

Evolving an identifiable face of a criminal

Charlie Frowd, Faye C. Skelton, Chris Atherton and Peter J.B. Hancock

How can we construct a suspect's face that people, such as police officers and members of the public, can accurately name? Traditional methods involve witnesses selecting individual facial features, but this process has limitations when witnesses cannot recall an offender's face in detail. EvoFIT was inspired by the natural processes of selection and breeding, but considerable work has been necessary for it to yield identifiable images. In particular, external features (ears, hair and neck) are a major source of distraction for constructors: blurring, or better still removing them entirely, greatly facilitates performance. Also, limiting recall or focusing on the holistic aspects of the face similarly promotes more identifiable images.

question

Which psychological factors are likely to affect the quality of a witness's composite?

resources

Frowd, C.D. (in press). Facial recall and computer composites. In C. Wilkinson and C. Rynn (Eds.) *Facial identification*. Cambridge University Press.
Frowd, C.D., Bruce, V. & Hancock, P.J.B. (2009). Evolving facial composite systems. *Forensic Update*, 98, 25–32.
www.EvoFIT.co.uk
www.uclan.ac.uk/cfrowd

references

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., Henderson, Z., Greenwood, K. et al. (1999). Verification of face identities from images captured on

video. *Journal of Experimental Psychology: Applied*, 5, 339–360.
Cutler, B.L., Penrod, S.D. & Martens, T.K. (1987). The reliability of eyewitness identifications. *Law and Human Behavior*, 11, 223–258.
Davies, G.M., van der Willik, P. & Morrison, L.J. (2000). Facial composite production. *Journal of Applied Psychology*, 85, 119–124.
Ellis, H.D. Shepherd, J.W. & Davies, G.M.

(1975). Use of photo-fit for recalling faces. *British Journal of Psychology*, 66, 29–37.
Ellis, H.D., Shepherd, J.W. & Davies, G.M. (1979). Identification of familiar and unfamiliar faces from internal and external features. *Perception*, 8, 431–439.
Frowd, C.D., Bruce, V. & Hancock, P.J.B. (2008a). Changing the face of criminal identification. *The*

Psychologist, 21, 670–672.
Frowd, C.D., Bruce, V., McIntyre, A. & Hancock, P.J.B. (2007a). The relative importance of external and internal features of facial composites. *British Journal of Psychology*, 98, 61–77.
Frowd, C.D., Bruce, V., Ness, H. et al. (2007b). Parallel approaches to composite production. *Ergonomics*, 50, 562–585.
Frowd, C.D., Bruce, V., Pitchford, M. et al.

Many crimes are committed for which the only record of the event is contained in the memory of an eyewitness. When this occurs, police use specialised interviewing techniques to recover an accurate description of the event and the people involved (e.g. Milne & Bull, 2006). For serious crimes such as sexual assault, burglary and murder, they may also invite witnesses (who may also be victims) to work with trained police personnel (or sketch artists) to construct a visual likeness of the offender's face. These pictures are known as facial composites and are used in public appeals for information, to locate suspects on whom police can focus their inquiries. The production of recognisable faces from eyewitness memory is therefore important for detecting and convicting offenders. Indeed, in the light of recent government cuts to public services, effective and affordable tools for policing have never been so important.

The nature of facial memory has intrigued psychologists for decades (e.g. Ellis et al., 1975, 1979; Hancock et al., 2000; Koehn & Fisher, 1997; Longmore et al., 2008). Part of this intrigue derives from evidence that human memory does not operate like a video camera: it does not make complete and accurate recordings, and memory is further compromised by the passage

of time (e.g. Shapiro & Penrod, 1986). It is possible, however, to harness technology to help bridge this gap.

Traditional 'feature' methods available to police for producing composites involve police officers or support staff asking witnesses first to describe a face in detail and then to build a composite of it by selecting facial parts – eyes, nose, mouth, hair, etc. In the UK, E-FIT and PRO-fit computer software is used, which allow features to be selected and positioned on the face in order to produce the best likeness of the offender's face that the witness can achieve (see Figure 1). Our 'gold' standard laboratory procedure to simulate face construction by real witnesses (Frowd et al., 2005b) revealed that participants can produce composites with PRO-fit and E-FIT that are correctly named (i.e. by other participants) fairly consistently, at a mean level of about 20 per cent – other labs report similar results (Brace et al., 2000; Davies et al., 2000). However, when construction occurs one or two days after seeing an unfamiliar face, the normal interval following a crime, naming typically falls to a few per cent correct (e.g. Frowd et al., 2005a, 2007b, 2010); only artists' sketches from people trained in portraiture appear to survive the passage of time, with naming rates of about 10 per cent after a two-day delay



Figure 1: Can you guess the famous face from these feature-based composites (Frowd et al., 2005b)? Participants saw the same picture (an unfamiliar identity) and, three-to-four hours later, described and constructed his face with (l-r) E-FIT, PRO-fit and sketch. Did you recognise David Beckham?

(Frowd et al., 2005a).

Four years ago, my colleagues and I (Frowd et al., 2008a) gave an overview in *The Psychologist* of composite systems and offered three solutions to improve their effectiveness. The first (Frowd et al., 2007c) progressively caricatures a composite in an animated sequence, first by exaggerating distinctive shape information, then by 'anti' caricaturing it, reducing features' distinctiveness (view an example at tinyurl.com/cuulu7e). Seeing a face presented in this way increases correct naming by tenfold for poor-quality images (manuscript under revision) and is also effective when the image quality is better.

Our second solution concerned the initial interview. The police use cognitive-based interviews to help recall: witnesses describe the offender's face in a 'free' or uninterrupted format and then focus on each feature to remember more detail – 'cued' recall. Using our 'holistic' interview, witnesses are then asked to reflect on the personality of the face and make seven whole-face judgements about it (e.g. masculinity, distinctiveness and honesty). Such holistic attribution improves a witness's ability to select facial features and promotes a much more identifiable image (Frowd et al., 2008b) – an advantage (Frowd et al., in press-a) which extends to our own evolutionary composite system (see below).

These developments were valuable, but one problem remained unsolved. Composite operators use the witness's description of the offender's face to locate a set of facial features from which the witness selects the best matches. Most



Figure 2: Manipulating the overall properties of the face using EvoFIT's 'holistic' tools. In this example, a composite of the footballer David Beckham was evolved (left) and then reworked (right) to improve the likeness by increasing perceived health and attractiveness.

people, however, cannot recall the face in detail, leading to a rather large number of matching features – far too many for a witness to inspect. This situation is common with confidence crimes – bogus officials, for instance, who con their way into people's homes and then steal money, jewellery and other valuables. Here, the victim is often unaware that a crime is taking place and tends to provide only a sketchy description of the offender's face. Following national police guidelines, creation of a composite is not permitted under these circumstances; our third solution is a face-production method that breaks the dependency on witness descriptions.

Our aim has been to develop an interface to human memory that is similar to how we naturally recognise faces, as whole entities. This has resulted in a system called EvoFIT that searches through the space of possible identities using natural processes of selection and breeding. In essence, arrays of intact faces are shown to witnesses, who select items

with the best overall likeness to a target, initially for shape and position of facial features, and then for colouring of features or 'texture'. Selected items are 'bred' together, to combine facial characteristics, producing further items for selection; when this process is repeated, the population of faces converges on a specific identity. Accuracy of evolution is enhanced by asking witnesses to identify the best individual face in each generation and then giving that face more breeding opportunities (Frowd et al., 2007b).

Two techniques have further improved accuracy of convergence on a target identity. The first is blurring of external facial features (hair, ears and neck) while witnesses select from face arrays.

This shifts their attention to the faces' internal facial features (eyes, brows, nose and mouth) – the region that is important for recognition later, when the composite is seen by police officers or members of the public. While blurring improves performance, evolved images are still somewhat error-prone.

The second technique enables inaccuracies to be reduced using 'holistic' scales: these include age, health, weight, and the extent to which the face is honest or masculine. Witnesses manipulate their evolved face along each scale while searching for a better likeness. An example manipulation is shown in Figure 2. Tested in Frowd et al. (2010), both techniques were independently effective; when used together, resulting EvoFITs yielded correct naming of 25 per cent: naming was just 5 per cent for composites produced under the same testing conditions using a feature system.

Since the original article in *The Psychologist* in 2008, our understanding of how we construct faces has markedly

(2009a). Evolving the memory of a criminal's face: Methods to search a face space more effectively. *Soft computing*, 14, 81–90.

Frowd, C.D., Bruce, V., Ross, D. et al. (2007c). 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. (2008b). 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. et al. (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. et al. (2005b). A forensically valid comparison of facial composite systems. *Psychology, Crime & Law*, 11, 33–52.

Frowd, C.D., Hancock, P.J.B., Bruce, V. et al. (2011b). 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. & 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., Nelson, L., Skelton F.C. et al. (in press-a). Interviewing techniques for Darwinian facial composite systems. *Applied Cognitive Psychology*.

Frowd, C.D., Pitchford, M., Bruce, V. et al.

improved. We now turn to how this has enabled creation of even more identifiable images.

External facial features

It is known that, for familiar faces, internal features comprise the most important region for recognition, but external features are valuable to some extent. Ellis et al. (1979) presented participants with photographs of intact (whole) faces, or part-faces containing either internal features or external features. Participants named intact faces best (with a mean of 80 per cent correct), then internal features (50 per cent), but external features were still recognised fairly-well (30 per cent). Using the same design (work currently submitted for publication), we found that external features were named at about 7 per cent correct. Given that 'feature' composites constructed after long delays tend to be very poorly named, this finding suggests that recognition of such images is driven mainly by external features – a worrying situation since offenders can easily change them, by getting a haircut, for example.

Research on the role played by context would suggest that accurate external features should facilitate face construction. Memon and Bruce (1983), for example, found that recognition memory for photographs of faces is enhanced when background context is consistent between encoding and retrieval – seeing a person again in a bank is superior to seeing them in a bank and then in a supermarket. A similar effect of congruence is found with hair (Cutler et al., 1987) and with clothing (Sporer, 1993). So, external features act as a contextual cue for recognising internal features. In contrast, other research suggests that external features promote error. Bruce et al. (1999) asked participants to select a target from an array of photographs of unfamiliar faces, and reported that participants tended to base selections on external features, hairstyle in particular. A similar result holds when

participants match composites to target photographs (Frowd et al., 2007a).

These seemingly conflicting results led to the novel development mentioned above: to blur but not remove the external features when witnesses select from EvoFIT face arrays. The idea was to reduce the negative impact of the externals while providing some context for face selection. The approach is very effective: without blurring, constructors produce EvoFITs that are virtually unidentifiable (Frowd et al., 2010).

The impact of external features was followed up over a series of experiments (Frowd & Hepton, 2009; Frowd et al., in press-b). In this work, internal features were constructed to a higher standard when external features in face arrays (1) more closely matched those of the target and (2) were shown with higher levels of blurring. These data suggest that hair, ears and neck are a distraction, but their influence is reduced when they better match the target or are blurred to a great extent. We also found that not presenting external features at all allowed users to produce images that were named twice as often as when externals are added at the start and then blurred – 23 per cent to 46 per cent. So, for face construction, the mere presence of external features is a distraction: users simply cannot ignore them, to the detriment of constructing accurate internal features. Ongoing research indicates that external features similarly interfere with construction of feature composites.

These results are interpretable in terms of holistic face-processing. We (Frowd et al., in press-c) have recently demonstrated that correct naming of a composite face is significantly higher than combining naming from internal and external features seen separately. Simply, recognising a whole face is more effective than recognising the sum of the constituent regions. Due to holistic face recognition, it is also difficult to process separate facial regions when they are brought into register to form an intact face. For face construction, it is therefore difficult to

select faces (in EvoFIT arrays) for their internal features when external features are present (blurred or otherwise).

Curiously, for this same reason, construction of external features may be inhibited by the presence of internal features (work as yet unpublished). At present, at the end of face construction, witnesses select appropriate hair by examining screens of different hairstyles applied to their evolved internal features (we originally thought this was a good idea). This implies that recognition of a composite may be less than ideal because the external features, in particular hair, are not constructed optimally. Current work is exploring this possibility by constructing external features when they are either seen with intact internal features (the current procedure) or with internal features that are blurred (or, better still, masked). In sum, theory would predict that the best procedure is to construct internal and external features independently before fusing them together into a single composite face.

Police field trials of EvoFIT

Laboratory research is an important stage in the development of a commercial product. For composite systems, this simple tenet has been repeatedly overlooked, leading to police being sold system after system and only later being told of product ineffectiveness (e.g. Ellis et al., 1975; Frowd et al., 2005a; Koehn & Fisher, 1997). This is a worrying tale given the importance of composites for catching rapists, murderers, thieves, etc. In 2007, after nine years of research, EvoFIT was promising enough to warrant a full field test; at the time, it produced images that were named with a mean of 25 per cent in the laboratory (Frowd et al., 2009). Then, the interview with witnesses involved free- and cued-recall, but not holistic recall (personality attribution), and EvoFIT presented face arrays with external-features blurring.

Field trials have their own problems, the very issues that laboratory research

(2010). The psychology of face construction. *Applied Cognitive Psychology*, doi: 10.1002/acp.1662
Frowd, C.D., Skelton F.C., Atherton, C. et al. (in press-b). Recovering faces from memory. *Journal of Experimental Psychology: Applied*.
Frowd, C.D., Skelton, F., Butt, N. et al. (in press-c). Familiarity effects in the construction of facial-composite images using modern software

systems. *Ergonomics*.
Hancock, P.J.B., Bruce, V. & Burton, A.M. [2000]. Recognition of unfamiliar faces. *Trends in Cognitive Sciences*, 4(9), 330-337.
Koehn, C.E. & Fisher R.P. [1997]. Constructing facial composites with the Mac-a-Mug Pro system. *Psychology, Crime & Law*, 3, 215-224.
Longmore, C.A., Liu, C.H. & Young, A.W. [2008]. Learning faces from

photographs. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 77-100.
Memon, A. & Bruce, V. [1983]. The effects of encoding strategy and context change on face recognition. *Human Learning*, 2, 313-326.
Milne, R. & Bull, R. [2006]. Interviewing victims of crime, including children and people with intellectual disabilities. In M. Kebbell & G. Davies

(Eds.) *Practical psychology for forensic investigations*. Chichester: Wiley.
Shapiro, P.N. & Penrod, S.D. [1986]. Meta-analysis of facial identification studies. *Psychological Bulletin*, 100, 139-156.
Sporer, S.L. [1993]. Eyewitness identification accuracy, confidence and decision times in simultaneous and sequential line-ups. *Journal of Applied Psychology*, 78, 22-33.



Figure 3: Example EvoFIT constructed in the field trials (left) and a photograph of the person identified from the composite and believed to be responsible for the crime (right). (The relevant police force has consented for these images to be published.)

seeks to control. Examples include length of time that an offender's face is seen, type of encoding (intentional, unintentional), time to construction (two days, one week, one month), circulation of composite (internal, public appeal), and so forth. In contrast, fieldwork does permit exploration of issues that cannot otherwise be evaluated – the effect of extreme stress on composite quality, for instance, as experienced by victims of rape. So, field experiments are also essential for evaluation of composite systems.

Police officers in Lancashire, Derbyshire, Devon and Cornwall, and Romanian forces were trained to use EvoFIT and then evaluated it for a nominal six-month period. EvoFIT was deployed 126 times across a range of crimes, mostly serious, including sexual assault, burglary and theft. Witnesses and victims came from different backgrounds with an age of 10 years and upwards. There were several notable successes in EvoFIT's deployment and one of these is illustrated in Figure 3 – see Frowd et al. (2011b) for more details. Following the trial period, forces audited the crimes in which EvoFIT had been used: names were put forward for around 50 per cent of the composites, and 32 suspects were arrested on the basis of identification from an EvoFIT. This represented an arrest from one in four composites. Despite the many uncontrolled variables, this figure bears obvious similarity to successful naming rates observed in the laboratory for this version of software (Frowd et al., 2010).

The field experiments also provided an assessment of one component or mnemonic of the cognitive interview used to recover a description of the offender's

face. As mentioned above, if witnesses are able, they normally give a free account of the face and then are prompted for more details using cued recall – for example, 'You mentioned that the eyes were oval in shape, but can you remember anything more about them?'. The field trials indicate that the arrest rate was 19.5 per cent when interviewers used cued recall, but this was very much higher when they did not, at 38.5 per cent. Probing in this way prompts witnesses to generate more information about facial features, increasing the likelihood of recalling incorrect details; this is problematic as witnesses rely on featural information to some extent when selecting from EvoFIT face arrays. Training given to police composite-officers has since been updated to reflect these findings; they are now advised to freely elicit a description of the face, but not probe for details about individual features.

Progress and the future

EvoFIT, in common with other composite systems, creates faces in 2D. While this method does not prevent production of a good-quality image, it has some issues. Readers may be aware of Trimensional, a face-scanning application for the iPhone. This app allows a snapshot of a face to be taken, rendered in 3D, and then rotated or made to look happy or 'mad'. While 3D technology is by no means new – scanning systems have been around for decades – expressional change is rendered more accurately in 3D than in 2D. In EvoFIT, 20 'holistic' scales also vary the overall appearance of a face in plausible ways, as Figure 2 illustrates, but sometimes errors are produced since

depth information is not taken into account.

A 3D EvoFIT model is a sensible technical solution to avoid rendering errors, but there are other benefits. In 3D, perspective can be readily modelled: up close, faces appear different from those seen from further away: presenting faces that match the witness's own experience seems likely to facilitate face selection. In addition, faces are not always seen front-on – an offender observed through the side window of a car, for instance – and so specifying the view in which faces are shown may also be advantageous.

As a final note, it is a pleasure to be able to report that the progress of EvoFIT has been extraordinary: the enhancements have allowed naming to more or less double every three years. Current work is establishing the value of combining techniques: holistic interview, omitting the cued-recall interviewing mnemonic, internal-features construction and animated caricature. It is likely that composite naming from such a version will be in the region of 60 per cent correct – that is, following construction from long delays tested using the gold-standard procedure.

Based on the field trials, EvoFIT now leads to the arrest of a suspect in about 40 per cent of cases – performance that seemed inconceivable a few years ago. The outcome of all this work becomes so worthwhile when a dangerous person is taken off our streets, such as in the conviction of rapist Asim Javed in 2010 after the use of an EvoFIT in a public appeal (see tinyurl.com/3hme7y5). The police now have access to technology that is effective in the fight against crime, and at a time when budgets demand effective tools. Our intention is to continue improving the system through cycles of laboratory research, police feedback and field trials.



Charlie Frowd
is Senior Lecturer in the School of Psychology at the University of Central Lancashire
cfrowd@uclan.ac.uk

Faye Skelton
is at the School of Psychology, University of Central Lancashire

Chris Atherton
is at the School of Psychology, University of Central Lancashire

Peter Hancock
is at Psychology, School of Natural Sciences, University of Stirling