











The impact of weapons and unusual objects on the construction of facial composites

William Blake Erickson ^a, Charity Brown ^b, Emma Portch ^c, James Michael Lampinen ^d, John E. Marsh ^e, Cristina Fodarella ^e, Anna Petkovic ^e, Carly Coultas^e, Amanda Newby^e, Louisa Date ^e, Peter J. B. Hancock ^f and Charlie D. Frowd ^e

^aDepartment of Life Sciences, Texas A&M University, San Antonio, TX, USA; ^bSchool of Psychology, University of Leeds, Leeds, UK; ^cDepartment of Psychology, Bournemouth University, Bournemouth, UK; ^dDepartment of Psychological Science, University of Arkansas, Fayetteville, AR, USA; ^eSchool of Psychology and Computer Science, University of Central Lancashire, Preston, UK; ^fPsychology, Faculty of Natural Sciences, University of Stirling, Stirling, UK

ABSTRACT

The presence of a weapon in the perpetration of a crime can impede an observer's ability to describe and/or recognise the person responsible. In the current experiment, we explore whether weapons when present at encoding of a target identity interfere with the construction of a facial composite. Participants encoded an unfamiliar target face seen either on its own or paired with a knife. Encoding duration (10 or 30 s) was also manipulated. The following day, participants recalled the face and constructed a composite of it using a holistic system (EvoFIT). Correct naming of the participants' composites was found to reduce reliably when target faces were paired with the weapon at 10 s but not at 30 s. These data suggest that the presence of a weapon reduces the effectiveness of facial composites following a short encoding duration. Implications for theory and police practice are discussed.

ARTICLE HISTORY

Received 21 September 2020
Accepted 10 May 2022

KEYWORDS

Facial composite; weapon; EvoFIT; law enforcement

The presence of a weapon during the commission of a crime can lead to a weapon-focus effect (WFE). This phenomenon is signalled by a marked decrease in eyewitness memory performance for crime-related details when a weapon is present at the time of encoding (see Fawcett et al., 2013 for review and meta-analysis). Laboratory testing has shown this deficit in performance results in reduced recall for details of the event, aspects of the crime scene and the perpetrator's clothing, general appearance, and face (e.g. Fawcett et al., 2013; Loftus et al., 1987; Maass & Köhnken, 1989). The most widely documented effect of weapon presence is on identification accuracy: performance in a line-up or identity parade reduces for a participant observer exposed to a weapon (cf. no weapon seen; e.g. Erickson et al., 2014; Fawcett et al., 2013; Loftus et al., 1987). In contrast, surveys of real

CONTACT Charlie D. Frowd  cfrowd1@uclan.ac.uk  School of Psychology and Computer Science, University of Central Lancashire, Preston PR1 2HE, UK

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

cases have demonstrated a weaker effect of weapon presence compared to other factors such as perpetrator race and duration of encoding (e.g. Tollestrup et al., 1994; Valentine et al., 2003; Yuille & Cutshall, 1986).

Theoretical mechanisms behind weapon focus

Witnessing or falling victim to a crime are stressful experiences. Some early research into eyewitness memory has its theoretical foundations in the effect of various levels of arousal on performance. Some stress is required for optimal performance at any given task, with lethargy and overwhelming perturbation each yielding worse performance (e.g. Yerkes & Dodson, 1908). Deffenbacher et al. (2004) applied Fazy and Hardy's (1988) catastrophe model of stress directly to eyewitness identification research, finding that increases in somatic anxiety result in gradual increases in performance up until the system reaches breaking point. After this point, increases in anxiety result in an abrupt and discontinuous drop in performance. In the brain, stress-related glucocorticoid levels modulate memory performance in a way that mirrors these patterns (Lupien et al., 2007). Perceptually, the arousal explanation is also supported by the cue-utilisation hypothesis (Easterbrook, 1959), proposing that attention allocates to central rather than peripheral details of a visual scene under heightened arousal. Weapons brandished during a crime are likely to increase arousal and stress, narrowing attention to themselves as central features of the scene. Paying more attention to a weapon means that less attention is allocated to a perpetrator's face, which would lie in the periphery, and so may not be encoded as effectively as when a weapon is absent.

It can be difficult to conclude what may be 'central' and 'peripheral' details in a real case given the variability in stress responses from person-to-person and from event-to-event (Christianson, 1992). Nonetheless, memory is less accurate for violent events than for nonviolent variants of the same events (Loftus & Burns, 1982), events wherein weapons are highly visible yield less accurate memory than events wherein weapons are less visible (Kramer et al., 1990), and, of most relevance, eyewitnesses are more likely to recall descriptions for and surrounding the weapon (e.g. the hands of the perpetrator) than eyewitnesses who do not see individuals holding threatening objects (Maass & Köhnken, 1989). All of these findings support the notion that the effect of weapons on attention may be mediated by a mechanism involving physiological arousal.

Another account of the underlying mechanism of the weapon-focus effect contends that weapons compete for visual attention since they are unusual stimuli in most contexts (e.g. Hope & Wright, 2007; Pickel, 1998; Pickel et al., 2008). This novelty-type explanation is based on early perceptual experiments wherein unusual features of a visual scene, such as *an octopus in a barnyard*, attract attention more quickly after onset and for longer intervals than contextually-relevant aspects, such as *a bucket in a barnyard* (Loftus & Mackworth, 1978). In theoretical terms, unusual objects are likely to disrupt voluntary control of attention and increase exogenous (i.e. stimulus-driven) control that may be difficult for a witness to overcome (Yantis, 2000). In the developed world, weapons are not common objects in most people's lives and so, when seen during the course of a crime, attract and hold attention at the expense of encoding a perpetrator's face and/or details of the crime.

Eyewitness experiments manipulating the WFE have tested this possibility by examining the context in which a weapon appears (e.g. a shooting range vs. a baseball park;

Pickel, 1999) and the unusualness of handheld objects themselves (e.g. perpetrators holding a non-threatening object such as a rubber chicken, or a gun; Erickson et al., 2014). Such research reveals that weapons in contextually-appropriate environments do not elicit a WFE, whereas unusual objects held during crimes do. Carlson, Pleasant et al. (2016) extended the paradigm by presenting participants with scenarios in which normally non-threatening objects were used as weapons. Contrary to the unusualness hypothesis, these items did not reduce line-up accuracy to the level observed with traditional weapons. It remains likely that the degree to which an object is threatening *and* unusual both impact on facial memory, albeit in differing ways (Carlson & Carlson, 2012).

Weapon focus and the search for suspects

Whether through elevating arousal or distracting attention with novelty, weapons diminish eyewitness memory. What does not appear to be in dispute is that the WFE is greater for recall than for facial identification (Fawcett et al., 2013; Kocab & Sporer, 2016). This result presents a forensically relevant problem that we investigate in the current study. Eliciting a description of a criminal event and its perpetrator during an interview relies upon verbal recall. Investigators rely on these accurate descriptions when searching for suspects (e.g. Technical Working Group for Eyewitness Evidence, 1999; Wells et al., 2020), and therefore a weapon's greatest effect in real cases may come before the police have even located a potential suspect for a witness to identify. This observation is particularly relevant to the current work as witnesses engage in both face recall and face recognition processes for the construction of facial composites.

Facial-composite images

When a perpetrator of a crime is not known, the police may ask witnesses to construct a facial composite of the perpetrator's face, the result of which is an image used as a reference for the police and as a tool to generate investigative leads from both the police and the public. Such images have historically been created by artists specially trained to generate facial images based on eyewitness reports (e.g. Davies & Little, 1990; Frowd, Carson, Ness, Richardson, et al., 2005; Lampinen et al., 2012; Laughery & Fowler, 1980). More standardised methods have used a feature-based 'kit' where witnesses work with a trained technician to assemble a composite from individual prefabricated features, commonly deployed now via a computer interface (e.g. E-FIT; FACES; Identikit 2000; PRO-fit) (e.g. Davies & Christie, 1982).

More recently, a radically different method has emerged for creating facial composites. The basic approach is that witnesses repeatedly select whole faces (or whole-face regions) from arrays of alternatives, with an evolutionary algorithm combining witness choices, to allow a composite to be 'evolved' (e.g. Frowd et al., 2004; George et al., 2008; Tredoux et al., 2006). Also, global aspects of an evolved identity can be changed, such as by altering face weight, age and health (e.g. Frowd et al., 2010). As such, faces are now created in line with holistic face perception – that is, perceiving a face as a whole entity – as opposed to the previous, non-natural approach of constructing the face by recalling and selecting individual features (for reviews see Frowd, 2011, 2017, 2021; Frowd, Skelton et al., 2012).

There are a number of holistic systems, such as EFIT-6 (previously EFIT-V), EvoFIT and ID, all of which have similar face construction procedures.

The system selected for use in the present study, EvoFIT, has been the subject of intense research and development (for reviews, see Frowd, 2017, 2021) and now creates composites that are more identifiable than those made by previous systems, both in the laboratory (e.g. Frowd, Bruce, Ness, et al., 2007; Frowd et al., 2010, 2013) and in tests of actual police investigations (e.g. Frowd et al., 2011; Frowd, Pitchford, et al., 2012). Therefore, this method shows considerable promise for theoretical investigation and forensic application, including exploration of potential moderators that alter composite effectiveness.

Composite construction can be influenced by myriad factors that affect eyewitness memory. To date, research on holistic face systems has considered what Wells (1978) terms *system* variables, factors over which law enforcement has direct control when investigating crimes. Examples include style of interview (Brown et al., 2020; Frowd, Nelson, et al., 2012; Giannou et al., 2021), use of image-enhancement techniques (Brown et al., 2019; Davis et al., 2011, 2015; Frowd, Skelton, Atherton, Pitchford, et al., 2012; Valentine et al., 2010), and procedural methods (e.g. Frowd et al., 2015; Frowd, Bruce, Ness, et al., 2007; Frowd, Park, et al., 2008; Frowd et al., 2009). Theoretically, these techniques improve the effectiveness of central, *internal features* of the face, the region including eyes, nose and mouth that is important for familiar-face recognition (e.g. Ellis et al., 1979) – that is, for recognition of the composite (Frowd, 2021). However, limited research has examined the influence of Wells's *estimator* variables (Frowd, White et al., 2014; Richardson et al., 2020), aspects of a crime over which law enforcement has no control. Presence of a weapon and exposure duration are two widely documented estimator variables in eyewitness research, the focus of the present study.

As discussed, the presence of a weapon during a crime should exert the strongest effect on recall. A 'feature' approach to face construction requires a witness to recall and assemble facial features, and so should be strongly influenced by presence of a weapon – although, correct naming of these composites is usually low under forensically relevant conditions, in particular following a long retention interval (a day or more) from face encoding to construction (e.g. Frowd, Carson, Ness, McQuiston, et al., 2005; Frowd et al., 2010), making assessment difficult. However, a holistic system such as EvoFIT requires minimal description to create its first generation of faces – screens from which witnesses select faces to begin construction – and correct naming of its composites is usually good. In a sense, a witness recognises best likenesses, and recognition memory should be affected by weapon presence to a lesser degree than recall. However, research does also indicate that face recall is still involved to some extent, particularly later in the construction process (e.g. Brown et al., 2020; Fodarella et al., 2021; Frowd, Bruce, Ness, et al., 2007); here, the ability to make fine adjustments to an evolved face can improve identification, such as by moving the eyes closer together or making the mouth larger.

Although presence of a weapon has itself not been investigated for a forensically-relevant design in relation to composite likeness, Hancock et al. (2011) examined the degree of stress at encoding. Half of their face constructors played a psychological action thriller computer game that included a briefly presented target; the other half took a passive 'onlooker' stance to the game, viewing the same content and target face. EvoFIT

composites created by game players were correctly named half as often as those constructed by passive observers. The authors suggested that elevated state anxiety for the players at encoding had a large negative impact on accurate recall and recognition required for effective composite construction – although, players' divided attention may also have been involved, resulting in higher cognitive load, or (as suggested above) use of a less global, less-effective type of face processing (e.g. Frowd, Bruce, Ness, et al., 2007). The current research manipulates arousal in a more overt weapon-focus paradigm, as described below.

The current study

The current study manipulated two estimator variables that are known to affect eyewitness memory accuracy: weapon presence and exposure duration. Our aim was to determine whether these variables would impact identifiability of facial composites. Additionally, we modelled the forensic situation as far as possible, and so included target faces that were unfamiliar to the person constructing the composite; a long retention interval from target encoding to face construction; and a naming task involving participants who were familiar with the targets, our main DV for assessing the effectiveness of the resulting composites. Whatever may be the mechanism, weapon presence was expected to produce less recognisable composite images by drawing witnesses' attention away from the face. More specifically, this could be due to the weapon becoming a more central detail as per the cue-utilisation hypothesis or due to the unexpected nature of the weapon in an experimental context. In either case, we predicted that the depiction of internal features would suffer most under the weapon condition, reducing recognition of the composite, and that this effect would be magnified under a shorter encoding duration.

We present the research as a series of stages, each with its own method and participant samples, an approach that mirrors established practice for facial-composite construction and assessment (e.g. Frowd, Carson, Ness, McQuiston, et al., 2005). Stage 1 comprised stimulus encoding and facial-composite construction by participants who were unfamiliar with the target faces. Stage 2 comprised a naming task where participants familiar with the targets were presented with the composites constructed in Stage 1 and asked to identify them by name. Stage 3 comprised a perceptual-similarity rating task where veridical images of the target faces were presented alongside the corresponding composites constructed in Stage 1 and involved another sample of participants who were unfamiliar with the targets.

Stage 1: face construction

Method

Participants

Forty participants, 36 female, from University of Central Lancashire constructed the composites; their ages ranged from 18 to 62 ($M = 26.8$, $SD = 10.7$) years. Participants were recruited on the basis of being unfamiliar with the targets: football players competing at an international level for England or Wales. Staff participated voluntarily in response

to an advert, while students received course credit for participation. Participants were randomly allocated in groups of 10 to each level of the two between-subjects factors: Encoding duration and Presence of knife.

Design

Participants each constructed a single composite face, and so the design was between-subjects for Presence of knife (absent vs. present) and Encoding duration (10 s vs. 30 s). In an attempt to maximise the chance of observing a forensically relevant effect, should one exist, we opted for these target exposure durations as they are not only plausible in the real world, but also because they produced the strongest weapon-focus effect in the Fawcett et al. (2013) meta-analysis (based on their 'intermediate' duration category from 10 s to 1 min).

Materials

Target faces were selected to be *unfamiliar* to participants who constructed the composites but *familiar* to those who would later name them. This objective was achieved with use of target photographs of 10 different male footballers who play at international level for England or Wales. For face construction, participants were recruited on the basis of not following this sport, whereas football fans were sought for composite naming. A colour photograph for each player was located on the Internet; all were of good quality, with photographic subjects presented clean-shaven, facing the camera and without glasses. Although most eyewitness research investigating the weapon-focus effect has employed simulated crime videos or slide presentations to maximise ecological validity when measuring dependent variables directly relating to memory (i.e. recall and identification/recognition), it is more common in composite construction research to utilise still images at the initial encoding phase as composite construction research often examines system variables (e.g. interviewing, construction methodology). This suggestion is supported by research which indicates that static and moving images of targets at encoding produce overall equivalent results (Frowd et al., 2015). Static images have also been used to investigate system variables in eyewitness research (e.g. Wooten et al., 2020).

Target photographs were each printed in the centre of a size A4 page to dimensions of approximately 8 cm (wide) × 10 cm (high), for participants to view and then to construct in the weapon-absent condition of the experiment. For presentation in the weapon-present condition, these targets were presented on top of a photograph of a clip-point knife pointed outwards (i.e. towards the participant) to simulate a threatening pose (see example, Figure 1). The knife was presented in colour to dimensions of 8 cm (wide) × 5.6 cm (high). This stimulus format was chosen as it mirrors the usual presentation of target images at encoding, as mentioned above, but also because of the unique challenge of obtaining a suitable number of images that depicted a target identity holding a threatening object.

EvoFIT version 1.6 was used to construct the composite faces.

Procedure

Following approval of ethics from the study's institution, participants were tested individually. They were randomly assigned, with equal sampling, to the two between-subjects factors of Encoding duration (10 s vs. 30 s) and Presence of knife (absent vs. present).



Figure 1. Example of the type of stimuli used in the weapon-present conditions in the experiment; actual materials cannot be presented here due to copyright. The footballer (Adam Lallana) and knife were obtained from Wikimedia Commons (2021). In the weapon-absent conditions of the experiment, the same facial photograph was shown (to different participants) without the weapon being presented.

Participants were given the general briefing that they might be shown a picture of a threatening weapon and would be involved in a two-part experiment involving a computer-based task. According to assignment, they were shown one of 10 target photographs (without replacement) first to briefly look at, and report whether the face was familiar. On two occasions, the participant reported that the identity was familiar and so another face was selected randomly and shown to these participants to briefly look at (which was then reported to be unfamiliar). Familiarity checks complete, participants then viewed the face alone or paired with the knife, for either 10 or 30 s, according to condition assignment.

Participants returned to the laboratory 20–28 h later. They were informed that they would recall the target face and any accompanying object seen on the previous day, and then construct a composite using EvoFIT. The procedure for carrying out the face recall interview and EvoFIT face construction is fairly involved and is described in detail in Fodarella et al. (2015). In brief, using cognitive interviewing techniques, participants were invited to think back to the time when the target had been seen, visualise the face and recall it in as much detail as possible, in their own time, without guessing and without prompts from the experimenter.

The experimenter started EvoFIT and participants chose a database for creating the face, selecting an appropriate age range to match the previously-seen target. Next,

participants were shown arrays of internal features faces (revealing the region around the eyes, brows, nose and mouth) and asked to select items with the best resemblance to the target face, focussing on the region around the eyes. Participants were asked to select for *smooth faces*, *facial texture* and overall match. This procedure (for smooth, texture and overall match) was repeated for a second cycle through the system, with selected faces being combined to evolve a face. Participants were then invited to improve the likeness using 'holistic' tools that changed age, weight, attractiveness and seven other overall properties of the face. During this stage, participants were asked to focus on the face as a whole rather than on the upper half. Next, participants were asked to select the best matching external features (hair, forehead, ears and neck), with these outer features shown on the constructed internal features. Finally, participants were given the opportunity to alter the size and position of individual features using a 'shape' tool and enhance the face again, if necessary, using holistic tools (followed by any further changes made using the shape tool). When participants reported that the best likeness had been achieved, the face was saved to disk as the composite.

Face recall and face construction were self-paced procedures. Ten composites were produced in each of the four individual conditions of the experiment, a total of 40 composites. The procedure took around 45 min per participant, including debriefing.

Stage 2: composite face naming

Method

Participants

Participants were recruited on the basis of being familiar with footballers who play at international level for England or Wales. They comprised students at the University of Central Lancashire, for course credit, and volunteer visitors at a local library in Cumbria, UK. There were 40 participants in total, 2 female, and their age ranged from 22 to 69 ($M = 31.2$, $SD = 9.9$) years. An additional six people were recruited to replace participants who did not meet the *a priori* rule of correctly naming at least 80% of the target pictures, as described below, to give this sample. Ten participants were allocated to each level of the two between-subject factors: Encoding duration and Presence of knife. This sample size (along with a similar sample size at face construction) is known to be able to



Figure 2. Composites of footballer Adam Lallana. They were constructed by a different person after (a) 10 s encoding duration without knife present, (b) 10 s with knife, (c) 30 s without knife and (d) 30 s with knife.

detect at least a *medium* effect size (Frowd, 2021), one that can reveal a forensically-useful difference, should one exist.

Materials

Materials were the 40 greyscale facial composites created in Stage 1 and the associated 10 target colour photographs, printed on separate pieces of A4 paper to dimensions of approximately 8 cm (wide) × 10 cm (high). Example composites are presented in Figure 2.

Design

The design was between-subjects, with each participant asked to name a set of composites from one of the four individual conditions in the experiment (comprised of two factors: Encoding duration and Presence of knife). After naming composites, participants were asked to name the target photographs to check that they were suitably familiar with the relevant identities. We applied an *a priori* rule: to be able to correctly name the majority of composite images, participants should correctly name at least 80% of the target photographs. If participants correctly named seven or fewer target pictures, as occurred on six occasions in total, the subsequent participant was presented with the same set of composites to name (i.e. was assigned to the same experimental condition as the person who did not meet the *a priori* rule).

Procedure

Participants were tested individually and the task was self-paced. Participants were told that composites would depict footballers who play at international level for England and Wales. Participants were asked to try to name each composite, if possible, or give a 'don't know' response. They were randomly allocated with equal sampling to the four individual conditions comprising the two factors of the experiment. The 10 composites were presented sequentially for participants to name, followed by the 10 target photographs, and participants were also asked to name these faces. Each person received a different random order of presentation of composites and target pictures. The procedure took around 10 min to complete, including debriefing.

Results

Responses to facial composites and target pictures were scored for accuracy.¹ Responses were coded as correct and assigned a value of 1 when participants gave the correct identity, and were coded as incorrect and assigned a value of 0 when a wrong name or a 'don't know' response was given. As participants were included who correctly named at least 80% of the target pictures, overall familiarity with the identities was very high ($M = 95.5\%$, $SD = 7.5\%$). For the few cases ($N = 18$) where a target picture was not correctly named (by 12 participants, 2–4 times by group), the associated composite also could not have been named correctly, and so responses to these composite items were removed prior to analysis.

Correct responses were much lower overall for facial composites than for target pictures ($M = 59.3\%$, $SD = 15.3\%$); this is the usual situation since composites are error-prone stimuli and therefore tend to not be named accurately. In more detail, 39 of the

Table 1. Correct naming of composites by duration of encoding and presence of knife.

Encoding (s)	Knife	
	Absent	Present
10	66.3 ^a (63/95)	47.4 ^{a,b} (46/97)
30	64.6 (62/96)	69.1 ^b (65/94)

Notes: Figures are expressed in percentage and calculated from participant responses in parentheses: summed correct responses (numerator) and total (correct and incorrect) responses (denominator). Data are presented for composites for which participants correctly named the relevant target photographs ($N = 382$ out of 400), leading to a overall mean of 61.8% correct. In the GLMM analysis of the significant interaction, there were two significant pairwise comparisons: ^a $p < .005$; the mean was greater (for 10 s encoding, knife absent) in six out of the 10 items, with two reversed and two ties. ^b $p < .001$; the mean was greater (for 30 s encoding, knife present) in eight out of the 10 items, two reversed.

40 composites were correctly named by at least one person; naming of these items (identities) spanned the entire naming scale, with six composites (all in 30 s encoding) named at 100%. As can be seen in Table 1, mean correct naming varied little by group except for encoding at 10 s in the presence of a knife, where, for these composites, correct naming was markedly lower (see also Table 1, Note).

Individual responses to composite items from participants were analysed using the regression technique, Generalized Linear Mixed Models (GLMM) in SPSS. This approach models IVs (*predictors* or *fixed effects*) in the context of *random effects*; the random effects are (i) participants in the naming stage of the experiment and (ii) composite items (stimuli). The experiment involved two fixed effects, Encoding (coded as 10 = 10 s; 30 = 30 s duration) and Weapon (coded as 0 = absent; 1 = present), both specified to have a descending sorting order (see Table 2, Note). The DV was individual participant responses, coded as above, with the model set to accommodate nominal responses using a binomial logistic link function.

Table 2. Model parameters for the interaction effect of weapon presence and encoding duration on correct composite naming.

Fixed Effects	<i>B</i>	<i>SE (B)</i>	<i>t</i> (378)	<i>p</i>	Exp(<i>B</i>)	95% CI(–)	95% CI(+)
Intercept	0.928	0.50	1.87	.06	2.53	0.95	6.71
Weapon							
<u>Present–Absent</u>							
<u>Encoding</u>							
10 s	–1.066	0.35	–3.06	.002	2.90†	1.46	5.76
30 s	0.220	0.35	0.63	.53	1.25	0.63	2.47
Encoding							
<u>30 s–10 s</u>							
<u>Weapon</u>							
Absent	–0.089	0.34	–0.26	.80	1.09†	0.56	2.15
Present	1.196	0.35	3.39	< .001	3.31	1.65	6.62

Notes: Comparisons are presented with reference to the lowest category (underlined); negative values of *B* indicate lower naming with respect to the reference. GLMM [IBM SPSS (Version 28) using the GENLINMIXED procedure (see Appendix)] final, Model-based full factorial Corrected model [$F(3, 378) = 4.86, p = .002$]. The model was specified with the lowest category of categorical predictors as reference (weapon absent; 10 s encoding), and target (DV) and predictors were sorted in a descending order; the sorting order of predictors was set to ascending to obtain all pairwise comparisons for the weapon \times encoding interaction. Information criteria are based on $-2 \log$ likelihood ($AICC = 1804.39, BIC = 1808.32$). Variance of random effects' intercept for items [$1.83, SE = 0.96, Z = 1.89, p = .06, CI (0.65, 5.14)$]. †For ease of interpretation, this effect size and its accompanying CIs are presented as values greater than unity (Osborne, 2017), calculated as $\text{Exp}(-B)$; based on Cohen's (1988) estimates, 1.5 can be considered as a 'small' effect size, 2.5 as 'medium' and 4.5 as 'large' (Sporer & Martschuk, 2014).

Estimation of parameters for GLMM involves linearization-based methods, also known as the pseudo likelihood approach, in SPSS (Garson, 2019; IBM, 2020), a standard iterative-fitting method. As the current sample has balanced data and is sufficiently large (by design), the residual method (e.g. cf. Satterthwaite approx.) was selected as degrees of freedom for computing tests of significance. Default settings for convergence criteria were used: parameter convergence with an absolute difference of 1E-6 and a maximum of 100 iterations for the algorithm's inner loop. For both fixed and random effects' models that were conducted, Beta (slope) coefficients (B), standard errors of B [$SE(B)$], effect sizes [$Exp(B)$] and confidence intervals (CI , all reported at 95%) were checked to be within sensible limits, neither too low nor too high, that might otherwise indicate an issue with the fit of the model.

The analysis initially assessed the composition of random effects. This assessment followed the 'gold' standard statistical procedure of Barr et al. (2013): for best generalisation, random effects should comprise the maximum number of terms, as justified by the structure of the data. For the current Independent Samples design, random effects involved random intercepts only (n.b., random slopes were not considered as none of the fixed effects were within-subjects). GLMM were duly conducted for each fixed effects' model, see below, each of which contained the best combination of random intercepts for participants and / or items, with each set of random intercepts included when there was sufficient variability in the response data to compute a non-zero estimate for this variable. For the final, full-factorial model, the variability remaining was not sufficient to compute random intercepts for *participants*, and so random effects included only random intercepts for *items* (see Table 1, Note).

The analysis also considered the most appropriate method to compute parameter estimates. There are two methods available in SPSS, a Model-based method and a Robust method that is sometimes preferable (e.g. when data are noisy). Both methods gave the same pattern of significant and non-significant differences, but Model-based was selected for presenting the results in the final model since the resulting standard errors for coefficients [$SE(B)$] of the interaction term were substantially lower (cf. Robust), thus providing a better fit of the data (also, Model-based is more often available (cf. Robust) in statistical software, and so this choice facilitates replication of results).

A hypothesis-testing (confirmatory) approach was conducted that initially comprised three models, each specified with different fixed effects (predictors) along with appropriate random effects (as described above). One model contained encoding only and a second contained weapon only. A third model contained the interaction between these two fixed effects; as it is standard practice to include individual predictors in a model that contains their interaction (e.g. Field, 2018), this third model was full factorial. Based on these results, a final model was then selected as (i) the full-factorial model (if the interaction was significant); otherwise (ii) a model containing both predictors (if both predictors were significant); or otherwise (iii) a model with a single significant predictor.

Based on the customary alpha of .1 for regression analyses, the first model was significant for Encoding [$F(1, 380) = 5.13, p = .024$], with better naming for the longer (cf. shorter) encoding; and the second for Weapon [$F(1, 380) = 3.20, p = .075$], with better naming when the knife was absent (cf. present). As the interaction between encoding and weapon was also significant in the third model [$F(1, 378) = 6.80, p = .010$], this GLMM was selected as the final model. It emerged without issues: (i)

there were only a few ($N = 7$ or 1.8% of total) cases (four in 10 s encoding, no weapon; and three in 30 s encoding, weapon present) where the standardised (Pearson) residuals (SR) fell (just) outside $2\ SD$ ($SR < -2.7$), and (ii) variances by group were within a reasonable spread both for residual errors (range of $\sigma^2 = 0.61$ – 1.06) and for predicted probability values (range of $\sigma^2 = 0.06$ – 0.10).

To explore the significant interaction, fixed coefficients were examined in the third (final) model (Table 2). This analysis indicated that presence (cf. absence) of a knife led to less-effective composites at 10 s encoding duration ($p = .002$), but a weapon-focus effect did not extend to the longer (30 s) encoding time ($p = .53$). It also revealed that 10 s encoding led to less-effective composites (cf. 30 s) when the weapon was present at encoding ($p < .001$), but there was no significant difference for encoding duration when the weapon was absent ($p = .80$).

We also assessed erroneous (mistaken) responses given by participants to composites, an assessment which emulates police investigations where names ('tips') turn out to be incorrect. Naming responses were re-scored (0 = no name given, 1 = mistaken name) and cases analysed (i) for which the target identity was known (as above) and (ii) for composites that had not been correctly named. GLMM conducted as above including both random intercepts were not significant for Presence of weapon [$p = .84$, $\text{Exp}(B) = 1.12$], Encoding duration [$p = .85$, $\text{Exp}(B) = 1.05$] and their interaction in a full model [$p = .89$, $\text{Exp}(B) = 1.23$]. Therefore, there is evidence that mistaken names were unaffected by the two predictors in the experiment.

Stage 3: composite likeness ratings

Method

Participants

The composites were rated for likeness by 20 volunteers, 13 female, with an age range from 18 to 74 ($M = 44.6$, $SD = 16.3$) years. Participants were sampled widely, those living in the Cumbria area, UK, recruited by word of mouth. They were recruited on the basis of *not* following football at an international level in the UK. This is because participants tend to rate composites harshly if they are familiar with the target identities (Frowd, 2021), reducing experimental power.

Materials

Materials were the composites constructed in Stage 1 and the 10 target photographs, all printed in the same way as for Stage 2.

Design and procedure

Participants rated the likeness of all 40 composites and so the design was within-subjects for the two factors of the experiment. Participants were tested individually and the task was self-paced. They were asked to give a rating of likeness (1 = very poor likeness ... 7 = very good likeness) for each composite in the presence of its relevant target photograph. All 40 composites were presented sequentially, and participants provided a rating as requested. The target photographs were then presented sequentially and participants asked to attempt to name the relevant footballers. To avoid familiarity effects,

we planned to exclude any participants who recognised three or more of the targets; however, no one met this criterion (and so all recruited participants were included in the sample). Each person received a different random order of presentation for composite and target pairs. The task was completed in about 20 min, including debriefing.

Results

The analyses followed the same coding scheme for predictors and procedure for composite naming, except that a multinomial probability distribution and a cumulative logit link function were used to model the ordinal-type participant data (composite likeness ratings). Participant ratings are summarised in Table 3 and indicate minimal differences by group except that mean rating was somewhat higher at 30 s (cf. 10 s) encoding duration.

GLMM conducted on individual participants' likeness ratings proceeded in the same way as for composite naming by including random intercepts for both participants and items. Estimates of variance for both of these random intercepts were able to be computed (cf. items only for naming), the result of which was a single combined model [$F(2, 792) = 4.78, p = .009$] that was significant for (i) Encoding duration [$SE(B) = 0.13, p = .010$], with higher likeness ratings for the longer (cf. shorter) encoding [$Exp(B) = 1.39$]; and (ii) Presence of weapon [$SE(B) = 0.13, p = .083$], with higher likeness ratings when a weapon was absent (cf. present) [$Exp(B) = 1.25$]; the full model was not significant for the interaction between encoding and weapon ($p = .11$).

Random slopes were then added to the random effects' model by including the same predictors that were specified as fixed effects. These models did not emerge significant ($p > .29$) for either fixed effect on its own or for the interaction in the full-factorial model. This result suggests that accounting for differences between each participant and between each item (identity), by including both random intercepts, produced a significant model of the data, but also including the within-subjects random effects – differences in the differences between means – by including both random slopes, did not provide a reliable model that was able to generalise fully to participants and items.

In Table 4, results are presented for the combined model that includes random intercepts for participants, and both random slopes; random intercepts for items was not included as no variance remained for items following inclusion of the other random effects. It is apparent that non-significant fixed effects emerged for likeness ratings as effect sizes [$Exp(B)$] were small for a study intended to be able to detect a larger, medium effect size. This outcome would appear to be due to the variable nature of random slopes, here leading to a three-fold increase in standard errors and a consequential reduction in statistical power of the tests.

Table 3. Likeness of composites by encoding duration and presence of knife.

Encoding (s)	Knife	
	Absent	Present
10	2.74 (0.09)	2.51 (0.09)
30	2.86 (0.10)	2.93 (0.13)

Notes: Rating scale (1 = very poor likeness ... 7 = very good likeness). Values are expressed using the mean (which gave a clearer pattern of results cf. median) and, in parentheses, by-participant SE of the mean.

Table 4. Model parameters for the effects of weapon presence and encoding duration on composite likeness ratings.

Fixed effects	<i>B</i>	<i>SE</i> (<i>B</i>)	<i>t</i> (792)	<i>p</i>	Exp(<i>B</i>)	95% CI(–)	95% CI(+)
Weapon							
Present–Absent	–0.254	0.37	–0.68	.50	1.29 ^a	0.62	2.69
Encoding							
30 s–10 s	0.409	0.36	1.14	.25	1.51	0.75	3.04

Notes: Comparisons are presented with reference to the lowest (underlined) category; negative values of *B* indicate lower ratings of likeness with respect to the reference. GEE [IBM SPSS (Version 28) using GENLINMIXED (see [Appendix](#))] in a Model-based Corrected model [$F(2, 792) = 0.89, p = .41$] with Threshold for rating values of *B*: [$1 = -1.76, 2 = 0.01, 3 = 1.47, 4 = 2.55, 5 = 4.06, 6 = 6.00$]. Sorting order of variables was ascending for target (DV) and descending for the two predictors. Variance of random effects for (i) participants' intercept [$1.43, SE = 0.68$], encoding [$0.50, SE = 0.22$] and weapon [$0.50, SE = 0.22$], and (ii) items' encoding [$0.30, SE = 0.15$] and weapon [$0.37, SE = 0.17$].

^aSee [Table 2, Notes](#).

Discussion

Understanding the role played by weapons on eyewitness memory is important for the apprehension and reliable conviction of criminals. This experiment examined the effects of weapon presence and facial encoding duration on the effectiveness of forensic composites made with EvoFIT, a holistic software system based on an evolutionary algorithm. Correct naming reduced significantly for composites constructed of target faces paired with a knife at 10 s encoding (cf. 10 s no knife). This means that a weapon-focus effect was evident, at least for the shorter encoding duration. This provides a novel extension of the weapon-focus effect from eyewitness recall and identification to downstream aspects of the search for criminal suspects. The effect of weapon presence and encoding duration on the overall appearance of the constructed composites, as assessed by ratings of likeness, was small and not reliable in a random effects' model that was most generalisable – that is, when including the best combination of random intercepts and random slopes.

To our knowledge, an encoding duration as short as 10 s has not featured in composite research that closely models the forensic situation and involves composite naming as a DV. With EvoFIT, as with the other holistic systems (e.g. EFIT-V and ID), constructors repeatedly select from face arrays to evolve a composite. This process is holistic in nature, with selection of whole faces or whole-face regions, and has been proposed (e.g. Frowd et al., 2004) to parallel natural holistic processing of upright, intact faces (e.g. Richler & Gauthier, 2014). We propose that such a brief encounter is likely to result in less detailed scrutiny of a target and encoding that is more global in nature, as evidenced here by only a small, albeit non-significant (cf. significant in a simpler model), reduction in ratings of likeness (from 30 to 10 s). This proposal is supported by research which reveals that methods designed to encourage global face processing, such as asking constructors to make personality judgments about a face after having described it, improve face construction for feature and holistic composite systems (e.g. Frowd, Bruce, Smith, et al., 2008; Frowd, Nelson, et al., 2012; Frowd et al., 2013, 2015; Skelton et al., 2019). Similarly, global procedures applied to finished composites also improve naming rates (e.g. Frowd, Skelton, Atherton, Pitchford, Hepton, et al., 2012; Frowd, Jones et al., 2014). Indeed, such procedures can be combined to give excellent results: in Frowd et al. (2013), with an encoding duration of 30 s and 24 h retention interval, mean naming for EvoFIT emerged at an astonishing 74% correct.

Although our experiment was not designed to directly test the theoretical explanations for the weapon-focus effect, our results do permit reasonable speculation. Presence of the knife at encoding reliably inhibited correct naming at 10 s encoding (cf. no weapon). It is worth mentioning that, in spite of a warning that a weapon may be seen, the experimenter observed that participants were noticeably shocked by the knife, and so it is possible that the decrement in correct naming ($MD = 18.9\%$, 10 s encoding) is in part related to the impact of stress and anxiety, as previously reported by Hancock et al. (2011). Therefore, our results fit more closely with the arousal hypothesis of the WFE because any unusualness or unexpectedness should have been abated by the (warning) instruction. This proposal is further supported by similar findings in eyewitness identification (e.g. Loftus & Burns, 1982; Maass & Köhnken, 1989). For composites, elevated stress at encoding also presumably interferes with identifiability (Davies, 2009; Hancock et al., 2011), and in real-world cases, composite constructors may be witnesses averse to the process due to reliving the trauma and feeling victimised by the individual whose face they are being asked to construct (Tredoux et al., 2021). However, note that some real-world victims experience extreme stress at encoding and yet produce a highly accurate likeness – such as the EvoFIT composite constructed of rapist Asim Javed as part of Operation Hatton (Frowd, 2017, 2021). We also note that, using a design similar to the one presented here, previous attempts at inducing a WFE with composites have been unsuccessful (Frowd, 2014), an outcome likely to be the result of the 30 s encoding duration that has been used (same as the null result found here). Finally, although the accuracy of external and internal facial features varies from composite to composite under the various experimental conditions, the lack of a reliable effect on perceptual similarity (likeness ratings) between weapon and no weapon conditions implicates no specific effect on either of these facial regions. Instead, it is likely that more holistic or structural codes required for familiar-face recognition (Bruce & Young, 1986; Young & Bruce, 2011) were disrupted by the presence of weapon. Further research could usefully establish the general effect of stress and anxiety on facial-composite production through measured physiology and follow-up surveys, as well as including the presence of a novel / unusual object.

The implication of the research is that a weapon is *unlikely* to have a measurable effect on the correct naming of composites when encoding duration is at least 30 s. For a much shorter duration, 10 s here, a negative effect of weapons would be expected, with a medium effect size [$\text{Exp}(B) = 2.9$]. The nature of an intermediate duration is unknown, but presumably somewhere over this range, the impact of weapons becomes negligible (and could be the focus of future research). This suggestion should be taken in the context of a real crime, for which one would expect greater attention to a weapon compared to a person seeing a weapon in a photograph (and so shorter delays may be more impactful to composite accuracy). Carlson, Young et al. (2016) previously found an effect of shorter exposure duration (3 s vs. 10 s) on eyewitness recall, but not identification or confidence-accuracy calibration, in a weapon-focus paradigm. Given that recall is important for composite construction (e.g. Frowd, Bruce, Ness, et al., 2007), we consider the encoding duration effect a replication and extension of this finding. More generally, given a similar procedure to construct faces for other holistic systems, an effect of weapons would also be expected. A similar argument could be made for traditional feature systems (e.g. E-FIT, FACES, Identikit 2000, PRO-fit). However, these types are believed to be based more on recall than

recognition (e.g. Frowd et al., 2005) and so, as weapons and objects influence recall to a greater extent than recognition (e.g. Loftus et al., 1987; Steblay, 1992), the effect could be stronger. We note though that, as feature-based composite accuracy is usually low following a long retention interval (e.g. Frowd, Carson, Ness, McQuiston, et al., 2005; Frowd et al., 2010, 2015), this proposal may be difficult to verify experimentally given that correct naming is likely to be suppressed further toward floor-level performance.

Forensic composite images are commissioned to help produce tips from police officers and members of the public so that an investigation may generate or eliminate suspects as well as corroborate victim and witness testimony. In some countries (e.g. Australia, UK, US, South Africa), composites may also be used as evidence in court. The experiment reported here modelled both aspects of composites as an investigative tool, by incorporating a naming task (e.g. when police and members of the public familiar with a fugitive may recognise that person) and a likeness rating task (e.g. when law enforcement officers may compare composites to suspects they investigate). Since the results reported here reflect what would appear to be the first formal assessment examining the combined contributions of weapon presence and encoding duration on composite naming accuracy, we make some tentative practical conclusions.

First, experts making recommendations based on our results should consider them in a way similar to how all estimator variable research are considered: namely, the aggregate reduction in composite accuracy due to the presence of weapon may produce false tips and therefore jeopardise innocent suspects and in turn lead to a failure to apprehend the actual guilty suspects who can go on to reoffend. That is not to say that police should not commission composite images for very briefly occurring crimes involving weapons, but that care should be taken when following up leads in such cases. It is worth mentioning, though, that mistaken names did not vary reliably in the experiment, and so false leads would appear not to be influenced (at least with a medium effect) by the presence of weapon or encoding duration. For the forensically-important measure of correct naming, though, composites constructed from 10 s encoding in the presence of a weapon still had very good utility for correctly identifying the target – about 50% correct naming. Also, although correct naming dropped by about 20% from 30 s to 10 s encoding duration in the presence of a weapon, naming of composites in the weapon-present condition was still nearly 60% correct overall, similar to mean naming of 56% reported in a recent meta-analysis of the system (Frowd et al., 2015). It has long been known that convictions should not be based on eyewitness evidence alone (e.g. Devlin, 1976), and so other evidence is important to support a reliable conviction later in an investigation (e.g. Greene & Loftus, 1984; Osborne & Davies, 2013), irrespective of the method used to generate a suspect in the first place.

Second, further research is needed to explore the various effects of weapons, unusual objects, arousal, and myriad other estimator variables on composite accuracy, as well as means of mitigating these effects at the interviewing and face construction stages. For example, given our current evidence (via likeness ratings) that a weapon's adverse effects on naming have a holistic mechanism, a 'holistic' tool designed to counteract this effect could conceivably be developed, as has been done for other global aspects of the face in holistic systems with a view to facilitate naming (Frowd et al., 2010). Also, composite research must overcome sampling and design limitations consequent to the

necessities of setting up these sorts of experiments, as previously recommended by Wells et al. (2005), to overcome failures to find small but potentially consequential effects. While the current design had good experimental power to be able to detect at least a *medium* effect size, to be of practical significance, researchers might like to consider avoiding the use of celebrity or otherwise well-known faces as targets for constructors, and then experimentally induce familiarity among participants tasked with naming composites, or (to achieve natural familiarity) consider a cross-site design (e.g. Frowd, Bruce, McIntyre, et al., 2007). These measures would allow researchers to use larger samples of constructors in accordance with recommendations for heavily powered experimental designs as well as allow the production of specially tailored stimulus materials including more naturalistic features like those seen in real crimes.

Note

1. These composite naming data as well as ratings of likeness are available from the corresponding author.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

William Blake Erickson  <http://orcid.org/0000-0002-2765-3699>

Charity Brown  <http://orcid.org/0000-0001-9697-4878>

Emma Portch  <http://orcid.org/0000-0002-9008-4142>

James M. Lampinen  <http://orcid.org/0000-0002-5854-521X>

John E. Marsh  <http://orcid.org/0000-0002-9494-1287>

Cristina Fodarella  <http://orcid.org/0000-0001-5551-3450>

Anna Petkovic  <http://orcid.org/0000-0002-1112-3482>

Louisa Date  <http://orcid.org/0000-0003-2290-4193>

Peter J. B. Hancock  <http://orcid.org/0000-0001-6025-7068>

Charlie D. Frowd  <http://orcid.org/0000-0002-5082-1259>

References

- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Brown, C., Portch, E., Nelson, L., & Frowd, C. D. (2020). Reevaluating the role of verbalization of faces for composite production: Descriptions of offenders matter!. *Journal of Experimental Psychology: Applied*, 26(2), 248–265. <https://doi.org/10.1037/xap0000251>
- Brown, C., Portch, E., Skelton, F. C., Fodarella, C., Kuivaniemi-Smith, H., Herold, K., Hancock, P. J. B., & Frowd, C. D. (2019). The impact of external facial features on the construction of facial composites. *Ergonomics*, 62, 575–592.
- Bruce, V., & Young, A. W. (1986). Understanding face recognition. *British Journal of Psychology*, 77, 305–327.
- Carlson, C. A., & Carlson, M. A. (2012). A distinctiveness-driven reversal of the weapon-focus effect. *Applied Psychology in Criminal Justice*, 8, 36–53.

- Carlson, C. A., Pleasant, W. E., Weatherford, D. R., Carlson, M. A., & Bednarz, J. E. (2016). The weapon focus effect: Testing an extension of the unusualness hypothesis [electronic version]. *Applied Psychology in Criminal Justice*, 12, 87–100.
- Carlson, C. A., Young, D. F., Weatherford, D. R., Carlson, M. A., Bednarz, J. E., & Jones, A. R. (2016). The influence of perpetrator exposure time and weapon presence/timing on eyewitness confidence and accuracy. *Applied Cognitive Psychology*, 30(6), 898–910. <https://doi.org/10.1002/acp.3275>
- Christianson, S. Å. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin*, 112(2), 284–309. <https://doi.org/10.1037/0033-2909.112.2.284>
- Davies, G. M., & Christie, D. (1982). Face recall: An examination of some factors limiting composite production accuracy. *Journal of Applied Psychology*, 67(1), 103–109. <https://doi.org/10.1037/0021-9010.67.1.103>
- Davies, G. M., & Little, M. (1990). Drawing on memory: Exploring the expertise of a police artist. *Medical Science and the Law*, 30(4), 345–354. <https://doi.org/10.1177/106002809003000412>
- Davies, S. (2009). The effects of stress levels on the construction of facial composites. *Diffusion*, 2. <http://atp.uclan.ac.uk/buddypress/diffusion/?p=1763>
- Davis, J. P., Simmons, S., Sulley, L., Solomon, C., & Gibson, S. (2015). An evaluation of post-production facial composite enhancement techniques. *Journal of Forensic Practice*, 17(4), 307–318. <https://doi.org/10.1108/JFP-08-2015-0042>
- Davis, J. P., Sulley, L., Solomon, C., & Gibson, S. (2011). A comparison of individual and morphed facial composites created using different systems. In G. Howells, K. Sirlantzis, A. Stoica, T. Huntsberger, & A. T. Arslan (Eds.), *2010 IEEE international conference on emerging security technologies* (pp. 56–60). IEEE.
- Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior*, 28(6), 687–706. <https://doi.org/10.1007/s10979-004-0565-x>
- Devlin, P. (1976). *Report to the Secretary of State for the Home Department of the Departmental Committee on evidence of identification in criminal cases*. HMSO.
- Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, 66(3), 183–201. <https://doi.org/10.1037/h0047707>
- 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. <https://doi.org/10.1068/p080431>
- Erickson, W. B., Lampinen, J. M., & Leding, J. K. (2014). The weapon focus effect in target present and target absent lineups: The roles of threat, novelty and timing. *Applied Cognitive Psychology*, 28(3), 349–359. <https://doi.org/10.1002/acp.3005>
- Fawcett, J. M., Russell, E. J., Peace, K. A., & Christie, J. (2013). Of guns and geese: A meta-analytic review of the 'weapon focus' literature. *Psychology, Crime & Law*, 19(1), 35–66. <https://doi.org/10.1080/1068316X.2011.599325>
- Fazey, J. A., & Hardy, L. (1988). *The inverted-U hypothesis: A catastrophe for sport psychology* (British Association for sports Sciences Monograph No. 1). National Coaching Foundation.
- Field, A. (2018). *Discovering statistics using SPSS* (5th ed.). Sage.
- Fodarella, C., Kuivaniemi-Smith, H., Gawrylowicz, J., & Frowd, C. D. (2015). Forensic procedures for facial-composite construction. *Journal of Forensic Practice*, 17(4), 259–270. <https://doi.org/10.1108/JFP-10-2014-0033>
- Fodarella, C., Marsh, J. E., Chu, S., Athwal-Kooner, P., Jones, H., Skelton, F. C., Wood, E., Jackson, E., & Frowd, C. D. (2021). The importance of detailed context reinstatement for the production of identifiable composite faces from memory. *Visual Cognition*, 29(3), 180–200. <https://doi.org/10.1080/13506285.2021.1890292>
- Frowd, C. D. (2011). Eyewitnesses and the use and application of cognitive theory. In G. Davey (Ed.), *Introduction to applied psychology* (pp. 267–289). BPS Wiley-Blackwell.
- Frowd, C. D. (2014, April 30). The impact of weapons and unusual objects on face recall and composite construction. *Symposium: Theoretical and Applied Aspects of Distraction*, University of Central Lancashire.

- Frowd, C. D. (2017). Facial composite systems: Production of an identifiable face. In M. Bindemann & A. Megreya (Eds.), *Face processing: Systems, disorders and cultural differences* (pp. 55–86). Nova Science.
- Frowd, C. D. (2021). Forensic facial composites. In M. Tolia, A. Smith, & J. M. Lampinen (Eds.), *Methods, measures, and theories in forensic facial-recognition* (pp. 34–64). Taylor and Francis.
- Frowd, C. D., Bruce, V., McIntyre, A., & Hancock, P. J. B. (2007). The relative importance of external and internal features of facial composites. *British Journal of Psychology*, 98(1), 61–77. <https://doi.org/10.1348/000712606X104481>
- Frowd, C. D., Bruce, V., Ness, H., Bowie, L., Thomson-Bogner, C., Paterson, J., McIntyre, A., & Hancock, P. J. B. (2007). Parallel approaches to composite production. *Ergonomics*, 50(4), 562–585. <https://doi.org/10.1080/00140130601154855>
- Frowd, C. D., Bruce, V., Pitchford, M., Gannon, C., Robinson, M., Tredoux, C., Park, J., McIntyre, A., & Hancock, P. J. B. (2009). Evolving the face of a criminal: How to search a face space. *Soft Computing*, 14(1), 81–90. <https://doi.org/10.1007/s00500-008-0391-z>
- Frowd, C. D., Bruce, V., Smith, A., & Hancock, P. J. B. (2008). Improving the quality of facial composites using a holistic cognitive interview. *Journal of Experimental Psychology: Applied*, 14(3), 276–287. <https://doi.org/10.1037/1076-898X.14.3.276>
- Frowd, C. D., Carson, D., Ness, H., McQuiston, D., Richardson, J., Baldwin, H., & Hancock, P. J. B. (2005). Contemporary composite techniques: The impact of a forensically-relevant target delay. *Legal & Criminological Psychology*, 10(1), 63–81. <https://doi.org/10.1348/135532504X15358>
- Frowd, C. D., Carson, D., Ness, H., Richardson, J., Morrison, L., McLanaghan, S., & Hancock, P. J. B. (2005). A forensically valid comparison of facial composite systems. *Psychology, Crime & Law*, 11(1), 33–52. <https://doi.org/10.1080/10683160310001634313>
- Frowd, C. D., Erickson, W. B., Lampinen, J. L., Skelton, F. C., McIntyre, A. H., & Hancock, P. J. B. (2015). A decade of evolving composite techniques: Regression- and meta-analysis. *Journal of Forensic Practice*, 17(4), 319–334. <https://doi.org/10.1108/JFP-08-2014-0025>
- Frowd, C. D., Hancock, P. J., Bruce, V., Skelton, F. C., Atherton, C., & Nelson, L. (2011). Catching more offenders with EvoFIT facial composites: Lab research and police field trials. *Global Journal of Human-Social Science Research*, 11, 35–46.
- Frowd, C. D., Hancock, P. J., & Carson, D. (2004). EvoFIT: A holistic, evolutionary facial imaging technique for creating composites. *ACM Transactions on Applied Perception (TAP)*, 1(1), 19–39. <https://doi.org/10.1145/1008722.1008725>
- Frowd, C. D., Jones, S., Fodarella, C., Skelton, F. C., Fields, S., Williams, A., Marsh, J., Thorley, R., Nelson, L., Greenwood, L., Date, L., Kearley, K., McIntyre, A., & Hancock, P. J. B. (2014). Configural and featural information in facial-composite images. *Science & Justice*, 54(3), 215–227. <https://doi.org/10.1016/j.scijus.2013.11.001>
- Frowd, C. D., Nelson, L., Skelton, F. C., Noyce, R., Atkins, R., Heard, P., Morgan, D., Fields, S., Henry, J., McIntyre, A., & Hancock, P. J. (2012). Interviewing techniques for Darwinian facial-composite systems. *Applied Cognitive Psychology*, 26(4), 576–584. <https://doi.org/10.1002/acp.2829>
- Frowd, C. D., Park, J., McIntyre, A., Bruce, V., Pitchford, M., Fields, S., Kenirons, M., & Hancock, P. J. B. (2008). Effecting an improvement to the fitness function. How to evolve a more identifiable face. In A. Stoica, T. Arslan, D. Howard, T. Higuchi, & A. El-Rayis (Eds.), *2008 ECSIS symposium on bio-inspired, learning, and intelligent systems for security* (pp. 3–10). CPS.
- Frowd, C. D., Pitchford, M., Bruce, V., Jackson, S., Hepton, G., Greenall, M., McIntyre, A., & Hancock, P. J. B. (2010). The psychology of face construction: Giving evolution a helping hand. *Applied Cognitive Psychology*, 25(2), 195–203. <https://doi.org/10.1002/acp.1662>
- Frowd, C. D., Pitchford, M., Skelton, F. C., Petkovic, A., Prosser, C., & Coates, B. (2012). Catching even more offenders with EvoFIT facial composites. In A. Stoica, D. Zarzhitsky, G. Howells, C. Frowd, K. McDonald-Maier, A. Erdogan, & T. Arslan (Eds.), *IEEE Proceedings of 2012 third international conference on emerging security technologies* (pp. 20–26). IEEE.
- Frowd, C. D., Skelton, F. C., Atherton, C., & Hancock, P. J. B. (2012). Evolving an identifiable face of a criminal. *The Psychologist*, 25, 116–119.

- 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(10), 1215–1241. <https://doi.org/10.1080/13506285.2012.743936>
- Frowd, C. D., Skelton, F. C., Atherton, C., Pitchford, M., Hepton, G., Holden, L., McIntyre, A., & Hancock, P. J. B. (2012). Recovering faces from memory: The distracting influence of external facial features. *Journal of Experimental Psychology: Applied*, 18(2), 224–238. <https://doi.org/10.1037/a0027393>
- Frowd, C. D., Skelton, F. C., Hepton, G., Holden, L., Minahil, S., Pitchford, M., McIntyre, A., Brown, C., & Hancock, P. J. B. (2013). Whole-face procedures for recovering facial images from memory. *Science & Justice*, 53(2), 89–97. <https://doi.org/10.1016/j.scijus.2012.12.004>
- Frowd, C. D., White, D., Kemp, R. I., Jenkins, R., Nawaz, K. (2014). Constructing faces from memory: the impact of image likeness and prototypical representations. *Journal of Forensic Practice*, 16, 243–256.
- Garson, G. D. (2019). *Hierarchical linear modeling: Guide and applications multilevel modeling in STATA®, IBM® SPSS®, SAS®, R, & HLM™*. SAGE Publications, Inc.
- George, B., Gibson, S. J., Maylin, M. I., & Solomon, C. J. (2008). EFIT-V: Interactive evolutionary strategy for the construction of photo-realistic facial composites. In *Proceedings of the 10th annual conference on genetic and evolutionary computation* (pp. 1485–1490). New York, NY: ACM. doi:10.1145/1389095.1389384
- Giannou, K., Frowd, C. D., Taylor, J. R., & Lander, K. (2021). Mindfulness in Face Recognition: Embedding mindfulness instructions in the face-composite construction process. *Applied Cognitive Psychology*, 35, 999–1010.
- Greene, E., & Loftus, E. F. (1984). Solving the eyewitness problem. *Behavioral Sciences & the Law*, 2(4), 395–406. <https://doi.org/10.1002/bsl.2370020406>
- Hancock, P. J. B., Burke, K., & Frowd, C. D. (2011). Testing facial composite construction under witness stress. *International Journal of Bio-Science and Bio-Technology*, 3, 65–71.
- Hope, L., & Wright, D. (2007). Beyond unusual? Examining the role of attention in the weapon focus effect. *Applied Cognitive Psychology*, 21(7), 951–961. <https://doi.org/10.1002/acp.1307>
- IBM. (2020). *SPSS modeler algorithms guide*. IBM Corporation.
- Kocab, K., & Sporer, S. L. (2016). The weapon focus effect for person identifications and descriptions: A meta-analysis. In M. K. Miller & B. H. Bornstein (Eds.), *Advances in psychology and law*, 1 (pp. 71–117). Springer International Publishing. doi.org/10.1007/978-3-319-29406-3_3
- Kramer, T. H., Buckhout, R., & Eugenio, P. (1990). Weapon focus, arousal, and eyewitness memory. *Law and Human Behavior*, 14(2), 167–184. <https://doi.org/10.1007/BF01062971>
- Lampinen, J. M., Neuschatz, J. S., & Cling, A. D. (2012). *The psychology of eyewitness identification*. NY Psychology Press.
- Laughery, K., & Fowler, R. (1980). Sketch artist and identikit procedures for generating facial images. *Journal of Applied Psychology*, 65(3), 307–316. <https://doi.org/10.1037/0021-9010.65.3.307>
- Loftus, E. F., & Burns, T. E. (1982). Mental shock can produce retrograde amnesia. *Memory & Cognition*, 10(4), 318–323. <https://doi.org/10.3758/BF03202423>
- Loftus, E. F., Loftus, G. R., & Messo, J. (1987). Some facts about weapon focus. *Law and Human Behavior*, 11(1), 55–62. <https://doi.org/10.1007/BF01044839>
- Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance*, 4(4), 565–572. <https://doi.org/10.1037/0096-1523.4.4.565>
- Lupien, S. J., Maheu, F., Tu, M., Fiocco, A., & Schramek, T. E. (2007). The effects of stress and stress hormones on human cognition. *Implications for the Field of Brain and Cognition*, 65(3), 209–237. <https://doi.org/10.1016/j.bandc.2007.02.007>
- Maass, A., & Köhnken, G. (1989). Eyewitness identification: Simulating the ‘weapon effect’. *Law and Human Behavior*, 13(4), 397–408. <https://doi.org/10.1007/BF01056411>
- Osborne, D., & Davies, P. G. (2013). Eyewitness identifications are affected by stereotypes about a suspect’s level of perceived stereotypicality. *Group Processes & Intergroup Relations*, 16(4), 488–504. <https://doi.org/10.1177/1368430212454927>

- Osborne, J. W. (2017). Ch. 5: Simple Linear models With categorical dependent variables: Binary logistic regression. In *Regression & Linear Modeling: Best practices and modern methods*. <https://doi.org/10.4135/9781071802724>
- Pickel, K. L. (1998). Unusualness and threat as possible causes of “weapon focus”. *Memory (Hove, England)*, 6(3), 277–295. <https://doi.org/10.1080/741942361>
- Pickel, K. L. (1999). The influence of context on the “weapon focus” effect. *Law and Human Behavior*, 23(3), 299–311. <https://doi.org/10.1023/A:1022356431375>
- Pickel, K. L., Narter, D. B., Jameson, M. M., & Lenhardt, T. T. (2008). The weapon focus effect in child eyewitnesses. *Psychology, Crime & Law*, 14(1), 61–72. <https://doi.org/10.1080/10683160701391307>
- Richardson, B., Brown, C., Heard, P., Pitchford, M., Portch, E., Lander, K., Marsh, J. E., Bell, R., Fodarella, C., Worthington, M., Ellison, L., Charters, P., Green, D., Minahil, S., & Frowd, C. D. (2020). The advantage of low and medium attractiveness for facial composite production from modern forensic systems. *Journal of Applied Memory and Cognition*, 9, 381–395.
- Richler, J. J., & Gauthier, I. (2014). A meta-analysis and review of holistic face processing. *Psychological Bulletin*, 140(5), 1281–1302. <https://doi.org/10.1037/a0037004>
- Skelton, F. C., Frowd, C. D., Hancock, P. J. B., Jones, H. S., Jones, B. C., Fodarella, C., Battersby, K., & Logan, K. (2019). Constructing identifiable composite faces: The importance of cognitive alignment of interview and construction procedure. *Journal of Experimental Psychology: Applied*, 26. <https://doi.org/10.1037/xap0000257>.
- Sporer, S. L., & Martschuk, N. (2014). The reliability of eyewitness identifications by the elderly: An evidence-based review. In M. P. Toglia, D. F. Ross, J. Pozzulo, & E. Pica (Eds.), *The elderly eyewitness in court* (pp. 3–37). Psychology Press.
- Stebay, N. M. (1992). A meta-analytic review of the weapon focus effect. *Law and Human Behavior*, 16, 413–424.
- Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness evidence: A guide for law enforcement [booklet]*. United States Department of Justice, Office of Justice Programs.
- Tollestrup, P. A., Turtle, J. W., & Yuille, J. C. (1994). Actual victims and witnesses to robbery and fraud: An archival analysis. In D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments* (pp. 144–160). Cambridge University Press.
- Tredoux, C., Nunez, D., Oxtoby, O., & Prag, B. (2006). An evaluation of ID: An eigenface based construction system. *South African Computer Journal*, 37, 90–97.
- Tredoux, C. G., Sporer, S. L., Vredeveltdt, A., Kempen, K., & Nortje, A. (2021). Does constructing a facial composite affect eyewitness memory? A research synthesis and meta-analysis. *Journal of Experimental Criminology*, 17(4), 713–741. <https://doi.org/10.1007/s11292-020-09432-z>
- Valentine, T., Davis, J. P., Thorner, K., Solomon, C., & Gibson, S. (2010). Evolving and combining facial composites: Between-witness and within-witness morphs compared. *Journal of Experimental Psychology: Applied*, 16(1), 72–86. <https://doi.org/10.1037/a0018801>
- Valentine, T., Pickering, A., & Darling, S. (2003). Characteristics of eyewitness identification that predict the outcome of real line-ups. *Applied Cognitive Psychology*, 17(8), 969–993. <https://doi.org/10.1002/acp.939>
- Wells, G. L. (1978). Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology*, 36(12), 1546–1557. <https://doi.org/10.1037/0022-3514.36.12.1546>
- Wells, G. L., Charman, S. D., & Olson, E. A. (2005). Building face composites can harm lineup identification performance. *Journal of Experimental Psychology: Applied*, 11(3), 147–156. <https://doi.org/10.1037/1076-898X.11.3.147>
- Wells, G. L., Kovera, M. B., Douglass, A. B., Brewer, N., Meissner, C. A., & Wixted, J. T. (2020). Policy and procedure recommendations for the collection and preservation of eyewitness identification evidence. *Law and Human Behavior*, 44(1), 3–36. <https://doi.org/10.1037/lhb0000359>
- Wikimedia Commons. (2021). Photograph of Adam Lallana, https://upload.wikimedia.org/wikipedia/commons/f/fe/Adam_Lallana_Southampton_FC_Manchester_United_2012%2C13.png. CC SA 3.0. Knife from .22 LR gun knife with revolving round magazine inside the handle, https://upload.wikimedia.org/wikipedia/commons/3/36/.22_LR_knife_gun_combat_ready.png

- Wooten, A. R., Carlson, C. A., Lockamy, R. F., Carlson, M. A., Jones, A. R., Dias, J. L., & Hemby, J. A. (2020). The number of fillers may not matter as long as they all match the description: The effect of simultaneous lineup size on eyewitness identification. *Applied Cognitive Psychology*, 34, 590–604.
- Yantis, S. (2000). Goal-directed and stimulus-driven determinants of attentional control. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII* (pp. 73–103). MIT Press.
- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, 18(5), 459–482. <https://doi.org/10.1002/cne.920180503>
- Young, A. W., & Bruce, V. (2011). Understanding person perception. *British Journal of Psychology*, 102(4), 959–974. <https://doi.org/10.1111/j.2044-8295.2011.02045.x>
- Yuille, J. C., & Cutshall, J. L. (1986). A case study of eyewitness memory of a crime. *Journal of Applied Psychology*, 71(2), 291–301. <https://doi.org/10.1037/0021-9010.71.2.291>

Appendix

SPSS GLMM Syntax for Correct Composite Naming: Full-factorial, Model-based Covariance Structure. The DV is *hits*, and the random effects are random intercepts for items (*itm*); fixed effects (predictors) are *encoding* and *weapon*. In practice, a GLMM model with random intercepts for participants (*ss*) was created initially using SPSS dialogs; this model was then pasted into the Syntax editor, to allow two RANDOM statements to be included, one for random intercepts for items (*itm*) and one for random intercepts for participants (*ss*). (Although, in this case, as described in the paper, random intercepts for participants (*ss*) were not included in the final model shown here.)

```
GENLINMIXED
/ DATA_STRUCTURE SUBJECTS=itm*ss
/ FIELDS TARGET=hits TRIALS=NONE OFFSET=NONE
/ TARGET_OPTIONS DISTRIBUTION=BINOMIAL LINK=LOGIT
/ FIXED EFFECTS=encoding*weapon encoding weapon USE_INTERCEPT=TRUE
/ RANDOM USE_INTERCEPT=TRUE SUBJECTS=itm COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
/ BUILD_OPTIONS TARGET_CATEGORY_ORDER=DESCENDING INPUTS_CATEGORY_ORDER=DESCENDING
MAX_ITERATIONS=100 CONFIDENCE_LEVEL=95 DF_METHOD=RESIDUAL COVB=MODEL PCONVERGE=0.000001
(ABSOLUTE) SCORING=0 SINGULAR=0.000000000001
/ EMMEANS_OPTIONS SCALE=ORIGINAL PADJUST=LSD.
```

SPSS GLMM Syntax for Composite Likeness Ratings for a combined model containing *encoding* and *weapon*: Model-based Covariance Structure. The DV is *rating*, and the random effects are random intercepts for participants (*ss*) and random slopes for both predictors; fixed effects (predictors) are *encoding* and *weapon*. The same procedure as described above was used to obtain two RANDOM statements, in this case including both intercepts and slopes, for both participants (*ss*) and items (*itm*). (Random intercepts for items (*itm*) were not included in this final model.)

```
GENLINMIXED
/ DATA_STRUCTURE SUBJECTS=itm*ss
/ FIELDS TARGET=rating TRIALS=NONE OFFSET=NONE
/ TARGET_OPTIONS DISTRIBUTION=MULTINOMIAL LINK=LOGIT
/ FIXED EFFECTS=encoding weapon USE_INTERCEPT=TRUE
/ RANDOM EFFECTS=encoding weapon USE_INTERCEPT=TRUE SUBJECTS=ss
COVARIANCE_TYPE=VARIANCE_COMPONENTS
/ RANDOM EFFECTS=encoding weapon USE_INTERCEPT=FALSE SUBJECTS=itm
COVARIANCE_TYPE=VARIANCE_COMPONENTS
SOLUTION=FALSE
/ BUILD_OPTIONS TARGET_CATEGORY_ORDER=ASCENDING INPUTS_CATEGORY_ORDER=DESCENDING
MAX_ITERATIONS=100 CONFIDENCE_LEVEL=95 DF_METHOD=RESIDUAL COVB=MODEL PCONVERGE=0.000001
(ABSOLUTE)
SCORING=0 SINGULAR=0.000000000001
/ EMMEANS_OPTIONS SCALE=ORIGINAL PADJUST=LSD.
```