Running head: Holistic face processing can impair composite identification

Holistic Processing Can Impair Identification of Forensic Facial Composites

Alex H. McIntyre, Peter J. B. Hancock, Charlie D. Frowd and Stephen R. H. Langton

University of Stirling

Author version submitted to Law and Human   Behavior

(This version - figures have been copied inline)

Author Note

Alex H. McIntyre, Psychology, School of Natural Science, University of Stirling; Peter J. B. Hancock, Psychology, School of Natural Science, University of Stirling; Charlie D. Frowd, School of Psychology, University of Central Lancashire; Stephen R. H. Langton, Psychology, School of Natural Science, University of Stirling.

Alex H. McIntyre is now at Psychology, School of Life, Sport and Social Sciences, Edinburgh Napier University.

Experiment 1 was completed in partial fulfilment of the doctoral thesis of the first author (2012).

Correspondence concerning this article should be addressed Alex H. McIntyre, Psychology, School of Life, Sport and Social Sciences, Edinburgh Napier University, Sighthill Court, Edinburgh. EH11 4BN.

E-mail: A.McIntyre@napier.ac.uk

Abstract

Facial composite systems help eyewitnesses to show the appearance of criminals. However, likenesses created by unfamiliar witnesses will not be completely accurate, and people familiar with the target can find them difficult to identify. Faces are processed holistically; we explore whether this impairs identification of inaccurate composite images and whether recognition can be improved. In Experiment 1 (n = 64) an imaging technique was used to make composites of celebrity faces more accurate and identification was contrasted with the original composite images. Corrected composites were better recognized, confirming that errors in production of the likenesses impair identification. The influence of holistic face processing was explored by misaligning the top and bottom parts of the composites (cf. Young, Hellawell, & Hay, 1984). Misalignment impaired recognition of corrected composites but identification of the original, inaccurate composites significantly improved. This effect was replicated with facial composites of non-celebrities in Experiment 2 (n = 57). We conclude that, like real faces, facial composites are processed holistically: recognition is impaired because unlike real faces, composites contain inaccuracies and holistic face processing makes it difficult to perceive identifiable features. This effect was consistent across composites of celebrities and composites of people who are personally familiar. Our findings suggest that identification of forensic facial composites can be enhanced by presenting composites in a misaligned format. (219 words)

Keywords: facial composite, face recognition, configural, featural, holistic, eyewitness.

Holistic Processing Can Impair Identification of Forensic Facial Composites

Following a crime, forensic facial composites make it possible to communicate the appearance of a perpetrator in a way that a verbal description cannot; they are particularly important when there is little, or no physical evidence. There are a number of forensic composite software applications but although they can produce good likenesses, this doesn't always translate to good rates of identification *(see Davies & Valentine, 2007 for a review of facial composite systems)*. Brace, Pike, Kemp, Turner, and Bennett (2006) found that witness ratings of composite quality did not predict successful identification. Likewise, Davies, van der Willik, and Morrison (2000), and Frowd, Carson, Ness, McQuiston-Surrett, Richardson, et al. (2005) reported that composites that were matched to target images around half of the time, obtained identification rates of only 2-3%. This suggests that even when facial composites appear to be good likenesses, identifiable information is not recognised during face perception. There are three core ideas in this paper: first, when composites that are produced by people unfamiliar with the target appear to be a good likenesses, they may contain identifiable information but the complete composite image will be incorrect: second, that holistic face perception of the inaccurate composite image (Carlson, Gronlund, Weatherford & Carlson, 2012) will inhibit recognition of any identifiable features by people familiar with the targets (Wilford & Wells, 2010): third, if facial composites are processed holistically (Carlson et al., 2012) and if the composite arrangements are typically incorrect, inhibiting holistic face perception should enhance recognition of any identifiable facial composite information (Wilford & Wells, 2010).

The way that faces are perceived depends on the familiarity of the face: unfamiliar face perception tends to be poor and is disproportionally influenced by attention to the external features, such as the face shape and the hairstyle (e.g. Bonner, Burton, & Bruce, 2003; Bruce et al., 1999; Ellis, Shepherd, & Davies, 1979). An eyewitness is generally unfamiliar with the perpetrator and had limited opportunity to encode his or her face. To create a composite the witness must retrieve the face memory and communicate it to the police (Brace, Pike, Allen, & Kemp, 2006): attention to the external features will influence how well the facial information was encoded and how effectively it can be reproduced. It is unavoidable that some parts of the composite will be poor and it is feasible that the internal features (i.e. the eyes, the nose and the mouth) may be reproduced less effectively. For the composite to be forensically useful, someone who is familiar with the perpetrator must identify it. Familiar face perception is much more effective but it is extremely sensitive to the internal features of the face and to their configuration (e.g. Bonner, Burton, & Bruce, 2003; Bruce et al., 1999; Ellis, Shepherd, & Davies, 1979). This means that there may be disparity between the facial information that will be generated by the unfamiliar witness and the quality of information that will be needed to achieve identification.

Faces provide information about the separate features, such as the eyes or nose, as well as information about the configural arrangement of the features. Typically, all of this is processed holistically as a single face stimulus (Farah, Tanaka, & Drain, 1995; Meinhardt-Injac, Persike, & Meinhardt, 2013). While this is efficient, it has consequences for how facial information is perceived: for example, individual features are recognised better within the context of the whole face image (Tanaka & Farah, 1993), but alterations to a facial configuration will impair recognition of the individual features (Tanaka & Sengco, 1997). Sensitivity to familiar faces (Haig,

1984; Hosie, Ellis, & Haig, 1988) means that any changes to the configuration will alter holistic perception of the whole image and the face will appear different (Tanaka & Sengco, 1997). Wilford and Wells (2010) showed that holistic face perception enhances detection of any face alteration but makes it difficult to identify what feature has changed. Facial composites can be processed holistically (Carlson et al., 2012) so if a composite is flawed, due to poor feature selection or incorrect positioning of features, the witness would be unable to correct it and identification of good features would be impaired (Wilford & Wells, 2010). There is evidence that even perfectly represented features would be harder to identify if an inaccurate composite is processed holistically (Young, Hellawell, & Hay, 1987).

Young et al. (1987) asked participants to name the top and bottom photograph halves of different famous faces, then they created composite images by aligning face halves of different people (i.e. the top half of one person's face with the bottom half of another person's face). The participants had difficulty identifying the aligned face halves even though the individual identities had been primed immediately prior to testing. Performance improved when the images were inverted and when the photograph halves were misaligned. Face inversion disrupts holistic face perception (e.g. Bartlett & Searcy, 1993; Rossion & Boremanse, 2008; Yin, 1969) so the authors concluded that holistic processing of the composite photographs caused perception of novel faces that impaired recognition of the familiar face parts. This effect is known as the composite face illusion: the paradigm has been used to evaluate holistic face processing in more than 60 psychological and neurophysiological studies (Rossion, 2013) and misalignment is widely accepted to be the most effective means of disrupting holistic face processing (e.g. de Heering et al., 2007; Maurer et al., 2002; Palermo et al., 2011). We propose that a similar effect is involved in perception of

forensic facial composites: the likeness may contain identifiable information but it is presented within the bounds of an inaccurate full-face image. Holistic processing of the composite makes perception of the identifiable component features too difficult and recognition fails.

To explore the effects of composite inaccuracy we compared identification for original facial composites with identification of composites that were manipulated to be more accurate. To explore the influence of holistic face processing and to determine whether the effects of composite inaccuracy could be reduced, the composite images were evaluated in both a full-face presentation and with a misaligned presentation. We speculated that misaligning the top and bottom parts of the original composite images could disrupt holistic face processing and allow good features to be recognised. We expected that performance would decline when the corrected facial composited were misaligned, thus demonstrating that holistic face perception is generally beneficial, but can inhibit identification of inaccurate facial composite images.

Our initial premise stemmed from observations that facial composites that *'appear'* to be a good likeness often fail to achieve good rates of identification (e.g. Brace et al, 2006; Davies et al., 2000; Frowd et al., 2005); as such, to evaluate the effects of inaccuracy and holistic face perception in Experiment 1, we selected a series of celebrity facial composites on the basis of good visual similarity to the target. For Experiment 2 we generated a series of non-celebrity facial composites to determine whether comparable effects would be observed with facial composites of personally familiar people that were produced in a forensically valid protocol.

EXPERIMENT 1

Participants

Sixty-four participants were recruited from the University of Stirling by opportunity sampling. There were equal numbers of males and females and all had normal or corrected-to-normal vision. Their ages ranged from 16 to 72 years ($M = 30.2$, *S.D.* $= 11.2$). Participation was voluntary except for one person who was awarded a course credit.

Materials

Facial composites. The rationale for this study was that composites that appear to be a good likeness often fail to achieve good rates of identification. Twenty-eight composites of male celebrities were selected from the University of Stirling archives on the basis of good visual similarity to the target. Celebrities are not personally familiar and caution should be adopted in generalizing results from celebrity composites to non-celebrity images. However, to obtain identification data for a sufficient number of personally familiar composites, both the composite targets and the participants would be sampled within an occupational setting. This means that the participants could employ a process of elimination rather than face recognition; for example, if the composite image has curly hair it must be 'Joe'. To assess the effects of inaccuracy on facial composite recognition it was important that natural face processing, rather than a process of deduction would be employed. We approached this by using celebrity composites to explore the effects of inaccuracy and misalignment in Experiment 1, and by replicating our evaluation of misalignment with non-celebrity composites in Experiment 2.

A diverse sample of celebrity targets was selected to be identifiable to a wide range of participants. The set comprised film stars, television personalities,

politicians, musicians and sportsmen: at a minimum, a subset would be identifiable to most people without priming identities from any particular domain. Each composite was produced in accordance with APA ethical standards using E-FIT or PRO-fit software for one of three previous studies. E-FIT and PRO-fit are highly similar and create likenesses by combining separate facial features within a face image, both are employed by UK police services and they show comparable performance in formal evaluations (*this study is not concerned with a comparison of composite systems, for a review of composite procedures see Frowd, Carson, Ness, Richardson, Morrison, et al. 2005b, and for a review of composite software see Davies & Valentine, 2007).* In each study, construction was preceded by a Cognitive Interview then witness participants created a celebrity composite from memory with the help of an experienced composite operator. The delay between viewing the target and creating the composite ranged from a few minutes (Frowd, Hancock, & Carson, 2004), and 3-4 hours (Frowd et al, 2005b) to 2 days (Frowd et al, 2005a). Similarity to target was determined by identification and matching data (mean correct matching rate 65.7%; mean identification rate 17.7%).

Corrected facial composites. To generate composites that more accurately represented the target faces, Psychomorph software was used to manipulate the original composites to show the shape information of the faces in target photographs (Tiddeman, Burt, & Perrett, 2001). Composite image transformations were achieved using templates created by tagging the facial landmarks of each facial composite, and the facial landmarks of the corresponding photographic face image. The facial composite templates were then aligned to the photographic templates: essentially, making the corrected composite images depict the surface texture of the original composites with the shape and proportional characteristics of the photographic image.

An example of the original and corrected facial composite stimuli is shown in   Figure 1. From these images it is apparent that while the Psychomorph procedure enhanced the accuracy of the facial representation, in some cases it also altered the angle of the head or imbued the composite with facial expression. Full-face expressionless photographs would have produced cleaner transformations but if we had aligned the composites to photographs the witness participants did not see, the image properties may have made features or proportional sizes look very different (Adini, Moses, & Ullman, 1994). The transformation would then have shifted the composites in a way that the witness-participants could not have intended and we would not   have evaluated the effect of inaccuracies in the attempts to achieve particular likenesses. Witness participants could be asked to create composites from memory of full face expressionless photographs; but improving the quality of the corrected composite images has no ecological validity and no forensic application, for this evaluation we were simply concerned with producing corrected versions of composites that are not named as well as visual similarity and matching data suggests they might be.

Misaligned composite images. The misaligned composite images were created from the original and corrected facial composites using Adobe Photoshop Elements 5.0. The images were split horizontally below the eyes and each generated two misaligned images: one with the lower portion moved left, aligning the nose with the left ear, and one similarly aligned to the right. The misaligned images portrayed all of the composite information, but in a format that could preclude holistic face processing. Examples of misaligned stimuli are shown in Figure 1.
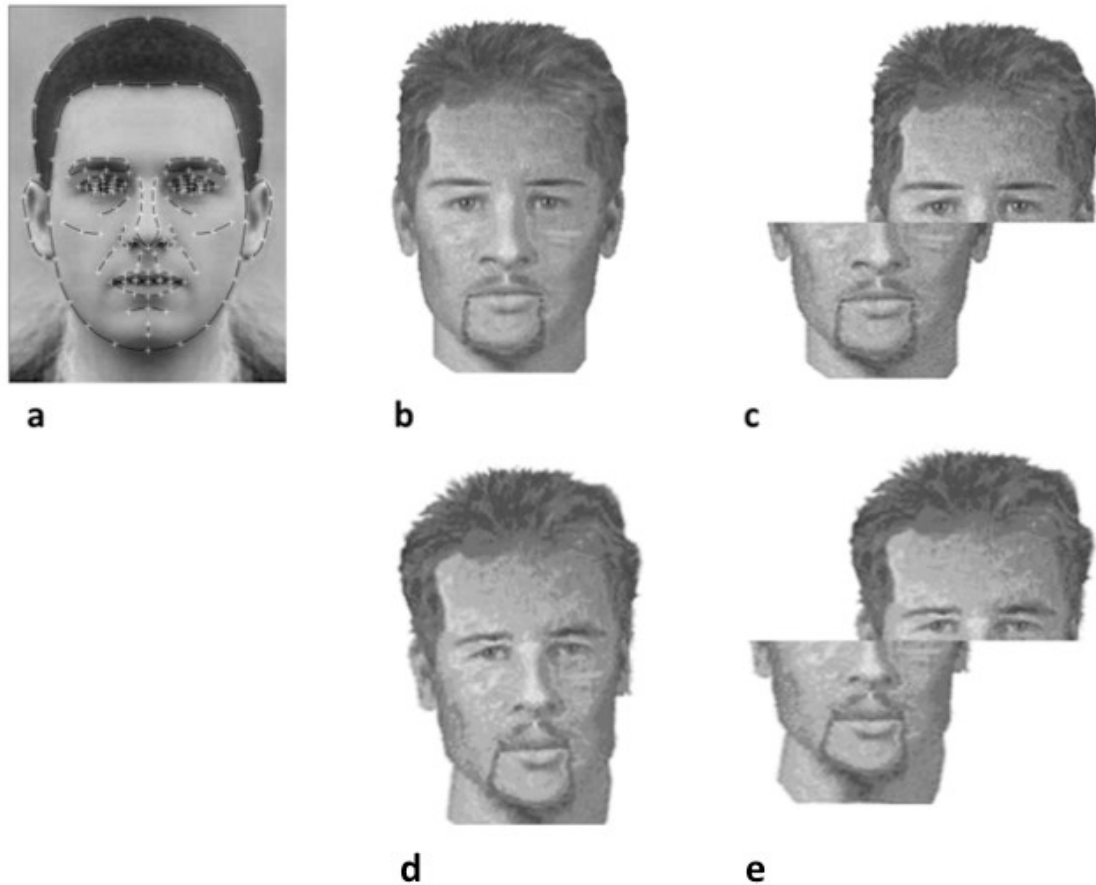
Figure 1. Examples of the facial composite stimuli with a Psychomorph template. Top row: (a) The Psychomorph template; (b) the original facial composite of Brad Pitt; (c) the misaligned original facial composite. Bottom row: (d) the corrected facial composite of Brad Pitt; (e) the misaligned corrected facial composite. The corrected composites was generated by tagging corresponding data points on the target photograph of Brad Pitt and on the original facial composite; the composite was then warped to the parameters of the target photograph. Due to copyright restrictions the target photograph cannot be shown. Misaligned facial composites were generated with Adobe Photoshop Elements 5.0. The direction of misalignment was counterbalanced across participants.

The facial composites and photographic images of the targets were cropped closely around the head and printed individually in the centre of white A4 paper in landscape orientation. Facial composites measured 8 cm in height and the target images measured 9 cm in height. The original facial composites were randomly allocated to one of two booklets and supplemented by misaligned composites, such that each booklet contained 14 original composites and 14 misaligned original composites, with each target represented once. This was replicated for booklets containing the corrected and misaligned corrected facial composites. The direction of misalignment was equally sampled and was counterbalanced across participants. A final booklet was compiled with the target photographs to control for familiarity.

Design and Procedure

A mixed factor 2 x 2 design was employed: the between participant factor was composite type (original facial composites; corrected facial composites), the within participant factor was presentation format (complete images; misaligned images). The participants were tested individually, and were informed that the images they would be shown were facial composites like those they would see on the "Crimewatch UK" television programme. They were advised that the composites were intended to portray famous males, and were then presented with each composite in turn and asked if they could identify the person. No additional information was provided regarding the misaligned images; if a participant commented on them, the request to identify the person was repeated. In cases where an image could not be named but the participant could provide unequivocal biographical information, responses were accepted as correct (e.g. "he's the actor who's married to Angelina Jolie" was an acceptable response to the composites of Brad Pitt, but "he's an American actor" was not). Presentation order was randomized for each participant

and each composite image was viewed once for an unlimited duration until a response was provided, or the image was rejected as unfamiliar. When all of the composites had been viewed, participants were asked to name the target photographs to take into account identities that were not known (see below for conditional naming rate).

Results

The celebrity targets were reasonably well known with a mean identification rate of 82.3%. Facial composite naming rates were calculated as a function of the number of targets that were known to each participant; for example, if a participant correctly named seven composites and 21 of the 28 targets, the identification rate was calculated as 33.3% (7 / 21), rather than 25% (7 / 28) for the full set. Conditional naming rates were calculated for the full-face images and for the misaligned images, the mean identification rates and standard errors are shown in Figure 2.

Conditional naming responses for participants were analysed with Mixed Factorial Analysis of Variance. As expected, there was a significant main effect of composite type, $F(1, 62) = 9.1$, $p = .004$, $\eta_p^2 = .13$, 90% CI [.03, .26]. The corrected composites were identified better ($M = 33.0\%$, $SE = 2.3\%$) than the original composites ($M = 18.2\%$, $SE = 2.3\%$). There was no significant main effect of presentation format ($F < 1$) but in line with our predictions, there was a significant interaction between composite type and presentation format, $F(1, 62) = 10.6$, $p = .002$, $\eta_p^2 = .15$, 90% CI [.04, .28]. Paired sample t-tests revealed medium sized but opposing effects for the original composites and the corrected composites: misaligning the original, inaccurate facial composites produced a significant improvement ($M = 26.9\%$, $SE = 2.5\%$), $t(31) = 2.5$, $p = .018$, $d = 0.6$, 95% CI [0.13, 1.13], while misaligning the corrected composites significantly impaired identification ($M = 26.8\%$, $SE = 2.5\%$), $t(31) = 2.1$, $p = .045$, $d = -0.5$, 95% CI [-0.95, 0.05].
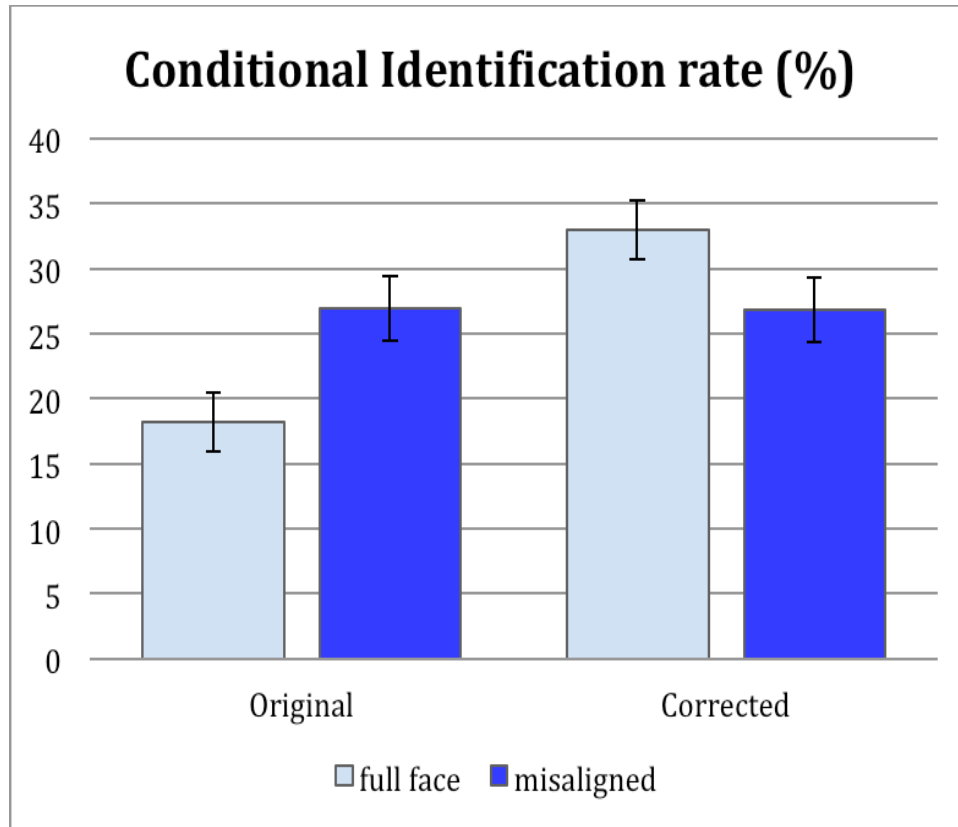
Figure 2. Mean composite identification rates and standard errors by group, and by presentation format. Performance was best for the composites that were corrected to show more accurate information, and was poorest for the original images that were generated by the unfamiliar witness-participants. Identification of the misaligned images was poorer for the corrected images, but an advantage for identification of the original facial composites was observed.

Repeated Measures ANOVA with composite items as cases found no significant main effect of composite type ($F < 1$) and no significant main effect of presentation format ($F < 1$). The interaction between composite type and presentation format was significant, $F(1, 54) = 7.93$, $p = .007$, $\eta_p^2 = .13$, 90% CI [.02, .27]. Planned comparisons confirmed that misalignment of the original composites enhanced identification, $t(27) = 2.52$, $p = .02$, $d = .21$, 95% CI [-0.31, 0.74], while at

the items level performance for the corrected composites did not decline, $t(27) = 1.48$, $p = .2$.

Discussion

In this evaluation we selected composites of celebrities that had previously been judged to be good likenesses. In spite of the apparent likeness, we confirmed that inaccuracies in the composites inhibit identification and our prediction that misaligning the images would enable recognition was supported. These results are encouraging but fall short of ideal on several important issues. By selecting composites that were judged to be good likenesses, we were able to address the issue of why composites that appear to be good quality are not better identified. However, we also ensured that identifiable information should be present in the misaligned images, and thus stacked the odds in favour of our prediction being supported. It might be that with poorer composites, or with images of more variable quality there would be no benefit of misalignment. It was important that we observed natural face recognition and participants did not use a process of elimination to guess the composite identities: to avoid any cueing we used celebrity targets from a diverse range of backgrounds. We are confident that we achieved our aim of evaluating natural face recognition for inaccurate composite images but there is a limit to what we can infer from these results. Celebrities are not personally familiar, and although there is no evidence that this influences results of facial composite studies, it is possible that participants recognise a memory for an iconic celebrity image, rather than the actual person. To determine whether misalignment would be effective for composites of people we meet in real life, Experiment 2 employed a forensically relevant protocol to create and evaluate facial composites of university lecturers,

Experiment 2

Our second evaluation was designed to explore whether the misalignment effect observed in Experiment 1 would also be present in composites of variable quality that were intended to portray personally familiar people. The protocol adheres to important criteria that would be inherent in real life: the composites would be created by unfamiliar witness participants following a 2 day delay and identification data would be collected from a sample of participants who would be personally familiar with the targets. The composites produced by the unfamiliar witnesses were unlikely to be completely accurate representations but as face images, they would be processed holistically. We predicted that misaligning the composite images would enable perception of any accurate information and enhance identification.

Method

Composite Construction

Witness participants. Twenty-four witness participants (6 male; 18 female) were recruited from staff and students at the University of Stirling. Their ages ranged from 18 and 46 years (M = *30.2,* S.D. = *11.2*). They all had normal or corrected to normal vision and they were paid £5.

Target Stimuli. Six lecturers (3 female) from the Psychology Department at Edinburgh Napier University agreed to be targets. They were matched to 6 members of staff from the Psychology Department at the University of Stirling on the basis of gender, age, weight and hairstyle. The target sample was controlled in as much that none of the volunteers had unique or distinctive characteristics that would bias identification processes and each composite could be mistaken for another unrelated individual within the psychology department. This control was important to ensure that the facial identification we were seeking would not be confounded by elimination

strategies. Each target sat for a series of photographs in various poses and displaying

a number of expressions. Photographs were captured at a distance of approximately 3

metres using a Sony Cyber shot digital camera with a resolution of 5.0 mega pixels

that was mounted on a tripod. Four images were selected of each target to be used as

stimuli. One was a full-face pose, one was ¾ profile, one was full profile and the

fourth portrayed the target looking up and away from the camera. Two of the images

displayed neutral expressions, one a smiling expression and one an angry expression.

All images also provided contextual background cues. This method was employed to

produce a range of images that might provide a richer memory representation for

witness participants than would be possible from one full-face image. Each image

was sized to a width of 10cm and all four were displayed in a word document on a

Dell Inspiron 6400 laptop. An example of the target stimuli is shown in Figure 3.

  Construction procedure. Witness participants were tested individually; each

was randomly allocated a target, if the target was familiar to them, an alternative was

presented until an unfamiliar person was found. They were then allowed to view the

target stimuli for one minute. They were thanked for their time and an appointment

was agreed for them to return to make a facial composite two days later. The

composite construction session was designed to closely approximate current police

practice and construction was preceded by a cognitive interview to enhance recall

(Geiselman et al., 1987). For each target two witness participants were asked to

create a composite likeness using ProFIT facial composite software. When each

participant was satisfied with the composite likeness they were debriefed, thanked and

paid for their time.

Figure 3. Example of target stimuli used in experiment 2. The series of photographs was designed to display a number of poses, expressions and retrieval cues in order to provide witness participants with a richer memorial representation.

Composite naming

Participants. Thirty participants (13 male; 17 female) were recruited from staff and final year students at the University of Stirling and 27 participants (7 male; 20 female) were recruited from final year students at Edinburgh Napier University. Ages ranged from 20 – 50 years (M = *22.91*, S.D. = *6.16*), they all had normal or corrected to normal vision and participation was voluntary.

Materials. Misaligned composite images were created as per Experiment 1 and four sets of presentation materials were compiled. Each set comprised 6 aligned and 6 misaligned facial composites from each institution. In each set a target was represented by both of the composites produced by the separate witness participants - one was aligned and one was misaligned. Alignment of each image was

counterbalanced across participants so that each composite likeness was presented in both formats. The composite images were shown using Microsoft PowerPoint and the order was randomised for each presentation.

Procedure. Participants were tested in groups (M = *8.1*). They were informed that they would be shown a series of facial composites, that some of them were intended to portray members of staff from the Psychology Department, and that they should attempt to identify each one. To prevent participants from employing elimination strategies, they were explicitly told that they would *not* be able to identify all of the composites as some of them portrayed lecturers from another university. No additional information was provided regarding the misaligned images. The importance of not sharing information during an experimental study was stressed to them and each participant was instructed to write down their demographic details and to record their own responses. If they thought they could identify a composite but could not provide a name, they were asked to write down unequivocal identifying information. The composites were presented sequentially on a wall screen and the participants attempted to name each composite in their own time.

Results

The targets from each institution were well known to the participants with a mean identification rate of 92.4%. Conditional naming rates were calculated for the full-face images and for the misaligned images. The mean identification rates and standard errors for each university sample are shown in Figure 4.

Conditional naming responses for participants were analysed with Mixed Factorial Analysis of Variance. The within participant factor was presentation format (complete images; misaligned images), the between participant factor was university sample (Edinburgh Napier; Stirling). There was a significant main effect of

misalignment, $F(1, 55) = 6.67$, $p = .012$, $\eta_p^2 = .11$, 90% CI [.01, .24], the misaligned

composite images ($M = 25.61\%$, $SE = 2.84\%$) were identified more successfully than

the original composites ($M = 17.22\%$, $SE = 2.89\%$). There was no significant main

effect of university group ($F < 1$) and no significant interaction between university

group and misalignment ($F < 1$). In both university samples facial composites of

personally familiar people were identified significantly better when the images were
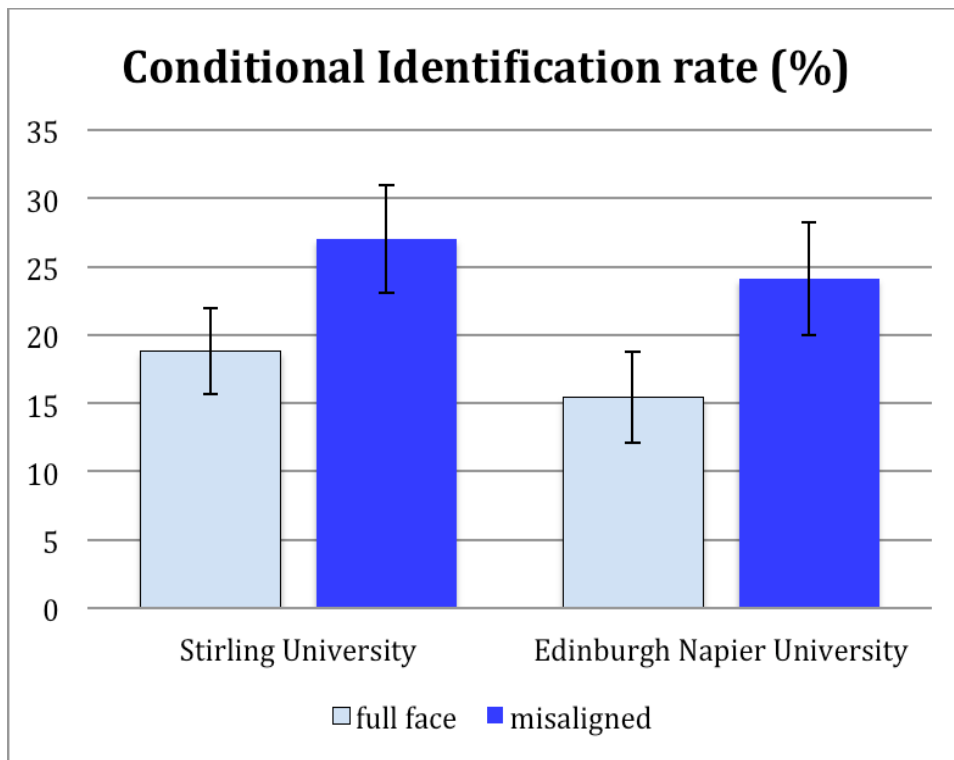
misaligned.

Figure 4. Mean composite identification rates by university group, and by presentation format. Performance was best for the non-celebrity composites that were misaligned. Performance between the university groups was comparable.

Repeated Measures ANOVA with composite items as cases failed to reach significance for misalignment, $F(1, 22) = 3.72$, $p = .067$, $\eta_p^2 = .15$, 90% CI [.01, .35]; however, the effect size indicates that presentation format accounted for 15% of the variance in facial composite recognition. There was no significant main effect of university sample ($F < 1$) and no significant interaction between university sample and misalignment ($F < 1$).

Unlike the facial composites that were evaluated in Experiment 1, the composites in this experiment portrayed people who were personally familiar to the participants. The construction of these composites was also carefully controlled to assimilate composite image production in real life: the witness participants were unfamiliar with the targets and had only a brief time to encode the facial information. They were then asked to return 2 days later to take part in a cognitive interview before being assisted by an experienced operator to construct a composite likeness. The resulting image set is not controlled for composite quality in any way and hence the replication of the misalignment effect with non-celebrity composites, that were not selected to be good likenesses is particularly striking. The results of Experiment 2 provide evidence that misaligning facial composite images can enhance identification of forensic facial composites: the composite misalignment effect is not confined to recognition of iconic celebrity images that were selected on the basis of similarity to a target image.

General discussion

Facial composites serve an important function in police investigations because language is grossly inadequate to convey another person's appearance. The composite systems allow witnesses to show others what a perpetrator looked like and they are capable of producing very good likenesses. However, the witness is

unfamiliar with the perpetrator and the task of creating a facial likeness is extremely difficult. The witness must encode the face, recall the face to provide a verbal description, communicate this information to another person, and then attempt to recognise facial attributes that match their visual memory (Brace et al., 2006). Given these parameters, it is remarkable that witnesses can and do create identifiable composite likenesses: it is also understandable that the quality of composite likenesses is highly variable.

This study was motivated by the observation that composites that can be matched to their target image and thus appear to be good likenesses, often fail to be recognised (Frowd et al., 2004; Frowd et al, 2005b; Frowd et al, 2005a). This suggests that when people familiar with the targets process the composite faces identifiable information is not recognised. We have proposed that composite faces will be processed holistically (Carlson et al, 2012) but as the images are likely to be inaccurate, holistic processing inhibits recognition of any identifiable information (Tanaka & Sengco, 1997). In two experiments we have shown that misaligning facial composites, a technique widely accepted to inhibit holistic face processing, can significantly elevate identification.

Conversely, misalignment impaired identification of the corrected composites and performance was comparable with identification of the misaligned original images. This could be taken as evidence that the misaligned original and the misaligned corrected images were processed in a similar way: in the absence of holistic face processing, identification was driven by perception of featural information. However, correction of the composites should have enhanced the appearance of individual features as well as the configuration, so we might have expected these to be more identifiable than the features in the original composites.

The absence of an advantage for the corrected images could reflect a reliance on holistic representations of familiar or iconic famous images, such that recognition from component information may not exceed a given threshold (e.g. Richler, Tanaka, Brown, & Gauthier, 2008). This interpretation is supported by findings that faces can be identified by their constituent features, and if features are obscured by blurring they can also be identified by the configuration: but, recognition is suboptimal in comparison with identification when both featural and configural facial information is available (Collishaw & Hole, 2000; Schwaninger, Lobmaier & Collishaw, 2002). Misalignment of a corrected image may reduce identification performance to a level that is achievable based solely on featural information; misalignment of the original composites may facilitate identification to a level that is achievable when featural information is not obscured by holistic perception of a flawed configuration.

Finally, positive facial expression facilitates face recognition (e.g. Garcia-Marques, Mackie, Claypool, & Garcia-Marques, 2004; Kaufmann & Schweinberger, 2004), and we have found that facial composites are more likely to be identified when they are manipulated to show positive affect (McIntyre, Hancock, Langton, & Frowd, in prep). If positive facial expression enhanced identification of the corrected composites, this effect was also lost with misalignment. This does not detract from the superior identification that was obtained for both celebrity and non-celebrity facial composites when the original images were misaligned.

Facial composites are identified better when they are misaligned, but a similar technique might also be useful for facial composite construction. The composite systems require witnesses to select features, or judge likeness within the context of whole faces (e.g. Davies & Christie, 1982; Tanaka & Sengco, 1997). It would be useful to determine whether holistic face processing is beneficial, or whether features

should be selected before generating the composite likeness. Police artists can produce better likenesses than featural composite systems (Frowd et al., 2005b); in the sketch artist protocol a witness identifies features from a manual before the artist combines them in a sketch (Gibson, 2008). It is possible that feature selection without holistic processing would be beneficial, particularly if a witness made a conscious effort to encode this information. An evaluation of this kind would be particularly timely as a new generation of composite systems are being developed specifically to generate likenesses from whole face images (e.g. Frowd, Bruce, Pitchford, Gannon, Robinson et al., 2011). Early evaluations indicate that misalignment is also beneficial to identification of these composite images (Frowd et al., unpublished).

In both of our evaluations misalignment significantly improved composite identification, but from Experiment 1 it is clear that when images are better quality holistic processing is beneficial. As we cannot know how accurate a facial composite image is, it is not possible to predetermine whether a conventional or misaligned presentation will be most effective in individual cases. While it would be possible to display both images, people might be inclined to ignore one or the other. A more promising option might be to include both in one presentation. We have previously had some success in using animation with caricature to search for the optimal level of distinctiveness for facial composite recognition (Frowd, Bruce, Ross, McIntyre, & Hancock, 2007; Frowd, Skelton, Atherton, Pitchford, Bruce, et al., 2012). Future evaluation of the misaligned facial composite effect will explore the efficacy of an animated sequence that misaligns the face image and then realigns it across a sequence of frames. Facial composite are increasingly being released to the public online and via television and in these contexts this might provide the optimal forensic application.

References

Adini, Y., Moses, Y., & Ullman, S. (1994). Face recognition: The problem of compensating for changes in illumination direction. In J. Eklundh (Ed.), *Lecture Notes in Computer Science: Vol. 800. Computer Vision—ECCV'94*, (pp. 286–296). Berlin, Germany: Springer-Verlag. doi: 10.1007/3-540-57956-7_33.

Bartlett, J. C., & Searcy, J. (1993). Inversion and configuration of faces. *Cognitive Psychology*, *25*(3), 281-316. doi.org/10.1006/cogp.1993.1007

Bonner, L., Burton, A. M., & Bruce, V. (2003). Getting to know you: How we learn new faces. Visual Cognition, 10(5), 527-536. doi:10.1080/13506280244000168

Brace, N. A., Pike, G. E., Allen, P., & Kemp, R. I. (2006a). Identifying composites of famous faces: Investigating memory, language and system issues. *Psychology, Crime and Law*, *12*(4), 351–366. doi:10.1080/10683160500151159

Brace, N., Pike, G., Kemp, R., Turner, J., & Bennett, P. (2006b). Does the presentation of multiple facial composites improve suspect identification? *Applied Cognitive Psychology*, *20*(2), 213–226. doi:10.1002/acp.1181

Bruce, V., & Young, A. (1986). Understanding face recognition. British Journal of Psychology, 77(3), 305-327.

Calder, A. J., Young, A.W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology, 26,* 527–551. doi: 10.1037/0096-1523.26.2.527

Carlson, C. A., Gronlund,S. D., Weatherford, D. R. & Carlson, M. A. (2012). Processing differences between feature-based facial composites and photos of real faces. *Applied Cognitive Psychology, 26*, 525–540. doi:

10.1002/acp.2824

Collishaw, S. M., & Hole, G. J. (2000). Featural and configurational processes in the recognition of faces of different familiarity. *Perception*, *29*(8), 893-909.

Davies, G. & Christie, D. (1982). Face recall: an examination of some factors limiting composite production accuracy. Journal of Applied Psychology, 67 (1), 103-109. doi: 10.1037/0021-9010.67.1.103

Davies, G. & Valentine, T. (2007). Facial composites: forensic utility and psychological research. In R .C .L. Lindsay, D. F. Ross, J. D. Reid & M. P. Toglia (Eds), *Handbook of Eyewitness Psychology*, Vol. 2 (pp 59-86). Mahwah, NJ: Erlbaum.

Davies, G., van der Willik, P., & Morrison, L. J. (2000). Facial composite production: A comparison of mechanical and computer-driven systems. *Journal of Applied Psychology*, *85*(1), 119-124. doi:10.1037/0021-9010.85.1.119

de Heering, A., Houthuys, S., & Rossion, B. (2007). Holistic face processing is mature at 4 years of age: Evidence from the composite face effect. *Journal of experimental child psychology*, *96*(1), 57-70.

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. Perception, 8(4), 431-439.

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995a). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 628-634. doi:10.1037/0096-1523.21.3.628

Frowd, C. D., Bruce, V., Pitchford, M., Jackson, S., Hepton, G., Greenall, M., McIntyre, A. H . & Hancock, P. J. B. (2011) The psychology of face construction: Giving evolution a helping hand. *Applied Cognitive Psychology,*

25(2), pp. 195-203. doi.org/10.1002/acp.1662

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

Frowd, C., Carson, D., Ness, H., Quiston-Surrett, D., Richardson, J., Baldwin, H., & Hancock, P. (2005a). Contemporary composite techniques: The impact of a forensically-relevant target delay. *Legal and Criminological Psychology*, *10*, 63-81. doi:10.1348/135532504X15358

Frowd, C., Carson, D., Ness, H., Richardson, J., Morrison, L., McLanaghan, S., & Hancock, P. J. B. (2005b). A forensically valid comparison of facial composite systems. *Psychology Crime & Law*, *11*(1), 33-52. doi:10.1080/10683160310001634313

Frowd, C. D., Hancock, P. J. B., & Carson, D. (2004). EvoFIT: A holistic, evolutionary facial imaging technique for creating composites. ACM Transactions on Applied Perception, 1(1), 19-39. doi:10.1145/1008722.1008725

Frowd, C.D., Skelton, F.C., Atherton, C., Pitchford, M., Bruce, V., Atkins, R., Gannon, C., Ross, D., Young, F., Nelson, L., Hepton, G., McIntyre, A.H., & Hancock, P.J.B. (2012). Understanding the multi-frame caricature advantage for recognising facial composites. Visual Cognition, 20, 1215-1241.

Garcia-Marques, T., Mackie, D. M., Claypool, H. M., & Garcia-Marques, L. (2004). Positivity can cue familiarity. *Personality and Social Psychology Bulletin, 30*(5), 585-593. doi: 10.1177/0146167203262856

Gibson, L. (2008). *Forensic art essentials a manual for law enforcement artists* (pp13-20). Burlington, MA: Academic Press.

Haig, N. D. (1984). The effect of feature displacement on face recognition. *Perception*, *13*(5), 505–512. doi:10.1068/p130505

Hosie, J. A., Ellis, H. D., & Haig, N. D. (1988). The effect of feature displacement on the perception of well-known faces. *Perception*, *17*(4), 461-474. doi:10.1068/p170461

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in cognitive sciences*, *6*(6), 255-260.

Meinhardt-Injac, B, Persike, M. & Meinhardt, G. (2013). Holistic face processing is induced by shape and texture. *Perception, 42*(2), 716-732. doi:10.1068/p7462

Palermo, R., Willis, M. L., Rivolta, D., McKone, E., Wilson, C. E., & Calder, A. J. (2011). Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. *Neuropsychologia*, *49*(5), 1226-1235.

Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition*, *21*(2), 139-253.

Rossion, B., & Boremanse, A. (2008). Nonlinear relationship between holistic processing of individual faces and picture-plane rotation: Evidence from the face composite illusion. *Journal of Vision*, *8*(4), 1-13. doi:10.1167/8.4.3

Schwaninger, A., Lobmaier, J. S., & Collishaw, S. M. (2002, January). Role of featural and configural information in familiar and unfamiliar face recognition. In *Biologically motivated computer vision* (pp. 643-650). Springer Berlin Heidelberg.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *46A*(2), 225-245. doi:10.1080/14640749308401045

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition*, *25*(5), 583–592. doi:10.3758/BF03211301

Tiddeman, B., Burt, D. M., & Perrett, D. I. (2001). Prototyping and transforming facial textures for perception research. *IEEE computer graphics and applications*, *21*(5), 42-50. doi.org/10.1109/38.946630

Wilford, M. M., & Wells, G. L. (2010). Does facial processing prioritize change detection? Change blindness illustrates costs and benefits of holistic processing. *Psychological science*, *21*(11), 1611-1615.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*(1), 141-145. doi:10.1037/h0027474

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, *16*(6), 747-759. doi:10.1068/p160747
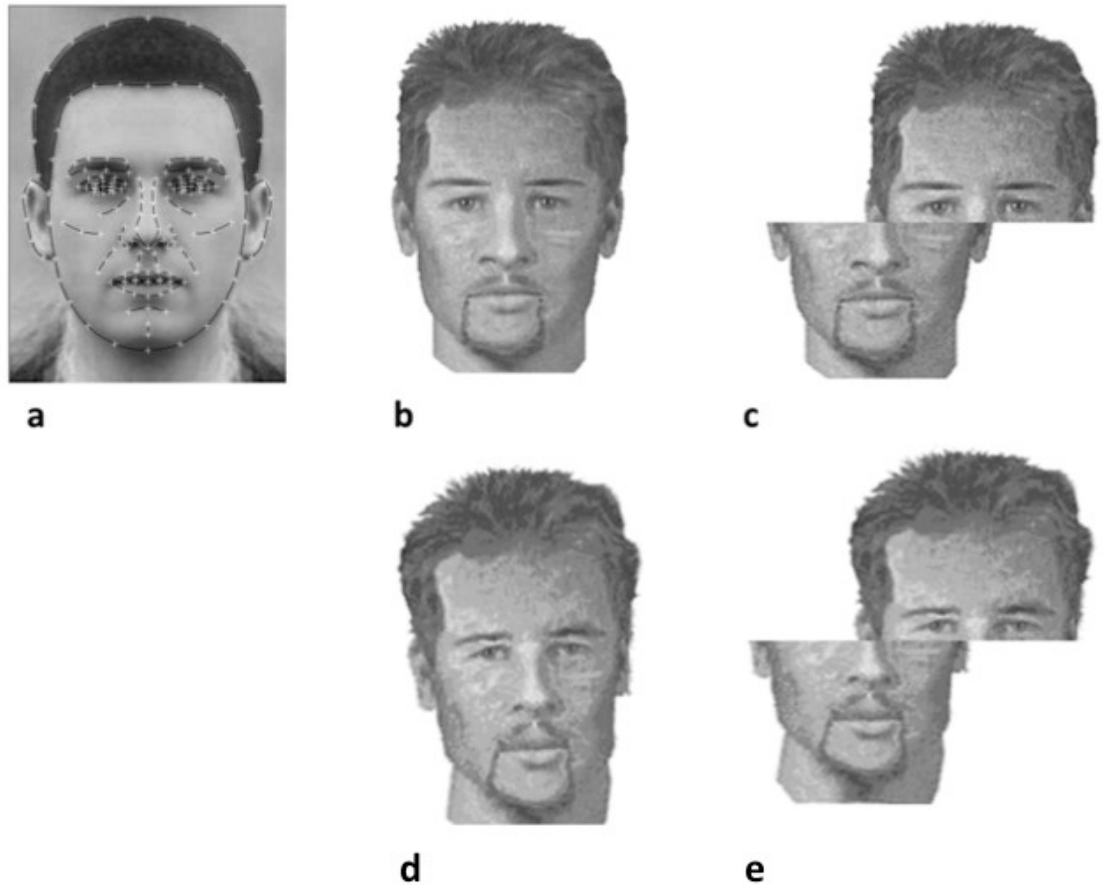
Figure 1. Examples of the facial composite stimuli with a Psychomorph template. Top row: (a) The Psychomorph template; (b) the original facial composite of Brad Pitt; (c) the misaligned original facial composite. Bottom row: (d) the corrected facial composite of Brad Pitt; (e) the misaligned corrected facial composite. The corrected composites was generated by tagging corresponding data points on the target photograph of Brad Pitt and on the original facial composite; the composite was then warped to the parameters of the target photograph. Due to copyright restrictions the target photograph cannot be shown. Misaligned facial composites were generated with Adobe Photoshop Elements 5.0. The direction of misalignment was counterbalanced across participants.
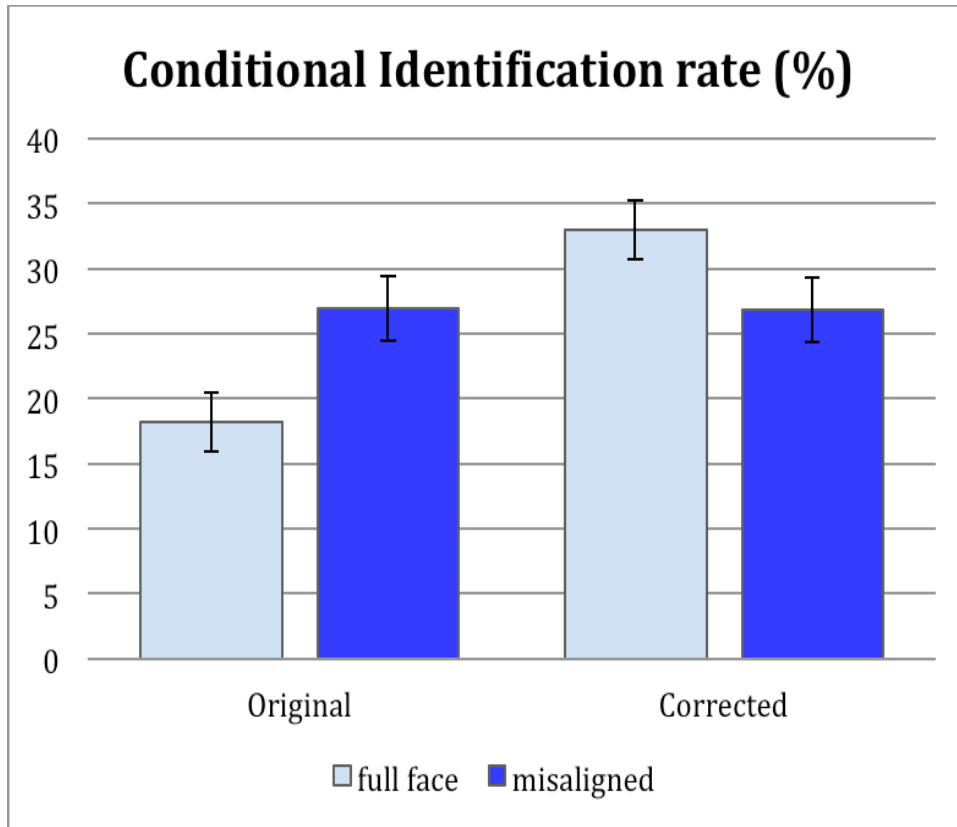
Figure 2. Mean composite identification rates and standard errors by group, and by presentation format. Performance was best for the composites that were corrected to show more accurate information, and was poorest for the original images that were generated by the unfamiliar witness-participants. Identification of the misaligned images was poorer for the corrected images, but an advantage for identification of the original facial composites was observed.

Figure 3. Example of target stimuli used in experiment 2. The series of photographs was designed to display a number of poses, expressions and retrieval cues in order to provide witness participants with a richer memorial representation.
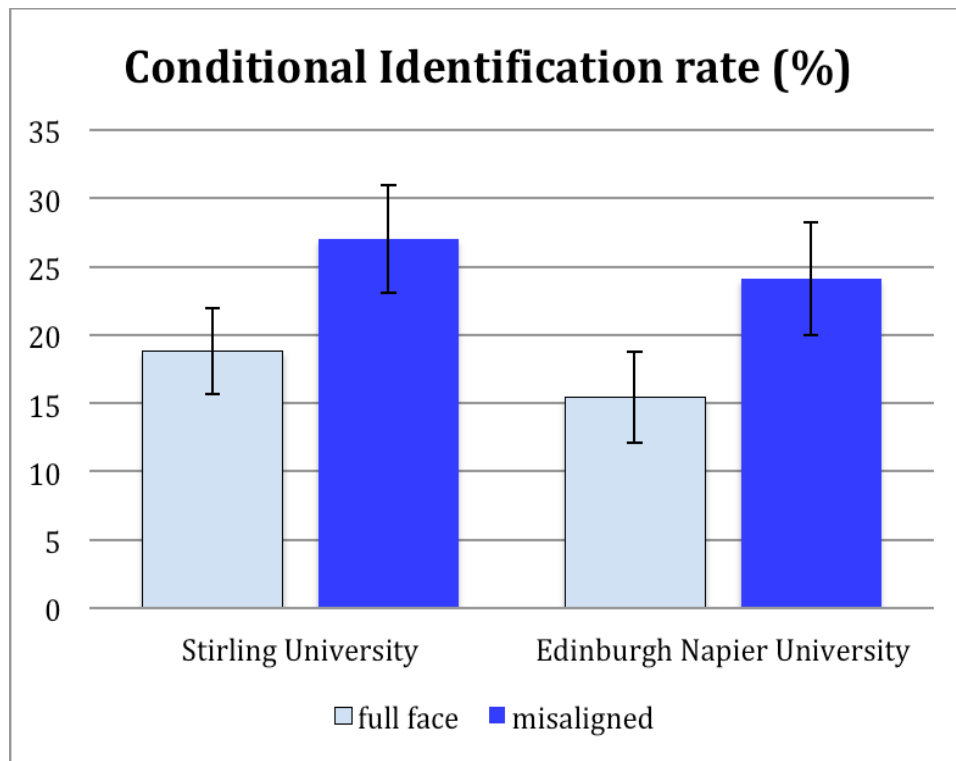
Figure 4. Mean composite identification rates by university group, and by presentation format. Performance was best for the non-celebrity composites that were misaligned. Performance between the university groups was comparable.