1 The additive effects of techniques

So, all three techniques are effective on their own and also when used in conjunction with each other. Each technique involves an overall or holistic procedure, which is why recognition—itsself a holistic process (e.g. Davies & Christie, 1982; Tanaka & Farah, 1993)—increased with each one. The techniques are likely to enhance recognition in different ways.

First, the H-CI encourages constructors to focus on the target face as a whole, mainly through character attribution, much in the same way as holistic encoding improves unfamiliar-face recognition (e.g. Shapiro & Penrod, 1986). The knock-on effect of this interview is that constructors select faces from EvoFIT arrays based on their overall appearance rather than on their individual features (Frowd, Bruce, Smith and Hancock, 2008).

Second, masking the external features interferes with processing of the face as a whole to a greater extent than does presenting blurred external features in the face arrays (Frowd, Skelton et al., 2012)—which in turn interferes with recognition to a greater extent than does presenting external features intact in the arrays (Frowd et al., 2010). In the current application, this interference is a good thing: when it is maximal (external-features masked), it enables a face constructor to produce identifiable internal features.

The H-CI and IF procedures alter the choices a constructor makes from the presented face arrays, shaping the course of the search through the face-space model<sup>2</sup>. Clearly, selections based on the overall appearance of the face are best achieved when both H-CI and IF procedures are used together (i.e. for the same constructor). Through the breeding process, which combines face parameters to provide alternatives for further selection, the final solution is more optimal in terms of the recognisability of the evolved face.

Third, viewing a composite sideways gives the appearance of a thinner face—that is it is an overall change in how the face appears. A proposed mechanism (Astbury, 2011) for this ‘perceptual’ stretch technique is that an observer’s face-recognition system must transform (normalise) this percept into a normal-looking face and, in doing so, the influence of error is reduced, facilitating recognition. The technique was found to have a similar benefit under different combinations of interview and type of construction. However, the three factors interacted: for IF constructions, the advantage of interview was reliably stronger for side-on than for front-on naming. Stated another way, the benefit of interview for IF constructions named side-on was greater than that predicted from the main effects.

But, why should this be? A plausible explanation is that these techniques together combine to improve performance in a non-linear way, principally because they involve (i) the region of the face that is particularly sensitive for familiar-face recognition, the eye region (Diamond & Carey, 1996; Leder & Bruce, 2000; McIntyre, Frowd, Bruce & Hancock, 2010; O’Donnell & Bruce, 2000), and (ii) individual facial features, which are also known to be important for recognition of faces (e.g. Cabeza & Kato, 2000) and composites (e.g. Frowd, Bruce, Smith and Hancock, 2008). The idea also resonates with a type of face model (two-

---

<sup>2</sup>There is also evidence that the IF procedure helps constructors to some extent after the evolving stage. This is when they use the holistic tool for improving the likeness of their evolved face: to change age, weight and other overall properties (Frowd, Skelton et al., 2012). Focusing on just the internal-features region in this task without the external features present similarly leads to a more identifiable image.

pool opponent coding) that can be linear as well as non-linear (e.g. Dakin & Omigie, 2009; Susilo, McKone, & Edwards, 2010). Relative to the CI, the H-CI promotes a composite that is more accurate in terms of the individual features (Frowd, Bruce, Smith and Hancock, 2008) as well as the internal-features region in general (ibid; Frowd, Skelton et al., 2012). Looking at the face side-on may reduce the appearance of error for individual features in the horizontal plane, as well as for errors between the eyes (i.e. eye spacing). Also, with side-on viewing, while residual errors relating to individual features and eye spacing may be somewhat high for IF constructions after the CI, these errors may be greatly reduced by the H-CI. The net result is a large increase in correct naming when conditions are combined.

### 3.2 Perceptual stretch

Perceptual stretch is an interesting technique, one which has general benefit: relative to front-on, it provided an increase in correct naming in all combinations of construction and interview type. While none of the 40 composites constructed decreased in mean correct naming from front-on to side-on, in 24 cases (60%) mean naming increased. Observing the face sideways gives the percept that the face is three-dimensional, which may in part account for why participants mistakenly named a face when looking at the face in that way—although this increase was infrequent as the effect size was very small [ $Exp(B) = 1.1$ ] (where 1.5 would be considered *small*).

Observing a face sideways is only an approximation to a physical image stretch. This is due to the active visual process of perspective. When looking down a street of parked cars, for example, perspective gives the impression that the cars are the same size—when, their physical size is actually getting smaller on the retina; the effect is known as *size constancy* (e.g. Coren, Ward & Enns, 1999). When viewing faces, as for differences in lighting and head pose (e.g. Hill & Bruce, 1996), changes in perspective influence unfamiliar-face recognition (e.g. Liu & Chaudhuri, 2006). In our case, when looking sideways, viewing distance across the face increases, visual angle decreases and observers experience perspective: the half of the face that is furthest from the viewer (e.g. the left side to us) when looking at the page from the (e.g. right) side appears to be increased. While the effect of perspective is smaller than the effect of stretching, the fact that perceptual stretch works to improve recognition of composites is all the more impressive. We have recently developed a technique which enables the face to be viewed front-on but gives a similar percept and recognition advantage to perceptual stretch (<http://tiny.cc/pbi-format>).

It is perhaps interesting to note that the underlying mechanism for perceptual stretch is believed to be different to the multi-frame caricature technique mentioned earlier for facilitating facial recognition. While the former appears to involve reducing the appearance of error, multi-frame caricature involves exaggerating distinctive features and relationships in the face (manuscript under revision). This does suggest that, as the methods

seem to have a different basis for enhancing recognition, their effects may be additive if both are used together at naming. Current work is exploring this possibility.

When perceptual stretch was initially evaluated, the design was between-subjects, with participants observing composites either front- or side-on (Astbury, 2011). Correct naming of these celebrity composites substantially increased with side-on relative to front-on (with incorrect names not differing reliably). The within-subjects design used here (front-on and then side-on naming for the same person) indicates that the technique generalises across designs. In fact, the within-subjects design is typically how it is used forensically: police officers and members of the public look at the face first from the front and, following written instructions accompanying the composite, then from the side. Our results also indicate another generalisation of the technique, indicating robustness, since perceptual stretch was found here to extend to another class of target identities, actors in the TV soap *Eastenders* (before, targets were well-known celebrities).

### 3.3 Consistency of results

It is worth considering consistency of our results with those reported elsewhere that have used the same conditions with EvoFIT. There are three relevant studies, see Table 2. In spite of differences in delay to construction (24 or 48 hours), presentation format of target (photograph or video) and target type (snooker players, shop staff, international footballers or actors from *Eastenders*), correct naming is remarkably similar (less than 2% difference) between the present study and the three others for composites constructed with a CI and blur (first column of data). Further, the potential pool of targets differs in size for these studies: one would imagine that fewer targets would facilitate composite naming. This pool size is much less for both *Eastenders* and shop staff than for international footballers, but consistency of results across these studies suggests that this is not an important factor. There may be more potential targets in police-work but the implication is that a larger search space is unlikely to influence composite naming.

Correct naming is also very similar between ours and Frowd, Nelson et al. (2012) for H-CI and blur (less than 4% difference), with both using video presentation. Arguably less similar (about 10% difference), although still clearly sensible, is ours and Frowd, Skelton et al. (2012) for CI and IF. It is possible that this larger difference has emerged as IF construction is more effective for static (photographic) targets than for moving (video) targets (used here)—possibly because motion attracts attention to the internal-features region, meaning that external-features masking is less effective. This finding has practical implications, since witnesses generally see offenders that move their head (either rigidly or otherwise); here, the effect is maintained for moving stimuli and indicates the likely real-life effect (and the potential need for follow-up work). Overall, there is very good consistency of techniques across studies.

Table 2. Percentage-correct naming of EvoFIT composites by interview and construction type in published research. Details are also presented for target characteristics and nominal delay between seeing a target and constructing a composite of the face.

Study	Interview and construction type				Target		
	CI Blur	H-CI Blur	CI IF	H-CI IF	Delay (hr)	Presentation	Category
Frowd et al. (2010)	24.5				48	Photo	Snooker players
Frowd, Nelson et al. (2012)	24.1	39.4			24	Video	Shop staff
Frowd, Skelton et al. (2012)	22.7		45.6		24	Photo	UK Footballers
Present	23.8	42.5	36.7	53.8	24	Video	Eastenders actors

### 3.4 Generalisation to other methods for accessing facial memory

So, the three techniques combine to promote more identifiable composites with EvoFIT, but would this benefit extend to feature-based composites? Existing data indicates that both the H-CI (Frowd, Bruce, Smith and Hancock, 2008) and perceptual stretch (Astbury, 2011) are effective on their own for a feature system (PRO-fit). There is no real reason to suppose that these two techniques should not produce additive effects, when combined, either for PRO-fit or for any of the other feature methods popular outside of the UK (e.g. FACES, Identikit 2000 and ComPhotofit). The research to confirm this suggestion is in progress, as is an equivalent technique to IF construction for feature systems that similarly masks external features (to improve construction accuracy of internal features). In the lab, feature systems occasionally recover an identifiable image from memory under forensically-relevant delays (e.g. Frowd, Carson, Ness, McQuiston et al., 2005; Frowd, Bruce, Ness et al., 2007; Frowd et al., 2010), with apparently similar identification for police practitioners (Frowd, Pitchford et al., 2012), and so improving their performance is likely to have increased utility for the detection of offenders using this type of technology.

Other recognition-type composite systems, E-FIT-V and ID, extract faces from memory using a similar method to EvoFIT (Frowd, 2012). They randomly sample items from PCA type models, and present them as arrays of faces for a user (witness) to select. As well as using a simultaneous-type presentation format, they utilise an Evolutionary Algorithm to combine user choices to carry out a search in the (PCA) face space. Post-evolving 'holistic' tools (e.g. age) are also used with the aim of enhancing the overall likeness. Existing research (Valentine et al., 2010) does suggest that E-FIT-V works fairly well under a short target delay, the same as for the feature methods, producing composites with mean naming

of about 20%, and so one would imagine that the techniques presented here would similarly promote a more identifiable image following forensically-relevant delays using this system.

### 3.5 Summary

In summary, we have evaluated three techniques which concern the appearance of the face as a whole and independently have been found to improve the recognisability of facial-composite images. Two of these techniques (H-CI and side-on naming) have been shown to be effective for both a feature- and a recognition-type method, while the third (IF construction) has been found to be effective for the latter type. In the current study, these three techniques successfully combined to improve the identifiability of composites produced after a 24 hour delay using EvoFIT. The result of using all three techniques together was a very recognisable composite, as evidenced by mean naming of 74% correct; if used in police investigations, they should prove valuable for the detection of offenders. We have argued that the techniques, if applied to other methods for accessing facial memory, should similarly promote more identifiable composites.

### 4.0 REFERENCES

ACPO (2009). Facial Identification Guidance. Produced by the National Policing Improvement Agency. Unpublished.

Astbury, R. (2011). *Another side to fighting crime*. Lancashire Evening Post, 02/11/2011, 17.

Barnett, A.G., Koper, N., Dobson, A.J., Schmiegelow, F., & Manseau, M. (2009) Selecting the correct variance–covariance structure for longitudinal data in ecology : a comparison of the Akaike, quasi-information and deviance information criteria. Downloaded on 7<sup>th</sup> August 2012, <http://eprints.qut.edu.au/19195/>

Benson, P.J. & Perrett, D.I. (1994). Visual processing of facial distinctiveness. *Perception*, 23, 75-93.

Berman, G.L., & Cutler, B.L. (1998). The influence of processing instructions at encoding and retrieval on face recognition accuracy. *Psychology, Crime & Law*, 4, 89-106.

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., 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.

- Cabeza, R., & Kato, T. (2000). Features are also important: Contributions of featural and configural processing to face recognition. *Psychological Science*, 11, 429-433.
- Coren, S., Ward, L.M., & Enns, J.T. (1999). *Sensation and Perception*. 5th ed. Fort Worth, TX: Harcourt Brace.
- Dakin, S.C., & Omigie, D. (2009). Psychophysical evidence for a non-linear representation of facial identity, *Vision Research*, 49, 2285-2296.
- 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.W., & Ellis, H.D. (1978). Remembering faces: acknowledging our limitations. *Journal of Forensic Science*, 18, 19-24.
- Davies, G.M., van der Willik, P., & Morrison, L.J. (2000). Facial Composite Production: A Comparison of Mechanical and Computer-Driven Systems. *Journal of Applied Psychology*, 85, 1, 119-124.
- Diamond, R., & Carey, S. (1986). Why faces, are and are not special: an effect of expertise. *Journal of Experimental Psychology: General*, 11, 107-117.
- 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, 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.
- 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. (2012). *Facial Recall and Computer Composites*. In C. Wilkinson and C. Rynn (Eds). *Facial Identification* (pp. 42 – 56). Cambridge University Press: New York.
- Frowd, C.D., Bruce, V., Gannon, C., Robinson, M., Tredoux, C., Park., J., McIntyre, A., & Hancock, P.J.B. (2007). 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. (2008). Changing the face of criminal identification. *The Psychologist*, 21, 670-672.
- Frowd, C.D., Bruce, V., & Hancock, P.J.B. (2009). Evolving facial composite systems. *Forensic Update*, 98, 25-32.



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

Frowd, C.D., Bruce, V., Ross, D., McIntyre, A., & Hancock, P.J.B. (2007). An application of caricature: how to improve the recognition of facial composites. *Visual Cognition*, 15, 1-31. DOI: 10.1080/13506280601058951.

Frowd, C.D., Bruce, V., Smith, A., & Hancock, P.J.B. (2008). 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. (2005). 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. (2005). A forensically valid comparison of facial composite systems. *Psychology, Crime & Law*, 11, 33-52.

Frowd, C.D., Hancock, P.J.B., Bruce, V., Skelton, F.C., Atherton, C., ... Sendrea, G. (2011). Catching more offenders with EvoFIT facial composites: lab research and police field trials. *Global Journal of Human Social Science*, 11, 46-58.

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., & Hepton, G. (2009). The benefit of hair for the construction of facial composite images. *British Journal of Forensic Practice*, 11, 15-25.

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

Frowd, C.D., Nelson, L., Skelton F.C., Noyce, R., Atkins, R., Heard, P., Morgan, D., Fields, S., Henry, J., McIntyre, A., & Hancock, P.J.B. (2012). Interviewing techniques for Darwinian facial composite systems. *Applied Cognitive Psychology*, DOI: 10.1002/acp.2829.

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., Pitchford, M., Skelton, F., Petkovic, A., Prosser, C., & Coates, B. (2012). Catching Even More Offenders with EvoFIT Facial Composites. In A. Stoica, D. Zarzhitsky, G. Howells, C. Frowd, K. McDonald-Maier, A. Erdogan, and T. Arslan (Eds.) *IEEE Proceedings of 2012 Third*

*International Conference on Emerging Security Technologies*, DOI 10.1109/EST.2012.26 (pp. 20 - 26).

Frowd, C.D., Skelton F., Atherton, C., Pitchford, M., Hepton, G., Holden, L., McIntyre, A., & Hancock, P.J.B. (2012). Recovering faces from memory: the distracting influence of external facial features. *Journal of Experimental Psychology: Applied*, 18, 224-238.

Frowd, C.D., Skelton, F., Butt, N., Hassan, A., & Fields, S. (2011). Familiarity effects in the construction of facial-composite images using modern software systems. *Ergonomics*, 54, 1147-1158.

Gibson, S.J., Solomon, C.J., Maylin, M.I.S., & Clark, C. (2009). New methodology in facial composite construction: from theory to practice. *International Journal of Electronic Security and Digital Forensics*, 2, 156-168.

Hancock, P.J.B., Bruce, V., & Burton, A.M. (1997). Testing principal component representations for faces. In J.A. Bullinaria, D.W.Glasspool and G. Houghton (Eds.). *Proceedings of 4th Neural Computation and Psychology Workshop*, London: Springer-Verlag, 84-97.

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

Hole, G.J., George, P.A., Eaves, K., & Rasek, A. (2002). Effects of geometric distortions on face-recognition performance. *Perception*, 31, 1221-1240.

Leder, H, & Bruce, V. (2000) When Inverted Faces Are Recognized: The Role of Configural Information in Face Recognition. *Quarterly Journal of Experimental Psychology: Section A*, 53, 513–536. doi:10.1080/027249800390583.

Lee, K.L., & Perrett, D.I. (1997). Presentation-time measures of the effect of manipulations in colour space on discrimination of famous faces. *Perception*, 26, 733-752.

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

Kovera, M.B., Penrod, S.D., Pappas, C., & Thill, D.L. (1997). Identification of computer generated facial composites. *Journal of Applied Psychology*, 82, 235-246.

Memon, A., & Bruce, V. (1985). Contexts effects in episodic studies of verbal and facial memory: A review. *Current Psychological Research & Reviews*, Winter 1985-86, 349-369.

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.
- Susilo, T., McKone, E., Edwards, M. (2010). What shape are the neural response functions underlying opponent coding in face space? A psychophysical investigation. *Vision Research*, 50, 300–314.
- Tanaka, J.W., & Farah, M.J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 46A, 225-245.
- Thomas, S.M., & Jordan, T.R. (2002). Determining the influence of Gaussian blurring on inversion effects with talking faces. *Perception & Psychophysics*, 64, 932-944.
- 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.
- Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion and race in face recognition. *Quarterly Journal of Experimental Psychology*, 43A, 161-204.
- Valentine, T., Davis, J. P., Thorner, K., Solomon, C., & Gibson, S. (2010). Evolving and combining facial composites: Between-witness and within-witness morphs compared. *Journal of Experimental Psychology: Applied*, 16, 72 – 86.
- Wells, G., Memon, A., & Penrod, S.D. (2007). Eyewitness evidence: improving its probative value. *Psychological sciences in the public interest*, 7, 45-75.