Global Journal of Human Social Science
Volume 11 Issue 3 Version 1.0 April 2011
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
ISSN: 0975-587x

Catching more offenders with Evofit Facial Composites: Lab Research and Police Field Trials

By Charlie D. Frowd, Peter J.B. Hancock, Vicki Bruce, Faye C. Skelton, Chris Atherton, Laura Nelson, Alex H. McIntyre, Melanie Pitchford, Rebecca Atkins,Andrew Webster, John Pollard, Beverley Hunt, Emma Price, Sandra Morgan,Roz Greening, Adrian Stoika, Romeo Dughila, Sergiu Maftei and Gabriel Sendrea

University of Central Lancashire

Abstract : Often, the only evidence of an offender's identity comes from the memory of an eyewitness. For over 12 years, we have been developing software called EvoFIT to help eyewitnesses recover their memories of offenders' faces, to assist police investigations. EvoFIT requires eyewitnesses to repeatedly select from arrays of faces, with 'breeding', to 'evolve' a face. Recently, police forces have been formally evaluating EvoFIT in criminal cases. The current paper describes four such police audits. It is reported that EvoFIT composites directly led to an arrest in 25.4% of cases overall; the arrest rate was 38.5% for forces that used a newer, less detailed face-recall interview. These results are similar to those found in the laboratory using simulated procedures. Here, we also evaluate the impact of interviewing techniques and outline further work that has improved system performance.

Keywords: facial composite, witness, victim, EvoFIT, recognition, memory, interface, crime.

Classification: GJHSS-A Classification: FOR Code: 170102,170104,170202

Strictly as per the compliance and regulations of:

Catching more offenders with EvoFIT Facial Composites: Lab Research and Police Field Trials

Charlie D. Frowd[1], Peter J.B. Hancock[2], Vicki Bruce[3], Faye C. Skelton[1] , Chris Atherton[1], Laura Nelson[1], Alex H. McIntyre[2], Melanie Pitchford[4], Rebecca Atkins[1], Andrew Webster[5], John Pollard[6], Beverley Hunt[7], Emma Price[7], Sandra Morgan[7], Roz Greening[8], Adrian Stoika[9], Romeo Dughila[10], Sergiu Maftei[10] and Gabriel Sendrea[10]

Abstract - **Often, the only evidence of an offender's identity comes from the memory of an eyewitness. For over 12 years, we have been developing software called EvoFIT to help eyewitnesses recover their memories of offenders' faces, to assist police investigations. EvoFIT requires eyewitnesses to repeatedly select from arrays of faces, with 'breeding', to 'evolve' a face. Recently, police forces have been formally evaluating EvoFIT in criminal cases. The current paper describes four such police audits. It is reported that EvoFIT composites directly led to an arrest in 25.4% of cases overall; the arrest rate was 38.5% for forces that used a newer, less detailed face-recall interview. These results are similar to those found in the laboratory using simulated procedures. Here, we also evaluate the impact of interviewing techniques and outline further work that has improved system performance.**

Keywords- facial composite, witness, victim, EvoFIT, recognition, memory, interface, crime.

I. Introduction

Witnesses and victims may be asked to construct a picture of a person they have seen committing a crime. This picture is known as a facial composite, and can help the police identify the offender.

Manuscript received on 10th February 2011.

About [1]: Dr Charlie Frowd, Department of Psychology, University of Central Lancashire, Preston PR12HE. cfrowd@uclan.ac.uk+441772893439
About [2]: Professor Peter J.B. Hancock, Department of Psychology, University of Stirling, Stirling, UK. pjbh1@stir.ac.uk
About [3]: Professor Vicki Bruce, School of Psychology, Newcastle University, Newcastle, UK. vicki.bruce@newcastle.ac.uk
About [4]: Melanie Pitchford, Department of Psychology, Lancaster University, Lancaster, UK. m.pitchford@lancaster.ac.uk
About [5]: DCI Andrew Webster. Force Major Investigation Team, Lancashire constabulary,UK. andrew.webster@lancashire.pnn.police.uk
About [6]: John Pollard, Family Support Unit, Lancashire constabulary, UK. john.pollard2@lancashire.pnn.police.uk
About [7]: Beverley Hunt, Forensic Services Unit, Derbyshire Constabulary, UK. beverley.hunt.6348@derbyshire.pnn.police.uk
About [8]: Roz Greening, Major Crime Investigation Team, Devon and Cornwall Constabulary, UK. roz.greening@devonandcornwall.pnn.police.uk
About [9]: Adrian Stoika, NASA Jet Propulsion Laboratory, NASA, USA. astoica@jpl.nasa.gov
About [10]: Romeo Dughila, Forensic Department of Iasi County Police Inspectorate, Romania. romeo_dughila@yaho

Various techniques are available to produce composite images. Traditionally, eyewitnesses described the face in detail and then built the composite by selecting individual facial features: hair, eyes, nose, mouth, etc. However, we do not perceive faces in such a piecemeal fashion, but instead process it more as a complete entity (Tanaka & Farah, 1993). For this reason, face construction using a 'feature' approach often results in poor quality images (e.g. Brace, Pike, Allen & Kemp, 2006; Bruce, Ness, Hancock, Newman & Rarity, 2002; Davies, van der Willik & Morrison, 2000; Frowd et al., 2005b), especially when carried out several days or longer after the face is originally seen (e.g. Frowd, Hancock & Carson, 2004; Frowd, McQuiston-Surrett, Kirkland, & Hancock, 2005; Frowd et al., 2005a); such delays are typical in eyewitness composite construction. More recent 'holistic' approaches (Frowd et al., 2004; Gibson, Solomon & Pallares-Bejarano, 2003; Tredoux, Nunez, Oxtoby & Prag, 2006) have attempted to overcome the feature issue by encouraging witnesses to focus on the face as a whole. In essence, users repeatedly select complete faces from screens of alternatives, with cycles of 'breeding' in between, to allow a composite to be 'evolved'. This method may be particularly useful when the witness has seen an offender's face, but cannot describe it in detail: under such circumstances, traditional methods are not effective, since these require good descriptions for locating accurate subsets of features to be shown to the witness. In contrast, 'holistic' systems only require fairly general information: age, gender and race; holistic methods therefore have the potential to facilitate construction even in the absence of detailed feature descriptions.

The focus of the current paper is on one of these methods, EvoFIT. This system has been the focus of considerable research and development in the laboratory (e.g. Frowd et al., 2007a, 2008b). For the last three years, EvoFIT has also been the subject of formal police field trials. These have explored the effectiveness of the system when used with actual witnesses and victims of crime. The results of three such evaluations

were presented as a conference paper in Frowd et al. (2010a); what follows is a revised version of this work that includes an evaluation by a fourth police force, a discussion on the use of interviewing techniques and police practice for face construction, and an overview of more recent developments.

a) Background to EvoFIT

The EvoFIT system has been comprehensively described in several published papers (Frowd et al., 2004, 2007a, 2010b)—for brevity, only an overview will be given here. EvoFIT was conceived in the 1990s by Peter Hancock (2000). He developed a computer program that presented arrays of whole faces. The faces were produced using a statistical technique called Principal Components Analysis (PCA) that captured variations in feature shape and greyscale colouring (or texture), and enabled further faces to be synthesized, initially with random characteristics. Users provided a goodness-of-fit rating for each face and a Genetic Algorithm (GA) combined their preferences (using proportional-fitness selection) to produce more items for selection. After a few iterations, faces in the set progressively resemble each other and the target face. The best likeness produced was saved as the composite. Peter's prototype was developed into a full system as part of the first author's Ph.D. work (Frowd et al., 2004). This led to development of a PCA model that generated white male faces between 18 and 35 years of age. Users would choose a hairstyle and then select from screens of complete faces. However, they found this procedure difficult, as some faces tended to be accurate by shapes of features, while others were more accurate by feature colouring and skin tone. These two aspects of faces are sometimes referred to as *shape* and *texture* (respectively). The solution was to present screens of shape for selection followed by screens of texture. Users then identified the most accurate likeness, a 'best' face, that was given twice the number of breeding opportunities in the GA and was also carried forward intact to the next generation as part of an 'elitist' strategy (to avoid 'damage' occurring to the face by genetic recombination and mutation operators). At this stage, EvoFIT was used in a police investigation in the Northants area. See Fig. 1 and Frowd, Bruce, Storås, Spick and Hancock (2006c) for details. The person responsible was later identified using 'familial' (family) DNA matching, and convicted. Early versions of EvoFIT did not reliably converge on an identifiable likeness in the laboratory (Frowd et al., 2004). This was in spite of adhering to UK working practices that aim to give optimal results, including use of a cognitive interview (CI) to help witnesses recall details of the face before starting face construction. A breakthrough emerged when selection of the 'best' face was refined: after users had selected shapes and textures, these were shown together in combination (each possible facial shape shown with each possible facial texture) for identifying the best likeness.

Fig. 1. The EvoFIT (left) and person (right) convicted in the 'Beast of Bozeat' case. Shortly after the crime, the perpetrator is believed to have changed his hairstyle in an attempt to conceal identity, as illustrated here.

An evaluation of this version of the software was carried out. Fifty laboratory-witnesses saw a photograph of a footballer whose face was unfamiliar to them, and two days later described the face (using a CI) and constructed a composite with EvoFIT or a traditional feature system. The resulting images were then given to football fans to name. Among witnesses who attempted to remember the face in detail, EvoFITs were correctly named at 11% and feature composites at 4% (Frowd et al., 2007b). In subsequent research (Frowd, Bruce, Plenderleith & Hancock, 2006b), we asked the same person to use the system more than once to construct a likeness of the same target face. There was good consistency of results, as Fig. 2 illustrates. When used in this way, the faces the user sees at the start change for each attempt—they are different random faces—and so the search process is also somewhat different each time,as is the resulting image.

Fig. 2. EvoFIT Images produced of TV presenter, Anthony (Ant) McPartlin. The same person used the system twice to evolve a composite of his face from memory, producing a consistent likeness.

Funding was sought from UK Government to further improve the software. We first sought to limit age expressivity, since sometimes faces were evolved that

portrayed age inaccurately. This work developed four databases of white male faces, segregated by age, to enable composite construction for offenders aged 17 years of age and older. Each of the databases was built using PCA as before, and in greyscale, as research suggests that face construction does not benefit from the use of colour images (Frowd et al, 2006b).

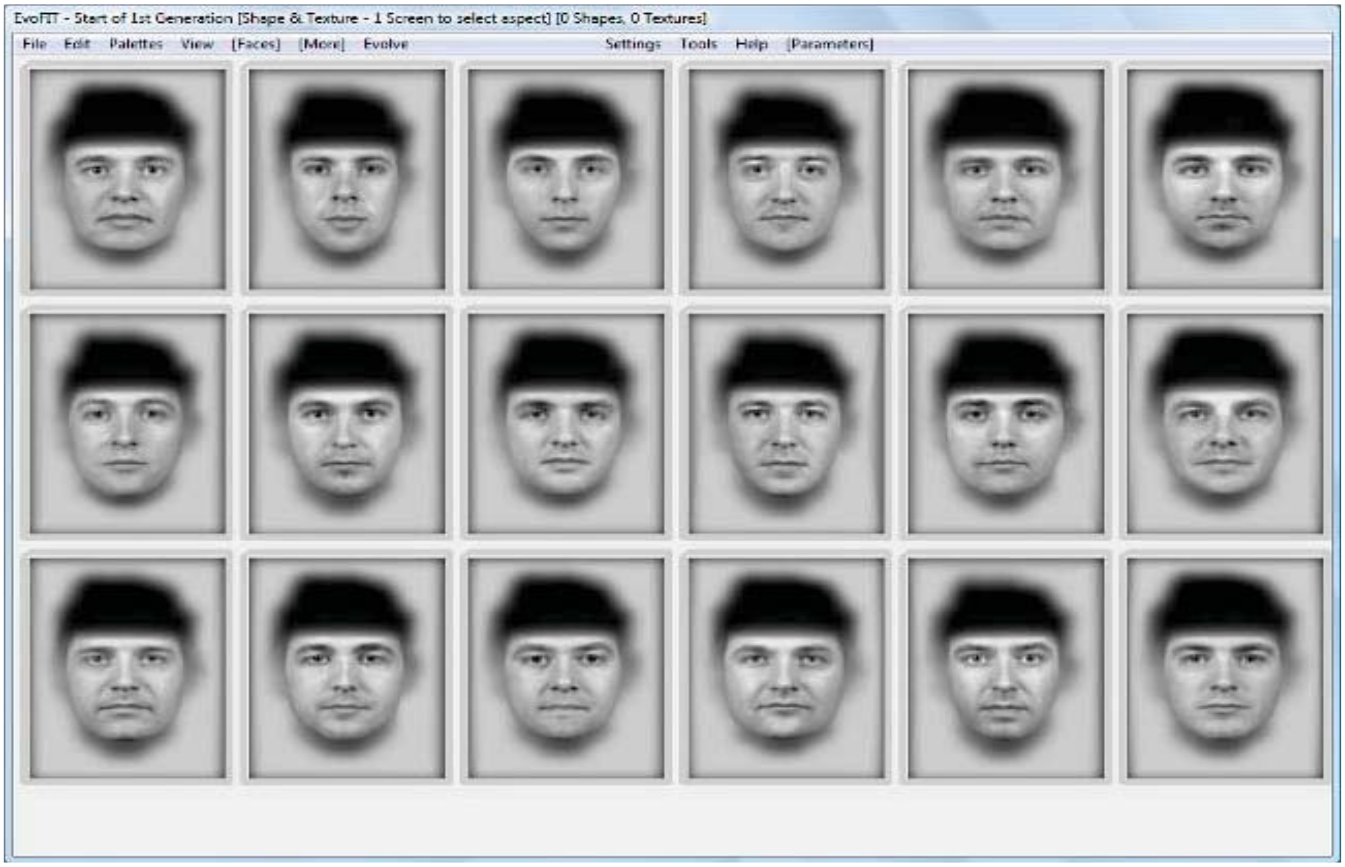

Fig. 3. Witnesses select from screens such as this. The external parts of the face are blurred to help witnesses concentrate on the central facial region.

Following development of these age-constrained databases, users still sometimes evolved faces with inaccurate ages, though to a lesser extent than before. We sought to overcome the problem by providing a sliding scale for adjusting composites' perceived age, and extended this facility to allow adjustment of other whole-face properties. These so-called *holistic tools* included face weight, masculinity, threatening, attractiveness, honesty and extroversion. See Fig. 4 for examples, and Frowd, Bruce, McIntyre, Ross and Hancock (2006a) for a description of how the scales were designed. Further scales were developed to add stubble, eye-bags and deep-set eyes, and to alter the greyscale levels of brows, irises, mouth creases, etc. These holistic tools are used at the end of evolving, after external feature blurring is turned off.

Fig. 4. An example of the ageing (top row) and pleasantness (bottom row) holistic scales. Manipulations increase in magnitude from left-to-right and are illustrated on an EvoFIT of TV celebrity, Simon Cowell.

Example celebrity EvoFITs constructed from memory using this version of EvoFIT are illustrated in Fig. 5.

Fig. 5. EvoFITs made from memory. From left to right (top row), they are of footballer, David Beckham; actor, David Tennant; former US president, George W. Bush; singer, Noel Gallagher; and (bottom row) footballer, John Terry; singer, UK politician, Gordon Brown; singer, Robbie Williams; and UK politician, Nick Clegg.

In an evaluation of this version of EvoFIT, in a similar design to Frowd et al. (2007b) that used unfamiliar target faces after a two-day delay, EvoFITs were correctly named at 24.5%, while those constructed from a feature system were named at 4.2% (Frowd et al, 2010b). Both external-features blurring and holistic scales were shown to be effective in improving composite naming. More recent research using similar face construction procedures has shown comparable levels of correct naming from EvoFIT: 24.1% (Frowd et al., Submitted-a) and 22.7% (Frowd et al., Submitted-b).

b) Enhancing performance further

There have been other attempts to improve the effectiveness of EvoFIT (Frowd et al., 2006b, 2007a, 2007c, 2008b). One of these involves changing the mode of presentation when publishing an image in the media. This is based on the idea that composites tend to appear quite similar to each other and that this lack of distinctiveness can make recognition difficult for members of the public, etc. Exaggerating facial distinctiveness may therefore help to overcome this problem. In a series of experiments, described in full in Frowd et al. (2007c), composite naming improved considerably when participants observed a composite while it was first progressively caricatured, by exaggerating the shape information in the face, then de-emphasized, by rendering this information more average. An example of the animation procedure can be found online by visiting http://www.uclan.ac.uk/animatedcomposite.

Correct naming using this technique was found to increase by more than 40% overall, and the benefits of caricature animation were shown to extend to sketch-based images and composites from feature systems, as well as composites from EvoFIT. Animated caricatures delivered the greatest benefit for poorly-named composites, which should allow this technique to be beneficial to traditional composites produced in criminal investigations. However, even good-quality images were recognised somewhat better using this technique. For the version of EvoFIT that was used in the following field trials, correct naming of its composites should increase from 24%, as mentioned above, to around 42% when viewed with caricature animation.

II. Field Trials

As can be seen from the above summary, considerable time has been spent developing EvoFIT in the laboratory, to ensure as far as possible that it operates effectively using police procedures: specifically, that it can produce a recognizable image from a person's memory of an unfamiliar face seen several days previously. Having taken about ten years to achieved this objective, we initiated formal field trials with the police. There are clearly aspects of system use that can be only tested in the field—for example, the effects of stress on composite production, such as those experienced by victims of stranger rape, cannot be properly established in a laboratory setting.

a) Measures of success

Measuring system performance in the field is not without its own difficulties; often these are the very issues that laboratory studies attempt to minimize. In the lab, users can see a target face for a fixed amount of time, under good lighting conditions and without distraction; these helpers can also be asked to construct a composite after a consistent interval of time. When a composite has been made, other people who are familiar with the target's identity can be asked to evaluate the quality of the face, by attempting to name it. All of these variables (and others, e.g. Frowd et al., 2007b) can affect whether a composite is recognised.

In police work, when a composite is published in the media, it is normally accompanied by other information that can help to trigger the correct identity: a description of the person (e.g. age, build and height), the modus operandi (e.g. assault, murder and deception), crime location, etc. In the lab, this information is generally not provided. In contrast, the more people who see a given composite, or 'wanted' poster, the higher the likelihood of correct identification; real-world composites may not be recognized if their circulation (police officers, members of the public, newspapers, TV) is poor.

To complicate matters even further, there are different definitions of 'success'. In the laboratory, success can be taken as the number of times a composite is correctly named. For example, if 20 people are shown a

composite and 6 of these observers correctly name it, this provides a correct naming level of 6/20, or 30%. Laboratory research can also consider the number of incorrect names given (e.g. Frowd et al., 2010b). Beyond the laboratory, a composite is valuable if it assists in some way in locating the perpetrator of a crime. This can generally be measured by: (1) the composite being named, (2) an arrest or (3) a successful conviction. In the UK, to limit wrongful conviction, convictions are not based on eyewitness evidence alone, since eyewitness identification and testimony can be inaccurate (e.g. Rattner, 1988). Evaluations based on composites that have triggered an arrest and which then lead to successful conviction would be the ideal measure. However, convictions can take considerable time to secure, thus making field evaluations rather lengthy. A sensible compromise, and an approach supported by Senior Investigating Officers (SIOs), is to base evaluations on (1) or (2). These two measures make good sense as they are what a SIO requires: a suspect on whom to focus enquiries. In the current work, the various EvoFITs constructed were audited within a census date of about a month of forces completing their trial.

b) Interviewing for producing composites

Witnesses and victims who construct composites are first given a cognitive interview to help them recover the memory of an offender's face. This interview is based on considerable work carried out by Ron Geiselman and his colleagues in the US (for a review, see Wells, Memon & Penrod, 2007). It is based on a number of cognitive techniques, *mnemonics,* adapted for obtaining accurate descriptions of faces (e.g. Frowd et al, 2005b). We have also developed the cognitive interview, specifically the face-recall interview used as part of composite construction (e.g. Frowd et al., 2008a, Submitted-a). The following paragraph provides an overview of how cognitive interviewing is typically used in police work; we outline our own developments later in this report.

The face-recall interview varies somewhat from operator to operator, but generally begins as a fairly informal conversation between witness and operator, with the aim of relaxing the witness and facilitating recall. Following this, witnesses are encouraged to think about the crime scene, their internal state (i.e., what they were thinking and feeling at the time, although this part is normally omitted for particularly traumatic offences such as rape), and some general characteristics of the offender (e.g. build, height, clothing)—a mnemonic technique known as *reinstatement of context.* Next, they are asked to describe the offender's face in their own time and in as much detail as possible, but without guessing. Police operators record this *free recall* and do not interrupt while it is taking place—except to ask a witness to slow down, if he or she is speaking too quickly for written notes to be made. Following this, operators may read back the given description for each feature and then pause, to request for further recall. This technique known as *cued recall*. For example, a witness might be reminded that they previously described the offender's eyes as 'small and light in colour'; when prompted, they might now also recall that the offender's eye shape was 'oval' and there were 'bags' under them. When the interview is complete, the session moves on to composite construction. During the field trials described here, instructions in cognitive interviewing for use with EvoFIT were provided as required.

c) Lancashire police trial

Prior to our involvement with them, Lancashire police force had used one of the UK's feature systems in twenty or so investigations, but had not found its composites helpful. The first formal evaluation of EvoFIT was carried out within this force, running from autumn 2007 to spring 2008. The project was assisted by funding from Crime Solutions, UCLan, UK. We used a version of EvoFIT containing the white-male database for constructing faces of offenders aged 17 years and older; a younger, teenage version was added during the trial. EvoFIT was used in conjunction with the PRO-fit composite system, to permit the inclusion of hats, glasses and other accessories.

A training course was developed and administered by the system designer (CDF) and the force's existing composite officer (JP). The course involved: training on the cognitive interview for obtaining facial descriptions from witnesses; EvoFIT system training; exhibiting of evidence for later use in court; software paint-package training, for the addition of shading, wrinkles, etc.; and considerable practice in all of these components. A total of 21 police officers and staff were trained, in order to provide representation at force headquarters and in each division; they were supported during the evaluation by the current composite officer as well as the system designer. After construction, composites were circulated within the force for identification, and some were published in the newspapers, on TV and on a 'wanted persons' webpage. The webpage also used the animated-caricature format.

The system was reported to work well with witnesses and victims, and feedback was used to improve EvoFIT's usability for police operators: e.g., improvements to information shown in the title bar, and the display of messages if digression occurred from the recommended procedure. Also, midway through the evaluation period, the construction procedure was improved to allow witnesses to set an appropriate facial aspect ratio—face width and length—from the start of

the construction procedure. This facilitated face selection generally. In addition, to allow better adjustment of an evolved image, an additional holistic scale was added to allow manipulation of face width.

During the trial, 30 EvoFITs were constructed, mainly for serious crimes such as sexual assault and distraction burglaries, but also for less serious offences such as minor theft. Six arrests were made, a success rate of 20.0%. The six-month trial of EvoFIT led to a number of notable successes, detailed below. The constabulary continues to use the software.

Case studies

Rape of female under 13 years: The first case occurred in August 2007. This involved a sexual assault on an 11 year old girl in Stanley Park, Blackpool. The assailant was described to be white male, 16 to 20 years of age with a slim build and dark, short, stubby hair, lighter at the tips. He was also described as wearing a dark blue tracksuit, pale blue vest and black trainers. During the crime, he was reported to have stolen a mobile phone. The week following the crime, two people were arrested, but were later eliminated from the enquiry. Due to the absence of further leads, an EvoFIT was constructed seven days later, see Fig. 6.

A public appeal was made in which police detectives and support officers attended the park at the same time as the crime had occurred the previous week. The appeal was based on showing members of the public the EvoFIT composite, and asking whether they recognised him. Two people named the EvoFIT as a local person, Ross Gleave, and placed him in the vicinity at the time of the attack. His name was also given following house-to-house enquiries. The police attended Gleave's home address and made an arrest. The description given by the victim was accurate and the stolen property was recovered from his address. Gleave was later identified by a number of other witnesses. These additional observers did not know him, but picked him out using VIPER, a system for conducting video line-ups (Video Identification Parade Electronic Recording). Gleave was convicted for the attempted rape of a child under 13 years, jailed for seven years and placed on the Sex Offenders Register for life.

Fig. 6. The EvoFIT produced in the Stanley Park assault, left, and the person subsequently identified and convicted of the offence, right.

Sexual assault: EvoFIT was valuable in solving another sexual assault case, this time in an unexpected way. The offence on a teenage male took place in Morecambe Bay; the victim subsequently produced an EvoFIT using the 50- year-old white-male database. A DNA sample of the offender was available and the police used this evidence to try to find a match. The DNA search, however, produced a dozen partial matches, but a photograph from one of them bore a strong resemblance to the EvoFIT. Consequently, police attended this person's home address first and were able to collect evidence linking him to the scene of the crime (a train ticket). The EvoFIT provided valuable intelligence to guide the investigation; it also reduced the amount of police time wasted following false leads.

d) Derbyshire police trial

For 12 months starting June 2008, Derbyshire police began a field trial of EvoFIT, with three composite officers being trained in its use. The version of EvoFIT used was similar to Lancashire Constabulary's, but with more databases. EvoFIT was found to work well and eyewitnesses reported being very satisfied with the likenesses produced. The force constructed 57 composites during the year of the trial, about twice the number made in Lancashire (perhaps sensible, since the Lancashire trial ran for half as long). Use of EvoFIT was considered successful when police obtained one of three outcomes: a name put forward, an arrest, or a person charged. These data are summarized in Table 1. Note that the actions depicted in the table are not mutually exclusive; for example, 7% of the figure for arrests (19.3%) also involve persons who were later charged. It can be seen that there were roughly twice the number of names put forward (by police officers, members of the public) than arrests made. The table also shows that about one-third of suspects were charged at the census date. The arrest rate was very similar to that found in the Lancashire trial (20.0%).

Table 1. Results of the Derbyshire police trial. Figures relate to successful actions arising from EvoFITs and are expressed as a percent of the total number of composites constructed (there were 57 in this evaluation).

Names put forward	Arrests	Charged
43.9	19.3	7.0

Case studies

Indecent exposure: early on in the Derbyshire field trial, EvoFIT was used in an indecent exposure incident. In this case, the female victim was pushing her newborn baby in a pram at the time of the offence, and afterwards reported having been terrified that the offender would harm her child.

The offender was described as a white male, approximately 30 to 35 years of age. The victim produced an EvoFIT of him two days after the offence using the 30 year (Western European) white-male database; the victim was very happy with the likeness produced. The image was taken by the police operative to the local police station for circulation within the force, where the face was recognised by local officers. Within four days, the offender had been arrested, charged and remanded in custody. He was sentenced to 16 months imprisonment at Crown Court and placed on the Sex Offenders Register.

Connected thefts: An EvoFIT image was produced by the victim of a 20- to 30-year-old Eastern European male. The man had approached her on the street and stolen a bank card from her purse. The victim had felt particularly vulnerable, being on a disability scooter. She was delighted with the likeness produced and was most impressed with the system.

Following a separate incident that occurred a week later with an almost identical *modus operandi* (method of operating), the victim of that crime produced a second EvoFIT image showing a 35-year-old Eastern European male. As there was a chance that these crimes might have been committed by the same person, a different police operative interviewed the second victim (different interviewers are used in such cases to avoid the possibility of subsequent images being unknowingly contaminated by the same interviewer). This incident had taken place 10 miles from the first, but the EvoFIT image produced were almost identical; for this reason, the crimes were linked, providing valuable intelligence to the investigation.

e) Devon and Cornwall police trial

Devon and Cornwall police have two officers who construct composites. One received EvoFIT training in January 2010 and used the system for an audited period of four months. Fifteen composites were constructed during this time. One of the images emerged as part of a bogus complaint, and helped to show that the complainant was lying. Of the remaining 14 EvoFITs, a name was put forward for 12 of them (80.0%) and an arrest warrant was issued for six (40.0%); only two EvoFITs remained unnamed (13.3%) at the census date.

Case studies

Sexual assault: one of the first EvoFITs constructed by this force was of a sexual assault offender. The incident was reported to have occurred in Plymouth, January 2010, on a female victim. Initial enquiries in the investigation were made to try to locate him, but these proved unsuccessful. CCTV also failed to provide useful leads and, despite a media campaign, no suspects could be identified. Three weeks after the incident, an EvoFIT was constructed using the Asian-male database. The victim was amazed by the lifelike image, and repeatedly said that 'it was just like him'. The EvoFIT was released in the media and several names were put forward: many people gave information about workplaces and addresses of the putative offender. Subsequent enquiries revealed that the likely culprit was an illegal immigrant who had "gone to ground" on the day that the image appeared in the media. His details have been circulated on the PNC (Police National Computer) by both UK Borders Agency and the OIC (Officer In Case). Enquiries to locate him are ongoing.

Sexual assault: late January 2010, a young female reported a serious sexual assault in Exeter. A description of the offender was circulated to local officers and a public appeal was made in the press. From this, several identifications were made by members of the public that resulted in a number of people being interviewed; however, these were all eventually excluded from the investigation. Four days after the incident, an EvoFIT was constructed by the victim. Although still distressed about what had happened, she found the procedure easy to follow and was able to complete a composite using the black-male database. The composite was circulated throughout forces in Exeter and then in the local press.

Two weeks later, a male contacted the enquiry. He said that he had been in the Exeter area where the offence had taken place, at the material time; he also said that the facial composite looked just like a photograph of him. It emerged that teammates with whom he used to play football had recognised him as the offender and had given him an ultimatum of contacting the police himself, or they would do it for him. Ultimately, no charge was brought against him, since he claimed that the sex had been consensual.

The case in general involved a great deal of time, money and effort. The alleged offender was not known to the police prior to the investigation, and so would not have been identifiable by DNA, description or *modus operandi*. Again, the composite was the valuable lead; without EvoFIT, the enquiry would have been even more protracted and costly, and the case may well have remained unsolved.

f) Romanian police trial

Further work was carried out to allow EvoFIT to be used in Romania. This was made possible by collaboration between authors CDF and AS, and the Forensic Department of Iasi County Police Inspectorate, Romania. Funding was provided by Crime Solutions, UCLan. At the start of the collaboration, the available EvoFIT databases allowed construction of Western European white (Identity Code 1, IC1), Afro-Caribbean black (IC3) and Asian offenders (IC4), plus offenders of

mixed-race parentage (e.g. white-black). It did not, however, accurately render the skin tones of Eastern European faces (IC2). While there are obvious similarities in skin pigmentation and facial features between Western and Eastern European faces, differences in physiognomies resulted in poor likenesses when constructing a face using the other race database.

In accordance with Romanian legislation, we entered into an agreement allowing an Eastern European male database to be created and then evaluated for use in criminal investigations in Romania. This involved photographs, taken by the Romanian police, of about 200 male faces, each showing a front-face view under controlled lighting. During the trial period, enhancements were made to increase the number of hairstyles available within the system and to initiate development of a female Eastern-European database. In addition, a mixed-race database was developed to cater for mixed-parentage offenders having both Eastern European and Asian ethnicity. This 'minority male' database was built with PCA using an equal number of faces from both of these racial types. The effectiveness of the newly-designed Eastern-European database was evaluated in the laboratory, as part of a research project by author RA. This involved asking people to construct Western and Eastern European male faces using the EvoFIT Western and Eastern male databases. It was found that better quality composites were produced when the race of the target matched the race of the database, as one would expect.

Author CDF traveled to Iasi in June 2009 to install the software and to provide training for two experienced police officers. Over a five-month period, EvoFIT was used 24 times, and this resulted in the location of nine suspects, corresponding to an arrest rate of 37.5%.

Case studies

Mobile phone thefts: Between May and August 2009, a series of very similar crimes was committed against minors, particularly against those aged 12 years and under. The offender in this investigation chose buildings with an elevator, to enable him to follow his victims into the elevator. Between floors, he stopped the lift and, under threat of violence, stole the young person's mobile phone.One of the victims, a 10-year-old girl, was interviewed to construct a composite of the offender. She could not describe the robber's facial features—a problem arising for many victims—but was able to produce an EvoFIT. The 23 to 35 year Eastern European male database was used; the resulting image is shown in Fig. 7.

Fig. 7. The left image is an EvoFIT produced by a 10 year old victim of robbery; the right image is a recent photograph of the person convicted.

The composite was released to local police forces. After a month, police detained a person with notable similarities to the composite. The man was later convicted and sentenced to 7 years in prison.

Shimano bike thief: EvoFIT also proved valuable for detecting a fairly-prolific bike thief. This involved four thefts of bicycles between May and August 2009, with the thief cutting safety locks. Two EvoFITs were constructed by eyewitnesses at Iasi Police Headquarters, leading to the arrest of the person shown in Fig. 8.

Fig. 8. These EvoFITs (left and centre) of a bicycle thief were constructed by separate witnesses over a period of two months. On the right is a photograph of the person believed to be respsonsible for committing these crimes.

Violent robbery: EvoFIT also proved valuable for detecting another offender who committed two robberies on one day. He threatened victims with a knife and then punched them in the face. Two of his victims constructed an EvoFIT at the Iasi Police Headquarters. One victim used the 23-to-30-years Eastern European male database (Fig. 9, left); the other, the 'minority male' database (Fig. 9, centre). The EvoFITs were produced between two and three days after the offences had taken place. The EvoFITs were released to the local police forces. Within a couple of hours, based on these images, the suspect was named by young people who lived in a neighbourhood near to where the robberies had been committed. The offender was convicted and sentenced to prison.

Fig. 9. The EvoFITs (left and centre) of an offender were constructed by separate victims. On the right is a photograph of the person believed to be responsible for two robberies.

Deception: A further noteworthy case involved a person who reported being robbed of a large sum of money. An EvoFIT was constructed as normal. As the session was nearing the end of completion, however, the complainant appeared to become rather agitated. It turned out that the 'victim' had both described and attempted to construct a composite of himself! He retracted the allegation of robbery.

III. Discussion

A range of techniques are available to law enforcement for constructing facial composites. Most use a feature-by-feature approach, which is an unnatural task for eyewitnesses, but new methods are emerging based on the selection and breeding of complete faces. The current work considers one such system: EvoFIT. This system presents arrays of whole faces for witnesses to repeatedly select and a composite is 'evolved' over time. EvoFIT is the result of considerable research and development, and performance in the laboratory is now consistently good; here, we report use and testing for effectiveness by four different police forces.

Feedback from the field trials improved both system ergonomics (e.g., better reporting of session status) and composite quality (e.g., facial aspect setting and new face-width holistic tool). The work revealed software bugs, allowing them to be rectified. As discussed below, the field trials have also provided insight into the most appropriate interviewing method for use with witnesses and victims.

Overall system effectiveness was also measured, based mainly on arrests arising from composite identifications. Reports across the forces for arrests were 20.0%, 19.3%, 40% and 37.5% of the total number of composites constructed—these totals were 30, 57, 15 and 24 respectively for Lancashire, Derbyshire, D&C and Romania. Based on the total number of arrests made (6+11+6+9=32) and the total number of composites constructed (30+57+15+24=126), the mean arrest rate was 25.4%. In spite of the large number of uncontrolled variables in field evaluations, this figure is comparable to 23.8% mean correct naming for EvoFIT measured in the laboratory (24.5%: Frowd et al., 2010b; 24.1%: Frowd et al., Submitted-a; 22.7%: Frowd et al., Submitted-b). So, based on two measures—arrest rate and mean correct naming—EvoFIT's performance is remarkably similar in the laboratory and in the hands of the intended user.

a) Interviewing styles

Arrest rates from the four forces indicate that EvoFIT's performance approximately doubled form earlier (19.5%) to later (38.5%) evaluations. While the basic procedure for operating EvoFIT changed little over the course of the field trials, with the exception of asking witnesses to select facial aspect at the start and adding a new holistic scale, there was a notable change in the interview method used by operators to help witnesses recover facial information about offenders.

In the early trials with Lancashire and Derbyshire, the interview aimed to help witnesses recover as much accurate information as possible and, as outlined above, this included free recall and cued recall. Police operators would then reflect on this information at the end of the session when a witness was making final enhancements to the face—when manipulating shape and placement of individual features using the Shape Tool, and when using the paint program. In later trials, less information was sought at the initial stage. Operators still requested free recall, but they did not proceed to *cued recall*, which would have involved prompting the witness for more accurate detail of each facial feature. Instead, this information was requested later in the session when required (during Shape Tool and artwork use). These two similar methods of interviewing both produce composites with good arrest rates, but the latter is clearly better. We now believe that we understand why.

It turns out that describing another person's face in detail can have an unfortunate side effect for that person: temporary interference in ability to recognise a face (e.g. Schooler & Engstler-Schooler, 1990). This rather unintuitive cognitive mechanism is known as the *verbal overshadowing effect* and has been extensively researched (see Schooler, 2002, for a review). There are several potential reasons why recognition is interfered with in this way. For example, after extensive recall, witnesses may continue to have considerable focus on individual features; this is likely to be problematic as faces are recognised more accurately when perceived as a complete entity rather than by their constituent parts (Tanaka & Farah, 1993). In addition, as inaccurate information tends to be recalled more often following extensive recall (e.g. Finger & Pezdek, 1999), witnesses may select facial parts (for a feature system) or whole faces (EvoFIT) that resemble these poorly-recalled features, promoting worse-quality composites.

The main part of what witnesses do when constructing faces involves recognition—they select individual facial features (eyes, nose, mouth, etc.) if building a 'feature' composite, or whole faces (from arrays) with EvoFIT. We ourselves have shown that asking a person to recall a face in detail does promote a less identifiable image from a feature system compared to when a person builds the face without having given any description (Frowd et al., in press). With EvoFIT, the issue seems to relate to extent of recall: we now know that recalling a face in detail does promote a more identifiable EvoFIT image than not recalling the face at all (Frowd et al., Submitted-a), but what is becoming apparent is that the level of detail being requested has previously perhaps been too great, potentially causing overshadowing-type effects. For this reason, as the field trials would suggest, asking very detailed information about individual features is probably best postponed until later in the session.

It is worth mentioning that we have recently developed a 'holistic' cognitive interview (H-CI) that appears to overcome some of the problems associated with face recall. In this interview, witnesses describe the face using free and cued recall, but are then asked to recall details of the *personality* of the offender's face. They may be asked, for example, "How intelligent was the face?" or "How masculine was the face?" In this final 'holistic' recall stage, cognitive processing is shifted from individual features (eyes, nose, and mouth) to the face as a whole; in doing so, witnesses focus less on that which was recalled during the cued phase, to more on holistic information, which is presumably useful when presented with EvoFIT whole-face arrays. In the research project (Frowd et al, Submitted-a), correct naming increased from 25% for the normal cognitive interview to 40% with the holistic-cognitive interview. The H-CI is currently being field trialed.

What do these results suggest about how best to use EvoFIT? It is clear that the type of interview administered—one involving free, cued and/or holistic recall mnemonics—does exert a strong influence on witnesses' face processing and on their ability to construct a composite. More specifically, we are seeing evidence that information witnesses recall towards the end of their recall tends to be what they focus on during face construction. Based on data available to date, it is probably sensible to avoid using cued recall. Whether it is best to use free recall followed by holistic recall (or even to take a short break between these two stages, as suggested by Finger & Pezdek, 1999) is the subject of ongoing research and field trials. Either way, current EvoFIT performance remains valuable for law enforcement: it is anticipated that we will be able to further optimize the interview for EvoFIT, and thus promote an even more identifiable image.

b) Deploying EvoFIT within a police force

There are two basic ways that composite systems are deployed within a police force, and EvoFIT is no exception. The first way is for officers and/or support staff within each division to be trained on its use. These personnel can be multi-skilled, including facial composites, and this approach has the advantage that a composite officer can be available locally, in police [regional] divisions, without someone having to travel potentially long distances to assist in witness interviewing and composite construction. The main issue, though, is that officers can become engaged in protracted investigations—for instance, a complicated murder—or reassigned elsewhere. Without regular use, de-skilling becomes a possibility. While EvoFIT is straightforward to use, and can be learned within a couple of days, skills need to be maintained for interviewing and use of a paint package—for the addition of scars, marks, etc.

While by no means a new idea, a deployment model that is gaining popularity in the UK (and one that is adopted elsewhere) is a dedicated facial identification unit, to provide a force-wide composite service. Personnel in these units are similarly multi-skilled, but their specialism tends to reside within the identification area, with roles typically including crime scene photography, and the production of identity parades (e.g. VIPER, PROMAT) and photospreads. Units typically contain two to four members of staff, depending not only on demand, but also on strategy: the type of crime for which a force ring-fences its composites. In spite of being deployed mainly for major incidents, there is no real reason, police resources aside, why composites should not also be used to solve less serious crime. For example, in the police trials reported above, EvoFIT helped officers locate an offender who had stolen a handbag: such use of a composite arguably has value in contributing to police-public relations; other, similar uses include addressing prolific cases of theft and vandalism. One type of crime for which EvoFIT has been rather successful has been for distraction burglary. Victims of these crimes tend to have poor recall of an offender's face, not having tried to remember it, thus rendering feature systems difficult to use. In general, police report that EvoFIT is not only much faster to use with victims than feature systems, and much more effective, but also that the range of applicable crimes is much greater. This provides many more opportunities for Senior Investigating Officers than was possible previously.

We provide annual training days on latest techniques emerging from research and field trials (see previous and following sections) including general updates (new databases, bug fixes and accessory packs). Such workshops are easier for staff to attend—and it is easier for staff to then support each other—when training is organized for a dedicated unit. Staff

within a unit can be given administrative support, which is valuable for booking appointments with witnesses, for maintaining publicity of composites within the force, and for auditing EvoFIT performance (as was carried out in the field trials here). This approach can also assist with procedure following composite interviews—for example, following domestic burglary, setting up an appointment with the safety officer.

c) Further developments

Recent research with EvoFIT has also been exploring the impact of external features on composite quality. Recall that the standard EvoFIT procedure is for witnesses to select the outer region of the face at the start—the hair, ears and neck—and for that region to then be blurred, to help the witness focus on the important internal features when selecting from face arrays. Blurring is disabled just prior to manipulation with the *holistic tools* (for changing age, masculinity, weight, etc.) and the *shape tool* (for changing size and position of individual features). However, this procedure may still not be optimal as the external features can still act as a distraction during tool use. In Frowd et al. (Submitted-b), it was found that the externals do interfere, even when blurred, and at each stage of face construction. We found that constructing internal features in their entirety, and then adding external features at the end of the session, doubled the rate in which the resulting composites were correctly named—naming increased from 23% (normal method of blurring used in the lab and field trials) to 45% (internals-only construction). This is an important finding and suggests that the mere presence of external features poses a distraction to the person building the face. The police are now in receipt of this new EvoFIT development and are field trialing it.

The impact of these developments, including interview (holistic-CI) and internals-only construction, if used together in the same session, should lead to performance of around 60% naming. In fact, even better performance might be possible if animated caricatures were used routinely in public appeals. The effectiveness of these combined developments are currently being established in the laboratory, with police field trials planned.

It is extraordinary that such performance is possible from a composite system. Only five years ago, EvoFIT was producing images that could be named barely more than 10% of the time: composites from feature systems appear to manage only half of this figure. It is now possible to produce a very identifiable composite from a person's memory of an unfamiliar face after an appreciable time delay: what was missing was simply an appropriate interface to human memory. It is the union of computing science, psychological procedures and field testing that have allowed such a system to be developed.

IV. Conclusion

The current paper described formal end-user (police) evaluations of EvoFIT. Overall system performance was found to average 25.4% across four police forces and increased to 40% for forces using an enhanced interview, a figure which suggests the system is effective in the battle against crime. There are promising areas of development to substantially improve performance, both in terms of the interview and the way in which EvoFIT is used.

V. Acknowledgements

We wish to thank the Engineering and Physical Sciences Research Council, UK, and Crime Solutions, UCLan, UK.

References Référence Referencis

1) Brace, N., Pike, G.E., Allen, P., & Kemp, R. (2006). Identifying composites of famous faces: investigating memory, language and system issues. Psychology, Crime and Law, 12, 351-366.
2) 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.
3) 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, 119-124.
4) 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.
5) Finger, K., & Pezdek, K. (1999). The effect of the cognitive interview on face identification accuracy: Release from verbal overshadowing. Journal of Applied Psychology, 84, 340-348.
6) Frowd, C.D., Bruce, V., Gannon, C., Robinson, M., Tredoux, C., Park., J., McIntyre, A., & Hancock, P.J.B. (2007a). Evolving the face of a criminal: how to search a face space more effectively. In A. Stoica, T. Arslan, D.Howard, T. Kim and A. El-Rayis (Eds.) 2007 ECSIS Symposium on Bio-inspired, Learning, and Intelligent Systems for Security, (pp. 3-10). NJ: CPS. (Edinburgh).
7) Frowd, C.D., Bruce, V., McIntyre, A., Ross., D. & Hancock, P.J.B. (2006a). Adding Holistic Dimensions to a Facial Composite System. Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition (pp. 183-188). Los Alamitos: Ca. Applied Cognitive Psychology.

8) Frowd, C.D., Bruce, V., Ness, H., Bowie, L., Thomson-Bogner, C., Paterson, J., McIntyre, A., & Hancock, P.J.B. (2007b). Parallel approaches to composite production. Ergonomics, 50, 562-585. Journal of Experimental Psychology: Applied.

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

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

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

12) Frowd, C.D., Bruce, V., Storås, K., Spick, P. & Hancock, P.J.B. (2006c). An evaluation of morphed composites constructed in a criminal investigation. Proceedings of the 16th Conference of the European Association of Psychology and Law, London: IP-PA Publishing, pp. 59-66.

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

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

15) Frowd, C.D., & Fields, S. (in press). Verbal overshadowing interference with facial composite production. Psychology, Crime and Law.

16) 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.

17) Frowd, C.D., Hancock, P.J.B., Bruce, V., McIntyre, A., Pitchford, M., Atkins, R., et al. (2010a). Giving crime the 'evo': catching criminals using EvoFIT facial composites. In G. Howells, K. Sirlantzis, A. Stoica, T. Huntsberger and A.T. Arslan (Eds.) 2010 IEEE International Conference on Emerging Security Technologies (pp. 36-43).

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

19) Frowd, C.D., Nelson, L., Skelton F.C., Noyce, R., Heard, P., Henry, J., Morgan, D., Fields, S., McIntyre, A., & Hancock, P.J.B. (Submitted-a). Interviewing techniques for Darwinian facial composite systems. Journal of Experimental Psychology: Applied.

20) Frowd, C.D., Park., J., McIntyre, A., Bruce, V., Pitchford, M., Fields, S., Kenirons, M. & Hancock, P.J.B. (2008b). Effecting an improvement to the fitness function. How to evolve a more identifiable face. In A. Stoica, T. Arslan, D. Howard, T. Higuchi, and A. El-Rayis (Eds.) 2008 ECSIS Symposium on Bio-inspired, Learning, and Intelligent Systems for Security, (pp. 3-10). NJ: CPS. (Edinburgh).

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

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

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

24) Rattner, A. (1988). Convicted but innocent: wrongful conviction and the criminal justice system. Law and Human Behavior, 12, 283-293.

25) Schooler, J.W. (2002). Verbalization produces a transfer inappropriate processing shift. Applied Cognitive Psychology, 16, 989-997.

26) Schooler, J.W., & Engstler-Schooler, T.Y. (1990). Verbal overshadowing of visual memories: some things are better left unsaid. Cognitive Psychology, 22, 36-71.

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

28) 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.