Article

Adjusting the focus of attention: Helping Witnesses to Evolve a More Identifiable Composite

Fodarella, Cristina, Frowd, Charlie, Warwick, Kelly, Hepton, Gemma, Stone, Katie, Date, Louisa Hannah and Heard, Priscilla

Available at http://clok.uclan.ac.uk/22022/

It is advisable to refer to the publisher's version if you intend to cite from the work.
10.15406/frcij.2017.05.00143

For more information about UCLan's research in this area go to http://www.uclan.ac.uk/researchgroups/ and search for <name of research Group>.

For information about Research generally at UCLan please go to http://www.uclan.ac.uk/research/

CLoK
Central Lancashire online Knowledge
www.clok.uclan.ac.uk

Adjusting the Focus of Attention: Helping Witnesses to Evolve a More Identifiable Composite

Abstract

When perceiving faces under normal circumstances, the focus of attention is likely to be on the upper (e.g., hair, eyes) than lower (e.g., mouth, chin) facial half [1]. If such a bias were to extend to face construction, then it may hinder the effectiveness of forensic evidence collected from witnesses and victims of crime. In Experiment 1, participants constructed a single face using the EvoFIT holistic (police) composite system 24 hours after having seen an unfamiliar target identity. When constructing the face, participants were asked to select items from face arrays based on the whole face, or for upper and lower facial halves separately; these faces were presented in arrays either intact or horizontally-misaligned [2], and with external features (hair, ears, neck) present or absent. More-identifiable composites were predicted from (i) selection of separate facial halves (cf. currently-used whole-face selection), (ii) presence of horizontal misalignment and (iii) absence of external features. Experiment 2 used the same basic design but participants were requested to select for (i) upper-face half during evolution, (ii) the same as (i) but also for subsequent adjustments of the face, and (iii) overall face (Control). The composites constructed in both experiments were named and rated for likeness. Experiment 1 unexpectedly revealed that the Control group produced the highest-named composites. In Experiment 2, upper face selection during the evolution stage produced more-effective composites. In terms of practical implications, for the EvoFIT composite system, and potentially for other holistic systems, witnesses should be instructed to select faces for the upper facial half during evolution, to maximise subsequent identification of their composites.

Keywords: EvoFIT; Upper facial features; Holistic facial composites; Composite-face effect; Lower facial features; External facial features

Cristina Fodarella[1]*, Charlie D Frowd[1], Kelly Warwick[2], Gemma Hepton[1], Katie Stone[2], Louisa Date[1] and Priscilla Heard[3]

[1]School of Psychology, University of Central Lancashire, UK
[2]Department of Psychology, University of Winchester, UK
[3]Department of Psychology, University of the West of England, UK

*Corresponding author: Cristina Fodarella, School of Psychology, University of Central Lancashire, Preston, UK, Email: CFodarella3@uclan.ac.uk

Introduction

The eye region plays an important role for the identification of familiar faces [3-5]. More attention is paid to the eyes than to any other facial feature [6,7] possibly due to the eyes' communicative value. Even in situations where individuals need to concentrate on lip movements in order to understand speech submerged in noise, eye-movement data indicate that participants attend to the eyes for the majority of the time [8]. It would seem that due to the fact that eyes reveal a considerable amount of information with regard to communication, mood and emotions, more attention is paid to them [7-9]. The fact that individuals with social issues, such as autism, show an impairment in judging eye manipulations [10], reinforces the theory that, typically, we attend to eyes more than any other internal feature due to their social communicative nature. In line with that research, upper facial features (eyes, brows and hair) tend to be more indicative of an individual's identity than lower facial features (mouth and chin) [1,11]. Not only adults, but also newborns, show a preference for fixating on upper as opposed to lower features [12]. Attention to upper or lower features can also be investigated by applying changes to different facial features and exploring participants' ability to detect these changes (by speed and accuracy) [13,14]. Results indicate that changes to upper features are detected faster than changes

to lower features [13], with hair alterations being detected the fastest, most accurately and most confidently, followed by eye alterations [14]. Changes made to the eye region also tend to be detected more readily in familiar than unfamiliar faces [3].

Although faces can be recognised by their individual features [4,14], face perception occurs holistically, as demonstrated with the so called composite-face effect (CFE[1]) [2]. In this effect, when top and bottom facial halves of two different identities are joined together, people have difficulty identifying the individual halves in the resulting face [2,15,16]. The finding indicates that the regions of the face (in this case the facial halves) are perceived differently when seen separately or in the context of a complete ('composite') face. By misaligning ('splitting') a compound face horizontally (as illustrated in Figure 1, right), both accuracy and speed-of-response increase [2,15-18]. This indicates that holistic processing is interrupted under this condition; that is, when a face is horizontally-misaligned. This holistic effect not only occurs for identification and same-different judgments of facial halves [2,15], but also in other face-processing tasks where judgment for the top half of a face is significantly slower and also influenced by the bottom face of a different half when aligned, but not when

---

[1]Note that the current paper makes a distinction between the composite face effect (CFE), as described by Young et al. [2], and facial composites, as used by forensic practitioners.

misaligned [19-23]. The CFE also occurs in faces where eyes are distorted by moving either one eye or two eyes upwards by an extensive amount [24], indicating that holistic processing persists even in unnatural-looking faces.

In a classic psychological study, Young et al. [2] found that the CFE diminishes when such novel faces are misaligned but also when they are inverted (i.e., turned upside down). Inversion generally leads to concealment of configural information (i.e., physical distances between individual facial features), and faces tend to be recognised by featural information [25,26]. Young et al. [2] theorised that the CFE occurs from interference to inaccurate configural properties created from fusing the two separate halves, and inverting or splitting a face assists by concealing this inaccurate information. It may seem that there is a conflicting issue: the general holistic processing of a face [2,21,24] versus the misalignment of a face for improved perception [2,15-18]. However, the misalignment of a face only appears to be beneficial when dealing with a novel face consisting of upper and lower halves from two different identities. Due to natural holistic processing of faces, judgment of the upper half of a face is worse when aligned with a different, lower half [19,21-23]. Therefore,

perception of a face is only improved in this instance, and not in normal faces, where both the upper and lower region is of the same identity, age, race, emotional congruence, etc.

The perceptual importance of upper features is also evident in facial-composite research. Upper features have been shown to be selected first, and for a longer period of time, during construction using the archaic Photofit system [27]. Upper features of constructed composites also seem to receive greater attention during naming, and are more helpful for identification than lower features [28-30]. Further research has revealed that same-different judgments of Identikit composites are made faster for upper than for lower facial regions [31]. More recent research suggests that recognition of finished composites is facilitated if observers perceive upper and lower features in a horizontally misaligned format [32]. Although evidence of this processing bias towards the upper face stems from old composite systems, it may extend to face construction of composites from the newer "holistic" systems, such as EFIT-V, ID and EvoFIT [33]. As to date, the upper face salience has not been investigated in these new systems, the current paper aims to plug this gap.
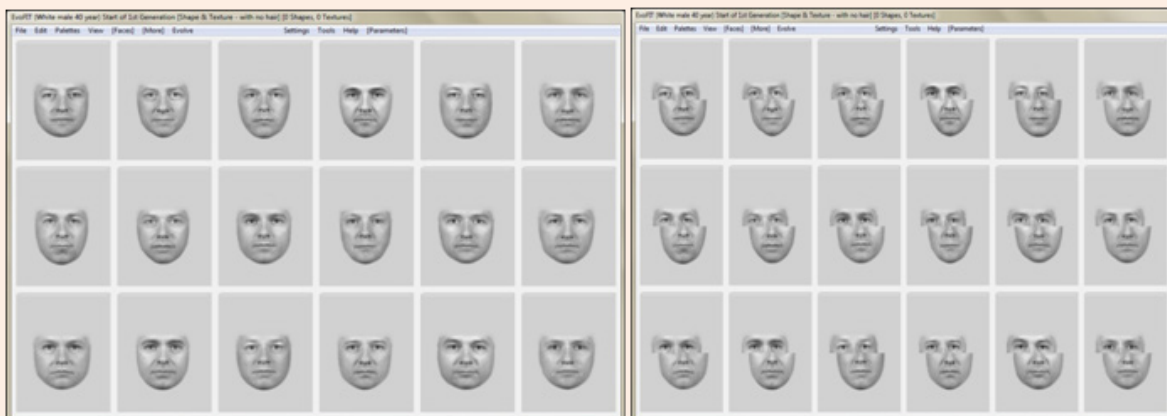

Figure 1: Example EvoFIT screens showing arrays of internal-only faces from which witnesses usually make selections (left), and arrays of horizontally-misaligned faces also presented to witnesses in Experiment 1 (right).

With holistic composite systems, constructors are requested to make repeated selections based on the overall likeness from arrays of faces. The characteristics of selected faces are combined and, over time, a composite is 'evolved'. These systems also include 'holistic' tools, to adjust the overall likeness of the face; and 'shape' tools, to manipulate facial features on demand. However, an unexpected finding is that presenting hair in these face arrays actually reduces composite quality (i.e., produces lower naming rates) [34,35], presumably as exterior information detracts constructors' attention from internal features (the region including eyes, brows, nose and mouth). Although hair can be a beneficial retrieval cue [36], it does not seem to be useful in this applied context. Part of the reason for this is that witnesses are required to recall and select the hair from memory, and so this feature is unlikely to be a perfect match [37], causing distraction and / or inadequate context for face selection [35]. Blurring

external features in the face arrays improves resulting composite likeness, but eliminating them entirely results in composites with higher correct and lower incorrect identification [35].

Accounts of participant-witnesses could provide a further explanation as to why hair is distractive during construction [34,35]. These accounts indicate that hair tends to be the main feature assisting participant-witness memory of the target face of which they are required to produce a composite [38]. Therefore, it would seem that witnesses may focus on hair to a greater extent than to internal features to the detriment of the latter. Indeed, hair tends to be described in greater detail when constructing composites than any other feature [30]; in general of course, we describe hair more often than any other facial feature, and it is established that this feature (along with other external features) is important for unfamiliar face recognition [39,40]. Consequently, EvoFIT presents face arrays that reveal internal features only

(Figure 1), with hair and the remaining external features added towards the end of the procedure. Even when presenting just internal features, it is possible that constructors may be biased to make choices on the upper facial half due to greater salience [1,11] and greater attention to the eye region [6,7]: a procedure that may result in inferior construction of the lower half of the face. Indeed, it does appear to be the case that the lower facial region is sometimes constructed inaccurately (for an example case, see http://tiny.cc/ManPRJ) and this effect could lead to a reduction in the ability to correctly name a witness's composite (e.g., by a police officer or member of the public). In the current work, over the course of two experiments, we investigated the impact of shifting participant's attention to upper and / or lower regions of the face during face construction. This investigation is of forensic importance, as increasing the overall likeness of a composite should lead to increases in suspect identification rates.

Experiment 1

In Experiment 1, we explored whether shifting constructors' attention to upper and lower facial regions separately would result in an overall more-identifiable representation compared with the current procedure of selecting items for the whole face (which may be biased towards the upper facial region). We followed a design that attempted to copy real-life procedures as far as possible, to be of direct relevance to forensic practice. This involved participants viewing an unfamiliar target and constructing a single EvoFIT composite of it after 24 hours; the resulting composites were then evaluated for effectiveness by asking other people to name them and to assign ratings of likeness in the presence of the target. There were four experimental conditions and a fifth, Control condition. The experimental conditions differed in the way in which faces were presented in EvoFIT arrays during face construction: they were arranged into two factors, horizontal misalignment (absent; present) and external features (absent; present).

Horizontally-misaligned faces (illustrated in Figure 1, right) were included with the aim of assisting face selection, as we are generally poor at perceiving upper and lower regions separately (cf. aligned facial halves). It is worth emphasising that constructors are likely to have difficulty selecting facial halves since the configuration (the spacing of features) in each face in the array is unlikely to be correct. Therefore, splitting the faces in this way should help them to make judgments about the separate regions [2]. It might seem contradictory that EvoFIT face selection is dominated by upper features, whilst upper and lower features are difficult to judge within an aligned face. However, if perception of upper features dominates during construction, then not only upper but also lower features should be judged better when misaligned. In addition, external features were included to investigate whether the disruptive effect of hair [35] would be removed if the face is no longer processed holistically (due to the presence of misalignment). In these four experimental conditions, to shift attention to upper and lower facial halves, participants were asked to select one face for the upper facial region and one face for the lower facial region. The authors do acknowledge that it is unclear whether a poorly constructed lower region is due to the upper region receiving more attention during face construction or whether the upper region is simply remembered better. In either case, with participants selecting for the lower region, the result may, in fact, be a worse quality composite.

To assess the general importance of these experimental conditions, a fifth condition was included, the current EvoFIT procedure used to construct a face: no misaligned faces, external-features absent and an explicit instruction to constructors to select all faces for overall likeness. This is referred to as the Control condition[2]. The following predictions were made for Experiment 1:

Attention by the constructor to upper region for one face and lower region for another will lead to more-identifiable composites than those in the Control condition (selection of two faces for overall likeness). This effect will be stronger when faces are horizontally-misaligned than aligned.

Lower features of the composites will be rated to be of worse likeness than upper features for the Control condition, and likeness ratings of lower features will be highest in the misaligned external-features-absent condition.

Method

We present details of the experiment as three successive stages of data collection: face construction, composite naming and composite likeness ratings.

Face construction

Design: Face construction was a between-subjects' design with participants constructing composites using EvoFIT in one of five construction conditions: 2 (horizontal-misalignment: absent; present) x 2 (external features: absent; present) + current EvoFIT procedure (as described above).

Participants : Fifty students from the University of Central Lancashire volunteered (27 males; 23 females; M_{age} = 21.4; SD_{age} = 4.2 years). Participants were recruited on the basis of being *unfamiliar* with targets in the study (characters from the ITV Coronation Street soap), to mirror the usual situation for composite constructors in the real world.

Materials: Ten photographs of characters from a popular UK TV soap (Coronation Street: Peter Barlow, Michelle Connor, Jason Grimshaw, Tracey McDonald, Karl Munro, David Platt, Gloria Prize, Kirk Sutherland, Sally Webster and Sophie Webster) were printed in colour (8 cm x 10 cm). As we are interested in overall effects for construction of all targets (male and female), we have included both genders in the experiment. Our intention is to analyse the overall effect rather than by gender (which may otherwise provide misleading results as we only include five targets per gender condition).

Procedure: Participants were tested individually and tasks were self-paced. A trained composite researcher with considerable experience worked with participants in all conditions to create the composites. Participants first viewed an unfamiliar target face of a Coronation Street actor, randomly selected, for 60 seconds in the

---

[2]We acknowledge that the inclusion of the Control condition creates a non-balanced design. However, this was intentional: the principal aim of this paper was to identify the effect of the two main factors (i.e., horizontal misalignment, and external features). Therefore, these two factors will be analysed first. Following this, in a separate analysis, we analyse how these manipulations fare against the current procedure for EvoFIT. Should either factor suggest a more-effective way to construct a face, we will attempt a replication to check for consistency of results.

knowledge that a composite would be required (i.e., intentional encoding was used). After 20 to 28 hours, each participant was interviewed by the researcher using a Cognitive Interview, to recall the appearance of the face, and created a single composite with EvoFIT. Participants were randomly assigned to one of five conditions within which the same 10 target faces were used for construction, to produce a total of 50 composites.

The Control condition of the experiment is the procedure taught to forensic practitioners who use EvoFIT [41]. This involved presenting arrays of 18 aligned (intact) faces without external features (Figure 1). Ten constructors first selected a single item for best match of facial width, a property of the face that was displayed on subsequent screens. Participants were then asked to select two faces per screen (12 screens in total) for overall likeness. The remaining 40 constructors were also presented with arrays of 18 faces, but the way these were presented differed: faces in arrays contained

i. External features present (selected at the start) or absent, and

ii. Intact or horizontally misaligned faces.

These 40 constructors were also asked to select two faces per screen; this time, however, they were they were asked to select one face of best likeness for the upper half of the face, and one face for the lower half of the face.

Once a face had been evolved in this way, horizontal misalignment, if used, was disabled. In all conditions, the face was now enhanced using holistic tools (to change age, weight and other overall properties of the face) [42] and shape tools (to adjust size and placement of individual features). Hair and external features were then added in conditions in which hair was initially absent (in the other conditions, these exterior features were added at the very start of face construction). For these holistic and shape enhancement tasks, constructors were asked to focus on the face as a whole irrespective of the condition to which they had been assigned. Once the best likeness had been reached, the face was saved to disk as the composite. Experimental sessions lasted between 60 and 90 minutes.

Composite naming

Design: Naming was a between-subjects' factor with participants invited to name composites created from one of the five conditions: 2 (horizontal-misalignment: absent; present) x 2 (external features: absent; present) + Control.

Participants : Fifty students from the University of Central Lancashire volunteered (3 males; 47 females; M_{age} = 24.6; SD_{age} = 6.4 years). Participants were recruited on the basis of being familiar with characters appearing on the Coronation Street soap and were assigned equally to the five levels of between-subjects factors of the experiment.

Materials: Ten target photographs and 50 EvoFIT facial composites (Figure 2) were printed in greyscale (the image mode of the composite system) to dimensions of 8 cm x 10 cm.

Procedure: Participants were tested individually, and randomly assigned to one of five conditions as per the experimental design. Participants were briefed that they would be shown composites of Coronation Street actors. They were then presented sequentially with composites from their assigned condition to name, or guess if unsure; it was also explained that they could, if desired, opt to not offer a name. Participants were then asked to name the 10 target photographs, as a check that they were familiar with the relevant identities. Participants received a different random order of presentation for composites and targets. The task was self-paced, and lasted for 10-15 minutes per person.

Composite likeness ratings

Design: Composite likeness ratings was a within-subjects' design, with further participants rating both upper and lower halves of composites constructed from all five conditions.

Figure 2: Example composites of actor Karl Munro constructed by condition:

i. No-misaligned external-features-absent,

ii. No-misaligned external-features-present,

iii. Horizontally-misaligned external-features-absent,

iv. Horizontally-misaligned external-features-present and

v. Current (Control) procedure. Due to reasons of copyright, we are unable to reproduce the relevant target photograph (but a simple search on the Internet should reveal his identity).

Participants: Eighteen students from the University of Central Lancashire participated voluntarily (8 males; 10 females; M_{age} = 25.4; SD_{age} = 8.5 years). They were recruited on the basis of being unfamiliar to Coronation Street actors and were different to those involved in Stages 1 and 2.

Materials: Facial composites were manipulated with Adobe Photoshop by highlighting and expurgating external features, resulting in internal-only composites. These were then separated in the middle of the nose, creating two sets of stimuli: one showing upper internal features, and one showing lower internal features (Figure 3). Target photographs and composites were printed in greyscale (8 cm x 10 cm).

Procedure: Participants were tested individually. Ten target photographs were presented in a different random, sequential order for each participant. Alongside each target photo, all upper and lower facial regions of composites associated with that identity (constructed across all five construction conditions) were

(a)   (b)

Figure 3: Example
  i. Upper and
  ii. Lower features for one of the composites of Karl Munro used in the likeness rating task.
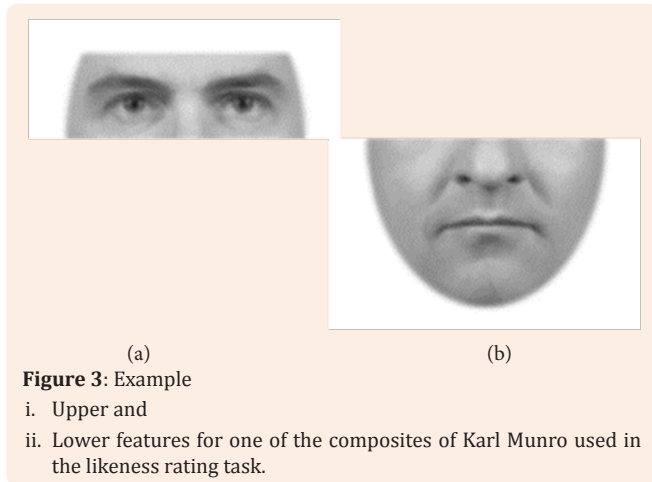
presented sequentially in a different random order for each target (that also changed randomly across participants). Participants were asked to rate the composite likeness (1 = poor likeness, 7 = good likeness). The task was self-paced and took about 15 minutes to complete.

Results

Composite naming: Participant responses to composites and targets were checked for missing data (of which no cases were found). Responses were then scored for accuracy with respect to the relevant identity, with a numeric value of 1 assigned if the item had been correctly named and 0 otherwise (for wrong name or no name given). Target photographs were named very well, at 99.0% correct overall, suggesting that participants appropriately had high familiarity with the relevant identities. For the cases for which the target was not correctly named, the associated composite could not have been correctly named (N = 5), and so responses to these composite items were recorded as 'missing data' and not analysed any further. The resulting composite scores are summarised in Table 1. Correct naming was 32.5% overall, which is of course much lower than correct naming of targets, but this is the usual situation as composites are constructed from

memory and are rarely named perfectly. It is also apparent that mean naming varied little by face presentation (construction using arrays with or without a horizontal split), or by external-features type (whether external features were or were not present during the main part of construction). In comparison, mean naming was much higher in the Control condition, the current EvoFIT procedure.

Table 1: Correct naming of composites by construction condition (face presentation and external features).

External Features	Face Presentation		Control
	Normal	**Misaligned**	
Absent	27.55	34.00	41.84†
	(27 / 98)	(34 / 100)	(41 / 98)
Present	30.00	28.28	-
	(30 / 100)	(28 / 99)	

Note: Values are correct naming scores calculated by dividing responses shown in parentheses and expressed as a percentage. Parenthesised values are summed correct responses (numerator) and total (correct and incorrect) responses (denominator) for composites that participants correctly named the relevant target (N = 160 out of 495).

†Greater than central tendency of the four other conditions, $p < .05$; Model's Constant [Beta (gradient) coefficient, B = -0.59; standard error of B, $SE(B)$ = 0.12; Wald $X^2(1)$ = 25.73; $p < .001$; Odds Ratio (effect size), $Exp(B)$ = 0.56].

Individual naming responses to composites were analysed using Binary Logistic Regression, to provide a combined by-participants and by-items model. All data were checked for appropriacy for carrying out this goodness-of-fit test (i.e., observed cell $f > 0$, and expected $f >= 5$ for at least 80% of cells). Two planned analyses were conducted. The first explored the effect of the two main factors, face presentation (*0* = normal and *1* = horizontal misalignment) and external-features type (*0* = external-features absent and *1* = external-features present), on the dichotomous Dependent Variable, correct naming. These two predictors were entered as a full-factorial model and subjected to Backward Stepwise elimination using Likelihood Ratio and probability of removal, $p = .1$. External-features type was removed at Step 1 ($p = .74$), face presentation at Step 2 ($p = .61$) and the interaction between these two factors at Step 3 ($p = .38$). The resulting model (which thus contained no predictors) suggested that none of the factors exerted a reliable effect on correct naming of composites.

In a second analysis, responses were collapsed over these two predictors and compared against responses given in the Control condition. This approach provides an indication of how the experimental conditions fare against the current EvoFIT procedure. A second logistic regression was run using a single predictor coded as *1* for Control and *2* otherwise. The resulting model emerged significant [$X^2(1)$ = 4.90, $p = .027$, R^2 = .01 (Cox and Snell) and .01 (Nagelkerke), Hosmer and Lemeshow $X^2(0)$ = 0] and simple contrasts revealed that the Control condition produced superior composites to the other conditions combined [B = 0.52, $SE(B)$ = 0.23, $X^2(1)$ = 5.00, $p = .025$, $Exp(B)$ = 1.68 (*95CI-* = 1.07, *95CI+* = 2.65)].

It is usual practice in composite research to consider the impact of wrong names offered to composites, to provide an indication of tendency to offer any name (i.e., a response bias); in a practical context, incorrect names may be considered undesirable as they can provide a false lead in a police investigation, but they are useful by allowing potential suspects to be eliminated from an enquiry. Participant responses were recoded, assigning a value of *1* for wrong (incorrect) name and *0* otherwise, and cases were labelled as 'missing' for which targets were incorrectly named (as above) as well as for composites that were correctly named. Responses were appropriate for a frequency-type analysis, and were 22.6% incorrect overall. The two analyses run in the same way as above both emerged with predictors that did not exert a significant effect on the DV, indicating that incorrect names given did not reliably change across the experiment.

Composite rating: Repeated-Measures (RM) ANOVA was run using composite ratings including four of the five composite conditions (i.e., excluding Control) for rating type (upper; lower), external-features type (absent; present) and horizontal-misalignment type (absent; present). Rating type was not significant ($F = 1.3$, $p = .30$), but there was a main effect of external features type [$F_{(1,8)} = 14.37$, $p = .005$, $\eta^2_p = .64$], indicating that composites constructed in the external-features absent condition were rated significantly higher overall than those constructed in the external-features present condition. There was also a main effect of horizontal-misalignment type [$F_{(1,8)} = 9.36$, $p = .016$, $\eta^2_p = .54$], suggesting (contrary to prediction) that composites constructed with no-misaligned faces were rated higher than those created with horizontally-misaligned faces.

There was one reliable interaction (Table 2): between external-features presence and rating type [$F_{(1,8)} = 9.25$, $p = .016$, $\eta^2_p = .54$]. Paired-samples t-tests (two-tailed) revealed that upper features of composites constructed without external features were rated significantly higher than upper features of composites constructed with external features present [$t(8) = 4.94$, $p = .001$, $MD = .43$ ($95CI- = 0.23$, $95CI+ = 0.64$), $d = 2.1$]: in contrast, lower features were rated similarly regardless of whether external features were present or not ($t(8) = 1.1$, $p = .30$). Further, whilst no reliable difference was found between ratings of lower and upper features of composites created with external features present [$t(8) = 1.2$, $p = .26$], when external features were absent, upper features were rated significantly higher than lower features [$t(8) = 2.40$, $p = .043$, $MD = .24$ ($95CI- = -0.01$, $95CI+ = 0.48$), $d = 1.3$]. Therefore, the initial finding of composites constructed without external features being rated higher than those with external features present was mediated by this interaction effect. That is, upper features were rated better when constructed without external features; this effect was not mediated by whether faces were aligned or misaligned.

A second RM ANOVA was run using by-participant composite rating data from all five conditions (Table 3). Results indicate a main effect of rating type [$F_{(1,8)} = 80.49$, $p < .001$, $\eta^2_p = .91$], with upper features rated higher than lower features. There was also a main effect of group [$F_{(4,32)} = 13.20$, $p < .001$, $\eta^2_p = .62$], and two-tailed simple contrasts indicated superiority of the Control condition relative to the four experimental conditions [$p < .01$, $MD > 0.25$ ($95CI- > 0.1$, $95CI+ > 0.4$), $d > 1.7$] except for split external

features absent ($p = .28$). The interaction between rating and group was also significant [$F_{(4,32)} = 32.94$, $p < .001$, $\eta^2_p = .81$]. Paired samples t-tests indicated that upper features were rated significantly higher than lower features in composites constructed in the Control condition [$t(8) = 12.60$, $p < .001$, $MD = 1.66$ ($95CI- = 1.35$, $95CI+ = 1.96$), $d = 6.6$], but not in any of the remaining conditions ($p > .05$). Relative to Control, ratings for upper features were significantly less [$MD > 0.75$ ($95CI- > 0.34$, $95CI+ > 1.19$), $d > 2.3$], and ratings for lower features were significantly higher [$MD > 0.38$ ($95CI- > 0.23$, $95CI+ > 0.54$), $d > 1.8$], for composites constructed in all remaining conditions ($p < .005$). Point-serial correlations (for all 50 composite items) were positive and large between mean correct naming and ratings of both upper [$r(48) = .45$, $p = .001$] and lower [$r(48) = .39$, $p = .005$] facial halves. Also, neither of these correlations changed substantially when controlling for ratings of the opposite facial half (with r decreasing by about .1), suggesting that both regions contributed equally to correct naming.

Table 2: Composite likeness ratings (and SD) for external features type (present; absent) and rating type (upper; lower). Data are collapsed over horizontal misalignment type (as this factor, and all of the interactions involving it) did not emerge reliable.

External Features Type			
Rating Type	**Absent**	**Present**	**Mean**
Upper	3.53[a,b]	3.10[a]	3.28
	(0.32)	(0.28)	(0.31)
Lower	3.29[b]	3.19	3.26
	(0.22)	(0.23)	(0.07)
Mean	3.41	3.14	
	(0.17)	(0.06)	

Note: Pairwise contrast significant, [a]$p < .005$, [b]$p < .05$.

Discussion

We aimed to improve correct naming rates of composites with arrays of faces presented to participants at construction in one of four experimental conditions. The aim was not successful. In fact, the current procedure (the Control condition) produced the best named composites, and likeness ratings indicated that this occurred due to a more-accurately constructed upper-face region. For the other conditions involving selection by individual facial halves, likeness increased for the lower face but decreased for the upper face. Also, based on correlations for ratings of internal composite likeness to correct naming, it was also apparent that both upper and lower regions of the face contributed somewhat equally to correct naming. This itself is an intriguing result, as one would expect a stronger association for the upper (cf. lower) face, but this is probably due to likeness rating being a less sensitive measure for the upper half (as we argue in the General Discussion). So, the instruction to select for overall likeness is clearly preferable to the one to select for different facial halves; this instruction also results in a tendency for witnesses to select for the upper half of the face—the region that is important for correct naming of the composite. However, it still may not be optimal for naming if witness' choices are not always based on the upper facial half. This idea is explored in Experiment 2.

Table 3: Composite ratings (and standard deviations) across rating type (upper; lower) and composite construction procedures (A = Current [Control]; B = no-misaligned external-features-absent; C = no-misaligned external-features-present; D = horizontally-misaligned external-features-absent; E = horizontally-misaligned external-features-present).

	Construction Procedure Type					
Rating Type	A	B	C	D	E	Mean
Upper	4.42[a,b]	3.66[b]	3.41[b]	3.22[b]	2.98[b]	3.54
	(0.29)	(0.34)	(0.27)	(0.27)	(0.24)	(0.50)
Lower	2.77[a,c]	3.34[c]	3.23[c]	3.22[c]	3.16[c]	3.14
	(0.20)	(0.21)	(0.22)	(0.28)	(0.17)	(0.20)
Mean	3.60	3.50	3.32†	3.22†	3.07†	
	(1.18)	(0.22)	(0.13)	(0.01)	(0.13)	

Note: Contrast [a] $p < .001$. Contrasts relative to Control (A)[b,c] $p < .005$.

†Less than Control (A), $p < .01$.

Experiment 2

Method (face construction and evaluation)

Design: In the current experiment, participants were asked to make selections based on the upper half of the face. Due to the relative importance of this region for later recognition of the composite, this approach should produce composites that are more identifiable than those produced in the usual way (selection for the face as a whole). In fact, if it is the case that the upper face is so important, then it is conceivable that even-more identifiable composites will be produced for constructors who focus on the upper facial half not just when selecting from arrays (to 'evolve' a face) but thereafter, when making adjustments using holistic tools and shape tools. This suggestion is similar to a finding by Frowd et al. [35]: EvoFIT composites with higher naming rates were constructed when external features were masked (cf. blurred) in presented faces during the evolving stage, but even-higher naming emerged when this masking procedure was extended to holistic and shape tool use.

In the current study, we therefore compared naming of composites created from participants who were asked explicitly to focus on

i. The upper face during evolution (Upper face I),

ii. The upper face during evolution and holistic- and shape-tool use (Upper face II), and

iii. The overall face throughout (Control).

The design used to construct and evaluate (name and rate) the composites was the same as in Experiment 1, except that now there was a single between-subjects' factor (method of face selection) and target photographs were 10 EastEnders's characters (Ian Beale, Lauren Branning, Max Branning, Stacey Branning, Danny Dyer, Carol Jackson, Billy Mitchell, Ronnie Mitchell, Alfie Moon and Kat Moon). For reason of efficiency, composite naming and likeness-rating data were collected during the same testing session. Ratings for the upper half of the composite were expected to be higher for Upper face I relative to Control, but be even higher for Upper face II; ratings for the lower half were expected to be higher in the Control relative to the two other conditions.

Participants and procedure: Participants constructing composites were sampled widely from staff and students at the University of Winchester, and staff working in High Street stores in Berkshire, UK (M_{age} = 24.1; SD_{age} = 7.0 years). As before, participants encoded one unfamiliar target, randomly selected, and then described and constructed a single EvoFIT composite of it after 20 to 28 hours. They were randomly assigned with equal sampling to one of three conditions:

i. Upper face I,

ii. Upper face II and

iii. Control.

In (i), once the evolving stage had been completed, participants were instructed to make overall changes to the face (i.e., from holistic tools onwards); in (ii), this instruction was issued after participants had added hair and other external features. Session times were similar to before; each person was offered a £10 honorarium.

Evaluation of the composites was carried out by recruiting a different group of participants with demographics the same as the face constructors; there were 30 volunteers (6 males, 24 females; M_{age} = 26.9; SD_{age} = 8.4 years). The procedure and presentation of materials for naming were the same as in Experiment 1, except that participants were recruited on the basis of being familiar with characters from EastEnders. They were presented with 10 composites from one of the three face-selection conditions, randomly selected with equal sampling. After attempting to name their assigned set of composites, as before, participants were asked to name the target pictures. Participants next completed a rating exercise. This involved rating one half of the assigned set of composites against the target photo, and then the other half (order of facial halves, upper and lower, was counterbalanced across participants, with random assignment, and for composites with a different random order of presentation for each person). The scale used was the same as before (1 = poor likeness, 7 = good likeness). Evaluation sessions took about 20 minutes per person.

Results

Composite naming: Composites and target photographs were scored, checked for missing data and appropriacy for goodness-of-fit tests, and analysed in the same way as Experiment 1. For correct responses, target naming was very high overall (M = 96.7%) and composite naming was fairly good (M = 29.3%). As can be seen in Table 4, relative to Control (focus on overall face), composites were named much better from constructors in Upper face I, who were asked to focus on the upper half during the evolving stage, but (unexpectedly) only slightly better named in Upper face II, from constructors who were asked to focus on the upper half for an extended period (i.e., for use of holistic and shape tools, up to addition of hair and external features).

Table 4: Correct naming of composites from constructors who were asked to focus on the whole face (Control) relative to those asked to focus on the upper facial half during evolving (Upper half I) and upper half for longer,

until addition of hair and other external features (Upper face II).

Whole Face	Upper Half I (evolve)	Upper Half II (Extended)
19.79	42.86†	25.00
(19 / 96)	(42 / 98)	(24 / 96)

Note: †Greater than the other two conditions, p < .01. Model's Constant [B = -0.93, SE(B) = 0.14, X2(1) = 47.65, p < .001, Exp(B) = 0.37].

Binary Logistic Regression conducted on the correct-naming responses for composites was significant for face selection (1 = Upper I, 2 = Upper II and 3 = Control) [$X^2(2)$ = 13.51, p = .001, R^2 = .05 (Cox and Snell) and 0.07 (Nagelkerke), Hosmer and Lemeshow $X^2(1)$ = 0, p = 1.0]. Simple contrasts revealed that while the Control produced worse composites than Upper face I [B = 1.11, $SE(B)$ = 0.33, $X^2(1)$ = 11.52, p = .001, $Exp(B)$ = 3.04 ($95CI$- = 1.60, $95CI$+ = 5.78)], there was no reliable difference to Upper face II [B = 0.30, $SE(B)$ = 0.35, $X^2(1)$ = 0.75, p = .39, $Exp(B)$ = 1.35 ($95CI$- = 0.68, $95CI$+ = 2.67)]. Upper face I was also superior to Upper face II [B = 0.81, $SE(B)$ = 0.31, $X^2(1)$ = 6.74, p = .009, $Exp(B)$ = 2.25 ($95CI$- = 1.22, $95CI$+ = 4.15)]. Binary Logistic Regression, carried out in the same way as before, revealed that incorrect names did not reliably change by face selection [$X^2(2)$ = 0.65, p = .72].

Composite rating: Mean participant ratings by face selection and facial half are summarised in Table 5. Mean values are very similar in all cells of the design except that it was much worse in Upper half II for ratings of the lower half of the composite. RM ANOVA was significant for neither face selection (F < 1) nor facial half (F < 2); however, the interaction between these two factors was reliable [$F(2,21)$ = 2.30, p = .026, η^2_p = .29], due to the inferior rating for lower facial halves in Upper half II relative to adjacent row and column means. Clearly, composites in Upper face II emerged with an inferior lower portion of the face. Correlations (for all 30 composites) were positive and large between mean correct naming and ratings of both upper [$r(28)$ = 0.45, p = .012] and lower [$r(28)$ = .53, p = .003] facial halves, as found in Experiment 1. Again, neither correlations changed substantially when controlling for ratings of the opposite half (r decreased by ~0.1), suggesting that both regions contributed equally to correct naming[3].

Lastly, we reflect on the relevance of correct naming of composites created using the standard (Control) procedure in Experiments 1 (M = 41.84%) and 2 (M = 19.79%). Clearly, for such homogenous samples, some individual differences are expected, but these differ from chance [$t(18)$ = 4.3, p < .001, MD = 24.0 ($95CI$- = 13.9, $95CI$+ = 34.2), d = 2.4]. Ratings of the upper half were also much greater in Experiment 1 than 2, with the opposite for the lower half, emerging as a significant interaction between experiment and facial region [$F(1, 15)$ = 6.3, p < .001, η^2_p = 0.87]. This result clearly fits with the idea that an upper half which is a better match to a target is associated with a more-effective composite (see below). In Experiment 1, while some benefit (cf. Experiment 2) is likely to be caused by properties of the target (e.g., an increase in distinctiveness) [43,44], it may simply be that these constructors paid greater attention to the upper half

anyway; such a tendency is reduced, with more consistent faces constructed, by asking witnesses to select for the upper half in the first place, which is exactly what Experiment 2 set out to consider.

Table 5: Mean ratings of individual facial halves for the constructed composites by method of face selection.

Facial Half	Face Selection		
	Whole Face	Upper Half I (Evolve)	Upper Half II (Extended)
Upper	3.85	3.93	4.01
	(3.15, 4.55)	(3.22, 4.63)	(3.31, 4.72)
Lower	3.91	4.06	3.20
	(3.27, 4.56)	(3.42, 4.71)	(2.55, 3.85)

Note: Rating scale (1 = poor likeness, 7 = good likeness). Values in parentheses are ±95% CI.

Discussion

The aim of Experiment 2 was to improve identification of composites by instructing (participant-) witnesses to select for the upper half of the face during the evolving stage of construction as well as for an extended period [cf. overall selection in the current (Control) procedure]. Correct naming rates of finished composites were significantly higher for instruction to select for the upper face during evolution compared to the other conditions. When faces were selected for the upper face for an extended period, the resulting composites had a significantly worse lower facial half than in other conditions, indicating the importance of upper-face selection during the evolution stage. Results also indicate that upper-face (cf. overall) selection leads to more consistently-effective composites.

General Discussion

The current studies investigated whether the quality or effectiveness of composites could be improved by manipulating the construction procedure to shift participant-witness's attention to separate upper and lower face regions (Experiment 1) as well as to the upper face region only (Experiment 2). The anticipated impact of this procedure was to improve the overall likeness and thereby facilitate naming. Contrary to expectation, composites created with the current (Control) procedure in Experiment 1 were correctly named better compared to other conditions. With the exception of one condition (horizontally-misaligned, external-features-absent), the Control also received overall higher likeness ratings than the four remaining conditions, where attention was directed to upper and lower regions separately. In Experiment 2, composites emerged with higher correct naming when participants were instructed to select for the upper face during the evolving process compared to the other two conditions.

The face-perception literature indicates that more attention is generally paid to upper than lower facial regions in photographs of faces [1,12-14] and during facial-composite construction [27]. Upper features are also more indicative of identity in photographs of faces [1,11] and facial composites [30]. Based on these results, we theorised that during normal EvoFIT composite construction, faces may be selected on their upper rather than lower facial

---

[3]The same finding emerged from these same correlations for when likeness and correct naming data were pooled across experiments.

regions, leading to worse construction of the lower region of the face and a less identifiable composite overall. Therefore, asking participant-witnesses in Experiment 1 to select faces based on individual regions rather than for overall likeness should increase attention to the lower region and improve identification of composites. This was not supported: the usual procedure of asking constructors to select by overall likeness led to better-named composites than asking them to select for individual-face regions.

Referring to likeness ratings of the separate composite regions indicate a plausible mechanism to explain the results. It was found that upper features were of a higher match in composites constructed in the Control relative to all other conditions. This indicates that during composite construction, when witnesses were asked to select faces based on their overall likeness, more attention had been generally paid to the upper-facial regions, leading to composites with more accurate upper (cf. lower) facial regions—a result that is in line with face-recognition research [1,13]. Also, when constructors were asked to pay attention to upper and lower facial regions separately, this led to no reliable difference in quality between their composites' upper and lower regions. In addition, although the accuracy of the lower-face region improved for composites constructed under these part-region-selection conditions, it was at a cost of the upper-face region. As the upper half was of better likeness in composites from the Control condition, this would seem to have promoted higher naming rates compared to all other conditions. Therefore, selection by whole faces promotes better-quality upper halves which also due to a bias of the upper face during naming leads to a more identifiable face.

Readers may remember that horizontally-misaligned faces in arrays were included in an attempt to assist participant-witnesses in making more accurate face half judgments; they may also conceal possibly wrong configural information of faces presented. For the latter reason, research also indicates that horizontal misalignment improves a person's ability to correctly name a composite [32]. Our findings on horizontal-misalignment type revealed that, overall, composites constructed with no-misaligned faces were against expectations of better likenesses than those created with horizontally-misaligned faces. The composite-face effect (CFE) [2] indicates that people have difficulty in mentally misaligning facial halves for individual judgments [2,15]. Physical misalignment can enable more accurate judgments [2,15-18], and it was therefore expected that horizontally misaligned faces would enable participant-witnesses to make more accurate judgments about upper and lower face regions, leading to better quality composites. However, as face selection based on individual face regions was found to lead to a worse composite likeness compared with Control, misaligned faces led to an even worse likeness. This may indicate that it is easier to process individual regions when in the context of a more complete face, similar to findings indicating that the selection of individual features is more accurate in the context of a complete face [45]. However, as the effect of horizontally-misaligned faces was not detectable in our naming task, it seems to be of little forensic importance. Current work is considering whether horizontally-misaligned

faces might be effective when the task for constructors is to select for the upper half throughout.

Experiment 1 also found that, compared to the current procedure, external features led to lower-named composites. Further, the quality of upper facial features was greater than the quality of lower facial features when composites were constructed without external features (as indicated by rating data). This difference vanished when composites were constructed with external-features present, with reduced quality of composites' upper facial regions. This is in line with past research which suggested that although hair can be a useful retrieval cue for face recognition [35], it is a distraction for face construction, leading to an overall less-effective composite [34]. Further, the findings indicate that the advantage of external features being absent during construction is restricted to improvements in the upper part of the face only. Such a suggestion is reasonable as there tends to be greater surface area of external features surrounding the upper (cf. lower) region, especially for hair, arguably a dominant feature for face perception [46].

Due to greater attention to [13,14] and possibly greater memory for upper facial features, it was theorised that more-consistent selection of upper features would improve identification of composites. As expected, Experiment 2 revealed the importance of selecting upper facial halves throughout during the evolving stage; this procedure trebled [$Exp(B)$ = 3.04] the composite naming rate compared to the currently used (Control) procedure. Shifting constructors' attention specifically to the upper facial half should have allowed them to provide important memory cues to the overall face: in particular, an improvement to the constructed shape and position of the eyes and eyebrows. This, in turn, should have resulted in an upper face that was detectable with better likeness. It is somewhat surprising therefore that rating for the upper face region was not higher compared to Control. However, while likeness ratings are a fairly good proxy to correct naming [47,48], they are not a perfect analogue. In both experiments, ratings only explained about one quarter of the variance (r^2, from correlations between naming and half-face ratings). Due to the importance of upper features for identification [1,30], it is conceivable that small improvements in construction of the eye region can lead to a marked improvement in recognition, but such differences may be difficult to measure using a likeness rating task.

It was expected that naming would be improved further when constructors specifically attended to the upper facial half beyond evolution of the face, using an approach followed by Frowd et al. [35]. This was not found in Experiment 2, as naming was the same as Control. However, likeness ratings indicated that the rated match reduced for the lower half, as one might expect if that was not the focus of attention. Naming did not increase for composites in Upper face II, presumably as the match of the upper half did not actually improve. It is worth mentioning that half of the constructors in this condition reported that it was difficult to continue focusing on the upper half during use of holistic tools; this feedback is entirely sensible as the lower region would no doubt have been difficult to ignore as it would have changed along with the upper half (as manipulations with this tool affect the face

as a whole). The project therefore indicates a much better method for forensic practitioners to construct composites with witnesses and victims of crime. Rather than instructing these observers to select for overall likeness, they should be instructed to select for the upper facial half during the evolving stage (during selection from face arrays); once a face has been evolved—that is, from holistic tools onwards—eyewitnesses should be instructed to focus on the face as a whole. This method should lead to much-more identifiable composites, such as the three-fold increase found here, and so many more offenders should be correctly identified using EvoFIT composites. Other components of the procedure should not change (i.e., presentation of arrays should still involve intact [not horizontally split] faces without external features). While these recommendations have yet to be tested on other holistic composite systems (e.g., ID and EFIT-V), the commonality of the approach [48] would seem to suggest that our findings should extend to these alternative methods of face production.

References

1. Pellicano E, Rhodes G, Peters M (2006) Are pre-schoolers sensitive to configural information in faces? Dev Sci 9: 270-277.

2. Young AW, Hellawell D, Hay DC (1987) Configurational information in face perception. Perception 16(6): 747-759.

3. O'Donnell C, Bruce V (2001) Familiarisation with faces selectively enhances sensitivity to changes made to the eyes. Perception 30(6): 755-764.

4. Sekuler AB, Gaspar CM, Gold JM, Bennett PJ (2004) Inversion leads to quantitative, not qualitative, changes in face processing. Current Biology 14(5): 391-396.

5. Vinette C, Gosselin F, Schyns PG (2004) Spatio temporal dynamics of face recognition in a flash: It's in the eyes. Cognitive Science 28(2): 289-301.

6. Arizpe J, Kravitz DJ, Yovel G, Baker CI (2012) Start position strongly influences fixation patterns during face processing: difficulties with eye movements as a measure of information use. PLoS One 7(2): e31106.

7. Henderson JM, Williams CC, Falk RJ (2005) Eye movements are functional during face learning. Memory and Cognition 33(1): 98-106.

8. Vatikiotis-Bateson E, Eigsti IM, Yano S, Munhall KG (1998) Eye movement of perceivers during audiovisual speech perception. Percept Psychophys 60(6): 926-940.

9. Nelson N, Mondloch C (2014) Is He Afraid or Looking at a Spider? Visual Attention to Facial Expressions Varies With the Task. Journal of Vision 14(10): 1389-1389.

10. Riby DM, Doherty-Sneddon G, Bruce V (2008) Exploring face perception in disorders of development: Evidence from Williams syndrome and autism. Journal of Neuropsychology 2(Pt 1): 47-64.

11. Goldstein AG, Mackenberg EJ (1966) Recognition of human faces from isolated facial features: A developmental study. Psychonomic Science 6(4): 149-150.

12. Turati C, Simion F, Milani I, Umilta C (2002) Newborns' preference for faces: What is crucial? Dev Psychol 38(6): 875-882.

13. Bukach CM, Le Grand R, Kaiser MD, Bub DN, Tanaka JW (2008) Preservation of mouth region processing in two cases of prosopagnosia. J Neuropsychol 2(Pt 1): 227-244.

14. Walker-Smith GJ (1978) The effects of delay and exposure duration in a face recognition task. Perception and Psychophysics 24(1): 63-70.

15. Hole GJ (1994) Configurations! Factors in the perception of unfamiliar faces. Perception 23(1): 65-74.

16. Richler JJ, Cheung OS, Gauthier I (2011) Holistic processing predicts face recognition. Psychological Science 22(4): 464-471.

17. de Heering A, Rossion B, Turati C, Simion F (2008) Holistic face processing can be independent of gaze behaviour: Evidence from the face composite effect. J Neuropsychol 2(Pt 1): 183-195.

18. Rossion B, Boremanse A (2008) Nonlinear relationship between holistic processing of individual faces and picture-plane rotation: evidence from the face composite illusion. J Vis 8(4): 1-13.

19. Abbas ZA, Duchaine B (2008) The role of holistic processing in judgments of facial attractiveness. Perception 37(8): 1187-1196.

20. Calder AJ, Young AW, Keane J, Dean M (2000) Configural information in facial expression perception. J Exp Psychol Hum Percept Perform 26(2): 527-551.

21. Hole GJ, George P (2011) Evidence for holistic processing of facial age. Visual Cognition 19(5): 585-615.

22. Michel C, Rossion B, Han J, Chung CS, Caldara R (2006) Holistic processing is finely tuned for faces of one's own race. Psychol Sci 17: 608-615.

23. Zhao M, Hayward WG (2010) Holistic processing underlies gender judgments of faces. Atten Percept Psychophys 72(3): 591-596.

24. Heering A de, Wallis J, Maurer D (2012) The composite face effect survives asymmetric face distortions. Perception 41(6): 707-716.

25. Leder H, Carbon CC (2006) Face specific configural processing of relational Information. Br J Psychol 97(Pt 1): 19-29.

26. McKone E, Robbins R (2011) Are faces special? In: Calder AJ & Rhodes G (Eds), The Oxford handbook of face perception, UK, pp. 149-176.

27. Ellis HD, Shepherd JW, Davies GM (1975) An investigation of the use of the use of the Photo-Fit technique for recalling faces. British Journal of Psychology 66(1): 29-37.

28. Ellis HD, Shepherd JW, Davies GM (1980) The deterioration of verbal descriptions of faces over different delay intervals. Journal of Police Science and Administration 8(1): 101-106.

29. Fisher GH, Cox RL (1975) Recognizing human faces. Appl Ergon 6(2): 104-109.

30. Laughery KR, Duval C, Wogalter MS (1986) Dynamics of facial recall. In: Young AW (Eds), Aspects of face processing 1986. Springer, Netherlands, pp. 373-387.

31. Matthews ML (1978) Discrimination of Identikit constructions of faces: Evidence for a dual processing strategy. Perception and Psychophysics 23(2): 153-161.

32. McIntyre A, Hancock PJB, Langton SRL, Frowd CD (2016) Holistic Face Processing Can Inhibit Recognition of Forensic Facial Composites. Law Hum Behav 40(2): 128-135.

33. Frowd CD (2015) Facial composites and techniques to improve image recognisability. In: Valentine T & and Davis J (Eds) Forensic facial identification: theory and practice of identification from eyewitnesses, composites and cctv, Wiley-Blackwell, UK.

34. Frowd CD, Pitchford M, Bruce V, Jackson S, Hepton G, et.al. (2010) The psychology of face construction: giving evolution a helping hand. Applied Cognitive Psychology 25(2): 195-203.

35. Frowd CD, Skelton F, Atherton C, Pitchford M, Hepton G, et.al. (2012) Recovering faces from memory: The distracting influence of external facial features. J Exp Psychol Appl 18(2): 224-238.

36. Cutler BL, Penrod SD, Martens TK (1987) Improving the reliability of eyewitness identifications: Putting context into context. Journal of Applied Psychology 72(4): 629-637.

37. Frowd CD, Bruce V, Ross D, McIntyre A, Hancock PJB (2007) An application of caricature: how to improve the recognition of facial composites. Visual Cognition 15(8): 1-31.

38. Goldstein AG, Chance J (1971) Visual recognition memory for complex configurations. Perception and Psychophysics 9(2): 237-241.

39. Bruce V, Henderson Z, Greenwood K, Hancock PJB, Burton AM, et.al. (1999) Verification of face identities from images captured on video. Journal of Experimental Psychology: Applied 5(4): 339-360.

40. Ellis HD, Shepherd JW, Davies GM (1979) Identification of familiar and unfamiliar faces from internal and external features: some implications for theories of face recognition. Perception 8(4): 431-439.

41. Fodarella C, Kuivaniemi-Smith H, Gawrylowicz J, Frowd CD (2015) Forensic procedures for facial-composite construction. Journal of Forensic Practice 17(4): 259-270.

42. Frowd CD, Bruce V, McIntyre AH, Ross D, Fields S, et.al. (2006) Implementing holistic dimensions for a facial composite system. Journal of Multimedia 1(3): 42-51.

43. Frowd CD, Carson D, Ness H, Richardson J, Morrison L, et.al. (2005) A forensically valid comparison of facial composite systems. Psychology, Crime and Law 11(1): 33-52.

44. Shapiro PN, Penrod SD (1986) Meta-analysis of facial identification rates. Psychological Bulletin 100(2): 139-156.

45. Davies GM, Christie DFM (1982) An examination of some factors limiting composite production accuracy. Journal of Applied Psychology 67(1): 103-109.

46. Ellis HD (1986) Face recall: A psychological perspective. Human Learning 5: 1-8.

47. Frowd CD, Bruce V, Smith A, Hancock, PJB (2008) Improving the quality of facial composites using a holistic cognitive interview. J Exp Psychol Appl 14(3): 276-287.

48. Valentine T, Davis JP, Thorner K, Solomon C, Gibson S (2010) Evolving and combining facial composites: Between-witness and within-witness morphs compared. J Exp Psychol Appl 16(1): 72-86.