

Video Article

Holistic Facial Composite Creation and Subsequent Video Line-up Eyewitness Identification Paradigm

Josh P. Davis¹, Andreea C. Maigut¹, Darrick Jolliffe², Stuart J. Gibson³, Chris J. Solomon³¹Department of Psychology, Social Work and Counselling, University of Greenwich²Department of Criminology, University of Greenwich³School of Physical Sciences, University of KentCorrespondence to: Josh P. Davis at j.p.davis@greenwich.ac.ukURL: <http://www.jove.com/video/53298>DOI: [doi:10.3791/53298](https://doi.org/10.3791/53298)

Keywords: Behavior, Issue 106, Eyewitness identification, repeated identification, holistic, facial configurations, facial features, video line-up, facial composite, memory, autobiographical memory, episodic memory, eyewitness testimony, genetic algorithm

Date Published: 12/24/2015

Citation: Davis, J.P., Maigut, A.C., Jolliffe, D., Gibson, S.J., Solomon, C.J. Holistic Facial Composite Creation and Subsequent Video Line-up Eyewitness Identification Paradigm. *J. Vis. Exp.* (106), e53298, doi:10.3791/53298 (2015).

Abstract

The paradigm detailed in this manuscript describes an applied experimental method based on real police investigations during which an eyewitness or victim to a crime may create from memory a holistic facial composite of the culprit with the assistance of a police operator. The aim is that the composite is recognized by someone who believes that they know the culprit. For this paradigm, participants view a *culprit* actor on video and following a delay, *participant-witnesses* construct a *holistic system* facial composite. *Controls* do not construct a composite. From a series of arrays of computer-generated, but realistic faces, the holistic system construction method primarily requires *participant-witnesses* to select the facial images most closely meeting their memory of the culprit. Variation between faces in successive arrays is reduced until ideally the final image possesses a close likeness to the culprit. Participant-witness directed tools can also alter facial features, configurations between features and holistic properties (e.g., age, distinctiveness, skin tone), all within a whole face context. The procedure is designed to closely match the holistic manner by which humans process faces. On completion, based on their memory of the culprit, ratings of composite-culprit similarity are collected from the participant-witnesses. Similar ratings are collected from culprit-acquaintance assessors, as a marker of composite recognition likelihood. Following a further delay, all participants — including the controls — attempt to identify the culprit in either a culprit-present or culprit-absent video line-up, to replicate circumstances in which the police have located the correct culprit, or an innocent suspect. Data of control and participant-witness line-up outcomes are presented, demonstrating the positive influence of holistic composite construction on identification accuracy. Correlational analyses are conducted to measure the relationship between assessor and participant-witness composite-culprit similarity ratings, delay, identification accuracy, and confidence to examine which factors influence video line-up outcomes.

Video Link

The video component of this article can be found at <http://www.jove.com/video/53298/>

Introduction

If the police have no suspect for a crime, an eyewitness, often the victim, may create a *facial composite* of the culprit from memory with the assistance of a police system operator¹. The aim is that someone familiar with the culprit will recognize that image. An identified suspect — who may not be the actual culprit — may be placed in a line-up to see whether the original eyewitness can identify them or not. Many eyewitnesses make misidentifications. From real police line-ups over 20% identify a foil in the UK² and the USA³, known to be innocent but included in the line-up to provide a test of the witness' memory of the culprit. Sometimes, witnesses select a factually innocent police suspect. This type of error may be the leading cause of wrongful convictions⁴⁻⁶. Eyewitnesses created facial composites in 46 of the first 250 US DNA-exoneration miscarriage of justice cases⁴, and many subsequently identified the innocent individual from a line-up. There may have been alternative inculcating evidence, so that the facial composite procedure was not necessarily responsible. However, line-up accuracy can be impaired^{e.g., 7-10}, unaffected^{e.g., 8}, or enhanced^{e.g., 11, 12} by composite creation, and the aim of applied research of the type described here is to suggest best practice procedures that may be adopted by the police, when facial composite construction is followed by a line-up.

Police forces in many countries employ computerized *holistic* facial composite systems^{e.g., 13-15}, replacing the previously ubiquitous *feature-based* systems (for reviews of composite systems see¹⁶⁻¹⁷). This is mainly because holistic systems more closely match the whole face 'Gestaltic' manner in which humans of all ages process and recognize faces^{e.g., 18-20}. With feature-based systems, witnesses are required to first, verbally recall the suspect's facial features (e.g., eyes, nose, and mouth), second, to locate these individual features from often limited feature databases and third, to assemble them into a whole composite. Each step is cognitively demanding — few people possess the vocabulary to provide a detailed facial description, feature-by-feature facial analysis is prone to error, and verbal recall tends to be less accurate than recognition, possibly due to a mismatch between the modality in which faces are encoded (visual mode), and the retrieval of the face from memory (verbal mode)²¹⁻²². Not surprisingly feature-based composites are often a poor culprit likeness¹⁶⁻¹⁷.

With holistic facial composite systems¹³⁻¹⁵, witnesses select from a series of realistic but computer-generated facial arrays, the image that most closely matches their memory of the culprit. Thus the primary retrieval mode is visual, matching the mode in which faces are encoded. The interfaces of the different systems vary, including whether images are in color or not, array numbers, and whether the witness views whole faces, or faces with the external features removed¹⁷. However, with all holistic systems, successive arrays of images are used to achieve a step-by-step improvement in composite-to-culprit likeness. At each step, the faces comprising an array are determined by the witness' previous selection(s) and an evolutionary algorithm. The witness selects a face from the array, and the evolutionary algorithm breeds those selections to create the new generation of faces in the following array. Additional tools can add clothing, tattoos, facial hair or other individuating marks, manipulate facial feature size and placement, and adjust holistic properties (e.g., age, distinctiveness). The process is complete when the witness is satisfied with their composite. Throughout, the police operator advises but does not influence the procedure.

In conjunction with an initial Cognitive Interview²³⁻²⁴, with an emphasis on the free recall component which facilitates information quality and quantity, holistic system production promotes whole face recognition, rather than as required with feature-based systems — analyses of isolated facial features. Manipulation of facial features and their configurations always occurs within the context of a whole face, and as a consequence, holistic system composites are often of higher quality, and recognized more readily by people familiar with those depicted than composites created using feature-based systems^{17, 25-26}. Police field surveys also suggest that holistic system composites generate higher rates of suspect identification^{e.g., 27-28}. Furthermore, even children as young as six-years-of age⁸, and adults with intellectual disabilities²⁹ can understand their task demands. For these reasons, particularly if a witness has difficulty in describing a face, the Association of Chief Police Officers (ACPO) in England and Wales recommend the use of holistic systems³⁰.

Once a potential suspect has been identified by the police, they may ask the composite-creating witness to view that suspect in a line-up or identity parade. Some research has found that composite creation enhances identification accuracy, when performance is compared with non-composite creating controls^{e.g., 11-12}. Other research, primarily employing feature-based systems has found that composite creation negatively impacts identification performance^{e.g., 7, 9-10}. If a composite is a poor likeness to the culprit, identification accuracy appears most susceptible^{8, 10}. This suggests that for creating witnesses, a facial composite may provide a more salient memory trace than that of the original suspect. Nevertheless, all other things being equal, the chances of a correct identification should be enhanced by the creation of a holistic system composite, as these are likely to be closer in likeness to the culprit than a feature-based composite.

The research paradigm described here closely replicates a design and procedure used with young adult^{8, 11, 31}, child⁸, and older adult³¹ participants. Composite creating *participant-witnesses* and non-composite creating *controls* view an initial 'culprit-actor' crime scene video. After a delay, the participant-witnesses construct a holistic system facial composite. Subsequently, after a further delay, all participants attempt to identify the suspect from a video line-up, the technology used in virtually all of approximately 110,000 formal identification procedures per annum in the UK^{2, 32}. The mean delay between crime scene video viewing and line-up will be equal for both groups. Based on a large body of empirical research, there is an ongoing debate as to the best procedure for conducting a line-up. Some researchers have argued in favor of sequential line-ups over the alternative simultaneous line-ups^{e.g., 33-34}. Others have an opposing view^{e.g., 35-36}. However, correct identification rates are higher from video line-ups than the often recommended US sequential procedure³⁷⁻³⁸. Regardless, the methodology described here has strong ecological validity as it closely matches procedures used in real police investigations, and could be adapted for use with different line-up types. Methods of ensuring that the line-up is 'fair' to the suspect are also described³⁹⁻⁴⁰. In addition, the holistic composite system employed is the standard used by the majority of UK police forces, the interviewing techniques are those normally used by police composite operators, and the video line-ups were constructed by the London Metropolitan Police Service as though a real investigation, following guidelines in the Police and Criminal Evidence Act (PACE) Codes of Practice (Code D), which prescribes police identification procedures in England and Wales⁴¹. Furthermore, the design is flexible, allowing for the measurement of alternative variables that might impact on composite quality and identification accuracy (e.g., composite system, culprit-actor gender, age, or ethnicity; delay between procedures).

In making decisions as to whether to publicize a facial composite or not, the police operator will likely ask the creating-witness to assess final composite quality. For this paradigm, following construction, each participant-witness rates their composite for likeness to their memory of the culprit. These ratings are compared to those provided by independent assessors — highly familiar with the person the composite is supposed to depict, so as to provide a more objective recognisability measure. These ratings serve as a proxy for the more ecologically valid method of asking people familiar with the culprit to attempt to name the composite. However, assessor ratings correlate with naming rates suggesting they are essentially measuring the same construct⁴². They also positively correlate with adult, but not child participant-witness creator assessments of quality⁸.

In summary, this paradigm employs an independent measures design. The first factor is participant role — participants are allocated to either the participant-witness composite creating group or the non-composite creating control group. The second factor is culprit presence — participants are presented with a culprit-present video line-up containing the culprit viewed in the original crime scene video, or a culprit-absent line-up in which the culprit is replaced by an additional foil. A third factor, if varied, is the delay between viewing the crime scene video and the line-up. The primary dependent variable is line-up accuracy. Within the participant-witness composite-creating group, a correlational design measures the relationship between ratings of composite-culprit similarity provided by participant-witnesses and independent assessors, some of whom should be acquaintances of the culprit. The relationship between the composite-culprit similarity ratings and video line-up accuracy is also examined.

Protocol

Procedures involving human participants were approved by the University Research Ethics Committee of the University of Greenwich, following guidelines issued by the British Psychological Society.

1. Viewing of Crime Scene Video Depicting the ‘Culprit’

1. Have the participant initial an information and consent form, to a study deceptively described as ‘Video Analysis’, but which lists their usual ethical rights as a research participant, and correctly states that they will view a video depicting a minor crime, and that outcomes of the study may assist future police investigations.
2. Have the participant provide an anonymous personal code, and demographic data of their age, gender, and ethnicity.
3. Have the participant view a randomly selected video clip on a laptop from a pool of videos of different actors, depicting good full body views and close-ups of the front and both sides of the face of the actor playing the part of a ‘culprit’ committing a minor crime (see **Figure 1A** and **1B** for example stills from the video).



Figure 1. Stills from the crime scene video. (A, B) Two stills from the crime scene video depicting full body and facial views of the culprit (see 1.3). [Please click here to view a larger version of this figure.](#)

4. Ask the participant whether they are unfamiliar with the culprit-actor (yes/no)?
NOTE: If familiar with the culprit-actor, the participant views a different culprit video.
5. Have the participant provide a verbal description of the culprit’s approximate age, gender, ethnicity and clothing.
6. Have the participant provide a verbal rating of prospective confidence in being able to recognize the culprit (0%: no confidence to 100%: absolutely certain).
7. Randomly allocate the participant to experimental condition (participant-witness vs. control), and ensure that composite creating participant-witnesses participate in Stages 2 and 7 below; controls in Stage 7 only. Ensure that the mean delay from Stage 1 to Stage 7 is equal for all participant groups.

2. Participant-witness Creation of a Facial Composite with the Assistance of an Operator

NOTE: This section of the protocol has been optimized for the holistic facial composite system, EFIT-V, but can be adapted for other software.

1. Have the trained composite system operator inform the composite-creating participant-witness that they will be creating a facial composite.

2. Have the operator ask the participant-witness how confident they are that they could recognize the culprit again (0% = not at all confident to 100% = highly confident).
3. Have the operator ask the participant-witness how confident they are that they could construct a facial composite of the culprit (0% = not at all confident to 100% = highly confident).
4. Have the operator ask the participant-witness for permission to audio record the session.
5. Have the operator interview the participant-witness using elements of the Cognitive Interview (e.g., rapport building), in which the operator primarily requests the participant-witness to provide a free-recall description of what they saw in the video. Have the operator ask the participant-witness to tell him/her everything they remember about the video and the person depicted, and inform them that when they have finished, they will be asked some additional questions.
6. Have the operator prompt the participant-witness with the questions listed in **Table 1**, but only if the description associated with that question is missing from the participant-witness' free recall account.

1	How old did the culprit appear to be?
2	What do you remember about the culprit's hair (length, type, style, color)?
3	What do you remember about the culprit's face (shape, length, breadth, complexion)?
4	What do you remember about the culprit's ears (shape, size, position, lobes)?
5	What do you remember about the culprit's nose (length, tilt, nostrils, shape, ridge)?
6	What do you remember about the culprit's eyebrows (thickness, space, shape, color)?
7	What do you remember about the culprit's eyes (shape, size, depth, space, shade, color)?
8	What do you remember about the culprit's mouth/lips (width, shape, upper, lower)?
9	What do you remember about the culprit's chin (shape, size, type)?
10	What do you remember about the culprit's facial hair (beard, moustache, stubble)?
11	Did the culprit wear spectacles?
12	Was there anything distinctive about the culprit (marks or scars)?
Note:	The operator should only ask any of these questions, if the description associated with that question is missing from the participant-witness' free recall account.

Table 1. Cued Post Cognitive Interview Questions.

7. Have the operator turn on the interface of the holistic facial composite system software on a laptop, which, as with a real investigation, stores data to ensure a reliable evidence chain.
 8. From the information gathered from the participant-witness during the Cognitive Interview, have the operator enter the gender, ethnicity, and age range of the described culprit into the appropriate boxes on the first screen of the composite system interface.
 9. Have the operator guide the participant-witness through the construction of the holistic facial composite following a procedure in which the participant-witness selects the best and rejects the worst matching images to their memory of the culprit from a 3 x 3 array of nine randomly displayed computer-generated images. If the participant-witness is not satisfied with any of the nine images, have the operator generate additional arrays.
- NOTE: The 'best' example selected from one array always appears in the subsequent array and the similarity between faces within an array increases automatically at each step of the process.

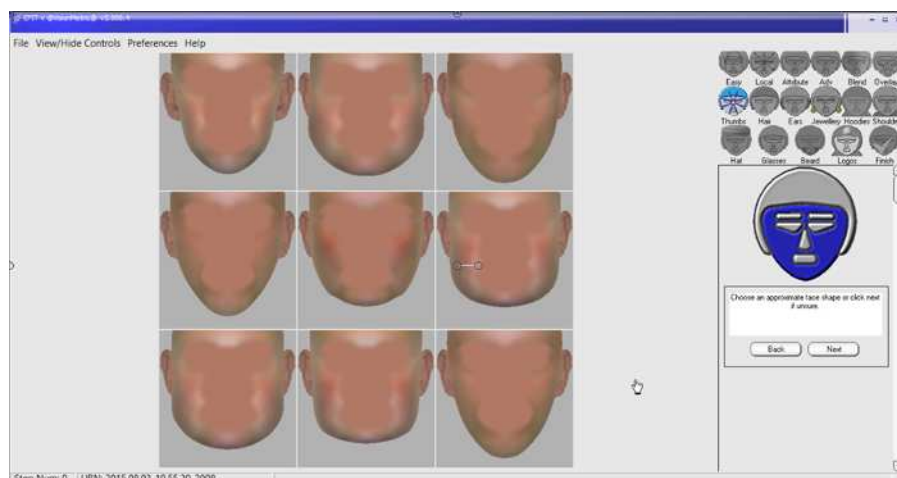


Figure 2. Facial composite construction method A: Face shape. At this stage in the facial composite construction procedure, after the operator enters basic description keywords into the holistic composite system, the participant-witness is asked to select an approximate face shape meeting their memory of the culprit from the nine images displayed on the screen, or to reject that array to produce a new display. As with the remainder of the construction process, this stage assesses recognition (see 2.9.1). [Please click here to view a larger version of this figure.](#)

1. Have the participant-witness select an approximate face shape matching that of the culprit from the automatically generated first displayed array of nine images (see **Figure 2**). Have the operator use the interface's TOOLS feature to enter that choice.
2. Have the participant-witness select the closest matching a) nose, b) mouth, c) eye and d) eyebrow shape from subsequent arrays using the interface's Tools feature in a similar manner to that described in 2.9.1.
3. With all array faces now possessing the features entered above, but with hair initially colored grey, have the participant-witness select an appropriate hairstyle and hair color from the large database in the interface's HAIR tool. Have the operator enter that choice (see **Figure 3**).

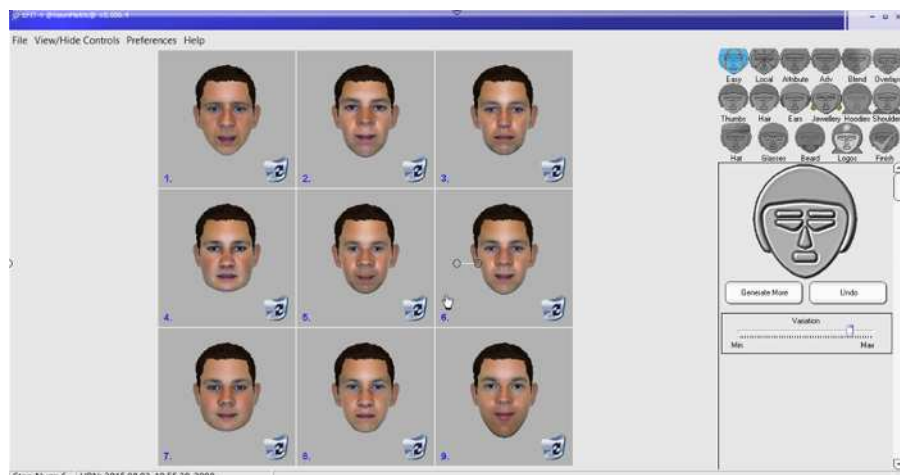


Figure 3. Facial composite construction method B: Hairstyle tool. Following selection of face shape, and facial features, the participant-witness is asked to select an approximate hairstyle from the nine images displayed on the screen, or to reject that array to produce a new display. The default hairstyle on all images is grey, until coloring is added (see 2.9.3). [Please click here to view a larger version of this figure.](#)

4. Have the participant-witness select appropriate shoulders with clothing and color from the palette available in the interface's SHOULDERS tool. Have the operator use the witness-directed controls to move, scale and rotate the neck and shoulders if necessary (see **Figure 4**).

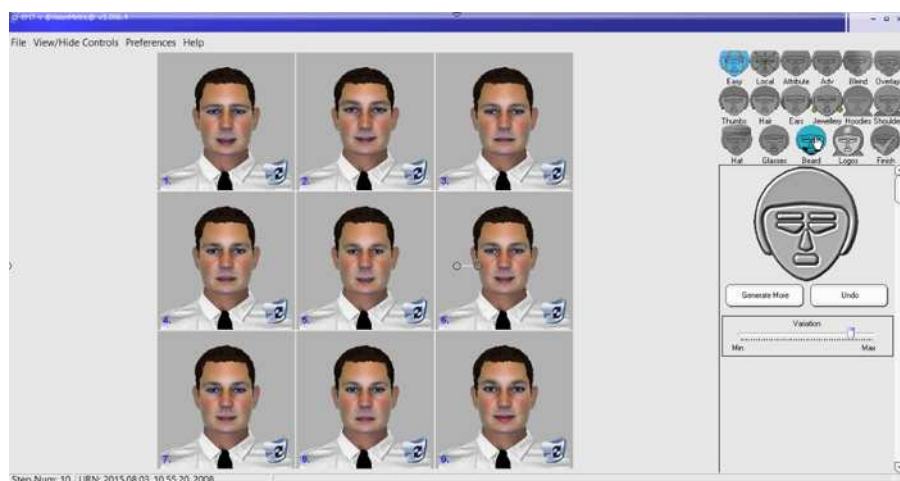


Figure 4. Facial composite construction method C: Shoulders tool. Following selection of face shape, and facial features, the participant-witness is asked to select shoulders from the nine images displayed on the screen, or to reject that array to produce a new display. Clothing color and style can be manipulated and company logos or other idiosyncratic features may be added (see 2.9.4). [Please click here to view a larger version of this figure.](#)

5. If they wish, have the participant-witness select clothing with or without logos (e.g., scarves, hoodies, spectacles, sunglasses), as well as facial hair (beards and moustaches) to the array faces using additional tools available on the interface. Have the operator enter these selections.
6. Have the operator demonstrate the interface's DYNAMIC OVERLAY tool to the participant-witness, which allows subtle changes to be made to the skin (e.g., wrinkles, age lines, eye-bags and shadows, prominent cheek bones, chubbiness, rough skin, acne, etc.), or to the overall face (e.g., shading). Have the operator make changes, if directed by the participant-witness.
7. Have the operator magnify the face on the screen in order for the participant-witness to inspect it more closely for editing using the LOCAL ATTRIBUTES tool, which allows systematic changes to be made to the shape of the individual facial features as well as the overall shape of the face and head (e.g., stretched, rotated and warped) (see **Figure 5**). Have the operator make changes, if directed by the participant-witness.



Figure 5. Facial composite construction method D: Local attributes tool. After the shoulders are selected, the participant-witness views a series of facial arrays possessing faces of differing variability from each other, although variability reduces in subsequent arrays, as each 'best' image is chosen. At this point in creation, the participant-witness may suggest changes to specific facial features, and compare the outcome to the original unmodified image on the screen. Even though changes are made to features, the methodology still accesses holistic processes as changes are made in the context of whole face comparison (see 2.9.7). [Please click here to view a larger version of this figure.](#)

8. Have the operator demonstrate the HOLISTIC ATTRIBUTES tool to the participant-witness, which allows holistic changes to be made to the face such as making it appear older or younger, more or less distinctive, and paler- or darker-skinned (**Figure 6**). Have the operator make changes, if directed by the participant-witness.



Figure 6. Facial composite construction method E: Holistic attributes tool. The participant-witness may also suggest changes to the holistic properties of the selected face (e.g., age, distinctiveness) by using a slider tool. Again, the outcome is compared to the original unmodified image on the screen (see 2.9.8). [Please click here to view a larger version of this figure.](#)

9. Have the operator display the final composite on the screen and have the participant-witness approve this in order to save the file by clicking on the SAVE IMAGE button in the interface's FINISH tool (see **Figure 7**).



Figure 7. Facial composite construction method F: Final image. In a police investigation this image would be printed, and a copy transferred to CD to be retained in the evidence bag (see 2.9.9). [Please click here to view a larger version of this figure.](#)

3. Collection of Post Composite Construction, Participant-witness Ratings of Culprit-composite Likeness

1. Based on their memory of the culprit, ask the participant-witness how close a likeness the facial composite is to the culprit they saw in the video (0%: poor likeness to 100%: exact match).
2. Based on their memory of the culprit, ask the participant-witness how confident they are that someone familiar with the culprit would be able to recognize them from the composite (0%: not at all confident to 100%: very confident).
NOTE: the mean rating from the scales described in 3.1 and 3.2 is calculated to produce a participant-witness self-rating of culprit-composite similarity.

4. Culprit-acquaintance Assessment of Culprit-composite Similarity

1. Have a close acquaintance of the culprit (culprit-acquaintance assessor), provide an independent assessment of composite quality by viewing the composite, adjacent to two video stills from the original crime scene showing close-up facial views of the culprit, displayed as a reminder of appearance in case the culprit's hairstyle etc. has since changed.
2. Have the culprit-acquaintance assessor provide an assessment of composite-suspect likeness (0% = no similarity; 100% = highly similar).
NOTE: The mean composite rating provided by a group of assessors is calculated to produce independent ratings of composite-suspect likeness.

5. Preparation of the Video Line-ups Containing the Culprit and Foils

NOTE: This section of the protocol has been optimized for the video line-up system PROMAT, although other systems are available.

1. Have a police officer create a video line-up of the culprit, at an identification suite in a police station.
 1. Have the police officer film a 15 sec video clip of the culprit consisting of a head-and-shoulders clip of the culprit facing the camera, turning to the left, then to the right before turning to face the camera again in standard environmental conditions (e.g., lighting, distance, camera, background). See **Figure 8** for example stills extracted from the video line-up of the culprit.





Figure 8. Video lineup stills. (A, [Please click here to view a larger version of this figure.](#) B) Frontal and profile facial image stills of the culprit in the culprit-present video line-up procedure (see 5.2). [Please click here to view a larger version of this figure.](#)

2. Have the police officer select videos of nine foils from a database of over 40,000, taken in the same environmental conditions and matched with the culprit for age, gender, ethnicity and 'general appearance in life'.
NOTE: Normally only eight foils are included in a video line-up. For experimental purposes, one of the nine foils is randomly selected to replace the culprit for the culprit-absent video line-up.
3. Have the culprit agree that the selected foils are suitable (e.g., they possess a reasonably similar appearance to the culprit), as a suspect would have this opportunity in a real police investigation.
NOTE: Alternatively, their legal representative could have this opportunity.
4. Have the police officer assemble the video line-up and copy it onto a CD to allow random playback later.

6. Mock Witness Paradigm Pilot Study to Test Video Line-up Fairness

1. Have a group of five pilot participants, unfamiliar with the culprit, and who do not participate in any other procedure, provide a written description of the culprit after viewing the crime scene video.
2. Have another pilot participant; blind to the study design and unfamiliar with the culprit amalgamate the descriptions collected in 6.1 into a single modal description by only including descriptions of features that are described consistently by the majority of the pilot participants, while disregarding those described by a minority of pilot participants.
NOTE: The interpretation of the instructions above is left to the judgement of the pilot participant.
3. Have a further group of 'mock-witness' participants, who have also never seen the culprit, or taken part in any other procedure of the research, view an array of full-face video stills of the nine line-up members – extracted from the video line-up and to select one member based on the modal description created in 6.2.

7. Presentation of the Video Line-up and Questionnaire

1. Have controls and composite creating participant-witnesses participate in this final phase of the study with the same delay between viewing the initial crime scene video for both groups.
NOTE: The controls can be provided with a distraction task (e.g., puzzles) during the period of time the participant-witnesses took to create a facial composite.
2. Randomly allocate the participant to view either a culprit-present or a culprit-absent video line-up.
3. Have the participant read the instructions on the Cued Description Form (see **Table 2**), and then complete the multiple-choice or cued questions.

Instructions					
The following is a cued description form, please try to enter comments in each section (if you can) appertaining to the particular aspect of the person (the culprit) that you saw in the original video clip. As describing a person is often a difficult task, it is important that you concentrate and stay focused for the next few minutes. Prior research has also demonstrated the importance of striving for accuracy and reporting only that which you are certain you remember.					
1	Ethnic appearance				
2	Height				
3	Apparent age				
4	Gender				
Please circle one or more responses to the following questions					
5	Build	Fat, Proportional, Thin, Stocky, Athletic, Heavy, Other			
6	Hair color	Dark brown, Light brown, Fair, Blonde, Grey, White, Black, Ginger, Auburn, Other			
7	Hair type	Bald, Thinning, Receding, Straight, Curly, Wavy, Dyed, Short, Collar length, Shoulder, Very long, Wig, Length, Other			
8	Eyes	Blue, Brown, Green, Grey, Cast, Staring, Other			
9	Complexion	Fresh, Pale, Ruddy, Tanned, Fair, Freckled, Dark tone, Mid tone, Light tone, Other			
10	Facial hair	Beard, Moustache, Bushy, Sideburns, Eyebrows, Other			
Clothing: Enter brief description (if appropriate)					
11	Shoes		12	Socks	
13	Trousers		14	Belt	
15	Shirt		16	Jacket	
17	Skirt		18	Dress	
19	Jumper		20	Topcoat	
21	Jewelry		22	Hat	
23	Other				

Table 2. Cued Description Form.

- Have the line-up administrator inform the participant that they will be attempting to identify the culprit they originally viewed in the crime scene video, in a video line-up displayed on a computer monitor.
- Have the line-up administrator warn the participant-witness that the culprit they saw in the initial crime scene video may or may not be present in the line-up.
- Have the line-up administrator start the video line-up procedure on a computer monitor consisting of a sequential display of the nine 15 sec clips which should be shown twice, with suspects and foils randomly ordered, and with a line-up member number (1–9) appearing with each video clip.
- Have the participant view the video line-up.
- On completion, have the line-up administer ask the participant whether they would like to view any part, or the whole of the line-up again.
NOTE: The participant may view part or the whole of the line-up as many times as they like.
- Have the participant respond in writing to a line-up questionnaire asking whether the culprit was present or not in the line-up (yes/no), and if the response is 'yes' to provide the line-up member number (1-9).
- If the participant has selected a line-up member, have the line-up administrator play the video clip of that member only to ensure the participant is satisfied with their response.
- Have the participant provide a confidence estimate in their line-up decision regardless of whether they made a selection or rejected the line-up in 7.5 (0% = no confidence to 100% = absolutely certain).

8. Data Analyses

NOTE: Some of the data collected (e.g., descriptions of the culprit) in this paradigm are included primarily to ensure that procedures conform to normal police practice in England and Wales and not specifically for later analyses. Nevertheless, it would be possible to analyze these data to test for relationships between perhaps description quality and quantity, and composite quality and identification accuracy. However, these would be supplemental analyses, and the data analyses listed here are those that are most likely to be employed to investigate important experimental hypotheses.

- Use an independent-measures t-test⁴³ to ensure that the ratings of prospective confidence in being able to recognize the culprit, which were collected shortly after the participants viewed the crime scene video, are equal in the two experimental conditions (see 1.3.3). Check that the outcome is non-significant.
- Test hypotheses concerning the objectivity of the participant-witness' self-assessments of their own composites, by using Pearson's correlation coefficient tests⁴³ to examine the relationship between these self-assessments to their individual composites (see 3.2), with

- the culprit-acquaintance assessor's ratings (see 4.5), and if collected the culprit-unfamiliar assessor ratings (see 4.6) to the entire set of composites.
3. Test hypotheses related to line-up performance, use hierarchical loglinear analyses⁴³, or chi-square tests⁴³, to examine the effects of experimental condition on line-up outcomes (see 7.9).
 4. Use culprit-present line-ups to provide an indication of the sensitivity of an identification procedure as measured primarily by correct culprit identification rates.
 5. Use culprit-absent line-ups to provide an indication of the fairness of the procedure, as measured by correct line-up rejection rates.

Representative Results

The data reported here are a subset of data collected in two studies in which the described experimental paradigm was partially followed^{8, 11}.

Line-up fairness checks

The pilot mock witness paradigm described in Section 6, is designed to ensure that the line-up is not biased against a suspect, in that they should not stand out in any manner to induce selection more often than would be expected by chance^{39,40}. From this procedure, a measure of line-up fairness is calculated by ensuring no member is selected significantly more often than would be expected by chance alone by the mock witnesses (e.g., $1/9 = 11.1\%$). Tredoux's E³⁹ measure of functional size is applied to assess the number of line-up members who are 'plausible', and in a real line-up would provide a suitable test of the witness' memory. Ideally this value should be close to the maximum (e.g., nine). For the representative data reported in¹¹ the line-up was found to be fair as mock witnesses selected the culprit at close to chance levels (10.9%), and the vast majority of foils were plausible (Tredoux's E = 7.05).

Matched condition checks

The next analysis uses an independent-measures t-test⁴³ to ensure that the ratings of prospective confidence in being able to recognize the culprit, collected shortly after the participants view the crime scene video are approximately equal in the two experimental conditions (see 1.6). The outcome should be non-significant. In¹¹, the responses on this scale described in 1.6 were as expected approximately equal, $t(266) = .57$, $p > .2$, indicating the participants were matched prior to any other procedure.

Composite-culprit similarity ratings

To test hypotheses concerning the objectivity of the participant-witness' self-assessments of their own composites, the second analysis examines whether there is a relationship between the composite-suspect similarity ratings provided by the participant-witnesses to their own composite only (see Section 3), and those provided by the independent culprit-acquaintance assessors (see Section 4). In⁸, a Pearson's correlation test⁴³ on the ratings provided to all 57 holistic facial composites was non-significant. However, when follow-up analyses were conducted with child- and adult-participant-witness data separated, there was a positive relationship between these ratings to the 26 adult composites, $r(26) = .46$, $p < .05$; but not to the 31 children's composites, $r(31) = .01$, $p > .2$; an indication that adult witnesses, but not children, can provide objective assessments of the quality of their own composites.

Video line-up responses

To test hypotheses related to line-up performance, hierarchical loglinear analyses⁴³, or chi-square tests⁴³ examine the effects of experimental condition on line-up outcomes (see 7.9). The reported effect size measure for these nominal data analyses is Φ , although odds ratios (OR) are also reported to provide a measure of the association between the two reported variables. As with most eyewitness research, culprit-present and culprit-absent video line-up data are separated. Each participant makes one line-up decision only.

The influence of composite construction on eyewitness identification is measured by comparing the line-up selections of participant-witnesses and controls. **Table 3** displays representative results taken from a subset of the data collected in Experiment 1 of¹¹ in which control line-up outcomes were compared with participant-witnesses who created a holistic facial composite using the system described in this protocol. The delay between viewing the initial culprit crime scene video and the video line-up in this experiment was approximately 2 hr.

Culprit-present line-ups provide an indication of the sensitivity of an identification procedure as measured primarily by suspect identification rates, which in this paradigm are correct culprit identifications. Other outcomes are incorrect foil identifications or incorrect line-up rejections.

Culprit-absent line-ups provide an indication of the fairness of the procedure, as measured by correct line-up rejection rates. Other outcomes are incorrect foil identifications. There was no designated 'innocent suspect' in this research, and therefore the first column in **Table 3** is blank for culprit-absent trials.

Culprit-present choosing behavior: From the data reported in **Table 3** originally presented in Experiment 1 in¹¹, the first analysis examines whether composite creation influences choosing behavior from a line-up as this may be indicative of a response bias. A 2 (participant role - participant-witness vs. control) x 2 (choosing behavior - chooser: culprit identification or foil identification vs. non-chooser: incorrect line-up rejection) chi-squared test⁴³ on the rates of selection of each outcome was not significant, $\chi^2(1, n = 108) < 1$, $p > .2$, $\Phi = .072$. Participant-witnesses (80.0%) were roughly equally likely as controls to be line-up choosers (73.1%, OR = 1.09).

Culprit-present correct identifications: The second and most critical culprit-present analysis examines response accuracy only. A 2 (participant role) x 2 (accuracy - correct: culprit identification vs. incorrect: foil identification or line-up rejection) chi-squared test⁴³ on the data from **Table 3** first presented in Experiment 1 in¹¹ was significant, $\chi^2(1, n = 108) = 5.48$, $p = .019$, $\Phi = .225$. Participant-witnesses (70.0%) made approximately one-and-a-half times more correct culprit line-up selections than controls (44.9%, OR = 1.56).

These results are consistent with a follow-up experiment (Experiment 2¹¹) in which newly qualified police operators were recruited, the crime scene videos depicted six different culprits, and the mean delay between crime scene video viewing and viewing the video line-up was

approximately 30 hr (participant-witness correct culprit identification rates = 48.8%; controls = 35.0%), and a meta-analysis finding positive effects of composite construction on line-up identification¹². However, other research⁸, using the same basic experimental paradigm, holistic facial composite system, and line-up type, but with different culprit-actors found no significant differences in correct identification rates between adult composite-creating participant-witnesses (34.6%) and controls (31.7%). Furthermore, in that research⁸, child controls between the ages of 6- and 11-years made more correct identifications (42.9%) than child participant-witnesses of the same age (19.4%). The latter result may be a consequence of the children's facial composite being significantly inferior to the adult's, children's initial memory for the culprit being worse or they struggled to understand the use of confidence scales. However, this result is consistent with research finding a positive relationship between composite quality and rates of correct identification from line-ups¹⁰. This explanation is also consistent with most previous research of this type, which has used the often inferior feature-based composite systems, finding that identification accuracy was reduced following composite construction e.g., 7, 9-10.

Culprit-present foil identifications: The third culprit-present analysis examines whether the proportion of foil identifications differs by condition. A 2 (participant role) x 2 (foil or not — foil identification vs. other decision: correct culprit identification or incorrect line-up rejection) chi-squared test⁴³ on the data reported in **Table 3** and¹¹ was significant, $\chi^2(1, n = 108) = 4.04, p = .045, \Phi = .193$. Controls (28.2%) made nearly three times as many foil selections as participant-witnesses (10.0%; OR = 2.82).

Culprit-absent lineup rejections: As there are only two outcomes associated with the culprit-absent data from **Table 3** as originally reported in¹¹, only one test is conducted. A 2 (condition) x 2 (accuracy — correct line-up rejection vs. incorrect foil identification) chi-squared test⁴³ was not significant, $\chi^2(1, n = 100) < 1, p > .2, \Phi = .055$. There were no differences in correct line-up rejection rates between participant-witnesses (44.4%) and controls (38.4%; OR = 1.16). These results are consistent with previous research finding similar null effects in culprit-absent trials¹⁰.

Relationship between composite quality and participant-witness line-up accuracy

A further analysis examines the relationship between the quality of the facial composites and the likelihood of correct culprit video line-up identifications. The representative data reported here are from the second experiment reported in¹¹ in which all line-ups were culprit-present. A point biserial correlation test⁴³ conducted on the relationship between line-up accuracy (1 = correct; 0 = incorrect) and culprit-acquaintance ratings of culprit-composite similarity, was not significant, $r(45) = -.05, p > .2$, suggesting that unlike some previous research^{8, 10}, there was no relationship between the quality of the participant-witness' facial composite and the accuracy of their video line-up responses. This unexpected non-significant finding may be the result of a number of extraneous variables (e.g., delay, multiple culprit-actor variables).

	Total		Suspect ID		Foil ID		Line-up rejection	
	n	%	n	%	n	%	n	%
Culprit-present								
Controls	78		35	44.9	22	28.2	21	26.9
Witnesses	30		21	70.0	3	10.0	6	20.0
Culprit-absent								
Controls	73		-		45	61.6	28	38.4
Witnesses	27		-		15	55.6	12	44.4

Table 3. Culprit-present and culprit-absent video line-up outcomes. Number of participants (n) and percentage of each type of line-up outcome as a function of culprit presence, and participant role from the subset of data originally published in Experiment 1 of¹¹, in which adult participant-witnesses used the same holistic composite system as described in the current protocol.

Discussion

The creation of a facial composite may provide the first lead in a police investigation¹. A composite constructing witness may subsequently be asked to view a line-up containing the police suspect. The police suspect may in fact be the guilty culprit, in which case the line-up will be culprit-present, or they may be innocent, and the line-up will be culprit-absent. The applied experimental paradigm described here has been employed in research demonstrating the positive influence of holistic facial composite production on correct culprit-actor selections from culprit-present sequential nine-person video line-up procedures, without having any impact on culprit-absent procedures¹¹. The paradigm possesses strong forensic and ecological validity as it closely replicates police procedures used in England and Wales. Indeed, although not a central component of the analyses, the protocol closely follows the type of interview procedures (e.g., Cognitive Interview²³⁻²⁴) that the police may employ in these circumstances, although it would be of interest to vary components of the Cognitive Interview in future research to examine this as a separate variable. Furthermore, even though a description is not actually required for holistic composite construction, an operator will still ask a witness to describe the culprit in order to demonstrate good practice, as if the description and appearance of the composite match, it suggests no undue influence on construction from the operator (for instance, concerns might be raised in court if the composite and police suspect possess blond hair, the description dark hair).

Other aspects of the procedure should be followed to enhance forensic validity. For instance, real witnesses would normally be unaware in advance that they will be witnessing a crime, and pre-warning participants that they will either create a facial composite or be asked to make an identification decision may influence the manner in which they attend to the initial video of the culprit. Therefore it is normal to avoid warning participants that the research is investigating eyewitness procedures, and for this purpose the title may often be slightly misleading (e.g., "video analysis study"). Furthermore, to avoid inadvertent experimenter bias, the various experimenter roles (e.g., composite system operator, line-up administrator) should ideally be conducted by different people. For similar reasons, the operator, who should be fully trained in the use of the composite system, and the line-up administrator, should both be unfamiliar with the culprit, and have never seen the culprit videos. Ideally, too, the line-up procedure should be conducted double-blind, to ensure no administrator bias⁴⁴⁻⁴⁶, although double blind procedures are not

prescribed in PACE Code D identification procedures for England and Wales. Similarly, PACE Code D does not require measures of identification confidence to be collected, even though as with this paradigm it is often collected for research purposes. Indeed, confidence may provide a marker for line-up accuracy. As such, there is normally a strong positive relationship between confidence and accuracy in line-up choosers, but not non-choosers⁴⁷⁻⁴⁹.

The basic paradigm could also easily be adapted for use with different types of facial composite system (see for example¹¹ in which the influence on video lineup outcomes of creating a feature-based facial composite was compared with creating a holistic system composite). There are a number of available holistic composite systems, and although all work on holistic principles, the interfaces differ¹³⁻¹⁵. The paradigm would also allow for comparisons with different types of line-up that may be standard in the UK or other legal jurisdictions. For instance, in the USA, most identification procedures are six-person sequential or simultaneous photographic line-ups³³⁻³⁶, whereas Poland mainly employs four-person simultaneous line-ups⁵⁰. Although nowadays very rare in the UK, some jurisdictions elsewhere may still regularly employ live line-ups with all members present in person. For the purposes of the experimental paradigm described here an additional foil was chosen by random from the nine selected by the police officer line-up administrator to replace the culprit in the culprit-absent line-up. However, some eyewitness identification paradigms, may, in advance, specifically select an 'innocent suspect' for this purpose. A culprit-absent line-up will therefore be created that might contain a number of different foils from those included in the culprit-present line-up^{e.g., 51}, as a real police line-up will be individualized to the specific suspect. To further examine the influence of composite creation on line-up outcomes, combining the results of culprit-present and culprit-absent line-ups would also inform as to whether memory sensitivity or response bias to make a selection is influenced (see for example⁵² which discusses the use of signal detection measures for this type of analysis).

In addition, there are many variables that are known to affect face recognition and eyewitness identification accuracy, and the design could accommodate the testing of these to examine whether they provide an additional influence when facial composite production is followed by a line-up (e.g., culprit-actor gender, age, or ethnicity; participant confidence, delay between procedures; foil selection methods, participant culprit view quality; witness description quality; for a review of 'estimator' variables of this type see⁵³). Nevertheless, one issue with this type of design is that large numbers of participants are required for sufficient statistical power, as line-up responses are primarily dichotomous in nature (e.g., correct vs. incorrect), and as with a real police investigation most participants will create one composite and view one line-up only. Furthermore, the design is time-consuming. To replicate police investigations, there should ideally be a delay between the three phases of the study, and composite construction alone can often take over an hour. Despite these issues, it is important that new technology that may be employed by the police during an investigation (e.g., a facial composite system), is empirically tested in the laboratory to investigate the effect this technology has on subsequent investigatory procedures (e.g., line-ups). The fate of a real police suspect, regardless of guilt, may partly depend on whether best practice is followed or not.

Disclosures

Solomon and Gibson are faculty members of the University of Kent and directors of VisionMetric Ltd. VisionMetric Ltd market the EFIT-V facial composite system. Solomon and Gibson's contribution to this work was to facilitate operator training that took place at the University of Kent, development of the software and provision of software support, and assistance with manuscript preparation. Data collection, analysis, interpretation, and manuscript preparation were performed by Davis, Jolliffe and Maigut who have nothing to disclose.

Acknowledgements

The research was partly funded by an internal University of Greenwich grant to the first and third authors. Thanks go to Dominic Goodchild, Chris Hughes, Adrian Ibanescu, Corrado Ranelli, and Charlie Shaw for acting as culprit-actors in the crime scene videos, and to Henry DC Williams for filming and directing the crime scene videos. Thanks also to Detective Chief Inspector Mick Neville, Inspector Barry Burnell and Sergeant Nick Milbourn from the Metropolitan Police Service for creating and supplying the PROMAT video line-ups for the project and advising on the design and video line-up procedure. Thanks also to Detective Constable Tony Barnes of the Metropolitan Police Service for providing advice on the use of the specific holistic facial composite system in police investigations, and allowing the second author of this paper, who is the composite operator depicted in the videos, to shadow him while working with eyewitnesses to create facial composites during investigations into real crimes.

References

1. McQuiston-Surrett, D., Topp, L. D., & Malpass, R. S. Use of facial composite systems in US law enforcement agencies. *Psych. Crime & Law*, **12**, 505 – 517 (2006).
2. Horry, R., Memon, A., Wright, D. B., & Milne, R. Predictors of eyewitness identification decisions from video lineups in England: A field study. *Law. Hum. Behav.* **36**, 257–265 (2012).
3. Penrod, S. D., & Kovera, M. Recent developments in North American identification science and practice. In: *Handbook of Psychology of Investigative Interviewing*. Bull, R., Valentine, T., & Williamson, T., eds., Chichester: Wiley-Blackwell, 257-283 (2009).
4. Garrett, B. L. *Convicting the Innocent: Where Criminal Prosecutions Go Wrong*. Cambridge, MA: Harvard University Press (2011).
5. Gross, S. R., Jacoby, K., Matheson, D. J., Montgomery, N., & Patil, S. Exonerations in the United States 1989 through 2003. *J. Crim. Law & Crim.* **95**, 523–560 (2005).
6. Gross, S. R., & Shaffer, M. Exonerations in the United States. *University of Michigan Public Law Working Paper*, No. **277**. Retrieved from papers.ssrn.com, 1989-2012 (2013).
7. Davies, G. M., Ellis, H., & Shepherd, J. Face identification: the influence of delay on accuracy of Photofit construction. *J. Police Sci. Adm.* **6**, 35–42 (1978).
8. Davis, J. P., Thorniley, S., Gibson, S., & Solomon, C. Holistic facial composite construction and subsequent lineup identification accuracy: Comparing adults and children. *J. Psych.* (2015).

9. Kempen, K., & Tredoux, C. G. 'Seeing is believing': The effect of viewing and constructing a composite on identification performance. *South African J. Psych.* **42**, 434–444 (2012).
10. Wells, G. L., Charman, S. D., & Olson, E. A. Building face composites can harm lineup identification performance. *J. of Exp. Psych: App.* **11**, 147–156, doi.org/10.1037/1076-898X.11.3.147 (2005).
11. Davis, J.P., Gibson, S., & Solomon, C. The positive influence of creating a holistic facial composite on video line-up identification. *App. Cog. Psych.* **28**, 634–639 (2014).
12. Meissner, C.A., & Brigham, J.C. A meta-analysis of the verbal overshadowing effect in face identification. *App. Cog. Psych.* **15**, 603–616 (2001).
13. Frowd, C., Bruce, V., Ness, H., Thomson-Bogor, C., Paterson, J., McIntyre, A., & Hancock, P. J. B. Parallel approaches to composite production. *Ergonomics.* **50**, 562-585 (2007).
14. Solomon, C. J., Gibson, S. J., & Maylin, M. EFIT-V: Evolutionary algorithms and computer composites. In: *Craniofacial Identification.*, Wilkinson, C., & Rynn, C., eds., Cambridge University Press, 24-41 (2012).
15. Tredoux, C. G., Nunez, D. T., Oxtoby, O., & Prag, B. An evaluation of ID: an eigenface based construction system. *South African Comp. J.* **37**, 90-96 (2007).
16. Davies, G. M., & Valentine, T. Facial composites: forensic utility and psychological research. In: *Handbook of Eyewitness Psychology, Volume 2., Memory for people*, Lindsay, R. C. L., Ross, D. F., Read, J. D., & Toglia, M. P. eds., Mahwah: LEA, 59-83 (2007).
17. Frowd, C. Facial composites and techniques to improve recognisability. In: *Forensic Facial Identification: Theory and Practice of Identification from Eyewitnesses, Composites and CCTV.* Valentine, T., & Davis, J. P., eds., Chichester: Wiley-Blackwell, 43-70 (2015).
18. Tanaka, J. W., & Farah, M. J. Parts and wholes in face recognition. *Qty J. of Exp. Psych.* **46A**, 225–245 (1993).
19. Tanaka, J. W., & Sengco, J. A. Features and their configuration in face recognition. *Mem. Cogn.* **25**, 583–592, (1997).
20. de Heering, A., Houthuys, S., & Rossion, B. Holistic face processing is mature at 4- years of age: Evidence from the composite face effect. *J. of Exp. Child Psych.* **96**, 57–70 (2007).
21. Schooler, J. W. Verbalization produces a transfer inappropriate processing shift. *App. Cog. Psych.* **16**, 989–997 (2002).
22. Wells, G. L. Verbal descriptions of faces from memory: Are they diagnostic of identification accuracy? *J. App. Psych.* **70**, 619–626 (1985).
23. Fisher, R. P., & Geiselman, R. E. *Memory: enhancing techniques for investigative interviewing: the cognitive interview.* Springfield, IL: Charles C. Thomas (1992).
24. Memon, A., Meissner, C. A., & Fraser, J. The Cognitive Interview: A meta-analytic review and study space analysis of the past 25 years. *Psy. Public Policy, Law.* **16**, 340–372, doi.org/10.1037/a0020518 (2010).
25. Davis, J. P., Sulley, L. Solomon, C., Gibson, S. A comparison of individual and morphed facial composites created using different systems. In: *2010 IEEE International Conference on Emerging Security Technologies.* Howells, G., Sirlantzis, K., Stoica, A., Huntsberger, T., & Arslan, A.T., eds., Canterbury: IEEE Computer Society, 56-60 (2010).
26. Frowd, C. D., Carson, D., Ness, H., Richardson, J., Morrison, L., Mclanaghan, S., Hancock, P. A forensically valid comparison of facial composite systems. *Psych. Crime. Law.* **11**, 33–52 (2005).
27. Frowd, C.D., et al. Giving crime the 'evo': catching criminals using EvoFIT facial composites. In: *2010 IEEE International Conference on Emerging Security Technologies.* Howells, G., Sirlantzis, K., Stoica, A., Huntsberger, T., & Arslan, A.T., eds., Canterbury: IEEE Computer Society, 36-43 (2010).
28. Solomon, C. J., Gibson, S. J., & Maylin, M. EFIT-V: Evolutionary algorithms and computer composites. In: *Craniofacial Identification.* Wilkinson, C., & Rynn, C., eds., Cambridge: Cambridge University Press, 24-41 (2012).
29. Gawrylowicz, J., Gabbert, F., Carson, D., Lindsay, W., & Hancock, P. Holistic versus featural facial composite systems for people with mild intellectual disabilities. *App. Cog. Psych.* **26**, 716–720 (2012).
30. Association of Chief Police Officers (ACPO). *Facial Identification Guidance*, Downloaded 3 January 2013 from <http://www.acpo.police.uk/documents/crime/2009/200911CRIFIG01.pdf> (2009).
31. Davis, J. P., Maitut, A. C., Jolliffe, D., Solomon, C. J., & Gibson, S. J. Holistic facial composite construction and subsequent video lineup identification outcomes: Comparing older and younger adults. *In preparation.* (2015).
32. Valentine, T., & Davis, J.P. Forensic facial identification: A practical best guide to best practice. In: *Forensic Facial Identification: Theory and Practice of Identification from Eyewitnesses, Composites and CCTV.* Valentine, T., & Davis, J. P., eds., Chichester: Wiley-Blackwell (2015).
33. Lindsay, R. C. L., & Wells, G. L. Improving eyewitness identification from lineups: Simultaneous versus sequential lineup presentations. *J. App. Psych.* **70**, 556-564 (1985).
34. Steblay, N. K., Dysart, J. E., & Wells, G. L. Seventy-two tests of the sequential lineup superiority effect: A meta-analysis and policy discussion. *Psych. Public Policy, Law.* **17**, 99-139, doi.org/10.1037/a0021650 (2011).
35. Clark, S. E., Moreland, M. B., & Rush, R. A. Lineup Composition and Lineup Fairness. In: *Forensic Facial Identification: Theory and Practice of Identification from Eyewitnesses, Composites and CCTV.* Valentine, T., & Davis, J. P., eds., Chichester: Wiley-Blackwell, 129-157 (2015).
36. Gronlund, S., Wixted, J., & Mickes, L. Evaluating eyewitness identification procedures using ROC analysis. *Curr. Dir. in Psych. Sci.* **23**, 3–10, (2014).
37. Valentine, T, Darling, S., & Memon, A. Do strict rules and moving images increase the reliability of sequential identification procedures? *App. Cog. Psych.* **21**, 933-949 (2007).
38. Wilcock, R., & Kneller, W. A comparison of presentation methods of video identification parades. *App. Cog. Psych.* **25**, 835-40 (2011).
39. Tredoux, C. G. Statistical inference on measures of lineup fairness. *Law Hum. Behav.* **22**, 217–237 (1998).
40. Malpass, R. S, Tredoux, C., & McQuiston-Surret, D. Lineup construction and measuring lineup fairness. In: *Handbook of Eyewitness Psychology: Vol. 2.*, Lindsay, R., Ross, D., Read, D., & Toglia, M., eds., Mahwah: Lawrence Erlbaum, 155-178 (2007).
41. *Police and Criminal Evidence Act (1984) Codes of Practice, Code D.* Retrieved from: <http://www.homeoffice.gov.uk/publications/police/operational-policing/pace-codes/pace-code-d-2011> (2011).
42. Valentine, T., Davis, J. P., Thorner, K., Solomon, C., & Gibson, S. Evolving and combining facial composites: Between-witness and within-witness morphs compared. *J. Exp. Psych: App.* **16**(1), 72-86 (2010).
43. Field, A. *Discovering Statistics using IBM SPSS 4th Edition.* London, SAGE Publications Ltd. (2013).
44. Greathouse, S. M., & Kovera, M. B. Instruction bias and line-up presentation moderate the effects of administrator knowledge on eyewitness identification. *Law Hum. Behav.* **33**, 70-82 (2009).
45. Wells, G. *Eyewitness Identification: A System Handbook.* Toronto, Carswell (1988).
46. Clark, S. E. Costs and benefits of eyewitness identification reform: Psychological science and public policy. *Pers. Psych. Sci.* **7**, 238–259 (2012).

47. Brewer, N., Weber, N., & Semmler, C. Eyewitness identification. In: *Psychology and Law: An Empirical Perspective*. Brewer, N., & Williams, K.D., eds., New York, Guilford, 177-221 (2005).
48. Brewer, N., & Wells, G. L. The confidence-accuracy relationship in eyewitness identification: Effects of lineup instructions, functional size and target-absent base rates. *J. Exp. Psych. App.* **12**, 11–30 (2006).
49. Sporer, S. L., Penrod, S., Read, D., & Cutler, B. Choosing, confidence and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin*. **118**, 315–327, (1995).
50. Rozporządzenie Ministra Sprawiedliwości. *Pozycja 981*. Retrieved from: http://static1.money.pl/d/akty_prawne/pdf/DU/2003/104/DU20031040981.pdf (2003).
51. Valentine, T., Davis, J.P., Memon, A., & Roberts, A. Showups and their influence on a subsequent video lineup. *App. Cog. Psych.* **26**(1), 1-23. (2012).
52. Mickes, L., Moreland, M.B., Clark, S.E., Wixted, J.T. Missing the information needed to perform ROC analysis? Then compute d' , not the diagnosticity ratio. *J App Res Mem and Cog.* **3**(2), 58–62. (2014).
53. Wilcock, R., Bull, R., & Milne, R. *Witness Identification in Criminal Cases: Psychology and Practice*. Oxford: Oxford University Press (2008).