Interviewing and Visualisation Techniques: Attempting to Further Improve EvoFIT Facial Composites

Emma Portch  
Division of Neuroscience and Experimental Psychology, University of Manchester, Manchester, UK  
Email: emma.portch@manchester.ac.uk

Karen Logan  
School of Psychology, University of Central Lancashire, Preston, UK  
Email: klogan1@uclan.ac.uk

Charlie D. Frowd  
School of Psychology, University of Central Lancashire, Preston, UK  
Email: cfrowd1@uclan.ac.uk

Abstract—Victims of and witnesses to crime are asked to describe an offender using cognitive interviewing techniques (CI), before constructing a visual likeness of the face. The aim of the current experiment was to investigate whether composite construction using the EvoFIT holistic system would benefit from parallel use of three enhancement techniques. The study manipulated the type of interview used to elicit a face description (CI vs. holistic cognitive interview, or H-CI) and trialled a visualisation technique for selecting faces (no visualisation vs. visualisation of external features). Also included was a new construction procedure for EvoFIT that requested constructors to focus on the region around the eyes when making face selections. Based on past research, it was anticipated that both the H-CI and external-feature visualisation would promote construction of a more identifiable composite (compared to when each technique was not used). Rather unexpectedly, the results revealed that neither technique improved correct naming of composites, yet an interaction was observed: visualisation of external features led to a benefit that approached significance when used in conjunction with the H-CI (cf. CI). However, when no external-feature visualisation was used, composites were better named following the more usual CI (cf. H-CI) protocol. Results are promising for the new method of face selection, which was used by all participants (focusing on the eye region). Importantly, this construction process mimics the whole-face (holistic) strategy commonly considered to underpin face recognition [1]. When witnesses experience a forensically-relevant delay of one or two days between witnessing an event and constructing a composite, EvoFIT has been shown to produce visual likenesses that are correctly identified around 50% of the time. This compares favourably to an identification rate of around 5% for composites constructed using feature-based systems [e.g. PRO-fit; 1-3].

First, we considered the technique used to obtain a description of an offender’s face from a witness. Police practitioners commonly make use of cognitive interviewing mnemonics since they encourage thorough and accurate face recall, whilst providing appropriate retrieval support [4]. Using this procedure, a witness is asked to think back to the incident, visualise the offender’s face (context reinstatement), and describe everything that can be remembered about the face (free recall). The interviewer can then consider asking the witness to focus on each facial feature in turn and prompt for further recall (cued recall). However, some researchers argue that detailed face recall can be detrimental to subsequent face recognition: a so-called verbal overshadowing effect [5-6]. Whilst a witness must adopt a ‘feature-based’ approach to describe different aspects of the face (e.g., whether the nose is long or short), recognition itself places considerable emphasis on consideration of the global properties of a face; it is these particular holistic elements that cannot be easily described (e.g., distances between features on a face). Approaching composite construction with a sub-optimal feature-based processing style may thus impede a witness’s ability to select appropriate facial features, or decide that a composite has reached the best visual match [7].
A modified interview technique can be employed to ‘re-instate’ an appropriate ‘whole-face’ processing style prior to face construction: the holistic interview protocol [8]. Using this technique, a witness is asked to reflect upon personality characteristics that are conveyed by the offender’s face, as a whole, and to rate the face on seven global or holistic dimensions (e.g., friendliness, masculinity, distinctiveness, health and pleasantness). When composites were constructed using feature-based systems, recognition rates were significantly higher when constructors had first completed a H-CI (41.2%) compared to a standard face-recall CI [8.6%, 8].

A similar pattern of results were obtained when composites had been constructed using the holistic EvoFIT system: composites constructed following a H-CI were recognised 39.4% of the time, whilst those constructed following a standard face-recall CI were recognised only 24.1% of the time [9]. As the H-CI (cf. CI) appears to improve the quality of composites constructed using both feature-based and holistic systems it is reasonable to assume that the implementation of this technique enhances both the accurate construction of internal- and external- facial features [8,9].

Frowd et al. [1] also found that the holistic interviewing protocol improved naming of EvoFIT composites when used in conjunction with other techniques during construction: methods that aimed to match holistic processing at interview, face construction and composite naming.

Second, we investigated strategies employed during face construction itself. Research using feature-based technologies (involving selection of individual facial features to construct a composite) suggest that witnesses typically construct poor representations of a target’s internal features (e.g., eyes, brows, nose and mouth), while doing better to accurately construct external features (e.g., hair, ears, face shape and neck [10]). These results are unsurprising given that we tend to focus on external facial features when processing unfamiliar faces. However, they are problematic, as an important aim of a composite is for it to be recognised by those familiar with the target identity, and familiar face recognition relies to a greater extent on the processing of internal facial features [11].

A technique recently trialled with EvoFIT aims to improve composite recognition by directing a constructor’s attention toward internal facial features. Specifically, participants are requested to focus on the internal eye region (rather than the whole internal-feature region), when initially selecting faces from arrays [12]. Preliminary findings suggest that this technique increases construction of recognisable composites compared to when a constructor receives standard instructions (to select faces based on the overall match of the internal features). Although some witnesses may naturally focus on the eye region for face selection, particularly when external features are masked (see below, 12, 13), others may prioritise the resemblance of other internal features, such as the nose. Providing a blanket instruction should reduce individual differences in construction strategy and improve performance.

As a second way to focus constructor attention on internal features, manipulations may be made to the face itself. External features may be obscured using Gaussian blur; results indicate that the correct naming of composites increases with higher levels of blurring [3, 10, 13]. However, composite recognition rates increase still further when external features are completely masked (45.6%), rather than heavily blurred (22.7%), perhaps suggesting that the physical presence of these features presents a distraction to witnesses during construction [12, 13]. In the latter case, external features are added to the face in the final stages of composite construction.

In spite of these findings, face recognition is a holistic process, and thus it should be the case that external features provide a useful implicit context for the accurate selection and recognition of internal features for unfamiliar target faces [14]. A newer strategy may present a useful compromise: during composite construction, a witness may be asked to visualise the external features of the face, to compensate for their physical absence in the initial EvoFIT face arrays. In this way, the witness actively reinstates a ‘whole-face’ context during construction, which may encourage the accurate selection of internal features. Indeed, when external features were masked during initial face arrays, more recognisable EvoFIts were constructed by participants who had been instructed to adopt external-feature only visualisation (63.3%), compared to those who had not [49.2%; 15]. Importantly, external-feature visualisation may work to secure complementary and additive benefits when used in conjunction with the H-CI, as both methods encourage a whole-face processing style during composite construction.

II. EXPERIMENT: INTRODUCTION

The present study investigated the potential for additive improvements in composite recognition rates when a combination of (established and newer) enhancement techniques were used during face construction [see also 1]. All our participant-constructors were instructed to focus on the internal eye region when selecting faces that most resembled the target identity. Type of interview and visualisation technique were manipulated, between-participants. One group of participants completed a standard face-recall CI prior to composite construction, whereas a second group completed a H-CI. Within these two groups, half of the participants approached initial face selection with the instruction to consistently visualise external facial features (which were masked for everyone), whereas the other half proceeded under standard instructions (no visualisation).

As in Frowd et al. [1], we expected each technique to improve identifiability in the anticipated direction: composites would be better named if they had been constructed after a H-CI (vs. CI) and under instruction to visualise external features (vs. no instruction). We also investigated whether the separate benefits to composite construction, afforded by each technique, were enhanced when used in combination. Frowd et al. [1] found a three-way interaction in their experiment as a combined effect was observed to be more effective than the sum of the effects for each manipulation. When compared to baseline conditions, composite constructed under an additive approach were named very successfully (74% correct). A similar outcome was anticipated here.
III. Method

A two-stage methodology was used [16]. One group of participants were interviewed and then constructed a single EvoFIT composite. These participants confirmed that they were unfamiliar with the pool of target identities (characters from the TV soap ‘Coronation Street’). A second group of participants attempted to name the composites. These participants were regular viewers of Coronation Street and should therefore be familiar with the pool of sampled identities.

A. Stage 1: Composite Construction

1) Participants

Participants were 21 female and 11 male staff and students from the University of Central Lancashire, Preston, UK. All reported to be over 18 years of age. Participants were assigned in equal groups of eight to the two between-participants factors: type of interview (CI vs. H-CI) and visualisation technique (none vs. external-features). Participation was voluntary or for course credit.

2) Materials

Eight good quality head and shoulder colour photographs of characters from the TV soap Coronation Street were located on the Internet and used as target identities. These target faces were photographed in a largely full-face pose with a neutral expression. Images were printed individually on A4 paper at 8cm (width) x 10cm (height). Composites were constructed using EvoFIT software (version 1.6).

3) Design

The design at face construction was factorial, fully crossed, for (i) type of interview (CI vs. H-CI) and (ii) visualisation technique (none vs. external-features). Participants created a single composite under one of these four conditions (with random assignment). The eight target identities were used once in each condition (with participants also randomly assigned to target identity).

4) Procedure

Participants studied a photograph of an unfamiliar target individual for 1 minute in the knowledge that they would later construct a composite of the face. They were randomly assigned, with equal sampling, to construct a composite under one of two types of interview (CI vs. H-CI), and one of two methods to visualise the face (none, the current procedure; vs. visualise external features in each array face).

After 20 to 28 hours, participants met with the experimenter to construct a composite of the target face. Participants were tested individually and worked at their own pace. The procedure to interview witnesses and construct a composite is involved and is described in detail in [17]. For the sake of brevity, we provide an overview here.

All participants completed a Cognitive Interview that included three stages: rapport building, context reinstatement and free recall of the face (see General Introduction). Half of the participants then underwent a Holistic Interview. Here, the participant was given one minute to think about the personality conveyed by the target face, before rating that face for seven holistic traits (e.g., friendliness, extroversion) on a three-point scale (low / medium / high).

Next, participants worked alongside the experimenter to construct an EvoFIT composite. Using this system, the participant first viewed 70 faces across successive screens and chose six that best resembled the target identity. The external features were not presented on any face. Participants made selections from arrays that sequentially displayed three types of face (smooth faces, textured faces, and combinations of smooth and textured faces) and also indicated which of the selected faces represented the best likeness. EvoFIT ‘bred’ the selected faces together, giving preferential weighting to the designated ‘best face’, to produce another similarly-sized face array. This process continued until the participant indicated that the ‘best face’ strongly resembled the target, upon which holistic tools were used to edit global properties (e.g. age, weight, pleasantness) and shape (size and position of individual features). Afterwards, external features were added, and holistic and shape tools offered to further improve the likeness, as required. When the process was completed, the best likeness was saved to disk.

During composite construction all participants were asked to focus on the eye region when they selected faces from arrays. In addition, half of the participants were also asked to visualise the external features of the face; for the other half, the visualisation protocol was not mentioned. Face construction took about an hour to complete per participant, including debriefing.

B. Stage 2: Composite Naming

1) Participants

Participants were 33 female and seven male staff and students from the University of Central Lancashire, Preston, UK. All participants reported to be over the age of 18. They were assigned in equal groups of 10 to view composites constructed under the two between-participant factors: type of interview (CI vs. H-CI) and visualisation technique (none vs. external features). Participation was voluntary or for course credit.

2) Materials

The 32 composites from Stage 1 were printed in greyscale (8cm x 10cm). See Fig 1. for examples. The composites were divided into four equal sets of eight: each set contained the eight composites that had been constructed under a different condition. The eight target photographs were also required.

3) Design and Procedure

Participants were tested individually and worked at their own pace. They were randomly allocated to one of four composite sets with equal sampling in a between-participants design for type of interview (CI vs. H-CI) and visualisation technique (none vs. external features). Participants sequentially viewed the eight composites in the assigned set. They were asked to name each composite where possible; “don’t know” type responses were permitted. Afterwards, to check that participants were actually familiar with the target identities, they were asked to name the eight target photographs, also presented sequentially. Each person received a different random order of presentation for
composites and target photographs. The naming procedure took about 15 minutes, including the time for debriefing.

Fig. 1. Example composites constructed in the experiment (from left to right): CI+No EFV, CI+EFV, H-CI+No EFV and H-CI+EFV. Each of these composites were constructed by a different participant who saw a picture of ITV Coronation Street character Kirk Sutherland. For reasons of copyright, a picture of the actor used cannot be reproduced here; however, an example of his appearance can easily be located on the Internet.

IV. RESULTS

The target photographs were named very well by participants ($M = 99.1\%$, $SD = 3.3\%$), indicating excellent familiarity with the relevant identities. As a composite was unlikely to be correctly named if the corresponding target photo was not correctly named, these cases ($N = 3 / 320$) were treated as missing data and were not subject to further analysis.

Spontaneous responses to composites were scored for accuracy (coded as ‘1’ if the participant had provided the correct name or description of the soap’s character, and ‘0’ otherwise). A summary of correct responses are presented in Table I. The overall mean correct naming appeared to be appropriate for this system and design, at 41.4\% ($SD = 49.4\%$) [1].

As can be seen in Table I, however, the results in general were not as expected. Correct naming was clearly superior in the baseline (Control) condition (face-recall CI and no visualisation of external features): the experimental condition that was expected to produce composites with the lowest correct naming rates. The H-CI led to much worse naming (cf. CI) when used without external-feature visualisation, although it seemed to be somewhat effective when participants were instructed to use this type of visualisation. External-feature visualisation (cf. none) resulted in much worse quality composites after a CI but a good improvement after a H-CI. The reliability of these effects is considered below.

Generalized Estimating Equations (GEE) was used to analyse the dichotomised naming responses to composites (scored as 0 and 1). This type of analysis provides a powerful by-participants and by-items model with target items coded as a within-participants’ variable. To provide a fit with the lowest SE model parameters values, an Exchangeable structure was used for the Working Correlation Matrix and a Model-based (cf. Robust-) estimator was selected for the Covariance Matrix. As the analysis involves Chi-Square, observed and expected frequencies were checked to ensure that $f$(observed) > 0, and $f$(expected) were not less than five for > 20\% of cells.

Once a model had been built, the associated $B$ and $SE(B)$ parameters were checked to be within sensible bounds (i.e., with values that were neither too low nor too high). As is normal for regression-type analyses where there is a possibility of suppressor (interacting) variables, a saturated model was adopted and predictors were subject to sequential backward removal if their presence was unimportant (based on the smallest Wald $X^2$ when $p > .10$). Using this procedure, visualisation (0 = none and 1 = external-feature visualisation) was removed at Step 1 ($X^2 = 0.0$) and type of interview (0 = CI and 1 = H-CI) at Step 2 ($X^2 = 0.8$).

The resulting GEE model was reliable (at the usual criteria of $p < .10$) for the interaction between these two factors [$\chi^2(3) = 7.1, p = .07$]. The interaction was significant as composites constructed after a H-CI (cf. CI) led to lower correct naming rates when no visualisation techniques were used [$Slope = -0.9, SE(B) = 0.4, p = .017$, Odds Ratio $1/\text{Exp}(B) = 2.3$, 95\% CI (1.2, 4.7)]. Other effects involved in the interaction approached significance ($p < .10$): there was a weak benefit of external-feature visualisation (cf. no visualisation) under H-CI [$B = 0.6, SE(B) = 0.4, \text{Exp}(B) = 1.8$ (0.9, 3.7)] but a weak deficit under CI [$B = -0.7, SE(B) = 0.4, 1/\text{Exp}(B) = 1.9$ (1.0, 3.8)].

No effects emerged as reliable predictors for names given to composites that were incorrect (a ‘mistaken’ name).

TABLE I. PERCENTAGE OF CORRECTLY NAMED COMPOSITES

<table>
<thead>
<tr>
<th>Visualisation technique</th>
<th>Interview</th>
<th>EF</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI</td>
<td>55.0abc</td>
<td>38.8abc</td>
</tr>
<tr>
<td>H-CI</td>
<td>33.8abc</td>
<td>47.5bc</td>
</tr>
</tbody>
</table>

* The term CI refers to a face-recall Cognitive Interview, H-CI to Holistic-Cognitive Interview, and EF to External Features. Values represent percentage-correct naming for composites for which participants correctly named the relevant target photograph ($N = 317 / 320$). GEE model parameters (QIC: 436.1; Intercept $B = -0.7, SE(B) = 0.2$, $p = .006, \text{Exp}(B) = 0.5$ (0.3, 0.9)) and indicated contrasts *$p < .05$ and **$p < .10$. See text for further details.
The present study examined whether the recognisability of EvoFIT composites could be improved when a combination of techniques were used to support construction, namely, holistic interview mnemonics, external-feature visualisation and directed focus on the eye region during face selection. When used alone in previous research, all three techniques have been found to produce reliable increases in correct naming of composites, relative to baseline conditions [e.g. H-CI vs. CI; external-feature visualisation vs. no visualisation; eye-region focus vs. internal-feature focus; 1, 8, 9, 12, 15]. While holistic interviewing techniques aim to re-instate a useful whole-face processing style before construction, visualisation and focusing techniques are believed to draw constructor attention towards internal features which, when accurately constructed, have the potential to enhance familiar face recognition [10]. In particular, we assessed whether a combined approach might lead to the construction of more recognisable composites compared with when just one approach is used (see [1 for a similar approach].

Several unexpected results were obtained. There was no clear evidence to suggest that more recognisable composites were constructed following (i) a H-CI (vs. face-recall CI), and (ii) instructions to engage in external-feature visualisation (vs. no visualisation). Instead, an interaction emerged: although composites constructed after a H-CI (cf. CI) were of significantly worse quality when no visualisation techniques were used, a standard, albeit marginal, H-CI benefit was observed when used in conjunction with external-feature visualisation. In contrast, composites constructed after a CI were better named when no visualisation techniques were used (cf. external-feature visualisation). If interviewing and visualisation techniques had acted to improve composite quality in the anticipated, additive fashion, then the best naming rates were expected to emerge under conditions in which the participant engaged in a H-CI, then constructed composites under external-feature visualisation. Instead, best performance was obtained under ‘baseline’ conditions, where participants constructed composites after a CI and engaged in no visualisation. Results then suggest that the current combination of techniques do not work harmoniously to support composite construction.

Failures to find (usually robust) benefits in composite naming as a result of a H-CI (cf. CI) are likely to stem from the parallel use of a newer, third technique: instructing the participant to focus on the eye region during initial face selection in EvoFIT [12]. Whereas the H-CI may encourage the witness to adopt an optimal ‘holistic’ processing style, which is likely to be useful for reaching the decision that a composite has reached a good level of visual likeness, the instruction to focus on the eye region may encourage the participant to revert back to a ‘feature-based’ processing style. Theories of transfer appropriate processing would predict this ‘processing conflict’ to have a negative impact on composite construction [5-7].

The H-CI (cf. CI) did lead to small, relative improvements in composite quality when combined with an external-feature visualisation technique. Arguably this technique complements a holistic processing style; by simply visualising external features throughout composite construction the participant is able to build internal features within the context of a whole face [14]. Perhaps then the alignment in the processing styles encouraged by both the H-CI and external-feature visualisation techniques (both holistic) mitigated against the conflict introduced when participants were asked to focus on the eye region (feature-based processing). Under these conditions a holistic processing strategy may have maintained dominance, particularly as participants were only required to adopt a feature-based strategy (focus on the eye region) during the early stages of face construction (when selecting faces from initial EvoFIT arrays).

In future, research participants may be provided with targeted instructions to lessen processing conflicts when a conjunction of techniques are employed during construction (namely the H-CI and eye region focusing technique). Current H-CI instructions encourage witnesses to consider the whole face when they make personality judgments. However, it may be possible to ask the witness to make the same holistic attributions whilst focusing just on the eye region. While the personality judgment itself may instate a holistic processing style, generally useful for feature selection and face recognition [6, 8, 9], being instructed to focus on the eye region while making these decisions may ease the processing transition when the witness receives similar instructions while selecting faces from initial arrays. We would thus predict an increase in composite identifyability as a result of using a combination of methods that encourage complementary processing strategies [1].

The transfer appropriate processing model may also help to explain why the best-named composites were (unexpectedly) produced in a condition where participants completed a face-recall CI with no instructions to visualise external features. Here, participants were arguably encouraged to adopt a consistent feature-based processing style throughout construction. Providing a detailed face description during the CI instates a feature-based processing style that remains unchallenged when the participant encounters internal-feature only faces during the initial EvoFIT arrays, and receives further instruction to focus on the eye region when making initial face selections. As this condition mirrors current procedure for forensic practitioners, we recommend that existing practices with EvoFIT remain unchanged.

In summary, the present work aimed to assess whether the recognisability of EvoFIT composites could be improved when using a combination of new and existing techniques to support construction; namely, holistic-cognitive interviewing, external-feature visualisation and a focus on the eye region when making selections from face arrays. We predicted that each technique would lead to separate increases in quality but that the best composites would be produced when all three techniques were employed; that is, additive improvements would be greater than summed, separate improvements [see 1]. Findings suggested otherwise: predicted effects were not found when using H-CI or external-feature visualisation techniques, and combining the two techniques did not lead to robust improvements. These findings suggest that combined techniques are only effective when they encourage
complementary processing strategies for face selection and recognition [1]. Promising results were afforded by a newer manipulation (asking participants to focus on the eye region during face selection), and current research attempts to align processing styles encouraged by this, and existing techniques, to improve composite quality. Findings of this ensuing work should have implications for current police procedures with EvoFIT, and potentially other holistic methods of face construction (e.g. [18]).

REFERENCES