JFP-08-2014-0022.R1

The benefit of context for facial-composite construction

Skelton, Faye (contact); Frowd, Charlie; Speers, Kathryn

Structured Abstract:

Purpose - The aim of this study was to investigate whether the presence of a whole-face context during facial composite production facilitates construction of facial composite images.

Design/Methodology - In Experiment 1, constructors viewed a celebrity face and then developed a facial composite using PRO-fit software in one of two conditions: either the full-face was visible while facial features were selected, or only the feature currently being selected. The composites were named by different participants. We then replicated the study using a more forensically-valid procedure: In Experiment 2 non-football fans viewed an image of a premiership footballer and 24 hours later constructed a composite of the face with a trained software operator. The resulting composites were named by football fans.

Findings - In both studies, the presence of the facial context promoted more identifiable facial composites.

Research limitations/implications - Current composite software was deployed in a conventional and unconventional way to demonstrate the importance of facial context.

Practical implications - Results confirm that composite software should have the whole-face context visible to witnesses throughout construction. Although some software systems do this, there remain others that present features in isolation and these findings show that these systems are unlikely to be optimal.

Originality/value - This is the first study to demonstrate the importance of a full-face context for the construction of facial composite images. Results are valuable to police forces and developers of composite software. (234 words, 250 max.)

The benefit of context for facial-composite construction

Witnesses to and victims of crime are often asked to describe the appearance of a criminal they have seen, and to construct a likeness of the face. These ‘facial composites’ are traditionally constructed by witnesses selecting individual facial features – eyes, nose, mouth, face shape, and so forth – to piece together an overall image. The police publish such images in newspapers or on television in order to generate lines of enquiry. Unfortunately, recognition of these ‘feature-based’ composites tends to be poor. For example, Frowd et al. (2005b) found correct naming rates of around 20% for feature systems (such as PRO-fit and E-FIT) used after a 3-4 hour delay; when a forensically-valid 2-day delay was inserted between

viewing the face and composite construction, naming rates were around 3% (e.g. Frowd et al., 2005a, 2007b, 2016). Research has demonstrated that such a delay negatively affects both face recall and recognition (e.g. Shapiro and Penrod, 1986; Shepherd, 1983), although it is more detrimental to recall and this is likely to impact upon face construction, which typically occurs around 2 days post-event.

Due to a general difficulty in recalling information, interview techniques have been developed that encompass different strategies to aid memory retrieval. Specifically, use of a Cognitive Interview (CI) (Fisher and Geiselman, 1992) is associated with more detailed and accurate witness statements than other types of interview; some studies have also reported a corresponding reduction in false information (see Köhnken et al., 1999 for a meta-analysis). The original version of the CI involved four stages: context reinstatement, recall everything, recall in different orders and recall from different perspectives (see also Milne and Bull, 1999). The first of these, context reinstatement, is based on the encoding specificity principle (Tulving and Thomson, 1973) and incorporates reinstatement of emotional, perceptual and sequencing aspects of an event. The rationale is that memories are linked to the context in which they were created, and so the more similar the encoding and retrieval conditions, the more complete and accurate the information recalled should be. There is a large body of empirical evidence supporting this theory (e.g. Davies and Thomson, 1988). Research into context-dependent effects shows that recall is better when tested in the environment in which the material was encoded rather than in a novel context; for example, Godden and Baddeley (1975) found that divers who both learned and recalled word lists underwater, or both learned and recalled word lists on dry land, recalled 46% more information than divers who learned lists in one environment but recalled them in the other.

Memon and Bruce (1983) also found that the benefit of context extends to face recognition: previously seen faces presented against their original background were recognised more quickly and accurately than those presented against new backgrounds. Previously unseen faces presented against ‘seen’ contexts were often falsely recognised as familiar, demonstrating the strength of context encoding. Furthermore, Rainis (1993) found that the semantics of the context are also encoded. In this case, faces presented against a different church to that at encoding, for example, were also recognised more quickly than those presented against unrelated backgrounds. Thus, the context need not be an exact match to assist memory retrieval.

Importantly, the face itself acts as a background context for identification of features: facial features are better recognised when presented in their original whole-face context than when presented as isolated features (Tanaka and Farah, 1993). The recognition advantage for facial features seen in context has been replicated a number of times (e.g. Campbell et al., 1995, 1999; Davies and Christie, 1982) and provides strong evidence for holistic face processing. Thus, research indicates that context benefits both recall and

recognition. As face construction requires both recall of facial features, and recognition of the likeness of features to build a face, context has the potential to benefit facial composites.

However, the potential benefit of context extends beyond acting as a cue for recall and recognition. One important factor potentially contributing to poor composite naming is the mismatch between familiar and unfamiliar face processing. Previous research has shown that familiar faces tend to be recognised more reliably from their internal-features (eyes, brows, nose and mouth) than from their external features (face shape and hair for example; Ellis et al., 1979). This is likely to be due to the generally stable appearance of internal-features over time, whereas external features may change, for example due to fluctuations in body weight or changes in hairstyle. On the contrary, research has shown that unfamiliar faces are recognised equally-well by internal and external features (e.g. Ellis et al., 1979); and, for this type of face processing, we are strongly influenced by the presence of external features (e.g. Bruce et al., 1999; Frowd et al., 2012).

For the current application, as the aim of publicising a composite image is to trigger a familiarity response in a member of the public, it is imperative for the detection of offenders that internal features of a composite are recognisable as the face it represents. However, research indicates that the internal-features of facial composites are generally poorly constructed. Frowd et al. (2007a) found that when composites had been constructed of unfamiliar faces, the internal-features were matched less accurately than the external-features. When they had been constructed of familiar faces, however, the internal-features were matched only slightly better than when constructed of unfamiliar faces. This indicates that face construction tends to naturally focus on the exterior parts, with the internal-features being poorly constructed regardless of target familiarity. Recent work using EFIT-V supports this. Valentine et al. (2010) found that morphing – a technique believed to reduce error as compared to individual veridical composites (Bruce et al., 2002) – benefits similarity ratings of internal-features more so than external features. This suggests that the external features of individual composites contain less error and had therefore been constructed more accurately. Further support for the role of hair and context was found by Frowd and Hepton (2009), who focused on EvoFIT, one of the newest types of composite system based on the repeated selection and breeding from arrays of complete faces (similar to E-FITV). They found that when participants evolved a composite from arrays where hair exactly matched a target, naming of the internal-features was superior to composites evolved with similar hair or poorly-matching hair. These findings suggest that good quality external features (as context) can improve the naming of the internal-features.

The above studies indicate that context is important for the construction of faces from memory, since different but related contexts (i.e. different backgrounds) can facilitate performance. They also suggest a benefit for the context provided by external-features, and for selecting individual features in the context of a

complete face. Traditional feature-based systems have varied in their method of construction. The archaic Photofit used isolated feature selection, with witnesses being referred to pages of eyes, noses, etc.; similarly the FACES composite software system also involves isolated feature selection whereas E-FIT and PRO-fit allow feature selection in the context of a complete face. With these latter software systems, individual facial features are selected by switching them in and out of an intact face. Although software companies have developed systems with the potential benefit of context in mind, research has yet to demonstrate whether this method actually helps to produce a more-identifiable image.

The current study set out to do just that: to examine whether context improves the quality of facial composites. The first experiment constructed composites under favourable conditions, famous face targets and a very-short delay, and the second used unfamiliar faces and an overnight delay, to more closely approximate the situation confronting eyewitnesses. In both cases, two groups of participants were required, one to construct the faces ('constructors' using whole-face or isolated feature selection) and the other to evaluate them by naming. It was expected that faces constructed by selecting features in the context of a whole-face would be better named than those constructed using isolated-feature selection. Also, it was expected that the internal-features of composites produced using the whole-face method would be more accurately named than the internal-features of composites produced via isolated feature selection.

Method: Experiment 1 – Familiar face composites

Stage 1 : Composite Construction

Design

A between-participants design was used, with constructors generating a composite with individual feature selection either in a whole-face context or in isolation. In the latter case, PRO-fit software was modified to allow just one feature to be viewed at a time, but to reveal the complete face when all features had been selected, to then allow each part to be sized and positioned on the face (as normal). Each person constructed a single composite in one of these two conditions (whole-face / isolated-feature).

Participants

Twenty M.Sc. Forensic Psychology students at the University of Central Lancashire (UCLan) participated, during a seminar on facial composites (16 females, 4 males, M_{age}= 24 years).

Materials

Photographs of 10 celebrities (Jennifer Aniston, Tony Blair, Pierce Brosnan, George W. Bush, Mariah Carey, Hugh Grant, Nicole Kidman, Madonna, Kylie Minogue and Brad Pitt) were gathered via online search engines. Familiar faces were used to maximise naming rates and verify whether the manipulation

worked in principle. Front-facing images were printed in colour to approximately 6cm (width) x 8cm (height). Each was placed in an envelope with written instructions for the relevant condition. Verbal description sheets were used for participants to note down what they could remember about the face prior to composite construction, with prompts for facial shape, hair, eyebrows, eyes, nose, mouth and ears. PRO-fit software version 3.5 was used. We note here that the experimenter was aware of the target identities, but did not know which identity each constructor had been randomly allocated, and was not involved in the construction process.

Procedure

Constructors completed the task in a classroom. They were initially divided into two groups, with each being briefed on the use of PRO-fit separately, according to their condition. The first author briefed those participants allocated to the whole-face condition, while the second author briefed those in the isolated feature condition. Both trainers had previously met and agreed upon the training procedure to ensure consistency. The procedure was the same for both groups, with the exception that the isolated-feature condition selected features in isolation, and had to click a box once they had selected their features, in order to switch on the whole-face context. Once briefed, participants returned to the testing room and were directed to the appropriate side of the room for construction. On one side, PRO-fit was set for use as normal, to allow individual features to be selected in the context of a complete face; to do this, features would be seen switched in and out of a single face. On the other, selection was made by seeing one feature at a time: a nose, a pair of eyes, etc. Though participants may have been aware that there were two conditions, as the class had been trained in two separate groups, they were unaware of the hypotheses and had not previously constructed a composite.

Participants were handed an envelope and asked to remove the picture and observe it for one minute, which was timed. Afterwards, they replaced the picture and wrote down what they could recall about the face on the verbal-description sheet. They were handed brief written instructions to guide them through the operation of PRO-fit. This prompted them to input their description for each feature in turn, to narrow down the options from which to choose. Once they had located around 12 to 20 examples per feature, they viewed each feature individually and selected the best match for their target. During this process, those in the context condition were able to see the features in the context of the full-face, before they decided on the best-matching exemplar for each feature. Constructors in the other group saw each feature in isolation.

Once feature selection was complete, those in the isolated-feature group switched on the whole-face context, allowing all features of the face to be seen together. All constructors then resized and positioned

their chosen features using the tools available in PRO-fit to produce the best likeness possible. The task took around an hour, and participants in both conditions took around the same time to complete their composites.

Stage 2: Composite Naming

Design

The composites produced in this study were unlikely to be of good quality, since the constructors constructed the images themselves rather than with a trained composite-system operator, and so a sensitive measure of composite quality was used (Frowd et al., 2007b): naming participants were shown composites from both context conditions and selected an identity for each from a list of written names corresponding to the identities. This so-called constrained-naming task aimed to facilitate performance. The design for context type was within-subjects.

Participants

An opportunity sample of 11 female and 7 male staff and students (M_{age} = 25 years) volunteered to name the composites.

Materials

Each of the composites was printed in greyscale (PRO-fit uses this image mode) to a size of about 6cm x 8cm. Example composites are shown in Figure 1. A sheet was prepared containing a list of relevant celebrity names.

Figure 1. Example composites of Brad Pitt constructed using feature selection in the context of a whole face (left) and by isolated features (right), correctly named at 70.8% and 66.7% respectively, and representing the best image in each condition. Each image was created by a different person.

Procedure

Participants were tested individually. They were shown each composite in turn and asked to select a name from the given sheet, if they believed the identity to be present. Participants were told to expect more

than one composite of each celebrity. Composites were presented in a different random order for each person. The task was self-paced and took about 10 minutes per person.

Results and Discussion

The raw data comprised the total number of correct names for each of the 20 composites (10 produced with whole-face context and 10 without). Out of a total of 360 possible correct (20 composites x 18 participants), 141 responses were correct and an incorrect name was chosen on 183 occasions. The mean correct naming was 42.2% (*SD* = 27.7%) with face context and was somewhat lower at 36.1% without (*SD* = 26.3%). For the inferential statistics, a by-items analysis was carried out. This type of analysis overcomes constructor outlier effects – exceptionally good or bad quality composites – that can easily skew a by-participants analysis. A one-tailed paired samples t-test confirmed benefit of facial context, $t(9) = 2.42$, $p < .05$, Cohen's $d = 0.22$.

As expected, composites constructed in the whole-face context were named significantly higher than those produced using isolated-feature selection. However, the study utilised target faces that were familiar to constructors, who themselves constructed a face. When a crime is committed, it is normal for a witness to describe an unfamiliar face using a CI and then be guided through the process of construction by an experienced police operative. It is also usual for witnesses to be interviewed and produce the composite after one or two days. In addition, police officers and members of the public attempt to recognise the face spontaneously – they do not have a list of names from which to choose. So, Experiment 1 verified that the context manipulation was effective; we attempted a replication in Experiment 2 with design changes made to improve ecological validity. Thus, Experiment 2 uses the 'Gold Standard' procedure for composite construction (Frowd et al., 2005b).

Method: Experiment 2 – Unfamiliar face composites

Stage 1: Composite Construction

Design

The design was basically the same as the previous experiment's, but was made more realistic. The target faces were chosen to be unfamiliar to constructors, but familiar to participants who would later attempt to name them (see Materials). A 22 to 26 hour delay was imposed between seeing the face and starting the construction session; and, as part of face construction, the Experimenter (the third author) administered a CI to elicit a description of the face using procedures typical of a UK police investigation (see Frowd et al., 2005a and Procedure below), and operated PRO-fit. To further reflect police procedures, the artwork package in PRO-fit was used to enhance the likeness, by adding shading, wrinkles, marks, etc.,

as requested by the constructor. The role of the Experimenter, who was suitably experienced in the use of PRO-fit, was to develop the composite under the direction of the constructor.

Participants

Constructors were 20 staff and students from UCLan (14 females, 6 males, M_{age} = 32 years). All were recruited on the basis of being unfamiliar with UK footballers and were paid £5 for their time.

Materials

Photographs of 10 UK international-level footballers (Joe Cole, Peter Crouch, Robbie Fowler, Ole Gunnar Solskjaer, Frank Lampard, Gary Neville, John O'Shea, Paul Scholes, Alan Smith and John Terry) were sourced using online search engines, and printed in colour to approximately 5cm x 5cm. Footballers were chosen as it is necessary to utilise targets that are not recognisable to those constructing the composites, but are widely recognisable for ease of recruiting for the evaluation tasks. Although some footballers may be known to non-football fans, we were careful to avoid using very well-known identities (e.g. David Beckham) and also checked that no-one constructed a composite of an identity with which they were familiar. Verbal-description sheets were used to record participants' descriptions of each facial feature prior to construction using PRO-fit.

Procedure

Constructors were tested individually and were randomly assigned to face construction by feature selection made in the context of a whole-face or in isolation. Each of the ten identities was constructed twice, once in each condition, with the experimenter remaining blind to these identities Participants were briefly shown a target face, also randomly selected, and asked whether they recognised the face. If the face was reported familiar, it was placed back in the envelope and another one selected randomly. When the first unfamiliar face was found, participants inspected it for one minute in the knowledge that a composite would later be made of the face. This method of encoding has been used in composite research (e.g. Bruce et al., 2002; Frowd et al., 2012), and gives rise to composites with naming rates similar to those constructed after a more realistic encoding, such as using moving stimuli (Frowd et al., 2016).

Between 22 and 26 hours later, participants returned to the lab. The procedure used to create a composite is fairly detailed and is described in Fodarella et al. (2016). In brief, an overview of the session was given, and a CI used to recall the face. Afterwards, the given description was repeated back for each feature and the participant was prompted to attempt further recall. For face construction, the Experimenter started PRO-fit and used the same basic procedure as constructors had used themselves in Experiment 1. In

addition, the Experimenter made use of the artwork package in PRO-fit to enhance the likeness, also under the direction of the constructor. Construction sessions took about an hour.

Stage 2: Composite Naming

Design

A within-participants design was used, with participants naming the composites from both conditions. Two naming tasks were used. The first required participants familiar with footballers to spontaneously name the whole-face composites. The second task required naming of only the internal-features (region encompassing the eyes, brows, nose and mouth) region of the composites. This was to check whether the whole-face selection procedure resulted in the construction of a more accurate set of internal-features: the important region for familiar-face recognition. As this task is more difficult than using complete images, it was made easier using the constrained-naming procedure of Experiment 1, by providing participants with a list of written names. While police would not use this procedure, the focus here was to examine how well the internal-features could be named without the potentially-distracting presence of external features.

Participants

Thirty-six participants volunteered from UCLan. All described themselves as football fans. Twenty-four were presented with complete faces (5 females, 19 males, M_{age}=27 years), with equal sampling to the two levels of context type, and 12 with internal-features composites (2 females, 10 males, M_{age}= 24 years).

Materials

The whole-face composites were printed in greyscale to a size of approximately 5cm x 5cm. The target faces used in the construction stage were printed in colour to the same size. For the internal-features naming task, electronic versions of the composites were edited into an oval in Adobe Photoshop, with the image cropped above the eyebrows to eliminate cues from hair. For these participants, a written list of the footballers' names was provided. Example composites are shown in Figure 2a and 2b.

Procedure

The constrained-naming procedure for the internal-feature composites was the same as that used in Experiment 1, except that participants were recruited on the basis of being familiar with footballers and were told that the composites were of UK international-level footballers. They were shown internal composite features (see Figure 2c and 2d for examples). The procedure was the same for those given complete images, except that intact composites were presented without the list of reference names. In both cases, participants were also asked to name the target photographs (in a different random order for each person) as a check that participants knew the relevant identities; all participants achieved at least 75% correct.

2a 2b 2c 2d

Figure 2. Example composites of the footballer Gary Neville constructed with the facial-context present (2a) and by isolated-feature selection (2b), correctly named at 79.2% and 12.5% of the time respectively; internal-features of these images are shown in 2c and 2d, correctly named at 50% and 41.7% of the time, respectively.

Results

Out of 480 possible correct responses on the whole-face naming task (24 participants x 20 composites), there were only 50 correct names on the spontaneous naming task; the remainder were incorrect names. For the internal-features constrained-naming task, out of 240 possible correct responses, the correct name was chosen on 63 occasions, with the remainder incorrect names. Means and standard deviations for whole-face spontaneous and internal-features constrained-naming tasks are shown in Table 1, and indicate somewhat higher means for construction using the whole-face context.

Table 1. Performance of complete and internal-features composites in the two naming tasks

	Spontaneous naming of complete composites	**Constrained-naming of internal-features**
Isolated feature selection	7.1	21.7
	(13.2)	(22.3)
Selection in whole-face context	13.8	30.8
	(24.0)	(22.6)

Note. Naming figures are expressed in percentage correct, and those in parentheses are standard deviations of the means.

A one-tailed paired-samples t-test on the spontaneous naming rates, by-items, revealed that complete composites constructed using a whole-face context were named significantly higher than those produced using isolated-feature selection, $t(9) = 3.84$, $p < .05$, $d = 0.35$. This test found the same reliable result for internal-features' naming, $t(9) = 2.79$, $p < .05$, $d = 0.41$. See also Footnote [1] for an additional measure.

Discussion

Previous research indicates that context assists memory retrieval. With facial composites, the whole-face acts as the context within which individual features are selected. However, while research indicates benefit when making judgements about a single feature in a whole-face context (e.g. Davies and Christie, 1982), there appears to be no formal studies that have verified whether context really does benefit face construction. This was the aim of the current study.

Experiment 1 used favourable conditions and evaluated composites using a sensitive 'constrained' naming task, with participants selecting from a list of written names. Experiment 2 increased ecological validity using a 22 to 26 hour delay between viewing an unfamiliar target and face construction. Both experiments revealed that context improved naming levels for the complete image, and in Experiment 2 this produced a more identifiable set of internal-features. The effect size found in Experiment 1 was small (Cohen's $d = 0.22$); however, in the more carefully-controlled Experiment 2 the effect was medium (0.35 and 0.41 for spontaneous and internal-feature naming, respectively), demonstrating the marked improvement in effectiveness for construction in a whole-face (cf. isolated) context.

Thus, results reveal that correct naming is significantly higher when composites were constructed with the face context present than when absent, and in realistic conditions context has a moderate effect size for naming. The findings suggest that it is easier to select a well-matching facial part in the context of a whole-face than selecting that part in isolation. This is consistent with previous studies demonstrating a recognition advantage for features seen in context (e.g. Campbell et al., 1995; Memon and Bruce, 1983; Tanaka and Farah, 1993). Our results suggest that the benefit is consistent whether the target is familiar or unfamiliar to the person constructing the face. Also, internal-features constructed in-context were named more accurately than those constructed via isolated feature selection, indicating that the internal-features in the context condition were a better likeness to the target's features than those in the isolated feature condition.

Taken together with Frowd and Hepton's (2009) findings, these data suggest that context is important for different composite software systems and methodologies. The presence of the whole-face context provides additional information to help witnesses select the components of the internal-features, and the better the match the facial context, the better the internal-features should be constructed and then

subsequently recognised. Indeed, recent work indicates that feature selection using a whole-face context also provides benefit over isolated-feature selection for feature-based composites when used after the Holistic-CI (Kuivaniemi-Smith et al., unpublished), presumably as isolated-feature selection is not closely aligned with face recognition. However, there is research to suggest that providing the entire face as context may not be optimal: Frowd et al. (2012) found strong evidence that the presence of external features distract witnesses during EvoFIT composite construction, even if the features are an exact match to those of the target. Composite quality benefitted from masking external features during construction of internal features, with hairstyle and the remaining external features selected thereafter. Frowd et al. (2013) provide further support for this external-feature effect as well as demonstrating benefit for varying facial context at naming.

Thus it appears that selecting features in the absence of any other features is detrimental to composite quality, however selecting features in the context of a whole-face leaves the witness open to distraction by the external features (e.g. hair). Future work could usefully explore exactly how much facial context is optimal for performance: we might expect that important retrieval cues are actually contained within the central portion of the face, and that internal-features need only be selected with other internal-features in view for optimal performance. Our current work is exploring the effectiveness of different interview techniques (e.g. CI, Holistic-CI, Frowd et al., 2008, 2013) with different construction methods in order to determine maximum effectiveness of modern composite software systems.

Our results reported here provide good evidence that, when using traditional feature systems such as PRO-fit and E-FIT, the police should ensure that the facial background is visible while witnesses select individual facial features, as this will increase the likelihood of the image being correctly named. Many more advanced feature-based systems have been designed in this way, and this research is the first to confirm that this is indeed the best procedure. We should acknowledge that changes made to PRO-fit for these studies necessitated using the software differently to how it was originally intended. However, developers of systems using isolated-feature selection (e.g., the FACES system that is popular in the US, Frowd et al., 2007b) should be aware that their software could be usefully improved based on these findings.

Implications for Practice

- Facial composite identification is significantly affected by the construction technique used.
- Selecting facial features without other facial features in view is detrimental to successful naming of composites.
- Developers of feature-based composite software should ensure that their systems allow selection of facial features in the context of other facial-features, to ensure optimal performance.

Footnote [1] We also collected likeness ratings, whereby different participants rated the internal-feature composites for similarity to the target images. The internal features of in-context composites were rated a significantly better likeness to the targets than those created using isolated-feature selection, $t(9) = 2.59$, $p < .05$, $d = 0.65$.

References

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P.J.B., Burton, A.M. and Miller, P. (1999), "Verification of face identities from images captured on video", *Journal of Experimental Psychology: Applied*, Vol. 5, pp. 339 360.

Bruce, V., Ness, H., Hancock, P.J.B., Newman, C. and Rarity, J. (2002), "Four heads are better than one: Combining face composites yields improvements in face likeness", *Journal of Applied Psychology*, Vol. 87, pp. 894-902.

Campbell, R., Coleman, M., Walker, J., Benson, P.J., Wallace, S., Michelotti, J. and Baron-Cohen, S. (1999), "When does the inner-face advantage in face recognition arise and why?", *Visual Cognition*, Vol. 6, pp. 196-216.

Campbell, R. Walker, J. and Baron-Cohen, S. (1995), "The use of internal and external features in the development of familiar face identification", *Visual Cognition*, Vol. 6, pp. 197-216.

Davies, G.M. and Christie, D. (1982), "Face recall: an examination of some factors limiting composite production accuracy", *Journal of Applied Psychology*, Vol. 67, pp. 103-109.

Davies, G.M. and Thomson, D.M. (Eds.). (1988), *Memory in context: Context in memory*, Chichester, Wiley.

Ellis, H.D., Shepherd, J.W. and Davies, G.M. (1979), "An investigation of the use of the photo-fit technique for recalling faces", *British Journal of Psychology*, Vol. 66, pp. 29-37.

Fisher, R.P. and Geiselman, R.E. (1992), *Memory enhancing techniques for investigative interviewing: The Cognitive Interview*, Springfield III, Charles C. Thomas.

Fodarella, C., Kuivaniemi-Smith, H.J. and Frowd, C.D. (2016), "Detailed procedures for forensic face construction", *Journal of Forensic Practice*.

Frowd C.D. Bruce, V. McIntyre, A. and Hancock, P.J.B. (2007a), "The relative importance of external and internal-features of facial composites", *British Journal of Psychology*, Vol. 98, pp. 61-77.

Frowd, C.D., Bruce, V., Smith, A. and Hancock, P.J.B. (2008), "Improving the quality of facial composites using a holistic cognitive interview", *Journal of Experimental Psychology: Applied*, Vol. 14, pp. 276-287.

Frowd, C.D., Carson, D., Ness, H., McQuiston, D., Richardson, J., Baldwin, H. and Hancock, P.J.B. (2005a), "Contemporary composite techniques: the impact of a forensically-relevant target delay", *Legal and Criminological Psychology*, Vol. 10, pp. 63-81.

Frowd, C.D., Carson, D., Ness, H., Richardson, J., Morrison, L., McLanaghan, S. and Hancock, P.J.B. (2005b), "A forensically valid comparison of facial composite systems", *Psychology, Crime and Law*, Vol. 11, pp. 33-52.

Frowd, C.D., Erickson, W.B., Lampinen, J.L., Skelton, F.C., McIntyre, A.H. and Hancock, P.J.B. (2016), "A Decade of Evolving Composites: Regression- and Meta-Analysis", *Journal of Forensic Practice.*

Frowd, C.D. and Hepton, G. (2009), "The benefit of hair for the construction of facial composite images", *British Journal of Forensic Practice*, Vol. 11, pp. 15-26.

Frowd, C.D., McQuiston-Surrett, D., Anandaciva, S., Ireland, C.E. and Hancock, P.J.B. (2007b), "An evaluation of us systems for facial composite production", *Ergonomics*, Vol. 50, pp. 1987-1998.

Frowd, C.D., Skelton, F., Atherton, C., Pitchford, M., Hepton, G., Holden, L., … Hancock, P.J.B. (2012), "Recovering faces from memory: The distracting influence of external facial features", *Journal of Experimental Psychology: Applied*, Vol. 18, pp. 224-238.

Frowd, C.D., Skelton F.C., Hepton, G., Holden, L., Minahil, S., Pitchford, M., McIntyre, A., Brown, C. and Hancock, P.J.B. (2013), "Whole-face procedures for recovering facial images from memory", *Science and Justice*, Vol 53, pp. 89-97.

Godden, D.R. and Baddeley, A.D. (1975), "Context-dependent memory in two natural environments: On land and underwater", *British Journal of Psychology*, Vol. 66, No 3, pp. 325-331.

Köhnken, G., Milne, R., Memon, A. and Bull, R. (1999), "The Cognitive Interview: A Meta-analysis", *Psychology, Crime and Law*, Vol. 5, pp. 3-27.

Kuivaniemi-Smith, H., Sanderson, H., Minahil, S. and Frowd, C.D. "Improving the effectiveness of sketch-based composite images", *unpublished manuscript.*

Memon, A. and Bruce, V. (1983), "The Effect of Encoding Strategy and Context Change on Face Recognition", *Human Learning: Journal of Practical Research and Applications,* Vol. 2, pp. 313-326.

Milne, R. and Bull, R. (1999), *Investigative interviewing: Psychology and practice,* Chichester, John Wiley and Sons, Ltd.

Rainis, N. (1993), "Semantic Contexts and Face Recognition", *Applied Cognitive Psychology,* Vol. 15, pp. 173-186.

Shepherd, J.W. (1983), "Identification after long delays", In S.M.A. Lloyd-Bostock and B.R. Clifford (Eds.), Evaluating witness evidence, Chichester, Wiley, pp. 173-187.

Shapiro, S. and Penrod, B.L. (1986), "Meta-analysis of Facial Identification Studies", *Psychological Bulletin,* Vol. 100, pp. 139-156.

Tanaka, J.W. and Farah, M.J. (1993), "Parts and Wholes in Face Recognition", *The Quarterly Journal of Experimental Psychology,* Vol. 46A, pp. 225-245.

Tulving, E. and Thomson, D.M. (1973), "Encoding Specificity and Retrieval Processes in Episodic Memory", *Psychological Review,* Vol. 80, pp. 353-370.

Valentine, T., Davis, J.P., Thorner, K., Solomon, C. and Gibson, S. (2010), "Evolving and combining facial composites: Between-witness and within-witness morphs compared", *Journal of Experimental Psychology: Applied*, Vol. 61, pp. 72-86.