EvoFIT facial composite images: a detailed assessment of impact on forensic practitioners, police investigators, victims, witnesses, offenders and the media

Abstract—This paper assesses use of EvoFIT facial composites by police practitioners in the UK and overseas. Results reveal that this composite system is used extensively: a total of 2,440 times since September 2013. With a suspect identification rate of 60% and a conviction rate of 17%, the impact of this forensic technique is appreciable for helping the police to identify and convict offenders. It was also found that empirically-driven enhancement techniques were used frequently by police practitioners—including use of detailed context reinstatement and holistic techniques during interview, asking the witness to focus on the eye-region during construction, or presenting a stretched image for identification. Research evaluating EvoFIT composites published by police in the media also reveals that composites were deployed almost exclusively for serious offences, and most often for sexual crimes. In addition, the vast majority were of male offenders, in their early 30’s; victims were female (aged mid 20’s to mid 30’s).

A similar outcome was observed for composites published from another recognition system, EFIT-V / 6. Although this overall exercise revealed that EvoFIT composites were sometimes presented to the public in an optimal, stretched fashion, it was also apparent that a worrying number of composites were presented in a way that did not faithfully represent the constructed image, specifically, with the external features cropped.

Keywords: EvoFIT; detailed assessment; witness; victim; identification; facial composite; E-FIT

I. INTRODUCTION

The reduction in funding to public services in the UK has had a huge knock-on effect for health, education and security. Assessing the effectiveness of procedures, policies and systems is more important than ever, both to shape and improve the quality of what is available and drive future research and development. Our focus in this paper is on the effectiveness of one aspect of policing, facial-composite images [1]. These are pictures created from the memory of an eyewitness, most often a victim of crime. Police use composite images to generate intelligence, in particular to identify suspects and, along with corroborative evidence, convict those ultimately responsible. When a composite has been constructed, police circulate the facial image within their force or across forces, in the hope that a member of police staff will be familiar with the person depicted, recognise the face and provide a lead to the investigation. If no names are forthcoming, or when suggested identities have been discounted as potential suspects, the police publish composite images in the media. This process appeals to members of the public using available outlets: TV, newspapers, and social media. Again, the hope is that someone, this time a member of the public, will be familiar with the identity and name the face to the police; see [1][9] for examples.

In the current project, our primary aim is to assess use of composites from one of the newer recognition-based systems, EvoFIT. This composite system has been in regular police use since 2007. We assess system use by forensic (police) practitioners and police investigations over a six-year period. The exercise was completed by asking police practitioners of this system to providing operational information as part of the UK government’s Research Excellence Framework (REF), set for 2021. The project considered requests emerging from investigating officers (as part of a criminal investigation) for a composite to be constructed by witnesses and victims as part of an interview by police practitioners. The project also considered impact of the composite after it had been constructed (e.g., circulated within a police force). Impact was also assessed in a novel way, by a review of composites published by police in the media. This additional exercise was considered for all production systems (EvoFIT and non-EvoFIT types) and aimed to explore characteristics of person’s sought (usually offenders) including crime type, location, media outlet and publication format. Evidence for the exercise was sourced from an internet search using composite-related search terms (listed below).

II. ASSESSING EVOFIT SYSTEM USE

There are three main types of composite system available to police practitioners [1][2]. First are ‘mechanical’ systems such as Photofit and Identikit. Eyewitnesses construct a face of an offender by selecting facial features (e.g., eyes, nose, mouth, hair) printed onto rigid card or acetate slides. Second are software feature systems, such as E-FIT, PRO-fit, FACES, Identikit 2000. They work in a similar way, with eyewitnesses selecting facial features, but better likeness are possible, in particular as facial features are selected in a whole-face context (as described in [4]). However, neither type of system produces identifiable faces especially when a long delay occurs between witnessing the face and constructing a composite [10]. Third, are recognition-based systems. These newer implementations were developed to overcome the problem of poor identification and include
EFIT-V [now called EFIT-6], EvoFIT and ID. Using these systems, eyewitnesses select from screens of whole faces (or whole face regions), with choices combined, to ‘evolve’ a composite. Various software tools are also used to enhance an evolved likeness both at a featural and at a holistic level.

Our focus of attention in the current work is on the EvoFIT implementation, the development of which has been described in several papers submitted to this conference series (e.g., [1]). In brief, EvoFIT was developed at the University of Central Lancashire and the University of Stirling, EvoFIT was commercialized for police use about eight years later, and the system went into regular police use in 2007. Other recognition type systems available to police include EFIT-V [now called EFIT-6] [6] and ID [7]. There does not appear to be published research on the performance of either of these specific implementations using procedures that simulate real life. To our knowledge, identification of 20% correct has been measured for the former, but only following a brief unrealistic interval between seeing a face and constructing a composite [8]—the expectancy would be lower identification after one or two days, as in real life [10].

For systems in general, research has reached a series of milestones to develop techniques aimed at improving the effectiveness of composites (see [9] for a recent review). Some of these developments relate to computerised feature systems, such as adding a character- or personality-based enhancement to the standard free recall of the face elicited prior to composite construction. This so-called Holistic Cognitive Interview (or H-CI) is effective for improving identifiability of feature-based composites, compared with a more standard face-recall interview (face-recall CI), increasing correct naming rates by four times [11]. It is also effective for EvoFIT [12], and there are reports from practitioners of EFIT-V that this interview technique is also being used with this recognition-based system [13].

Further development has led to the H-CI in effect ‘co-evolving’ with the composite system. Research indicates that constructors produce more reliable composites with EvoFIT if, during the initial evolving stage, they are asked to focus on the upper half of the face, the region around the eyes that is important for recognition of a composite [14]. However, while effective with the previous (whole-face) method of face selection in EvoFIT, the H-CI no longer enhances composite naming (cf. face-recall CI) when the focus of eyewitness attention is switched to the eye region. When face-array selection is based on the eye-region, facilitation may only be achieved when the interview technique encourages a similar focus. Specifically, the H-CI continues to provide an advantage when it is followed by a modified script that asks the witness to make character attributions based solely on the eye region. The outcome of this work is that the ‘enhanced’ H-CI, in effect, allows alignment of cognitive processing for a witness from one stage of face construction to the next. These techniques, when used in combination, can lead to very effective composites: use of H-CI, EvoFIT, and naming of the face viewed from side-on have been measured to produce composites that were correctly named at an astonishing 74% correct [5]. Police field trials of the system also reveal a high level of suspect identification (60%) in criminal cases [1].

Further research on the initial interview stage indicates that naming of PRO-fit and EvoFIT composites increases if the constructor is asked to recall the environment in which a target face had been seen [16][17]. This is because there is benefit to face recall and construction from additional cues that a witness remembers from the environment (crime scene). In another project, EvoFIT composites emerge more identifiable if constructors are asked to ignore width of face rather than an alternative (previously-used) method where witnesses selected a single face for width at the start and the system generated faces based on this selection [18]. Further research, as mentioned above, has developed post-production techniques that improve naming of finished composites. The approach includes presenting a composite (i) in a moving caricature (facial feature exaggeration) format, (ii) by turning the face side-on (to produce a stretched image) and (iii) by adding a hat, sunglasses and possibly other accessories such as facial hair (to conceal inaccuracies in the composite) [14][19][20]. These techniques can be very effective. For example, a range of plausible appearances were created for a recent criminal investigation to help the police identify a murderer at large [22], as illustrated in Fig. 1.

Fig. 1. EvoFIT facial composite (top far left) with plausible likenesses created to brief officers hunting for murder suspect. bottom far right. Upon arrest, the suspect had grown a beard: his appearance resembled the third depiction, top row.

In the following sections, we assess use of EvoFIT and associated interviewing and post-production procedures from September 2013 to March 2019, a timespan associated with the UK government’s period of assessment for university research. In the first section below, A, forensic use of EvoFIT is considered over this interval, including uptake of milestones of development (mentioned above). In the second section, B, we assess the impact of composites produced both from EvoFIT and from non-EvoFIT systems, that have been published by police in the media.
A. Forensic Use of EvoFIT Facial Composites

EvoFIT has been used by 27 police forces since its creation, of which 20 out of 43 (47%) are in England and Wales, plus Police Service of Northern Ireland (PSNI). We were able to assess current usage in an anonymised opportunity sample from the following forces: An Garda Síochána, Bedfordshire, Cambridgeshire, Dyfed Powys, Greater Manchester, Hertfordshire, Israel, Lancashire, North Wales, Police Service of Northern Ireland, Romania and Staffordshire. We requested estimates of total use in the assessment period, number of different investigating officers who have requested a composite, and information about circulation both within and outside the relevant force. We also requested information about the use of the aforementioned milestones in development for interview, system and post-production stages.

(a) Overall EvoFIT system usage from September 2013 to March 2019

Figures for individual system usage of EvoFIT include the 13 forces mentioned above plus the number of composites constructed on a case-by-case basis from consultancy. Total usage was 2,440. These composites were produced by a total of 56 forensic practitioners who interviewed witnesses and victims of crime. They were requested by a total of 1,540 police investigators (Investigating and Senior Investigating Officers)—that is, requests from different people. The overall mean requests per investigator was 1.4, a figure that suggests that in general composites tended to be requested by different investigators rather than multiple requests from the same person. We were also able to calculate impact using another metric, total time EvoFIT licenses had been purchased and were in use across forces, a total of 1,277 months.

Circulation figures were also requested. Practitioners pass on composites to investigating officers, and so the ensuing outcome can be difficult to know. For the seven forces who were able to make an assessment, 2,002 out of 2,085 (96%) of their composites were circulated internally within a force, and circulation was to an estimated 34,834 police staff. These figures suggest that composites were sent to a total of 17.7 million police staff. Using mean circulation rate per force ($M = 5805.7$), we estimate that the 2,440 composites were circulated to 18.7 million police staff.

In cases where a composite was not identified internally, police circulate images to the media. It is particularly difficult for practitioners to assess the number of composites published in this way. However, for practitioners who were able to make an estimate, a mean of 18% of their composites were published on the force’s website. We also asked about media outlets in which composites were published. Forces tended to use about five different outlets in total ($M = 4.7$), but the average using the median calculates as three ($M_{n} = 3.0$), and that figure is probably a more accurate estimate. The split between paper ($M = 48\%$) and online ($M = 52\%$) publication formats was roughly equal. One force in the UK and another overseas reported that the stretched (perceptual-backdrop) format was used for police appeals; accordingly, we found cases where composites had been published in the media using this format for robbery, burglary and trespass (see below).

Given that the assessment period for the UK Government’s research assessment exercise runs to July 2020, we estimated the total number of composites that should have been constructed by then. This was based on the mean number of composites constructed by each police force per month projected ahead by 16 months. A further 709 composites should be constructed in this time, giving a total estimate of 3,149 EvoFITs produced from September 2013 to July 2020.

(b) Uptake of Procedural Enhancements

As mentioned above, each of the three main stages involved in the construction and recognition of an EvoFIT has been enhanced through research and distributed to police practitioners, and so all changes should have generated impact in the assessment period. These include improvements to the initial interview (to give an ‘enhanced’ H-CI), EvoFIT system (mechanism for setting facial width, selection by focusing on the eye region, new face and accessory databases) and post-production (incl. stretched composite and hats-and-sunglasses format) [5][12][14][16][17][18][19]. Enhancements were discussed with practitioners, to elicit feedback, and included in (i) training courses for new practitioners and (ii) annual workshops for existing practitioners. This process allowed police practitioners to take advantage of best practice. Some of these enhancements were related to software, all of which aimed to facilitate a more accurate composite, increasing suspect identification. These include hat-sunglasses and stretched post-production formats plus other enhancements: new face databases, improvements to hair and accessories and new / improved system features. Other improvements were procedural: focus on eye region for (i) (enhanced) H-CI and (ii) EvoFIT; and (iii) detailed recall of the context (environment) prior to recall of the face. Some enhancements included both software and procedural changes: e.g. a mechanism for setting facial width at the start of composite construction.

While software changes form part of the normal operating procedure to construct an EvoFIT composite, we were interested in assessing uptake of other developments which emerged from research. For these enhancements, their use depends on circumstance. For example, the H-CI tends to be less effective [23] when an eyewitness’s recall of the face is sketchy. Similarly, eyewitnesses may not be asked to recall the environment in detail (prior to face recall) if it is apparent to practitioners that witnesses had not encoded the crime scene (as per the research [14][15]). Estimates were requested as a percentage of the number of composites constructed. Results indicate overall mean use was (i) 88% for use of the enhanced version of the H-CI, (ii) 88% for focus on the eye region when selecting matches in EvoFIT, and (iii) 50% for detailed recall of the crime scene.

B. Impact on the Media

Police investigators publish composites in the media when no suspect has been identified. Outlets that police use include both paper and online formats, and here we assessed the impact of composites on the latter. The assessment itself is fairly straightforward to conduct as the objective of publishing a composite is for as many people as possible to see the image.
In this part, we considered all production systems: E-FIT, PRO-fit, EFIT-V / 6 and forensic sketch composites. Our initial plan was to classify results into EvoFIT and non-EvoFIT types, but the latter only turned out to involve composites from the recognition system, EFIT-V / 6.

Google Images was used throughout as the search engine to locate cases. We used various search terms including EFIT, E-FIT, EvoFIT, sketch and facial composite. The resulting composite was checked for type given that some composites are referred to by the generic name of “efit” or “composite”; we also note that some composites were incorrectly stated as an E-FIT, even though they were from EvoFIT, or vice versa (this is usually obvious as most composites include either a logo or written name). To facilitate the search process, if a composite led to a specific website, be it newspaper or police site, that location was then followed up by searching for other composites that might have also been published there. Lastly, to get a long-term view of impact, we accepted publication dates of composites over the last decade, producing results from October 2008 to March 2019. For all composites located, the vast majority (86%) were published since September 2013. While our expectation was that a few somewhat ‘ancient’ composites would be found, this result indicated that many public appeals for information involving composites (14%) have been left on the internet for long period of time, seemingly without good cause.

In the following sections, we report results separately for published composites from EvoFIT and non-EvoFIT (specifically EFIT-V / 6) systems. The aim was to locate a good-sized sample, ideally between 100 and 200 cases. The exercise readily revealed 150 EvoFITs, but we struggled to locate more than a hundred cases for non-EvoFIT systems, stopping at 111. We collected a range of data published from each police appeal including incident date, publication date, type of crime, location, police force, number of offenders sought, and offender and victim characteristics. We note that some information was not always available. For example, victim’s gender and / or age was not always mentioned, but we compiled as much information as possible. As offenders’ age was usually an estimate, we converted that reported into an average. For example, a person sought in his “30’s” was coded as 35 years; similarly, we coded “14-15” as 14.5, “late 20’s” as 27.5, “about 30” as 30, “50-60” as 55, “teenage” as 16 and “elderly” as 65. In most cases victims’ age was mentioned specifically, but when only a range was provided we used the same averaging methods described above.

(a) Public appeals involving EvoFIT composites.

The 150 EvoFITs we located comprised police appeals from 17 forces in England, Scotland, Wales, Northern Ireland and overseas (incl. Republic of Ireland and Romania). Distribution by type of crime is illustrated in Fig. 2. With the exception of one case of theft, all crimes for which composites were published were serious. Most were for sexual offences including rape, sexual assault and indecent exposure, then with roughly equal occurrence for assault, robbery and burglary. There were various low-occurrence crimes that we classified under ‘other’: abduction, forced entry, grabbing, hate crime, hijacking, misdirection (to gain entry into property), murder and stabbing; also included were eight crimes that were ‘attempted’—that is, were not successful—for abduction, assault, rape and robbery.

In the following analysis, we considered frequency of occurrence for main characteristics. We were cautious not to over-analyse data, an exercise that might otherwise provide an unrepresentative view of the situation. As such, to avoid statistics based on infrequent values, we avoided sub-dividing categories.

The vast majority of persons sought by police were male ($M = 95.3\%$), from 12 to 85 years, average of about 30 years ($M = 31.7, Mn = 30.0$ years); there was one additional case where a male and female were sought, and in another, more than one male. The majority of offenders were reported as being white (76.7\%). In contrast, where specified (95.2\%), victims were mainly female (80.0\%). Age ranged from 10 to 87 years; these data are positively skewed: victim’s average age was mid 20’s using the median ($Mn = 26.0$ years) and mid 30’s using the mean ($M = 35.8$ years).

![Fig. 2. Publication of EvoFITs by type of crime](image)

All composites were presented in greyscale, appropriate as this format is the image modality for the EvoFIT system (cf. non-EvoFIT composites below). There were 44 cases (29.3\%) where a published composite included one or more accessories (cap, glasses, hat, hood and sunglasses). There were four appeals (2.7\%) where a composite was presented using the stretched (PBI) format (these cases relate to the REF period, from September 2013). We did observe, however, some cases where a composite had been published not as constructed: (i) 22 cases (14.0\%) where the face had been cropped, to reveal the central part of the face, and (ii) in one case, the face had been obviously stretched horizontally.

(b) Public appeals involving Non-EvoFIT composites.

The 111 non-EvoFIT type composites were published by 17 police forces in England and Scotland. As far as we could tell, all composites were created using EFIT-V / 6. Distribution by type of crime is illustrated in Fig. 3. As for EvoFIT, about one third of all cases relate to sexual offences (indecent exposure, rape and sexual assault). Assault, burglary and robbery had next prevalence, similar in distribution to before, although the category for ‘other’ is higher here, mainly as attempted crimes (abduction, burglary, kidnapping, murder, robbery and sexual assault) occurred...
more frequently overall (14.8%) than for EvoFIT composites (5.3%). This (other) category also included arson, entering a property without permission, fatal shooting, grabbing, hijab pulled from face, impersonation of a police officer, stabbing and suspicious behaviour.

Offenders sought were most often white (71.2%), ranging from estimates of 16 to 60 years, average age mid 30’s ($M = 33.6$, $Mn = 35.0$ years), and were almost entirely male (97.3%). Victims were mostly female (77.9%), ranging from 6 to 85 years, on average in their 20’s; again the central tendency for age was subject to positive skew ($M = 28.4$, $Mn = 22$ years).

All composite faces were presented in colour (cf. greyscale for EvoFIT), and included at least one accessory for about one-third (30.6%) of appeals. As found for EvoFIT composites, a small percentage (four cases; 3.6%) of published images did not appear as constructed.

Fig. 3. Publication of Non-EvoFITs by type of crime

III. GENERAL DISCUSSION.

It is important to investigate the quality of outcomes of forensic procedures. In the current assessment we considered police use of EvoFIT composites between September 2013 and March 2019, as well as EvoFIT and non-EvoFIT composites that have been published in the media in the past decade. It is clear that there has been considerable use of the EvoFIT system: during this period 2,440 EvoFITs have been produced, a figure which both accounts for composites constructed by police practitioners, in conjunction with witness and victims, and through external consultancy. Requests were generally made by different Investigating and Senior-Investigating Officers. Software updates were used when released (via training courses and annual police workshops), and uptake of more procedural advancements emerging from research (Enhanced H-Cl, focus on eye region with EvoFIT, detailed recall of environmental context) all had high reported usage—for between 50% and 88% of cases. There was also evidence that the side-on (PBI, Perceptual Backdrop Image format) post-production technique was in use in the UK and overseas. These results suggest that there is very good impact of the underlying published research.

An investigation of composites published in the media over the past decade revealed further interesting findings. Both EvoFIT and non-EvoFIT systems impacted on the media, although it would seem that police have published relatively more EvoFIT images (given the difficulty we experienced in locating more than about a hundred appeals involving non-EvoFIT composites). For both recognition-type composite systems, their use was almost entirely for serious offences—only one case of more minor crime (theft) was identified. While it is sometimes the case that a composite is used to locate an eyewitness, a person who then might help the police to locate an offender, no such instances were found in the media. For both composite systems, sexual crimes represented about one third of all appeals, with about 15% each for assault, burglary and robbery. Offenders’ age ranged greatly, but was similar for both types, at around 30 – 35 years. Age of victim also varied greatly but (positive) skew in the data makes analysis more tricky; the best estimate would be average age of around mid 20’s to mid 30’s.

Two issues emerged that were notably surprising. The first is that some composites had been presented horizontally stretched or cropped: examples of image cropping were fairly high for EvoFIT (14%), less so for EFIT-V / 6 (4%). What is possible here is that media outlets cropped the image for space, but another possibility is for stylistic reasons (to show the central part of the face); however, provisional research from our laboratory indicates that revealing less than a complete face may not be best for identification. Much better, as our research has shown recently [19], would be to present the face with alternative representations including hat and sunglasses. Clearly, additional research is needed to clarify what should be best practice, and to identify the reason why published images have not faithfully represented the original composite, as constructed by a witness.

The other issue concerns image modality. For EvoFIT, the system deliberately produces faces in greyscale as it is established that colour information does not reliably boost identification rates [21]. However, while EFIT-V / 6 can use either mode, practitioners would appear to be using colour all of the time (or at least the image being given to the press is in colour). One of the issues with this approach is that it can be very hard to carry out artwork in colour, and there have been some notable (even laughable) cases where this process has gone awry, leading to a composite with green hair for example. Further work is clearly needed to promote best practice on image modality.

In summary, the project sought to explore use of EvoFIT composites in current police practice. Practitioner feedback reveals that this system has been used extensively, some 2,400 times, since September 2013. Projected use by July 2020, the end of the assessment period for the UK Government’s university research exercise, indicates construction of about 3,100 composites during this timeframe. There was also evidence of good uptake of research protocols for several new techniques—including detailed recall of environment, enhanced H-Cl and focus on eye region for EvoFIT. There are clearly many composites published in police appeals in the media over the last decade, and 150 were readily located. The range of crimes was almost exclusively serious (esp. assault, burglary and sexual offences) and committed mainly by males, with the vast majority of victims being female. A similar situation occurred for EFIT-V / 6 composites, although these composites turned out to be less prevalently published in the media. For both systems, some images were presented cropped or non-proportionally stretched, and so not shown as...
had been constructed by the eyewitness; further research and education would appear necessary here. For EvoFIT, its facial composites are quite clearly having a marked impact on police investigations, witnesses, victims, police investigators and the media. The impact is appreciable both in terms of reach (the number of facial composites constructed with use both in the UK and overseas) and significance, for identification of offenders of serious crime.

IV. REFERENCES


[23] C. Brown, C.D. Frowd & E. Portch, “Constructing Facial Composites: Increasing the Forensic Relevance of Laboratory Research”, Research grant from the UK’s Economic and Social Sciences Research Council (ESRC) ES/I002022/1. (http://grt.rcuk.ac.uk/projects?ref=ES%2FI002022%2F1.)