Featural and holistic processing in facial composite construction: the role of cognitive style and processing sets

Donna A. Taylor
School of Social Sciences, Humanities and Languages

Featural and holistic processing in facial composite construction: The role of cognitive style and processing sets

Donna A. Taylor

A thesis submitted in partial fulfilment of the requirements of the University of Westminster for the degree of Doctor of Philosophy

July 2012

Abstract

When a serious crime has been committed, eyewitnesses may be required to assist a police investigation by constructing a facial composite of the perpetrator of the crime with the help of a police operator. A large body of research has investigated the utility of composite construction systems and the ways in which they are implemented with eyewitnesses. There has been less research conducted on individual differences which might have an impact on the accuracy of facial composites which eyewitnesses produce. The first aim of the research presented within this thesis was to investigate whether individual differences in stable cognitive style have an effect on the accuracy of the facial composites they produce. The second aim of the research was to investigate whether manipulating the temporary cognitive processing state of individuals during face encoding and prior to facial composite construction affects the accuracy of the facial composites they produce. These issues were investigated using two facial composite construction systems currently in widespread use by UK police forces, E-FIT and EFIT-V.

Study One investigated, for the first time in the facial composite literature, individual differences in the cognitive style of field dependence/independence (Witkin, Oltman, Raskin & Karp, 1971). Results indicated that field independent individuals produced more accurate composites than field dependent individuals. Study Two investigated individual differences in holistic/analytic cognitive style (Riding & Cheema, 1991). Results indicated that individuals with a holistic cognitive style produced more accurate composites than individuals with an analytic cognitive style.

Study Three manipulated the way in which faces were encoded by individuals, and introduced a Navon (1977) task into the composite construction process using E-FIT. Results showed that the Navon task had an effect on the accuracy of the facial composites that individuals produced which was mediated by the way in which the target face had been encoded. Study Four introduced a Navon task prior to composite construction using the EFIT-V system. In addition, the field dependence/independence cognitive style of the participants

who created an EFIT-V was measured. Results showed that the Navon task had an effect on the accuracy of the EFIT-V composites that individuals produced, which was mediated both by the way in which the target face was encoded, and by the cognitive style of the individual.

Overall, the findings indicated that there is a strong featural cognitive processing element to facial composite construction which is at odds with the way in which faces are processed and represented in memory. Collectively, the results indicate that featural cognitive processing prior to the composite construction process may lead to more accurate facial composites. In addition to this, if an individual does not have a natural featural processing cognitive style, then inducing a featural cognitive processing state may also lead to more accurate facial composites.

List of Contents

List of Tables

List of Figures

ACKNOWLEDGEMENTS

Firstly I would like to thank my Director of Studies Dr. Mark Gardner. Mark has been an excellent mentor and guide throughout the whole of my PhD experience. I understand now why he took so long to decide to take on the role of my Director of Studies – Mark has given freely of his time and expert knowledge and helped make the whole process run in a smooth and efficient manner. I have enjoyed my PhD experience, which is in no small part due to the dedication and care that Mark has exhibited to ensure that I have reached this point. Thank you Mark, very much.

I would also like to thank Dr. David Groome who was on my original supervisory team for five years. David did not have as much of an active role in my research as my other supervisors, but he always maintained an interest in it.

I would like to thank my external expert supervisor Dr. Jim Turner. Jim suggested that I should do a PhD when I hadn't even considered it, so it is partly his fault that I haven't had a life for the last six years. On a serious note, Jim has been a fantastic supervisor. Despite the geographical distance Jim has offered invaluable expert advice in the field of facial composite research both by email and in many long team meetings. Thank you Jim.

Although a late addition to my supervisory team, I owe a big thank you to Professor Tom Buchanan. I was initially worried about how a new supervisor would fit into our established order of working. I need not have worried; Tom has been excellent. He offered me great advice when I needed it, was always approachable, and took the time and effort to get to know my research rather late in the whole process, which cannot have been easy. Thank you Tom.

Extra thanks go to the Psychology department technicians Lejla Kanlic-Mandlukic and Haulah Zachariah. They very kindly helped me to set up the online evaluation of Study One and have offered advice, support and friendship throughout my studies.

Thanks to Dr. Chris Solomon and Dr. Stuart Gibson for the conversations about E-FIT and EFIT-V. Chris also welcomed me on an EFIT-V training course free of charge, which was rather nice of him.

Thanks to Dr. Elizabeth Peterson for granting me a licence to use the E-CSA-W/A and VICS tests in my second Study. Thanks also to Professor Tim Perfect who supplied the Navon materials used in Studies Three and Four of this thesis.

Thank you to colleagues and friends at Westminster who have supported me throughout my studies. Dave Williams deserves a special mention for spending hours sorting my word files and data files into a manageable system. Thanks to Lisa Thorn and Andrea Oskis who gave me a lot of support and advice at the beginning of the PhD. Thanks to Nina Smyth for bringing the PhD room back to life after Lisa and Andrea left, and for all the laughs. Thanks to Karen Ryder who has become a good friend over the last few years working together at Westminster. The long phone chats about PhD things have helped more than you realise.

Thanks to Joan Veale for paying my fees for the first four years. Without that help I wouldn't have been able to do the PhD at all. Thanks to Hazel Dewart for waiving the PhD fees in the penultimate year, which was also a huge help.

Last, but by no means least, I would like to acknowledge the help and support of my very special family. My sister Michelle Weller, who has looked after my children every other weekend, all weekend, for the last six years and given me the time to do my reading and writing. My mum Jo Brockey has never stopped reminding me of how proud she is of me, and has been a constant source of love, support and encouragement. My brilliant children Dan Courtney, Emilie Veale and Hailie Veale, who have also supported and helped me all the way. A special thanks to Emilie, who assured me when I was contemplating giving up after one year of study, that she didn't want me to have any extra time to spend with her so I had better get on with it!

DEDICATION

This thesis is dedicated with all my love to Dan, Emilie and Hailie.

Declaration

I declare that all the material contained in this thesis is my own work.

Chapter One: Overall Aims of the Research

In cases where a serious crime has been committed eyewitnesses may be required to assist a police investigation by constructing a facial composite with the help of a police composite construction officer. This is a visual representation of the perpetrator's face which can be used by the police to help generate leads when there is no suspect readily apparent. The research presented in this thesis investigated the role of holistic and featural cognitive processing in the encoding and recall of faces, and how this processing relates to the construction of target-accurate facial composites. These findings may potentially contribute to two forms of practical applications. First, they may contribute to recommendations about which witnesses are likely to construct target-accurate facial composites on occasions where several eyewitnesses are available. Second, they may help inform changes to the protocol for using composite construction systems through the introduction of evidence-based interventions designed to maximise the performance of eyewitnesses.

Facial composites are generally poor likenesses of a perpetrator's face (Frowd, Bruce, Smith and Hancock, 2008) and this has implications for their practical use in criminal investigations. Not only can they prevent members of the public from identifying the perpetrator of a crime if they are inaccurate, they can also lead to the prosecution of an innocent person who resembles the inaccurate facial composite produced. Several facial composite systems have been used by police forces in the last forty to fifty years, each with limited success, and a considerable amount of psychological research has been conducted with these systems in an attempt to modify and improve them. This research is of practical benefit in terms of enabling eyewitnesses to produce facial composites which are more similar to the target face they are supposed to depict, and also of theoretical interest with regard to the psychological limitations in how faces are processed and recalled.

The majority of research on facial composite construction has been devoted to the development of the systems used to create composites, and to the optimum conditions for witnesses to create composites. Research into these 'system' variables (e.g. Davies & Christie, 1982; Frowd et al., 2008) has contributed to the successful modification of facial composite systems and the way in which they are implemented. In the terminology of forensic psychology, system variables are post-event variables which can be controlled or manipulated, and include the way in which eyewitnesses are interviewed prior to composite construction for example. However, even if system variables were to be successfully optimised, eyewitness variability in the ability to produce facial composites with a high degree of similarity to the target face would remain.

Knowledge of eyewitness variables which impact on memory performance would allow prediction of conditions under which witnesses may be more likely to make errors. Therefore this knowledge could increase confidence in the composites produced by witnesses who possess characteristics which are correlated with accurate composite construction performance. Additionally, knowledge of eyewitness variables helps in understanding the likely impact of system variables. This was demonstrated by Emmett, Clifford and Gwyer (2003) who showed that the cognitive style (field dependent or independent) of an individual has an impact on the effectiveness of the cognitive interview (context reinstatement elicits significantly more correct information from field dependent individuals than field independent individuals).

The literature acknowledges large individual differences in the ability to produce accurate facial composites (Ellis, Shepherd & Davies, 1975; Laughery & Fowler, 1980) although these individual differences have received less attention than system variables. This is possibly because witness variables cannot be manipulated in the same way as system variables. However according to Laughery and Fowler (1980) if there is a straightforward way of measuring characteristics of the witness, and if these characteristics correlate with the quality of facial composite produced, then more or less confidence could be placed in that image by the police when a decision needs to be made on whether to publish the composite.

The research reported in this thesis investigated the role of holistic and featural processing in the construction of facial composites from two perspectives. Firstly, the relationship between stable individual cognitive style and the ability to construct accurate facial composites was explored in two studies. Secondly, two further studies examined the possibility that cognitive processing is a malleable state that can be induced prior to composite construction in order to promote optimal memory performance in individuals. The roles played in composite construction by cognitive style as a generally stable trait, and cognitive processing as a malleable temporary processing state, were assessed using two of the most popular composite construction systems currently in use by police forces in the UK, E-FIT (Aspley, 1993) which is predominantly a feature-based system, and EFIT-V (Visionmetric, 2004) which is a whole-face recognition-based system.

In summary, the research conducted in this thesis examined the role of holistic and featural cognitive style both as a stable trait and a temporary processing state in the production of facial composites using both a feature-based and a whole-face recognition-based composite construction system. Chapter Two outlines psychological theories of face perception, recall and recognition which are relevant to the task of producing a facial composite. Chapter Three outlines the history and development of composite construction systems and details how these systems which were originally very feature-based have progressed over time into whole-face recognition-based systems. Chapters Four and Five introduce research on cognitive style as a stable trait and as a temporary processing state and assess how cognitive style and cognitive processing may impact on the construction of accurate facial composites. Chapter Six describes some general methods used in both the construction and evaluation of the facial composites in the research presented in four subsequent empirical chapters.

Chapter Two: Theories of Face Processing, Recall and Recognition

This chapter describes research relating to face perception and recall, with emphasis on aspects which are relevant to the process of facial composite construction. The configural (holistic) way in which faces are generally perceived and stored in memory is at odds with the featural way in which composite systems up to and including E-FIT are used, and has particular relevance for the way in which the latest composite systems such as EFIT-V have been developed. Another issue highlighted is the distinction between face recognition and face recall, and how these separate cognitive processes may relate to the two different composite construction systems used in the research presented in this thesis: E-FIT which incorporates both recall and recognition, and EFIT-V which works mainly on the principle of the recognition of whole faces.

2.1: Configural processing and representation of faces in memory

Prior to the introduction of EFIT-V around 2010 to several police forces in the UK, most composite construction systems up to and including E-FIT required the witness to describe the individual features of the target face, and then build a facial composite feature by feature. However, there is a lot of evidence indicating that faces are not perceived or remembered as a collection of isolated features. Rather, faces are processed in a holistic manner.

The holistic manner in which faces are processed and represented in memory has implications for the use of facial composite systems, which require eyewitnesses to externalise their mental representation of a previously seen face. Computerised composite systems such as E-FIT were designed to capitalise on the fact that configural processing is essential for face recognition. Although witnesses still have to choose each feature individually within the E-FIT program, they view the features within the context of a whole face and not in isolation as with older composite systems. The following research highlights the importance of configural processing in face processing, recall, and recognition and how this impacts on facial composite construction.

It has long been established that faces are processed and remembered in a different way to other types of visual stimuli. There is a specialised mechanism for the processing of faces which was first identified in a seminal study by Yin (1969) who found that if a picture of a face was inverted (displayed upside down) then recognition of that face suffered far greater disruption than that caused by the inversion of other objects. There is some debate about whether this specialised processing mechanism is exclusive to face processing or extends to other classes of visual stimuli for which a person shows expertise (Diamond & Carey, 1986; Valentine, 1988). However, the way in which faces are processed and represented in memory is argued to be qualitatively different from the representation of most other visual stimuli for which no expertise is shown (Kanwisher, 2000; Yovel, Paller & Levy, 2005). An influential idea is that faces are disproportionately sensitive to the effects of inversion because inversion interrupts *configural* processing.

Configural processing is an umbrella term used to describe three different processes which combine when faces are perceived (Maurer, Le Grand, & Mondloch, 2002). People are sensitive to first order relations of faces, the fact that all faces have two eyes situated above a nose situated above a mouth. All faces display this configuration, so in order to be able to distinguish between thousands of different faces that are encountered, we are also sensitive to second order relations of the face, the individual features *and* distances between individual features which contribute to making each face unique. Finally faces are processed in a holistic way. There is a fusion of a whole face into a perceptual gestalt from which it is difficult to extract individual featural information.

Evidence for the primacy of holistic encoding of faces was demonstrated by Farah, Tanaka and Drain (1995). They constructed a series of dot patterns which differed in the degree to which the patterns could be perceived as a whole or as a series of parts, based on the position and colour of the dots. It was found that the dot patterns which had been encoded in a holistic way were much more affected by inversion than dot patterns which had been encoded in a featural way. Therefore, it was suggested that faces are differentially affected by inversion because they are encoded primarily in a holistic way.

Not only is the whole face more than the sum of the individual features, but the whole face can even be recognised without perception of those individual features. Faces can be recognised from their configural (holistic) properties even when individual features have been degraded in some way so that they are not perceptible (McKone, Martini, & Nakayama, 2001; McKone, 2004), which supports the idea that faces are primarily processed in a holistic manner. The ability to process faces configurally (holistically) in the absence of featural information demonstrates a double dissociation between configural (holistic) and featural processing, because features can also be processed in the absence of the whole face (McKone, 2004). This dissociation supports the idea that it is holistic processing which is more important in face perception than the perception of second order relations of the face.

Several research studies point to the primacy of holistic face processing, all of which show strong effects for upright faces which are absent or much weaker for inverted faces. Young, Hellawell and Hay (1987) conducted a series of experiments in which faces were divided horizontally into two halves and rejoined with different faces to create ‘new’ combinations of faces (called composite faces within their study, but not to be confused with facial composites produced using compositing systems). Famous faces were used which would be familiar to participants, and composite images were produced which were two different top and bottom halves of faces either fully aligned or horizontally misaligned. Recognition of the top halves of the famous faces was seriously impaired when the composites were fully aligned, which indicated that the perception of a whole face image interfered with the identification of individual features. When the top and bottom halves of the faces were misaligned then an image of a new holistically processed face was not induced and the different halves were much easier to identify. This was termed the ‘composite face effect’ and applies not only to familiar faces, it is also observed in the processing of unfamiliar faces (Hole, 1994).

A further experiment within the same study (Young et al., 1987) showed that this ‘composite face effect’ disappeared for inverted faces. Furthermore, the task of identifying parts of a composite face was made easier by inversion. This demonstrated that when two halves of a composite face are inconsistent (but perfectly aligned) holistic processing of the upright face leads to a decrement in recognition of individual features. This may have implications for the E-FIT system where individual features are chosen within the context of the whole face. It may be that there is no significant advantage in choosing individual features within the context of a whole face, as perception of the whole face image may well disrupt the perception of individual features.

However Tanaka and Farah (1993) conducted an experiment assessing how accurately facial features could be identified in isolation or as part of a whole face. Participants were required to become familiar with faces which had been constructed using a computerised facial composite system (Mac-a-Mug Pro). Participants then had to identify different features of the learned faces when they were presented in isolation or displayed on the original face. It was found that it was much easier to make similarity judgements about individual features when they were presented within the context of the whole face rather than in isolation (part-whole recognition effect). This advantage for whole face presentation was not found for scrambled faces, inverted faces or houses. The results of this study suggest that judging the accuracy of a particular facial feature is much easier within the context of the whole face (Homa, Haver & Schwartz, 1976). When a feature is removed from its facial context and viewed in isolation then information concerning the relationship to other features is lost.

Sergent (1984) suggested that face perception produces a holistic image in which featural and configural (relational) information is combined interactively, and that configural information is as important as the individual features of the face for face recognition (Rakover, 2002; Farah et al., 1995; Tanaka & Sengco, 1997). Although there is much research to support the idea that faces are generally processed in a holistic way, there are conditions in which each type of processing would appear to be used predominantly. For instance these may depend on the information processing task being performed (see section 2.2 – holistic vs. featural encoding), and also on the visual stimulus presented. It is also the case that there are individual differences in the general propensity to perceive visual stimuli in a holistic or a featural way (see Chapter 4 – Cognitive Style), and this may also have an impact on which type of processing is used and subsequently on the ability to produce an accurate facial composite.

2.2: The effect of holistic vs. featural encoding on face processing

Research into face recognition and recall has revealed that the way in which faces are encoded can have an impact on how they are subsequently recalled and/or recognised. Coin and Tiberghian (1997) conducted a meta-analysis of studies which looked at different types of face encoding on subsequent face recognition. The major finding from the meta-analysis was that faces that are encoded while making judgements about aspects of their personality are subsequently better recognised than if featural judgements are made about a face during encoding. Faces that are encoded with no judgements required about either personality or features are less easily recognised than faces encoded with personality judgements, and more easily recognised than faces encoded with featural judgements.

There are several theories proposed to account for the superiority of making personality judgements of faces for subsequent face recognition. Following Craik and Lockhart's (1972) levels of processing theory, it is possible that judging the inferred character of a face promotes a semantic (deep) level of processing, as opposed to making featural (shallow) judgements about a face, and that it is this deeper processing which promotes better subsequent recognition. Bower and Karlin (1974) found that faces which were encoded with judgements about personality were better remembered than faces which were encoded with judgements about gender. They argued this was because the faces which required judgements about personality were processed at a deeper level than the faces which required judgements about gender.

Winograd (1976) found that judgements of personality or occupation led to significantly greater face recognition than judgements about hair type or size of nose. The 'elaboration hypothesis' (Winograd, 1981) argued that superior recognition performance following ratings for personality was due to the fact that more features are processed in order to make trait judgements than to make individual featural judgements. However, Wells and Turtle (1988) argued that if more features are processed following personality judgements then verbal descriptions should also be better following personality judgements. The opposite result was found, and has since been replicated by Finger and Pezdek

(1999); individuals who made featural judgements about faces gave more accurate descriptions, inconsistent with predictions of the elaboration hypothesis.

Wells and Hyrciw (1984) found that making personality judgements about a face led to better face recognition. Their ‘matching superiority’ hypothesis argued that making personality judgements about a face requires an individual to adopt a global or holistic processing strategy in which the overall face is considered when making such personality judgements, and that it is the *match* between global/holistic processing at encoding and holistic processing at recognition that facilitates this advantage. The matching superiority hypothesis received further support within the same study with the finding that a featural face encoding strategy was associated with the production of more accurate facial composites using Identikit, which requires individual features of the face to be selected in the process of composite construction. Therefore the match between featural processing at the encoding and featural processing in composite construction was argued to facilitate the advantage for featural encoding of faces in Identikit construction.

It is difficult however, to generalise the results from studies which look at the effect of encoding on face recognition. The encoding instructions that participants receive only apply to a very limited set of strategies used for encoding faces in the real world. In reality faces are encoded in a spontaneous manner without imposed judgements of personality or certain features as they are within an experimental setting. Other studies have examined the influence of inducing holistic or featural processing *after* faces have been encoded and *before* a subsequent recognition task. These studies offer the possibility of an intervention that can be of practical use with eyewitnesses to elicit the best memory of a previously seen face.

Macrae and Lewis (2002) gave participants a Navon task (Navon, 1977) after they had viewed a face, and before attempting to pick out the previously seen face from a line-up. The Navon task is a visual task where individuals view large letters which are constructed from smaller different letters (for example, a large letter T constructed from small N’s). The global Navon task consists of repeatedly identifying the large letter which induces a holistic processing mode. The local Navon task consists of repeatedly identifying the small letters from which the large letter is constructed and induces a featural processing mode.

Results from Macrae and Lewis (2002) showed that participants who completed a global Navon task in the time between viewing a target face and subsequent recognition test were significantly more accurate at face recognition compared to a control group of participants. Those participants who completed a local Navon task between viewing a target face and recognition test were significantly impaired in their ability to recognise the target face relative to the control group. These findings partially support the matching superiority hypothesis that better performance will be achieved if there is a match between cognitive processing state (holistic) and cognitive task to be performed (whole-face recognition), compared to other conditions where there is a mis-match. The difference in the Macrae and Lewis (2002) study was that processing was manipulated at retrieval and not encoding. To date, there are no published studies examining the effect of the global or local Navon task on face recall and subsequent facial composite accuracy using any composite construction system.

Other studies have examined the effect of inducing different types of cognitive processing at the retrieval stage of memory and found different results to Macrae and Lewis (2002). Berman and Cutler (1998) used judgements of personality, judgements of features and a control group who made no judgements about the faces they viewed prior to a face recognition task to assess their effect on face recognition. A 'featural inferiority' theory was proposed to account for the finding that the featural processing condition impaired face recognition performance relative to the holistic processing and control conditions. A further experiment by Berman and Cutler (1998) where processing was manipulated at both the encoding and retrieval stages of face recognition found that personality judgments at the encoding stage were beneficial for face recognition. There was no interaction between encoding and retrieval processing conditions, a finding which does not support the matching superiority hypothesis suggested by Macrae and Lewis (2002).

However, different methodologies between face processing studies make absolute comparisons difficult. In some studies participants are informed there will be a perception test to follow and this may influence their encoding strategy. Laughery, Duval and Wogalter (1986) found that participants tend to remember a face by studying the individual features of the face if they know a face perception task will ensue. This strategy may be detrimental to performance if the perception task is one of recognition of a previously seen face (Coin & Tiberghian, 1) but advantageous if the perception task is one of facial composite construction (Wells & Hyrciw, 1984).

Another methodological issue may be the different exposure times to the target face between studies, in that longer exposure times may lead to more successful subsequent recognition of the studied face. Longer exposure time could be confounded with encoding instruction as it arguably takes longer to make a personality judgement about a face than to focus on a feature (Bloom & Mudd, 1991). Additionally there is no way of ascertaining that asking an individual to encode a face in a certain way will actually lead to them doing so. If a person is asked to give personality ratings for a face they are looking at, there is no way of measuring which type of encoding has predominantly been used. One way to judge how a person encoded a face during a recognition study is to ask them to self-report how they did so, a strategy adopted by Olsson and Juslin (1999).

Olsson and Juslin (1999) asked participants who had watched a video clip of a staged crime to self-report the spontaneous encoding strategy they had used when viewing the face on-screen. Nearly two thirds of participants reported using a holistic encoding strategy when encoding the face in the clip, and a quarter reported using a featural strategy when encoding the face. Participants who reported using a holistic encoding strategy had more correct identifications and fewer false identifications when picking out the face they saw in a line-up, than those who reported using a featural encoding strategy.

Other studies which have found that character attribution may influence composite quality include Shepherd, Ellis, McMurran and Davies (1978) and Davies and Oldman (1999). Shepherd et al. (1978) reported that eyewitness-participants' beliefs about whether the person who was the target face was a murderer or local hero influenced the attractiveness ratings of the facial composites that they produced (using Photo-FIT). If the participants who constructed a Photo-FIT thought that the target face they viewed was a murderer, their composites were rated as being more unattractive than if they thought they were constructing a Photo-FIT of a local hero. However, no information about the accuracy of the facial composites produced was contained in their report. In a follow-up experiment, Davies and Oldman (1999) asked participants to construct facial composites (using E-FIT) of faces that the participants were known to either like or dislike, in order to check for the influence of liking and disliking on composite construction accuracy. They found that the faces of disliked characters were more accurately portrayed using E-FIT, a finding which fits well in an applied setting where a witness will arguably have similar feelings of dislike for a previously seen perpetrator of a crime. This finding was attributed to the fact that positive judgements may encourage more global evaluations of faces and, if this is the case, it might be predicted that positive attitudes toward a person would facilitate recognition but conversely impair recall.

2.3: The Distinction between Face Recognition And Face Recall

Davies and Christie (1982) argued that any facial composite construction system will only be successful insofar as it has the ability to tap into the human capacity for face recognition. This is because recall is effortful and slow, and decays rapidly over time, in contrast to recognition, which tends to be fast, automatic, and relatively stable over time. Construction of a facial composite by an eyewitness is not purely a face recall or a face recognition task, but essentially a combination of both recall and recognition if a witness is using the E-FIT system. However, the EFIT-V system has made the task of producing a facial composite one of mainly recognition if this is preferred by the witness. Davies, Shepherd and Ellis (1978) pointed out that in memory for faces recognition is a far easier task than recall. This is because to recall some information memory has to be searched, and then a decision has to be made on whether what is recalled is correct, whereas for recognition, memory search is not required. Information can be presented and it can be readily identified if there is a match in memory for that material. This has implications for any facial composite system where the witness must recall a face from their memory in order to construct a facial composite (as with E-FIT).

A further problem with long-term memory is that it tends to be general rather than specific, in that the gist of information tends to be remembered rather than perceptual detail, as this would be too much for memory to store and is generally unnecessary. Facial composites are not considered to be an exact likeness of a perpetrator's face. They are supposed to depict a 'type-likeness', a depiction which is accurate for more global aspects of a face, such as sex, age, hairstyle and colour, and whether the face is broad or thin. It is hoped that such 'type likeness' visual information can be combined with other information about a crime such as the geographical area in which the crime was committed, the time and date of the offence, and so on, to enable witnesses to make a connection between a facial composite and possible perpetrators of a crime. The EFIT-V system was designed to overcome the difficulties posed by remembering specific details of a face, particularly after a period of time has passed. However no research to date has studied the context in which E-FIT and EFIT-V composites are constructed, and whether cognitive style (trait) or processing (state) within individuals would have an effect on the accuracy of composites produced.

2.4: Processing Familiar and Unfamiliar faces

There is a difference in the way in which familiar and unfamiliar faces are processed and recognised. Bruce and Young's (1986) model of face recognition suggested that there are independent routes for processing familiar and unfamiliar faces. Familiar faces are recognised through a stored 'face recognition unit' which fires automatically when a familiar face is encountered. Unfamiliar faces are processed through early structural representations where facial information is encoded selectively.

When a facial composite is constructed the eyewitness has viewed an unfamiliar face - if the face was known to the witness there would be no need to construct a composite. Memory for familiar faces is strong, even in substandard viewing conditions such as poor quality video or pixelated faces, whereas memory for unfamiliar faces is poor (Hancock, Bruce & Burton, 2000). The relative importance of different areas of the face changes with familiarity of that face – faces that are familiar are better recognised by the internal features of the face (the eyes, nose and mouth) and facial expression is better recognised. Faces that are unfamiliar are better recognised by the external features of the face, the hairstyle, shape and size of the head (Young et al., 1987; Ellis, Shepherd & Davies, 1979). Bruce et al. (1999) found that the matching of unfamiliar faces is driven by external facial features.

Facial composites are constructed by people unfamiliar with the face they are constructing, in the hope that someone who is familiar with the target face will recognise them from the composite. Consequently, there might be an issue of mismatch between the areas of a face that people are looking at when they study a composite, and areas that others are likely to have concentrated on when making one. However, it has been found that even when someone constructs a composite of a face they are familiar with, the external features of the face are more accurately represented than the internal features of the face (Frowd, Bruce, McIntyre & Hancock, 2007). This suggests that in general, individuals are poor at representing the internal features of a face, whether it is familiar or not, and this poor performance is not a function of face familiarity. These results were found using PRO-fit (a similar featural type construction system as E-FIT), and it was suggested that the presence of the external features dominates the witness's perception and detracts from the selection of accurate internal features.

2.5: The Verbal Overshadowing Effect

There is a small but significant 'verbal overshadowing effect' (VOE) observed when people describe visual stimuli such as faces prior to identification. Schooler and Engstler-Schooler (1990) conducted a study in which participants viewed a target face in a videotaped scenario of a bank robbery. Following on from this, half the participants described the target face they had viewed and half did not. In a subsequent identification task using an 8-person line-up it was found that those who had verbally described the face were much less accurate at making successful identifications (38% accurate) than those participants who had not described the face after viewing the videotape (64% accurate). This indicated that there is something about the act of verbally describing a face that is detrimental to the subsequent identification of that face. Fallshore and Schooler (1995) attributed the verbal overshadowing effect to a 'transfer-inappropriate processing' shift; verbalising visual stimuli causes a shift to a featural style of processing which is at odds with the holistic/configural processing which is optimal for whole face recognition. There are three pieces of evidence that have been cited in support of this account of verbal overshadowing: firstly, the effect occurs even if the stimulus being described is not the target face (Dodson, Johnson & Schooler, 1997) therefore removing the possibility of a switch to verbal coding of the target face that interferes with the internal visual representation. Secondly, the effect is not observed in inverted faces, and finally the effect is not observed in other-race faces (Fallshore & Schooler, 1995). There has been a great deal of subsequent research looking at the processes involved in the VOE.

Finger and Pezdek (1999) found that the verbal overshadowing effect was strongest when a recognition test immediately followed a verbal description of a face (ten minute delay between description and identification task). Participants who gave a detailed description of a face during a cognitive interview were significantly poorer at subsequent face recognition (48% accuracy) than participants who had been given a standard police interview (73% accuracy). The difference between the standard and the cognitive interview is that in the cognitive interview questions are more open-ended; free recall and cued recall is used more extensively. The cognitive interview also includes elements such as reinstatement of context and recalling events in different orders. It was found that if the recognition task immediately followed the verbal description, then recognition was poorer because participants were relying on their verbal descriptions when attempting face recognition, and this was interfering with the stored visual memory of the face. Introducing a delay of one hour was sufficient for face recognition in the cognitive interview condition to be increased significantly.

The fact that the one hour delay was successful in mediating the effect of the VOE is support for the idea that it does not matter if more incorrect details are elicited during the description phase, these do not affect the ability to recognise the face at a later date, given a time delay between description and identification task. The VOE therefore seems to be mediated by a delay between verbal description of a face and a face recognition task. When a witness is constructing a facial composite using E-FIT there is no delay between description and composite construction. How verbal overshadowing may relate to composite construction where, within the E-FIT system at least, witnesses are required to describe the features of the face in order to produce an initial composite remains to be investigated. Brown and Lloyd-Jones (2002) found that the verbal overshadowing effect was stronger in participants who were encouraged to focus on describing specific features of a face, and weaker when participants were not explicitly encouraged to describe features.

Brace, Pike, Allen and Kemp (2006) investigated the role of the verbal component of constructing an E-FIT by requiring some participants to build a composite with an operator and by requiring operators to build composites alone, therefore removing the verbal element of the process. In the condition where composites were constructed from memory using a describer and an E-FIT operator, composites were rated significantly less similar than composites in all other conditions. This indicates that there could be a verbal overshadowing element in describing a face from memory during composite construction. Memory did not seem to play a role in the E-FITs which were constructed by operators without a describer, as they were rated similarly whether the E-FIT was constructed in the presence of a photograph of the target face or completely from memory. The finding that verbalisation has an effect on facial composite construction has implications for the way in which the E-FIT system is used, as verbalisation is an element which cannot in practical terms be removed. However with EFIT-V there is a considerably reduced verbal element to composite construction. Following an initial very brief description of global factors such as sex, race and age of the target face, the witness can proceed to construct an EFIT-V composite with no further description required.

2.6: Summary

Research on face processing and recognition suggests strongly that faces are generally processed in a holistic way: we see faces and represent them in memory as a perceptual gestalt from which it is difficult to extract featural information. Whole faces are more than the sum of their individual features, and familiar faces can in fact be recognised easily even when the features are blurred or obscured in some way.

Research has also demonstrated that the way in which faces are encoded and processed is influenced by the information processing task to be performed. This is related to the subsequent ability to recall/recognise them. Faces which are encoded in a holistic way using judgements of personality are subsequently better recognised, and faces which are encoded using featural judgements are better recalled for the purpose of facial composite construction, at least when using featural systems such as Identikit and Photo-FIT (Wells & Hyrciw, 1984).

The recognition of faces is an easier cognitive task than the recall of faces, and this has implications for the ability to produce an accurate facial composite of a previously seen face. Unfamiliar faces are better recognised by the external features of the face, therefore the task of facial composite construction where an eyewitness is required to recall *featural* details of an *unfamiliar* face, is an extremely difficult one.

However individuals differ in the degree to which they generally rely on holistic and featural processing, and these differences might be a contributory factor in the individual differences in performance observed in facial composite studies (see section 3.3). Chapter Three describes research on facial composite construction systems, with emphasis on two of the main systems in use in the UK today. The issue of individual differences in cognitive style is addressed in Chapter Four.

Chapter Three: Composite Construction Systems

This chapter describes the development of facial composite systems and how this has been informed by theories of face processing. Psychological research has suggested that there is a mismatch between the way in which faces are processed (holistically), and the way in which facial composite systems are implemented (featurally). Therefore research has been driven by attempts to lessen the impact of this mismatch by developing composite systems which capitalise on the greater capacity for whole-face recognition as opposed to the more difficult cognitive task of face recall, and in particular recalling individual features of a face. The following sections briefly describe early facial composite systems and research which led to the development of the present day systems which are the subject of investigation within this thesis: E-FIT and EFIT-V.

3.1: Sketch Artist

Prior to the introduction of composite systems the most widely used way of constructing a facial likeness was with the police sketch artist (Frowd et al., 2005a). This person was a skilled portrait artist who would draw a pencil image of a perpetrator's face with the help and guidance of the eyewitness. Although this method enjoyed some success and is still used to some extent today, composite systems were created because the sketches required specialist artistic skills, and a more readily accessible method of face construction was required for police use (Ellis, 1986). Composite construction systems required less skill and training to implement, and offered a method of both increasing the number of and standardising the facial composites produced by police forces.

3.2: Identikit

Identikit was introduced in 1959 as an alternative to the police sketch artist. This consisted of transparent celluloid sheets of line drawings of parts of the face which could be assembled to produce a whole face image. The individual features of the face were initially presented in a booklet which witnesses viewed in order to select the features they required to build up a whole face image. The completed face image could be modified by using a pencil to draw over the finished composite.

Laughery and Fowler (1980) conducted a study which directly compared composites produced using the Identikit system with likenesses created by the sketch artist technique. Facial composites were made with both systems in conditions where the target face was in view and where the target face was not in view. Sketch artist composites received significantly higher accuracy ratings than Identikits, and all composites created with the target face present were higher rated than composites created from memory. It would be expected that when constructing a composite with the target face on view to the witness, that the resulting composite would be more similar to the target face it was supposed to depict. However, within the Identikit system, there was very little difference in the ratings of composites made from memory and when the target face was in view, which suggested that the system was limited in the ability to produce accurate composites. A major restriction for Identikit was the limited number of suitable features with which to construct an accurate facial composite.

Laughery and Fowler (1980) also found a positive correlation between scores on Gordon's (1949) Test of Visual Imagery Control (TVIC) and accuracy ratings for composites constructed from memory using sketch artists. The TVIC is designed to measure the ability to manipulate or control mental images, irrespective of their vividness. This fits in well with further data in the study which showed that participants who produced composites using sketch artists spent much more time moving between features and returning to features, rather than working on completing one feature at a time before moving on to the next. It was suggested that the moving around between features may result in better relationships between features, which are as important as the features themselves for face recognition (Sergent, 1984). Laughery and Fowler (1980) also found a positive correlation between scores on the verbal element of the Scholastic Aptitude Test (SAT) and accuracy ratings for Identikit composites created from memory, a result which is harder to interpret.

3.3: Photo-FIT

The Photo-FIT system invented by Penry was first used in Britain on a trial basis in 1969 (Ellis et al., 1975). Photo-FIT differed from the original version of

Identikit, in that black and white photographs of individual facial features were used instead of line drawings. The Photo-FIT system consisted of numerous examples of the forehead and hair, eyes (with eyebrows), nose, mouth and chin arranged in a booklet from which eyewitnesses chose the closest likeness to their memory of a face. These individual features were then placed together to make a face.

A great deal of research was conducted into the efficacy of the Photo-FIT system as a tool for composite production, and several implications for the psychological limitations of face recall were identified. In practical terms, Photo-FIT appeared to suffer from the same inherent limitation as had been found in Identikit. An early survey of police officers' experience of the Photo-FIT system suggested that the range of features offered was potentially inadequate (King, 1971).

One of the first lab-based studies of Photo-FIT was conducted by Ellis et al., (1975) who examined participants' ability to reproduce target faces which had been constructed using the photographed features available within the Photo-FIT booklet. This made it possible in theory to produce an exact match of the target face as seen by witness participants. A comparison was made between a target-present condition in which participants viewed the target face throughout composite production, and a target-absent condition in which participants constructed a composite from memory. Results indicated that participants had difficulty in making an accurate reconstruction of a face even when the face was present throughout the construction, and the exact features were available within the Photo-FIT kit to do so. The fact that participants could not construct a perfectly accurate reproduction of a face, even when the face was present illustrated a further limitation of systems that use books of different features that witnesses must select from prior to constructing a facial composite. Namely that selecting facial features in isolation (picking them out from an array of similar looking features in the Photo-FIT booklet) is an extremely difficult cognitive task, and highly prone to errors and mis-judgements. This may be because features are usually perceived in the context of a whole face, and how each feature is perceived depends partly on its spatial relation to other features in the face (Tanaka & Farah, 1993).

Ellis et al. (1975) also reported that participants who constructed a face from memory tended to produce even less accurate Photo-FITs. However, the difference in ratings given for target-present and target-absent Photo-FITs was not statistically significant. This is surprising given that when participants made a composite from memory, they were only allowed 10 seconds in which to view the target face. The absence of an effect of availability of the target might reflect the difficulties that participants encounter when attempting to select features in isolation.

A second experiment within the study by Ellis et al. (1975) using ‘poor’ and ‘good’ participants (as denoted by the ratings their first composite received) confirmed that composites made by the participants whose first composites received high ratings were successfully identified more often than those made by the participants whose first composites had received low ratings. The accuracy of composites produced was stable over two Photo-FIT attempts in this study, lending weight to the idea that, notwithstanding limitations in composite systems, witnesses differed in their ability to overcome these limitations and produce accurate composites. Ellis et al. (1975) suggested that the cognitive style of field dependence/independence might account for the different abilities of eyewitnesses to produce accurate facial composites. They did not test the cognitive style of the participants in their study.

Christie and Ellis (1981) compared Photo-FITs of faces to verbal descriptions of faces and found the verbal descriptions to be significantly more accurate than the Photo-FITs. Even descriptions made from memory were judged to be more accurate than Photo-FITs made with the target face present. This suggested that the limitation in the effectiveness of Photo-FIT constructions was not due to a limitation in the descriptive abilities of participants. McQuiston-Surrett and Topp (2008) found a similar result, namely that descriptions of faces were more useful for identification of faces than facial composites.

The generally poor quality of facial likeness found in the above studies was attributed to the imprecision of the Photo-FIT system itself, which prevented the accurate physical representation of a facial image (Ellis et al., 1975; Christie & Ellis, 1981). There was no correlation between the accuracy of the verbal description and the accuracy of the composite that participants produced. This contrasts with the findings of Laughery and Fowler (1980) using Identikit, where higher rated verbal ability correlated with higher rated composites. There was no correlation between the amount of time spent constructing a composite and accuracy rating of that composite, indicating that time is not a factor in the production of accurate facial composites using Photo-FIT. However, there was a correlation between how long a participant took to verbally describe the face, and the accuracy of that description. In addition to highlighting the major limitations of early composite construction systems, the studies above indicate there may be a variable verbal ability element within individuals which may account for some of the variation found in the accuracy of facial composites that individuals can produce.

According to Christie and Ellis (1981) the feature by feature approach to face construction may be inappropriate if the face is perceived and stored as a gestalt, that is a whole entity from which it is difficult to perceptually extract featural information (see section 2.1). This approach may be particularly problematic when alternative features are looked at in isolation, and it was suggested that the development of an accurate facial composite system is "likely to be measured by the degree to which face recall can be made to resemble face recognition" (Davies & Christie, 1982, p. 108). Davies, Ellis and Christie (1981) sought to explore whether allowing witnesses to construct a face by being presented with a whole face in the first instance, and then working to amend the face accordingly would be better than the fully featural approach. It was found that this made no difference to the accuracy of Photo-FIT composites produced. Allowing a witness to begin with a preferred face 'type' produced similar results, no advantage over the traditional Photo-FIT method. Further limitations of the Photo-FIT system included the fact that it was virtually impossible to change the shape of the constructed facial image, and the presence of lines over the finished image acted as a distracter when attempting to use the image for identification purposes (Ellis, Davies & Shepherd, 1978).

Technological advances allowed for the development and introduction of computerised composite building systems. The most well known of these is E-FIT (Electronic Facial Identification Technique) (Aspley Limited, 1993) and is the most commonly used system both in the UK and several countries worldwide. In the last two to three years, a newer composite system called EFIT-V (Electronic Facial Identification Technique – Volume) (Visionmetric Ltd, 2004) has been adopted by around 20 UK police forces, and both of these systems are used in the research described in this thesis. The following section describes some research that has been conducted with E-FIT and with EFIT-V and highlights the crucial differences between the two systems in terms of the cognitive task required of the witness using the systems to construct a facial composite.

3.4: Theory behind the development of E-FIT

According to Christie, Davies, Shepherd and Ellis (1981) a number of law enforcement agencies were drawn to the idea of using new computer graphics techniques to produce a more accurate facial likeness than previous systems could achieve. The British Home Office sponsored the development of a computerised composite system called E-FIT.

E-FIT was introduced in the UK in the early 1990's and was a computerised alternative to the old Photo-FIT system. The E-FIT system was designed in order to address some practical limitations which had been identified within Photo-FIT, and also to incorporate the psychological theory of holistic face processing. Eyewitnesses were still required to select individual features of the face in order to produce a composite image but the new E-FIT system allowed the features to be viewed within the context of the whole face, and not in isolation. Davies and Christie (1982) found that selecting features in isolation and not as part of the whole face within the Photo-FIT system, led to lower accuracy ratings for that feature. The E-FIT system therefore had both practical and theoretical advantages over the Photo-FIT system.

Firstly, unlike Photo-FIT, there were no lines on the finished E-FIT image; the resulting facial composite was thus more realistic and should have been more easily recognisable, as the presence of lines on a Photo-FIT image was found to be detrimental to face recognition (Ellis et al.,1978). Secondly, E-FIT contains a far greater number of photographed features for hair, face shape, eyes, eyebrows, mouth and ears, and early research on the Photo-FIT system suggested that low number of features could be a problem (Ellis et al., 1978). These features are selected from a large database of individual pictorial facial features and assembled for viewing as a whole face. Thirdly, features are not viewed in isolation as they were in the Photo-FIT system, but are first described and then finally viewed within the context of a whole face.

Furthermore, the selected features were able to be resized and/or repositioned as required. If alternative features are required, these are also selected within the context of the whole face. Finally, the image can be amended further using a software drawing package to enhance the image, such as by adding facial marks. Therefore the E-FIT system allows for the production of a greatly increased number of different faces, having a much larger library of features than previous systems such as Photo-FIT and also having the ability to easily perform subtle changes to any given feature including face shape.

3.4.1: E-FIT Research Findings

Research on the efficacy of the E-FIT system can be divided into two main categories. Firstly, as E-FIT was designed to capitalise on the holistic way in which faces are processed and represented in memory there are comparisons with older composite systems such as Photo-FIT to assess the relative merits of these different systems. Secondly there is research which concentrates on ways in which use of the E-FIT system or individual witnesses can be manipulated in order to improve the accuracy of facial composites that are produced.

3.4.2: Research comparing E-FIT with earlier facial composite systems

Despite the technological advances employed in the production of the E-FIT system itself, early research suggested that the facial composites produced with E-FIT were no more accurate than those produced by the old Photo-FIT system. There was only an advantage for composites made with E-FIT when they were made with the target face present during composite construction, which demonstrated E-FIT's superiority in terms of range of available features and ease of manipulation of those features.

Davies, van der Willik and Morrison (2000) conducted a study comparing the E-FIT system with Photo-FIT. Participants constructed a total of four composites. Two were constructed using Photo-FIT with the target face both absent and present, and two were constructed using E-FIT with the target face also both absent and present. Two target faces were familiar to participants and two were unfamiliar faces. Results showed an advantage for E-FIT over the Photo-FIT system in only one of four conditions, where the target face was familiar and when a photograph was present. This advantage in the least ecologically valid condition was interpreted by Davies et al. (2000) to indicate that E-FIT was in practice no better at producing a good likeness of a target face from memory than its predecessor the Photo-FIT system.

However, the similarities in composites produced with the E-FIT and Photo-FIT systems could possibly be attributed to a number of methodological issues within the study itself. There was no artistic enhancement of the composites produced by participants but research has shown that artistic enhancement improves likeness quality of images produced (Gibling & Bennett, 1994). This coupled with the potential for artistic enhancement that the E-FIT system enjoys with the use of computer graphics packages may have had an impact on the results found. Furthermore witness-participants were allowed a time limit of only twenty minutes to complete each composite. This time limit could have masked any differences in composite quality that may have emerged if the participants were allowed all the time they needed to construct a facial composite. Finally, the naming task that was used to evaluate the composites used multiple composites of each target face, potentially elevating naming rates.

Frowd et al. (2005b) conducted a comparison of several composite building systems including both E-FIT and Photo-FIT. This study addressed many of the ecological validity issues inherent in the Davies et al. (2000) study. Participants constructed only one composite each with no time limit, using a cognitive interview (standard police practice) and a delay of between 3 and 4 hours between viewing the target face and constructing a composite (a 3-4 day delay would be the norm in practice). They found that E-FIT performed the best in terms of naming of the composites (around 20% correct), together with PRO-fit, (a system which is very similar in composite construction method to E-FIT, except for the fact that PRO-fit contains an internal artistic enhancement package, whereas E-FITs are exported to an external package such as Picture Publisher for artistic enhancement).

Photo-FITs were correctly named around 5% of the time within the study and their poorer performance could not be attributed to distracting boundary lines on the final composites because these were removed electronically “to allow a fairer comparison with other techniques” (p. 41). Therefore, under conditions where one facial composite was constructed from memory with no time limit, E-FIT composites performed significantly better than Photo-FITs in both a naming and a sorting task. A supplementary sorting task revealed a similar pattern of results with both E-FIT and PRO-fit composites being successfully matched around 75% of the time, compared with a 50% match rate for Photo-FIT.

The cognitive interview immediately prior to composite construction in the study by Frowd et al. (2005a) would potentially have had the effect of inducing featural processing. This is because during the cognitive interview the participant is encouraged to recall as many details about the face as possible. Describing individual facial features would arguably induce a featural or piecemeal way of trying to remember the to-be-described face. These are then repeated back feature by feature in their own words as a form of cued recall. It might be expected therefore that a more featurally based system such as Photo-FIT would benefit from the participant being primed towards featural processing prior to the composite task, but this was not the case. This finding that higher rated E-FITs were made after featural processing was arguably induced (by using a cognitive interview prior to composite construction) therefore indicate that the switch to a more holistic based system such as E-FIT is not detrimental in terms of allowing witnesses to produce more accurate facial composites. However, it may also demonstrate that although the E-FIT system is based on holistic processing principles, there is still a strong featural element to composite construction within a relatively more holistic system such as E-FIT.

A study was conducted by Frowd, et al. (2005a) which compared E-FIT (and PRO-fit) with artist sketches and EvoFIT (a whole face system being developed by the authors of the paper) in which the method of constructing composites was very similar to the methodology used by Frowd, et al. (2005b). The only difference in methodology was a two day time delay between the viewing of the target face and construction of the facial composite. Results of a naming task showed a floor effect, with naming rates at a mean level of 2.8% correct across all systems. This was attributed to the two day delay between viewing of the target face and construction of the facial composite and accounted for by a cognitive processing shift within that two day period. It was suggested that after a two day delay, even if participants had encoded the faces in a featural way, their memory of the face would be more of an overall impression (a gestalt) as they would have switched naturally over time to a holistic method of processing the face.

The supplementary sorting task used showed statistically equivalent sorting rates for all composite systems apart from the sketch artist, whose sketches were sorted significantly more often than composites from any composite construction system. Therefore, although naming may be considered to be the most forensically valid method of assessing the efficacy of facial composites, perhaps the very low naming rates within studies suggest that it is not an ideal experimental method for assessing differences between controlled conditions or composite construction systems.

Another method of assessing the efficacy of facial composites is the sorting task. This is a forced choice task where participants are required to match facial composites they are shown to target faces which are displayed simultaneously. The sorting task clearly elevates levels of correct identification (Frowd et al. (2005) where Photo-FITs received 50% successful identification rate compared to a 5% correct naming rate), and although participants are not always told how many composites should be placed with each target face, in a completely target-present line up of faces, it would seem intuitive that the composites would be spread fairly equally among the available target faces. This being the case it could be argued that the composite sorting task is also not an ideal method for evaluating facial composites in experimental conditions. This is because when the most accurate composites have been sorted by participants the degrees of freedom are reduced for subsequent matching of the remaining poorer quality facial composites. This increases the possibility of matching the poorer quality facial composites to the original target face and does not provide a measure of the relative accuracy between individual facial composites. In addition, the number of correct matches for each composite to the target face does not give a large enough range of scores to appropriately assess the differences between experimental groups in the same way as subjective likeness ratings.

3.4.3 Research on ways in which use of the E-FIT system or individual witnesses can be manipulated

In addition to research which compares the relatively more holistic E-FIT system with its primarily feature based predecessors, there has also been research looking at ways in which procedures with the system and with eyewitnesses can be modified in order to produce composites of a higher standard of similarity to a previously seen face. This sub-section focuses on research which has attempted to manipulate the cognitive processing set of individual participants in order to influence the quality of facial composite they produce.

Creating a facial composite using any system involves the different cognitive processes of recall and recognition, and there have been found to be different circumstances in which either can be enhanced prior to composite construction. However, these manipulations sometimes counteract each other. For instance, face recognition can be enhanced by holistic encoding of a face, but recall can be enhanced by featural encoding of a face. It has been found that recognition is enhanced by viewing facial features within the context of a whole face shape, something which E-FIT was specifically designed to capitalise on (Tanaka & Farah, 1993). Face recognition is an easier cognitive task than face recall; therefore it would seem likely that a system based on recognition such as EFIT-V would be easier for eyewitnesses to use.

Frowd et al. (2008) conducted a study which investigated whether inducing both featural and holistic processing at different points in the composite construction process would lead to the construction of more accurate facial composites. This was achieved through the development of a ‘Holistic Cognitive Interview’ (H-CI) which was compared to the standard Cognitive Interview using PRO-fit, a computerised system which works in a similar way to E-FIT. Participants viewed a video clip of the target face which lasted for less than one minute. Three to four hours later participants then provided a detailed description of the features of the face using the standard Cognitive Interview. Half of the participants then proceeded to construct a facial composite using PRO-fit, the other half were subjected to a five minute Holistic Cognitive Interview in which they were invited to think about aspects of the personality of the face they had viewed. This was followed by prompting the participants to make a series of judgements (low/medium/high) on seven holistic (personality) aspects of the face.

The five minute holistic intervention prior to constructing a facial composite was designed to 'switch' the mode of processing within the participants from a featural one which would have been induced during the standard Cognitive Interview to a holistic mode more beneficial for face recognition. The holistic cognitive interview arguably works in a similar way to the global Navon task. The finding that the global Navon task leads to better face recognition (Macrae & Lewis, 2002) was attributed to the inducement of a holistic cognitive processing style similar to the cognitive processing state that is used for processing and remembering faces.

Facial composites in the Cognitive Interview condition were correctly named by 9% of participants, and in the Holistic Cognitive Interview condition were correctly named by 41% of participants. This finding suggests a clear advantage for inducing holistic processing prior to construction of a facial composite. Female faces were named significantly more often than male faces, although this may be due in part to female faces being more identifiable because females tend to have a wider range of hairstyles than males, and hairstyle may have made the females in the study more distinctive in appearance.

3.4.4: Summary of E-FIT research findings

Research with the E-FIT system has followed two main strands, comparison with other composite systems and manipulation of system variables in an attempt to create more favourable conditions for eyewitnesses to produce accurate facial composites. Early research found comparable levels of performance between E-FIT and manual featural systems such as Photo-FIT (Davies et al., 2000) but this research was methodologically flawed. Subsequent research which addressed some of these limitations revealed a significant advantage for E-FIT over Photo-FIT with a 20% naming rate for E-FIT as opposed to a 5% naming rate for Photo-FIT (Frowd et al., 2005b). Research into system variables has shown that varying the cognitive interview procedures witnesses are exposed to has an impact on the quality of the composites they produce. Using a standard cognitive interview (inducing featural processing) prior to face description, and then requiring the witness to give personality judgements (inducing holistic processing) prior to composite construction produced the most accurate facial composites (Frowd et al., 2008).

3.5: Theory behind the development of EFIT-V

The variability observed between eyewitnesses in their ability to produce accurate facial composites has been attributed in the literature to a mis-match between the cognitive demands of producing a facial composite and the holistic way in which faces are processed and remembered (Wells & Hasel, 2007). New facial composite systems such as EFIT-V (Gibson, Solomon & Pallares-Bejarano, 2003) have attempted to address this mis-match by allowing the eyewitness to construct a composite using only whole faces rather than by a semi-piecemeal construction method. EFIT-V therefore was developed so that there could be holistic face matching procedures at the composite construction stage to match the holistic processing of faces at the face encoding stage. This follows on from Tulving and Thomson’s (1973) encoding specificity principle that memory performance will be enhanced if conditions are the same at encoding and retrieval. Also, face recognition is a far easier cognitive task than face recall. Therefore witnesses should find it easier to state which whole face is most recognisable to them when presented with several faces to choose from on screen, than to construct a face in a piecemeal way feature by feature as with the E-FIT system.

3.5.1: Research findings on holistic facial composite systems

There has been only one published study using EFIT-V, as it is a relatively new composite construction system. Valentine, Davis, Thorner, Solomon and Gibson

(2010) used EFIT-V in a series of experiments which investigated the usefulness of morphing composites made by four separate individuals, and compared them to composites which were morphs of 4 separate composites made by one individual. Therefore they compared between-witness morphs with within-witness morphs which were created both from memory and while the target face was in view. The composites were ranked for similarity to the target face and between-witness morphs were judged to be the most similar to the target face. However this could be confounded by the fact that the between-witness morphs combined 2 composites that were made from memory with 2 composites that were made while the target face was in view. Between-witness morphs were more similar to the target face than within-witness morphs which in turn were more similar than individual composites.

Surprisingly, individual composites produced from memory were ranked as better likenesses than individual composites produced while the target face was in view. This was attributed to the fact that participants working with a face in view would be concentrating more on individual features of the target face, which would not be beneficial when working with a system such as EFIT-V which exploits holistic processing of faces, and that holistic processing would be more prevalent in witnesses constructing a composite from memory. When naming was used as a measure of composite quality, five participants (0.8% of the total) gave the correct name of the person depicted in the composites without any cues to naming. When cues to naming were introduced this rose to around 20% of participants correctly naming individual composites, 44% naming between-witness morphs, and 32% naming within-witness morphs.

Between-witness morphs would be expected to be better likenesses than within-witness morphs if one assumes that the mistakes made between-witnesses would not be correlated, whereas the same witness making several composites might be expected to repeat their own mistakes. Finally the study replicated the finding that the external features of the composites, when presented in isolation, were more recognisable as corresponding to the original target face than internal features presented in isolation, or indeed complete composites. This finding demonstrates that external features of unfamiliar faces are more salient in perception than internal features. It also implies that when whole faces are presented, the lack of similarity of the internal features meant that composites were judged as less similar than if the external features alone were presented.

There has however been some research conducted on a system called EvoFIT (Frowd, Hancock & Carson, 2004) which works on the same 'whole face' principle of allowing witnesses to view a selection of whole faces on screen as opposed to selecting individual features of the face. EvoFIT differs slightly in that 18 faces are presented on screen at any one time (compared to nine faces for EFIT-V) and witnesses choose more than one face, these are then bred together to form a new set of faces. The breeding of faces selected for their similarity to the target face works along the theory that multiple composites, even if slightly wrong, when morphed together will produce something that is increasingly similar to the target face.

Frowd, et al. (2007b) manipulated the encoding of faces and the cognitive processing style of participants prior to constructing a composite using EvoFIT, and two different versions of PRO-fit in a series of experiments. Participants were instructed to attend to the individual features of a target face followed by a cognitive interview (both of which promote featural processing), while further participants were instructed to attend to personality judgements of a target face followed by personality ratings during a holistic interview (promoting holistic processing). Following on from this, all participants constructed a facial composite using the EvoFIT system and a version of PRO-fit (called parallel PRO-fit) where participants could view six whole faces on screen at a time, where only one feature was different between each face (e.g. six identical faces with six different hairstyles, and participants choose the closest hairstyle to their memory).

Frowd et al. (2007b) predicted that as holistic encoding and a holistic interview is beneficial to face recognition, that this would be the best condition for eyewitness participants to produce an accurate composite using a whole-face system such as EvoFIT. This is because with a whole-face based system, witnesses do not have to verbally describe the different features of the face, and do not have to mentally generate or *recall* the face they are attempting to construct, they just have to recognise something similar when they see it on screen.

However it was found that featural encoding led to more accurate composites as assessed by a sorting task, a finding that was contrary to expectation using a holistic system such as EvoFIT. Overall, it was also found that although encoding was a significant factor in all instances, the type of interview conducted prior to construction did not have an effect. Since featural encoding has such a large influence on composite construction, this study was not a direct appraisal of the effect of a holistic interview prior to composite construction, as the holistic interview condition followed on from a holistic encoding condition. More recent research demonstrated that the administration of a cognitive interview designed to aid recall of individual features, followed by a holistic interview designed to induce holistic processes that complement face recognition processes led to more accurate facial composites using EvoFIT (Frowd, et al., 2012).

3.5.2: Summary of research findings on holistic facial composite systems

The only study with EFIT-V to be published to date revealed that morphing composites from four different witnesses produced better likenesses than morphing composites from four separate attempts from the same witness. However, morphing several composites produced by the same witness produced better likenesses than any individual composite (Valentine et al., 2010). It seems unlikely that in reality a witness would be asked to produce four facial composites for the purpose of blending them into one, as this would be likely to place huge stress on the witness.

Research on another whole-face based system (EvoFIT) demonstrated that the way in which faces are encoded seems to play a role in facial composite construction (Frowd et al., 2007b). Featural encoding of faces is superior to holistic encoding of faces when constructing facial composites, even when the task is completed using a system where whole faces are viewed and no features are looked at in isolation.

3.6: General Summary

The development of facial composite systems over time has taken into account the relative ease with which faces are recognised, even though the process of creating a facial composite also involves face recall. Even under optimal conditions where composites are constructed by experienced operators with a target face in view performance appears limited to around 20% accuracy for naming the target face with any composite system. Efforts to improve and modify facial composite systems have taken into account the fact that there are psychological constraints on the ability to recall specific features of a face. However research on whole-face based systems has yet to categorically demonstrate that they produce more accurate facial composites than feature-based systems.

Individuals differ in their propensity to perceive visual stimuli in a holistic or featural manner, and this may impact on the way in which they process faces and construct facial composites. Additionally, there are no published studies to date which have investigated the effect of temporarily switching mode of cognitive processing using the Navon task with participants on the accuracy of the facial composites they produce. Studies One and Two of the empirical work presented in this thesis address the issue of inherent differences in the propensity to perceive stimuli in a holistic or featural manner, and their impact on accurate facial composite construction using E-FIT. Study Three addresses the impact of inducing cognitive processing style using the Navon task on facial composite construction using E-FIT. Study Four investigates the effect of the Navon task on facial composite construction using EFIT-V.

The next chapter on cognitive style raises the issue of individual differences in cognitive style and the possible interaction of cognitive style with face processing, recall and recognition, with emphasis on perceptual cognitive styles which may impact on the use of featural based and whole-face based composite construction systems.

Chapter Four: Cognitive Style as a Stable Trait

The previous chapters described the ways in which faces are generally processed, and research on the two main facial composite systems in use in the present day by police forces in the UK; E-FIT and EFIT-V. There is much evidence to suggest that in the main, faces are processed in a holistic manner and this may have an impact on the production of facial composites by eyewitnesses. This is because the way in which faces are generally processed is suggested to be at odds with the featural way in which the E-FIT program is implemented with eyewitnesses in the production of facial composites.

The present chapter introduces the concept of cognitive style, which is the way in which individuals prefer to process and organise information. The term cognitive style was first suggested by Allport (1937) to describe an individual's typical mode of problem-solving, thinking, perceiving and remembering. Within the first few decades of research on styles up until the 1970's, a large collection of styles and ways of measuring them were proposed by different theorists. The number of cognitive style labels identified within the psychological literature has led to a lack of coherent theory within the area of cognitive style research (Rayner, 2011). However, there are four major integrative models which have attempted to unify all cognitive styles which have been identified (Zhang & Sternberg, 2005). The most influential of these models, and the one that has generated the most subsequent research is that of Riding and Cheema (1991) who suggested that all cognitive styles could be subsumed within two major style constructs, the holistic/analytic and the verbaliser/imager styles. According to Zhang and Sternberg (2005) there is good supporting empirical evidence for the two style dimensions. Field dependence/independence is suggested to be subsumed within the holistic/analytic cognitive style family, although this is an area of debate among cognitive style researchers (Kozhevnikov, 2007).

The field dependence/independence cognitive style measured in Study One and the holistic/analytic cognitive style measured in Study Two within this thesis are suggested to indicate the basic visual perceptive preferences of individuals. The verbaliser/imager cognitive style measured in Study Two is thought to represent the way in which information is represented in memory, either predominantly in word or picture form. Interest in cognitive style is low among cognitive researchers, many of whom believe that their influence is overshadowed by other factors such as general ability and cognitive constraints all humans have in common. However, according to Kozhevnikov (2007), in numerous applied settings cognitive style can be a better predictor of performance than intelligence or situational factors, (particularly in the fields of education and industrial/organisational psychology).

At the most basic level, the holistic/analytic and field dependence/independence cognitive styles are associated with the visual perceptual preferences of an individual and, as such, may be an influential factor in the way faces are perceived by individuals, how they are represented in memory, and how an individual interacts with facial composite construction systems. The verbaliser/imager cognitive style may be an influential factor in facial composite construction, a process which contains elements of both verbalisation and visualisation.

Miller (1987) developed another influential model within which to describe all cognitive styles and grouped holistic/analytic styles according to the cognitive processes of perception, memory and attention. According to Miller different cognitive styles can be identified within these different cognitive stages, and all cognitive styles are subsumed within the holistic/analytic dimension. Within this model both holistic/analytic and field dependent/independent cognitive styles are identified at the visual *perception* stage of processing. The main difference between the two dimensions of cognitive styles is that holistic/analytic relates to pattern recognition and part/whole relations whereas field dependence /independence relates to selective attention. There is however a continuing debate within the literature about whether different style labels represent different style constructs or are similar constructs but with different words to describe those styles (Zhang & Sternberg, 2005). Both of these styles may therefore be important mediators in the accuracy of facial composites that individuals produce, as the way in which faces are perceived/encoded has an influence on the ability to recognise/recall them.

The present chapter describes research on the Field Dependence /Independence (Witkin et al., 1971) Holistic/Analytic, and Verbaliser/Imager (Riding & Cheema, 1991) cognitive styles. Consideration shall be given as to how these cognitive styles may relate to face perception and recall for facial composite construction.

'Cognitive style' has sometimes been used interchangeably with the term 'learning style' because the main focus of 'style' research has been concerned with the implications for and applications relevant to education and training. Riding and Cheema (1991) differentiated between the two terms and suggested that learning style is a fluid concept related to individual strategies used for learning which can be adapted and modified. By contrast, cognitive style is considered to be a fairly constant and fixed characteristic of an individual. It is suggested that cognitive style is a stable trait which interacts with personality and situational factors. However, more recent research has demonstrated that an individual's cognitive processing mode can be temporarily manipulated to produce memory performance typical of the induced mode, although the mechanism behind the observed effect is still a matter of debate (Perfect, Weston, Dennis & Snell, 2008). There is controversy surrounding the issue of whether cognitive styles can be considered to be fixed traits or temporary states, possibly because they have been found to interact with personality and situational factors (Zhang & Sternberg, 2005).

In a review of the literature on different cognitive styles, Riding and Cheema (1991) concluded that all of the styles identified could be grouped into two principal cognitive style dimensions; the holistic/analytic and verbal/imagery styles. The holistic/analytic style relates to how an individual habitually processes information either in wholes or parts, and the verbal/imagery style relates to how an individual tends to represent information in memory and during thinking, whether primarily in words or pictures. The first two empirical studies presented in this thesis investigated these two major cognitive style constructs and their possible role in facial composite construction. In addition to this the field dependence/independence style (Witkin, 1950) which is suggested to reside within the holistic/analytic category of cognitive styles was assessed for any possible impact on facial composite construction.

4.1: Field Dependence/Independence

The field dependence/independence (FDI) cognitive style used in Studies One and Four reported in this thesis was first identified by Witkin (1950), and can be used to distinguish an individual's preferred style of processing visual stimuli. Ellis et al. (1975) were among the first researchers to note that individuals differ quite considerably in their ability to build an accurate facial composite from memory. They suggested that FDI might be a witness variable likely to influence the ability to construct target-accurate facial composites. There are no published studies to date which assess the influence of field dependence /independence on the accuracy of facial composites that eyewitnesses produce.

FDI is measured by performance on an Embedded Figures Test, which places individuals on a value-free continuum running from extreme field dependence at one end to extreme field independence at the other. Scores on the Embedded Figures Test range from zero to eighteen, and some individuals' scores fall in the middle of the scoring scale indicating no preferred field dependent or field independent visual processing style. The majority of research that has used the Embedded Figures Test has performed a median split on test scores to identify individuals as either field dependent or independent (Emmett et al., 2003). The two main tests for FDI are the 'Rod and Frame' test (measuring perception of the true vertical or horizontal) and the '(Group) Embedded Figures Test' (Goldstein & Blackman, 1978), which involves the dis-embedding of a shape from its surrounding context (see Chapter Six, Section 6.5 for a description of the administration and scoring of the Group Embedded Figures Test, or GEFT).

Field dependence/independence is a category of style said to be subsumed within the holistic-analytic group of cognitive styles as identified by Riding and Cheema (1991). According to the field dependence/independence cognitive style, individuals who can attend to or successfully locate the small details of a visual stimulus and ignore the distracting or prevailing outer context in which those details are embedded are classed as field independent (or analytic). Conversely, individuals who attend to the whole visual stimulus and are not so detail-focused are classed as field dependent (or holistic). This ability to extract detail from context is exemplified in the GEFT where individuals locate a simple two-dimensional shape which is embedded (at the same size and in the same orientation) within a complex shape.

These two distinct ways of perceiving visual stimuli possibly relate to the task of face recall for the purpose of facial composite construction using E-FIT, as this requires an individual to mentally 'dis-embed' the features of the previously seen face in order to initially describe the face to an E-FIT operator, and later to focus on each feature separately in order to produce the final image. The E-FIT program has taken the fact that faces are processed in a holistic manner into account, and allows the eyewitness to choose facial features within the context of the whole face. Nonetheless it is still a piecemeal feature-by-feature approach to composite construction because the witness builds the facial composite by selecting the features of the face individually. It may be the case that field independent individuals would be conferred an advantage in facial composite construction using the E-FIT program, because they could view the whole facial composite on screen and perceptually isolate individual features in order to assess their similarity to the target face and change or amend them as required.[1]

According to Tiedemann (1989) field dependent individuals are better equipped to deal with situations requiring perceptiveness and interpersonal skills, and pay more attention to the faces of others in a social setting than field independent individuals. Riding and Wigley (1997) suggested that as cognitive styles affect the way in which an individual encodes information this may in turn affect how an individual internally represents situations in the external world. If this is the case then it can be expected that cognitive styles may also be related to aspects of social behaviour.

---

[1] The concept of FDI is not however restricted to visual perception as first construed by Witkin (1950). Individual performance on the GEFT was found by Witkin et al. (1971) to be 'related to performance on non-perceptual intellectual tasks' (Goldstein & Blackman, 1978, p175). A discussion of the non-perceptual aspect of FDI falls outside the scope of this thesis. However, for the purpose of facial composite construction it may be interesting to note that according to some researchers, field dependent individuals are relatively more influenced by others in a social setting and, as such, tend to pay more attention to the faces of others (Witkin & Goodenough, 1977).

Further research supports these theories and suggests that field dependent individuals make more use of information provided by others during ambiguous situations when others are perceived as a source of information that will eliminate this ambiguity (Endler, 2000). The more socially oriented behaviour exhibited by field dependent individuals as a group leads to the question of whether they might produce more accurate facial composites through being more generally practiced at taking notice of the faces of others in a social context. However, there is controversy in the literature surrounding the theoretical basis of field dependence and whether the construct can indeed be extended beyond visual perception to encompass other areas of behaviour. "Extension of Witkin's theory.... into more general realms of personality and cognition is only made possible by conceptual leaps of extremely dubious validity" (Griffiths & Sheen, 1992, p 137). Despite this, FDI has previously been used in research which examined the performance of mock witnesses in a forensic context.

Emmett et al. (2003) examined the effect of FDI style on the efficacy of context reinstatement in a cognitive interview after participants viewed a live (non-threatening) interruption to one of their lectures. Context reinstatement involves recreating at retrieval, the environmental context in which a previous event was encoded. The study was conceived because previous studies into the effect of context reinstatement on memory performance showed mixed results, indicating a mediating factor as yet unidentified in previous studies. Smith and Rothkopf (1984) first suggested that FDI could potentially determine an individual's susceptibility to context reinstatement. This is because the increased susceptibility of field dependent individuals to external contextual information suggests they may engage more intensely with the environmental context surrounding an event being encoded within their memory, and therefore gain greater benefit from reinstatement of that environmental context. It would be expected that field independent individuals would not benefit from reinstatement of context as they would be more likely to have concentrated originally on the focus of the event and not the surrounding context.

The study by Emmett et al. (2003) produced some interesting and initially mixed results. In a free recall condition where participants were required to write down all information they could remember without being prompted by cues, field dependent individuals (as defined by a median split on GEFT scores) performed significantly better with context reinstatement than without it. By contrast context reinstatement made no significant difference to the memory performance of field independent individuals. Context reinstatement was operationalised by placing some participants back in the lecture theatre in which the original event had taken place. However, in a cued recall condition where participants had to answer specific questions posed about the event they had witnessed, field independent participants answered significantly more questions correctly than did field dependent participants. Field independent individuals therefore respond particularly well to cued recall, where they are essentially being asked to scan their memory for details that were encoded when viewing the original event. This finding may have implications for the way in which eyewitnesses interact with facial composite systems. It may be possible that field independent individuals produce more accurate composites using E-FIT which is feature-based, and that field dependent individuals produce more accurate composites using EFIT-V which is whole-face based.

In a further study by Emmett et al. (2003) which included the holistic/analytic dimension of Riding's (1991) Cognitive Styles Analysis (CSA) test in addition to the GEFT, it was found that scores obtained on Riding's CSA did not correlate with scores on the GEFT, nor did they predict the differences in memory performance across conditions as did the GEFT. Similar to the first experiment by Emmett et al. (2003), field dependent individuals performed significantly better using context reinstatement with free recall, and field independent individuals performed better using cued recall (although the difference was not significant in experiment two). Therefore, although the GEFT has been criticised by Riding for only positively assessing one end of the FDI continuum it continues to be useful for producing a meaningful split among individuals which has been shown to apply to many areas of study, including eyewitness memory performance.

4.2: Holistic/Analytic

As previously stated, a number of different labels have been given to cognitive styles, and it has been argued that many are different conceptions of the same dimensions (Riding & Sadler-Smith, 1992). Riding and Cheema (1991) were among many researchers in the 1990s who attempted to unify the research on different cognitive styles into a model that would be coherent and practically useful. They suggested that all cognitive styles could be placed into two principal groups, the first of which is the holistic/analytic group. The concept of FDI is located within this group, with natural parallels being drawn between holistic and field dependent individuals and between analytic and field independent individuals. The holistic/analytic style is therefore conceptualised as an individual's preference for processing information either in complete wholes (holistics) or in discrete parts (analytics) (Davies & Graff, 2006).

Riding (1991) developed a computerised Cognitive Styles Analysis (CSA) test which positively assesses both ends of the holistic/analytic dimension. Similar to the GEFT, the CSA measures field independence with a task which requires a simple shape to be disembedded from a more complex shape. However, unlike the GEFT (where field dependence is inferred from poor performance on the disembedding task), the CSA uses pairs of complex geometrical shapes which the individual must decide are the same or different. As this task requires perception of the whole of the complex shapes and not parts of them in isolation, and measures the time taken to complete the task, it is thought to tap into preference for field dependence.

Davies and Graff (2006) proposed that a problem arises from the lack of counterbalancing in the CSA; the matching figures sub-test for field dependence is always presented first and this produces inflated scores for field dependence. This is because there is a general tendency for individuals to produce slower reaction times in the first sub-test and faster reaction times in the second sub-test which measures field independence. Emmett et al. (2003) used the CSA in the study looking at individual differences in susceptibility to context reinstatement and found that results on the CSA which designated participants as either holistic or analytic did not predict or correlate with susceptibility to context reinstatement in the way that the GEFT measure of FDI did.

Peterson, Deary and Austin (2003) doubled the length of the CSA test and renamed it the Extended Cognitive Styles Analysis – Holistic/Analytic (E-CSA-W/A). This increased the test-retest reliability of the test, and individual scores followed a normal distribution. Items in the holistic and analytic halves of the test are counterbalanced. The E-CSA–W/A is the test used to designate participants as holistic or analytic in Study Two of this thesis, and administration of the test is described in the procedure sub-section of methods for Study Two.

4.3: Verbaliser/Imager

The second major cognitive style identified by Riding and Cheema (1991) is the verbaliser/imager style and refers to an individual's preferred mode of representing cognitive information, either in words or pictures. According to Riding and Pearson, "when imagers consider information, they experience fluent, spontaneous and frequent mental pictures either of representations of the information itself or of associations with it" (1994, p. 416). This ability to habitually form images possibly relates to the task of facial composite construction, because imagers perform relatively better on pictorial tasks than do verbalisers (Riding & Douglas, 1993).

Riding and Douglas (1993) tested the effect of 'text-only' or 'text plus picture' computer presentations on learning performance in relation to verbaliser/imager cognitive style. They found that imagers learned significantly more information than verbalisers when pictures were part of the learning process, and verbalisers learned significantly more information in the text-only condition. Laughery and Fowler (1980) looked at the verbal and imagery abilities of participants as a possible means of distinguishing between 'good' and 'poor' witnesses. They found that high scores on Gordon's (1949) TVIC (Test of Visual Imagery Control) which measures the ability to manipulate or control mental images, correlated with high scores given to facial composites produced using a sketch artist. This raises the possibility that imagers may produce higher rated facial composites using facial composite systems. This possibility is explored in Study Two using the E-FIT composite construction system.

Peterson, Deary and Austin (2005b) developed the Verbal/Imagery Cognitive Style (VICS) test to address concerns about the reliability of Riding's (1991) CSA test. The VICS is the test used to designate participants as verbaliser or imager in Study Two reported in this thesis, and administration of the test is described in the procedure sub-section of methods for Study Two.

4.4: Summary

This chapter has introduced the concept of cognitive style, which is an individual's preferred way of processing information and representing that information in memory. There are two major cognitive styles, the holistic/analytic and the verbal/imagery (Riding & Cheema, 1991) in which most other styles identified in the literature can be subsumed.

The field dependence/independence cognitive style appears to be part of the holistic/analytic category of cognitive style, as both FDI and the holistic/analytic styles relate to whether information is processed in either wholes or parts. The tendency towards whole or part-based processing may have implications for an individual's ability to construct an accurate facial composite. An important real world issue is whether this occurs in different ways for different composite systems. The E-FIT composite system requires a witness to choose individual facial features in order to produce a full face image (the individual features are however chosen within the context of a whole face), so this may be an easier task for individuals with an analytic or field independent processing style. The newer EFIT-V composite system, and related systems such as EvoFIT work on the basis of whole face recognition and not the recall of individual features, and this may be an easier task for individuals with a holistic or field dependent style.

The verbal/imagery style relates to whether information is represented in memory primarily in pictures or words. The tendency towards picture based or word based representation may also have implications for an individual's ability to construct an accurate facial composite. It could be argued that individuals whose typical mode of representation is image-based might have a better memory for previously seen faces, at least in their own 'mind's eye'. However, the task of recalling a face for the production of an E-FIT involves an important element of verbalisation, and the clarity of facial description and interaction with an E-FIT 'operator' could be as important a factor as being able to accurately visualise the target face. While there is a case for cognitive style being a stable individual difference, there is also evidence that the cognitive style of an individual interacts with the situation they are in and the demands of the cognitive task to be performed. The following chapter highlights some research which demonstrates that the temporary manipulation of an individual's cognitive processing set can have an effect on their memory of a previously seen face.

Chapter Five: Cognitive Processing as a Temporary State

Although individuals do tend to have a preferred cognitive style, research has demonstrated that the cognitive processes associated with stable cognitive styles can be manipulated or induced in order to suit the memory task an individual will be performing (Dunning & Stern, 1994; Macrae & Lewis, 2002). This research may be potentially useful in practical terms because these cognitive processing manipulations might be used with witnesses *after* a target face has been encoded, and prior to an identification task, so therefore could be used as part of a procedure to elicit best memory performance from a witness.

Research has demonstrated that manipulating the cognitive processing set of an individual can have an effect on their recognition of a previously viewed unfamiliar face. This effect can be one of impairment or enhancement depending on the processing task undertaken by the individual (and the processing set induced by the task). For example, the verbal overshadowing effect (Schooler & Engstler-Schooler, 1990) where verbalising a face (or other complex stimuli) can impair face recognition, is believed to be driven by a sub-optimal (featural) processing set being activated prior to memory test. Fallshore and Schooler (1995) used a 'Transfer-Inappropriate-Processing' (TIP) theory to explain the verbal overshadowing effect, suggesting that the act of verbalisation causes a transfer to a featural processing strategy which is at odds with the holistic way in which faces are processed, which in turn leads to a detriment in face recognition performance. As reviewed in section 2.1, research such as Young et al. (1987) and Tanaka and Farah (1993) demonstrated that face recognition is enhanced in conditions in which faces are processed holistically. This chapter describes some different ways in which cognitive processing has been manipulated in research conducted on facial composite construction and also on face recognition.

5.1: Manipulating cognitive processing using different types of interview

The process of constructing a facial composite is one of both recall and recognition (Frowd et al., 2008). Recall is necessary for the description of a previously viewed target face, and usually involves a description of individual facial features. One reason for collecting a description of individual facial features is to limit the number of possible features a witness must view in systems such as PRO-fit and E-FIT in order to select a feature which is similar to the features of the target face in their memory. The composite construction process gradually becomes one of recognition, as the witness must ultimately decide when the optimum likeness to the previously seen face they are attempting to recreate has been reached. However, recall and recognition are two separate cognitive processes, and what may be beneficial to one type of cognitive process may hinder the other.

In order to utilize both featural and holistic processes prior to facial composite construction using the PRO-fit system, Frowd et al. (2008) developed a 'Holistic/Cognitive Interview' (H-CI) to be used with eyewitness participants. The cognitive interview (CI) which is usually used with real eyewitnesses is thought to induce a featural processing strategy (Schooler & Engstler-Schooler, 1990) which is at odds with the holistic way in which faces are recognised. The H-CI involves giving the CI to participants during the face description phase of composite construction, but then following this with a holistic interview where the participant considers the personality attributes of the target face and gives ratings on seven personality dimensions. The administration of a holistic interview induces the participant to switch their mode of processing from a featural mode induced by the cognitive interview, back to a holistic mode which is optimal for face recognition.

Frowd et al. (2008) found that the PRO-fit facial composites which had been constructed following an H-CI were correctly named 41% of the time. Facial composites which had been constructed following the CI were correctly named 8% of the time. Three types of supplementary sorting task where the inner features alone, the outer features alone and the whole composites were matched to the target face also displayed an overall advantage for the H-CI. Composites which were constructed following an H-CI were correctly sorted 38% of the time and composites constructed following the CI were correctly sorted 23% of the time. The advantage of the H-CI over the standard CI in facial composite construction was attributed to a shift away from the featural processing state induced by asking participants to describe a face, to a more holistic processing state which is the optimum state for face recognition. More recent research by Frowd et al. (2012) has extended the finding that the H-CI is beneficial to facial composite construction to EvoFIT, a whole-face based composite construction system.

Although the H-CI has been found to be beneficial for facial composite construction across both featural and holistic composite systems (Frowd et al. 2008; Frowd et al. 2012), other research has demonstrated that the way in which faces are encoded can be an influential factor in the quality of composites that participants produce. In practical terms, the way in which real witnesses encode a to-be-remembered face cannot be manipulated after the event has occurred. However, Olsson and Juslin (1999) found that a quarter of participants who watched a video clip of a staged crime reported using a featural encoding strategy to remember the face of the 'criminal' depicted in the video. Therefore it may be important to assess the likely influence of different cognitive processes utilised at the encoding stage by witnesses. Frowd et al. (2007b) manipulated featural and holistic processing at the encoding stage by asking participants to concentrate either on individual features of a face or personality aspects of a face during encoding. Results showed that better quality composites were produced following featural encoding, even when using a holistic composite construction system such as EvoFIT. This was a finding contrary to the prediction that utilising holistic processing both at encoding *and* at the interview stage would be beneficial when using a holistic composite construction system.

At the present time there is no published research examining the effect of manipulating cognitive processing using the Navon (1977) task on facial composite construction. However a great deal of research has focused on face recognition performance following the administration of a Navon (1977) task. The following section describes research on the temporary manipulation of cognitive processing set using the Navon (1977) task, with particular emphasis on studies where the manipulation of cognitive processing set has affected the face recognition memory performance of eyewitnesses.

5.2: Manipulating cognitive processing using the Navon (1977) task

Macrae and Lewis (2002) investigated the effect of activation of a global or local processing set on face recognition using the Navon (1977) letter identification task. Participants viewed a 30 second video of a robbery and then completed either a global or local Navon task for ten minutes. In a subsequent face recognition test, participants who had completed a local Navon task were significantly poorer at recognising the face they had viewed in the video (30% accurate) relative to a control group who had completed an unrelated filler task (60% accurate). Participants who had completed a global Navon task were significantly better at face recognition (83% accurate) than the control group.

The results of Macrae and Lewis (2002) support the Transfer-Inappropriate Processing (TIP) theory (Fallshore & Schooler, 1995) and demonstrate that a sub-optimal processing strategy can be induced without verbalisation, using a local Navon task to induce featural processing prior to a memory test. An additional finding from Macrae and Lewis (2002) was that completing a global Navon task was beneficial to face recognition, which potentially provides a simple means of improving eyewitness performance in a face recognition task. It may also be possible that the global Navon task will be particularly beneficial to witnesses constructing a facial composite using EFIT-V, which is a recognition based system. Macrae and Lewis attributed these findings to a shift in processing style to either a holistic (global Navon group) style beneficial for face recognition or a featural (local Navon group) style detrimental to face recognition.

However, the theory of the mechanics behind the Navon 'effect' was open to interpretations other than a shift in cognitive processing set. This is because the face identification task used in the Macrae and Lewis (2002) study does not provide a measure of processing. Laboratory tasks involving the presentation of single faces like those used in the research by Young, Hellawell and Hay (1987) provide a straightforward measure of cognitive processing. The results of a line-up task such as that used in the Macrae and Lewis (2002) study are open to several ways of interpretation. This is because there are a number of cognitive strategies which may be used in making a decision to select a face from a line-up, so therefore picking out the correct face in a line-up task may be due to factors other than the inducement of holistic processing. The differences in face recognition observed after administration of the Navon task could, for example, derive from differences in difficulty between the global and local Navon tasks, or from differences in motivation or arousal induced by the Navon tasks.

Weston and Perfect (2005) investigated the effect of a global or local Navon task on the identification of individual facial features to address the issue of interpretation of the Navon effect as a shift in processing set. Weston and Perfect (2005) used Young et al.'s (1987) 'composite face' paradigm to assess the degree to which participants could identify individual facial features in face halves which were fully aligned and slightly misaligned. It was hypothesised that for fully aligned composite faces, a local Navon task prior to identification would lead to faster and more accurate identification of facial features. The misaligned faces already evoke a more featural strategy needed for recognition of features, whereas fully aligned faces evoke holistic processing which make recognition of individual features relatively more difficult.

The results confirmed that a local Navon task was beneficial for feature recognition in the fully aligned faces relative to a control group, thus supporting Macrae and Lewis's assumption that a shift in cognitive processing style drives the Navon effect. However, the benefit of the local Navon task was confined to response latency; those who were in the local Navon task group were quicker at choosing the correct feature, but overall accuracy did not differ between Navon task groups. Therefore, there is no substantial evidence that the local Navon task is beneficial for the correct selection of features. Accordingly, the results of this study partially supported transfer-appropriate processing (TAP) theory, where the featural processing evoked by the local Navon task carried over to the processing task where individual features (eyes or mouth) had to be identified.

Further studies exploring the effect of the Navon task on face recognition have found a different pattern of results which question the transfer-inappropriate processing account adopted by Macrae and Lewis to explain the results of their 2002 study. Weston, Perfect, Schooler and Dennis (2008) compared the effects of Navon task (or verbalisation) on face recognition using the way in which the faces were originally encoded as an additional variable. In the second of three experiments they found that the positive effects of global or local Navon task on face recognition differed depending on the face encoding task that participants undertook. The global Navon task group were better at face recognition following holistic encoding of the target face (deciding which face looked most 'honest' at encoding) and the local Navon task group were better at face recognition following featural encoding of the target face (participants were asked to concentrate on the 'eyes' of the face at encoding).

In a contradiction to previous studies which showed a face recognition advantage following a global Navon task, this study showed that the administration of *either* type of Navon task was beneficial for face recognition - and this was context-dependent on the original encoding task employed by participants when viewing the target face. If the target face was encoded in a holistic manner then the global Navon task led to greater face recognition. However, if the target face was encoded in a featural manner then the local Navon task led to greater face recognition. According to transfer-inappropriate processing (TIP) theory, the local Navon task group should have shown impaired performance on the face recognition task, because face recognition is believed to be facilitated by holistic/configural or global processing of faces, and the local Navon task is purported to induce a featural processing set at odds with holistic processing. However these results are in line with *transfer-appropriate processing* theory (Morris, Brandsford & Franks, 1977; Roediger, 1990) which states that optimal memory performance is achieved when there is a *match* between the processes used at encoding and retrieval.

The findings of Wells and Hryciw (1984) are also consistent with TAP theory. Wells and Hyrciw induced holistic or featural processing strategies at the face encoding stage by getting participants to make judgements about personality or individual features of a face. They found that face recognition was enhanced following personality encoding and also that featural encoding of faces led to participants producing more accurate facial composites using the Photo-FIT system. It would be expected that participants who were required to produce a facial composite using Photo-FIT would benefit from featural encoding of the target face because the Photo-FIT system requires the witness to select individual facial features from a booklet containing isolated individual examples of each feature. In line with TAP theory therefore, Wells and Hyrciw highlighted the importance of *similar encoding and retrieval processes* for optimal memory performance.

The differential findings for the benefits of the Navon task for face recognition, with some studies finding a benefit for the global Navon task (Macrae and Lewis, 2002), and others finding a benefit for *either* Navon task relative to a control group (Weston et al, 2008) demonstrate very similar TIP and TAP theories of memory. Transfer- *inappropriate* processing (TIP) theory does not take into account the way in which stimuli are originally encoded, just the general cognitive processing set active prior to recognition test, and whether there is a match between current cognitive processing set and the cognitive task to be performed. For example, the verbal overshadowing effect is supposed to be a result of inappropriate (featural) processing being induced prior to face recognition test, for which holistic/configural processing is best utilised. Transfer-*appropriate* processing (TAP) theory takes into account the cognitive processing used at encoding, and accounts for memory performance by the match between processing set at encoding and retrieval.

Lewis, Mills, Hills and Weston (2009) conducted two experiments which provided support for the TAP hypothesis that the match between encoding process and retrieval process allows for optimal memory performance. Participants completed a global or local Navon task prior to studying a series of 14 faces. The Navon task was then repeated prior to the face recognition task. The results showed that face recognition performance was best when there was a match between Navon processing at encoding and retrieval. TAP theory could therefore account for some instances in which the local Navon task produces superior face recognition performance relative to a control group, namely those instances where local or featural *encoding* of a target face has been employed. However, in some studies (Weston et al., 2008) there is no interaction between encoding and retrieval in the Navon task, yet still the advantage over control group for local Navon processing remains.

Weston et al (2008) conducted a further experiment with the Navon task which showed an improvement in face recognition following *either* Navon task compared to a control group, regardless of whether a featural or holistic encoding strategy had been used. These results question the claim that the effects of Navon processing can be explained using a holistic and featural processing framework. The improvement found in this experiment for either Navon task regardless of encoding strategy used cannot be explained by transfer-appropriate processing theory, and suggests an effect of Navon task which has yet to be specified. According to Weston et al. (2008), there is no direct evidence to suggest that the global Navon task elicits the same cognitive processes as the holistic style used for face recognition, and "the claim that global and local Navon processing influences the holistic and featural processing styles necessary for face recognition remains to be tested" (p 609).

5.3: Automatic and controlled processing

The question of what exactly is transferred between the Navon tasks and subsequent face recognition tasks was addressed by Perfect et al., (2008) who examined the effects of precedence within the Navon stimuli in relation to their effect on face recognition. They suggested that the emphasis in previous Navon task studies has been on the aspects of Navon stimuli that participants *respond* to, rather than the aspects they *inhibit*. All previous studies examining the Navon task have used Navon stimuli that has a global 'precedence' – it is easier or more automatic for the participant to pick out the large letter (global task) than it is to pick out the small letters which make up the large letter (local task). In this sense, global precedence Navon letters are similar to Stroop (1938) stimuli, in that for the local Navon task the initial automatic immediate response (perception of the large letter) must be inhibited in order to respond to the small letters.

Perfect et al (2008) used two types of Navon stimuli originally devised by Parmentier and Andres (2003) – global precedence Navon stimuli where the large letter has perceptual dominance, and local precedence Navon stimuli where the small letters are more spread out, and it is easier to identify the local letters than the global letter they form (therefore the small letters have perceptual dominance). These stimuli are illustrated in Figure 5.1.

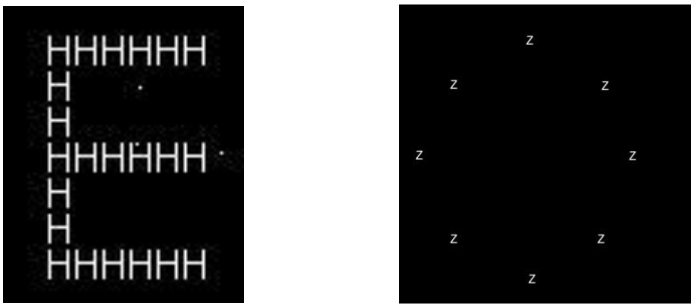

Figure 5.1 – Example of a Global Precedence Navon letter (E made of small H's) and a Local Precedence Navon letter (O made of small Z's).

In the first experiment by Perfect et al. (2008), participants viewed a short videotape of a robbery (the same as that used by Macrae & Lewis, 2002), and then completed a Navon task for around three minutes. There were four conditions of Navon task - global precedence Navon letters where participants attended to either the large letter or the smaller letters forming the large letter, and local precedence Navon letters where again, either the large letter or the smaller letters were attended to. Results showed a significant interaction between the precedence of the Navon stimuli and face recognition performance. In the global precedence Navon task group 65% of responses in the line-up task were accurate, and in the local precedence Navon task group 55% of responses in the line-up task were accurate. The accuracy rate for responses in the line-up task for the two groups in which precedence and responding did not match was 35%.

Perfect et al. (2008) conducted a further experiment which was a replication of Weston and Perfect (2005) using composite face stimuli (both aligned and misaligned) and both local and global precedence Navon stimuli. They found the reverse pattern of results to the first study; for global precedence Navon stimuli, participants in the local responding group were quicker at identifying individual features of a face, and for local precedence Navon stimuli, participants in the global responding group were quicker at identifying individual features. In other words, performance was facilitated when participants responded to the non-dominant stimulus dimension. The global precedence Navon stimuli conditions in both experiments support previous research findings, where responding to the global aspect of the Navon letter leads to more accurate face recognition and responding to the local aspect of the Navon letter leads to quicker identification of individual facial features. The local precedence Navon stimuli conditions in both experiments however, do not support previous research findings, because responding to the local aspect of local precedence Navon letters led to greater face recognition accuracy, and responding to the global aspect of local Navon stimuli led to quicker recognition of individual facial features.

Perfect et al (2008) concluded that the question of what is transferred between the Navon tasks and subsequent face recognition tasks could possibly be attributed to a *form* of transfer-appropriate processing, but not within a global/local processing framework. They suggested that responding to the dominant aspect of Navon stimuli, whether this was global or local, requires no inhibition and promotes *automatic* processing, whereas responding to the non dominant aspect of Navon stimuli requires controlled or analytic processing. This automatic/controlled processing account of the effect of Navon stimuli fits the pattern of results observed in the Perfect et al (2008) study, and is supported by previous research which showed that using an automatic processing strategy in eyewitness memory tasks leads to greater accuracy in those tasks (Dunning & Stern, 1994). The researchers do acknowledge however that “further research is necessary to establish the robustness of this account” (p. 1485).

5.4: Summary

It is arguably possible to manipulate an individual’s processing set temporarily in order to facilitate their memory performance. Research by Frowd et al. (2008; 2012) has suggested that shifting an individual’s cognitive processing state from featural to holistic prior to composite construction can have a beneficial effect on the accuracy of facial composites that individuals produce. The beneficial effect of inducing holistic processing prior to composite construction was demonstrated using both a featural and a holistic composite construction system. When featural processing is induced at the encoding stage however, there is an advantage for facial composite construction even when a holistic interview and a holistic construction system is used (Frowd et al., 2007b).

Further research has demonstrated that the manipulation of cognitive processing set using the Navon (1977) task can facilitate or hinder memory performance, depending on the type of Navon task used, and the nature of the memory task. Several theories have been proposed to account for the Navon task effect, but the nature of the effect is still the subject of some debate in the literature.

The beneficial effect of the global Navon task for face recognition has implications for the process of facial composite construction. With the E-FIT program, composite construction begins as a process of recall and later becomes one of recognition (when the features of the face have all been selected). The process of composite construction using EFIT-V is entirely one of recognition however, so the global Navon task might enhance the accuracy of composites produced by EFIT-V if it were to help witnesses to make better initial choices of faces to be evolved by the system. It may be the case that the match between cognitive processes induced by the global Navon task, and the holistic cognitive processes involved in successful face recognition could lead to more accurate facial composites using EFIT-V. Similarly, cognitive processes induced by the local Navon task may lead to more accurate facial composites using E-FIT, because inducing a featural processing strategy may be advantageous for the selection of individual facial features. There are no published studies to date which investigate the effect of the Navon task on facial composite construction.

Chapter Two described psychological theories of face processing. The holistic manner in which faces are processed and represented in memory has led to the development of facial composite systems, described in Chapter Three, which are designed to match the cognitive processes involved in face perception. Chapter Four explored the issue of cognitive style and presented research which demonstrated that the cognitive style of an individual can be a mediating factor in the memory performance of individuals in a forensic setting (Emmett et al., 2003). The research presented in Chapter Five demonstrated that it is possible to manipulate temporary cognitive processing state, and that this can have an impact on face recognition. Chapter Six describes some general methods used in the empirical research reported in this thesis. Chapters Seven to Ten report the four research studies which explore some issues of cognitive style and cognitive processing in relation to facial composite construction.

Chapter 6: General Methods

Chapter Six describes some general methods and procedures which are used in the empirical research presented in Chapters Seven to Ten. Methods of composite construction and evaluation of facial composites are described herein to avoid repetition within the empirical research chapters.

6.1: E-FIT Construction

For Studies 1, 2, and 3, participants constructed a facial composite using the E-FIT system. The following procedure was adopted by the experimenter who underwent the official training, as given to police operators, in the use of E-FIT for Windows (version 5.1) and cognitive interview techniques in 2007 at the University of Kent.

Description of the target face

In standard police procedure a cognitive interview is conducted with an eyewitness prior to giving a description of the previously seen face and constructing a facial composite. The purpose of the cognitive interview is to reinstate the context in which the face was seen in an attempt to enhance the internal representation the eyewitness has of that face. The cognitive interview procedure was not used in any of the studies described here as it would be too time-consuming for participants, and unnecessary due to the fact that participants were constructing a facial composite within a relatively short time of viewing the target face. In real life circumstances, eyewitnesses usually have a delay of at least 48 hours and sometimes considerably longer before constructing a facial composite. In addition to this, eyewitnesses construct a facial composite in a different context from that in which the original target face was viewed, so context reinstatement is important, whereas the participants in the research presented within this thesis constructed the facial composites in the same research room in which they had viewed the target face.

Participants therefore began the construction process by giving a description of the target face they had viewed under free recall conditions. Participants were asked to provide as much detail as they could possibly recall about the target face, and this information was noted down and used later with the drop-down menus within E-FIT in order to narrow down the selection of features the participants would view before finding a close match to the representation of the face in their memory. This was followed by cued recall in instances where the experimenter required clarification of the initial description, for example a participant sometimes described a facial feature in terms of personality such as 'mean-looking lips', or in terms of reference to a person they knew such as 'his eyes looked like my brother's eyes'.

<u>Composite Construction</u>

When the target face had been described in as much detail as the participant could, they were shown the E-FIT system by the system operator (experimenter) and given a brief description of the procedure they would follow in order to construct a composite. There are two possible ways of constructing a facial composite using E-FIT. For the first method, all of the descriptions given by the participant for each feature can be entered into the E-FIT system, and a whole face will be produced on the screen for viewing. This can then be amended by changing any feature for a different one in the database, and/or by resizing or repositioning the existing features on the face. The second method is for participants to view a schematic face (see Figure 6.1) on the screen and add features to the schematic face in any order they choose to, (known as the 'minimum face approach'). The second method of using the minimum face approach was the one used with all participants, as the minimum face approach is the approach recommended in the E-FIT training course for E-FIT composite construction.

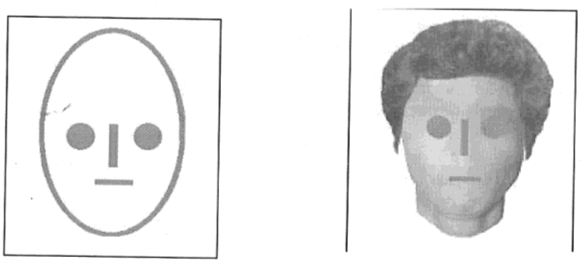

Figure 6.1: Example of the schematic face and how the features are added to that face. Here, face shape and hair have been added to the schematic face.

Before beginning to choose facial features participants selected which database of features to use, in this instance the white male database was used by all participants as all the target faces used in the studies were white male faces (with the exception of Study 2 where some participants returned to produce a composite of a white female face). Without exception all participants chose to put a face shape on the schematic face before adding any features to the face. Face shape was selected by using the original description the participant gave for the shape of the face, and applying that description to the drop-down menu box for face shape. This gave participants a number of blank face shapes on the screen from which they could choose the one which was the closest match to their representation of the shape of the target face in their memory. Following selection of the face shape, drop-down boxes were selected which contained descriptors for the feature they wanted to add to the face shape. The descriptors that were the closest match to the original description given by the participant for each feature were selected. If the participant did not give a description for a particular feature then there was an option within the drop-down menu which could be left as 'unsure', (see Figure 6.2).

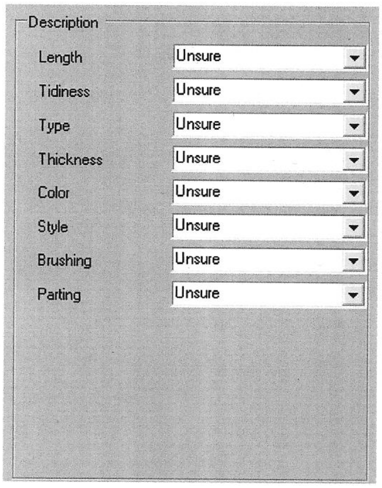

Figure 6.2: Drop down menu for the hair feature within E-FIT

The E-FIT system then sorted all of the examples of the selected feature into an order which most closely matched the description given by the participant. The feature exemplars were then viewed sequentially within the context of the whole face shape until the participant selected the best match to their memory of that feature. This process was repeated for each facial feature until a whole face was produced, (see Figure 6.3).

Figure 6.3: Example of the addition of features to create a whole-face image using E-FIT

<u>Modification and Artistic Enhancement</u>

When all the features of the face had been selected, participants had the opportunity to change the size of any feature (including the face shape). Participants also had the opportunity to resize any of the features they selected during the composite construction process, and although they often used the option to do so, when they had the opportunity to view the whole facial composite on screen and to see the size of the features in relation to one another, many opted to make additional changes. Features can be resized both vertically and horizontally, or can be replaced by a different example from the database. The position of the features can also be manipulated, for example the eyes can be positioned closer together. When the participant was satisfied with the likeness produced, the Picture Publisher image manipulation package was used for artistic enhancement if it was required, such as adding shadows under the eyes or small marks to the face such as spots or freckles. The final facial composite image was saved on the computer using an individual code number that had been assigned to each participant. Each of the E-FITs produced within the first three studies of this thesis took between forty and sixty minutes to construct (see Appendix I for an example of an E-FIT facial composite).

6.2: E-FIT Construction with Navon Task

In Study Three, the standard method of composite construction using E-FIT as described in Section 6.1 above was used as a control with one group of participants. The standard method of E-FIT composite construction was measured against the performance of two further groups of participants who produced a facial composite in conjunction with a global or local Navon task (see section 6.4). The respective Navon tasks were incorporated into the standard composite construction process at the end of choosing each feature of the face, and prior to choosing the next feature. For example, in the global Navon task condition, participants gave an initial description of the target face and chose the white male database at the start of the composite construction process. However, before proceeding to choose their first feature (all participants chose to begin with the face shape) participants completed a global Navon task on the computer for one minute. After the face shape was chosen, participants completed another minute of the global Navon task before proceeding to choose the next feature of the face. This process was repeated for each feature of the face, therefore adding approximately seven minutes to the composite construction procedure (a global Navon task was completed before the selection of face shape, hair, eyes, brows, nose, lips and ears). Participants in the local Navon task condition followed the same procedure, the only difference being that 7 local Navon tasks were completed during composite construction.

6.3: EFIT-V Construction

In Study 4 the facial composites were constructed using the EFIT-V system with the following procedure. An initial description of the target face was not required for construction of an EFIT-V composite as the participant does not build the face from individual features. EFIT-V capitalises on the relatively easier cognitive task of face recognition as opposed to face recall (required for construction of an E-FIT). In the first instance, the participant was required to indicate the gender, ethnicity, and approximate age of the target face. Following on from this a hairstyle and colour was selected, and the process of random generation of whole faces within the EFIT-V program began. The participant viewed 9 whole faces on the screen each with the same pre-selected hairstyle, and chose the face which bore the closest resemblance to the target face, (see Figure 6.4).

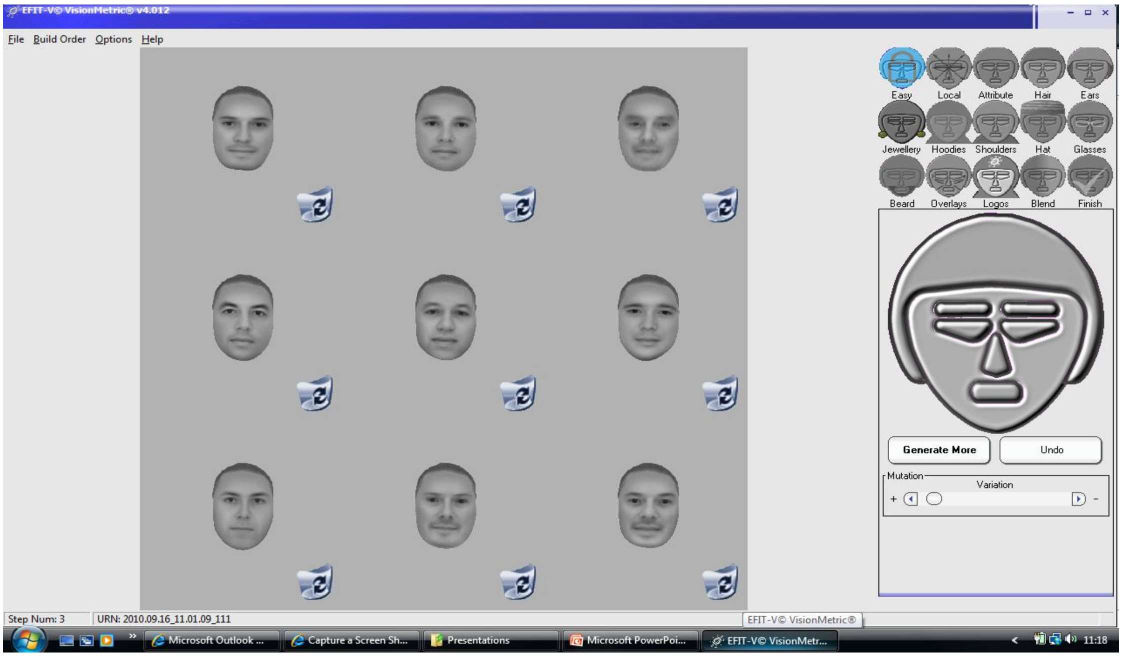

Figure 6.4: Example of the EFIT-V system where the face most similar to the target face is chosen

In the event that the participant indicated that no faces on screen were appropriate, there was a 'generate more' button which can produce another 9 completely new faces to the present ones on the screen (still retaining the chosen hairstyle). Once the participant selected an appropriate face, the next set of nine faces was produced by EFIT-V. The next set of nine faces generated always contained the original face that was chosen from the previous screen, so that it was never lost. If there was a particular feature of any face which the participant wanted to keep this could be locked (in much the same way as the hairstyle) so that all future generations of new faces varied but retained the locked feature. As each new generation of nine faces are produced, the act of choosing a face from each screen means that the faces all become more similar until the participant chooses to end the process and select a face as being the closest match to their memory of the target face. Similar to the E-FIT program, features on the final composite image can be resized, repositioned or replaced either during composite construction and/or towards the end of the process depending on the preference of the eyewitness. The final stage in the composite construction process is artistic enhancement where additional marks or shadows can be added to the image. The process of producing an EFIT-V composite is generally much quicker than producing an E-FIT, and participants took between five and twenty minutes to construct each of the EFIT-V composites used in Study Four (see Appendix II for an example of an EFIT-V facial composite).

6.4: Navon Task

In both Studies Three and Four participants completed either a global or a local Navon task using global-precedence Navon letters (see section 5.2 for an explanation of the difference between global-precedence and local-precedence Navon letters). The difference in the presentation of the Navon task was the duration and number of times it was administered to the participants.

In Study Three the Navon task was used for one minute, and for a total of seven times, interspersed in between the choosing of facial features in the production of E-FIT facial composites. In Study Four the Navon task was used for five minutes just once, immediately prior to producing an EFIT-V facial composite. A Navon letter is a large letter that is made up of smaller different letters (see Figure 6.5).

Figure 6.5: A global precedence Navon letter

In the global Navon condition, participants were required to identify the large letter presented on the screen. In the local Navon condition, participants were required to identify the small letters which were used to make the large letter. In the example shown in Figure 6.5, a participant completing a global Navon task would say ‘E’ when this letter was presented on screen, and a participant completing a local Navon task would say ‘H’ when this letter was presented on screen. The Navon letters were presented sequentially for one second each within a power-point slide presentation and were presented as white letters on a black background. These materials were provided by Professor Tim Perfect from the University of Plymouth, and had previously been employed in a study by Perfect et al., (2008).

6.5: Group Embedded Figures Test

In Studies One and Four participants completed a Group Embedded Figures Test (GEFT) as used by Emmett et al. (2003) to assess their relative levels of field dependence or independence. Materials consisted of a demonstration sheet in which a 5 x 5 square grid pattern was displayed (complex figure) alongside a simple figure (an outline letter ‘E’) which could be found embedded within the complex figure in the same orientation and scale. Participants were also provided with a single training sheet for practice purposes. Lastly two booklets were provided, each of which contained a set of nine complex figures printed two to a page (see Figure 6.6)

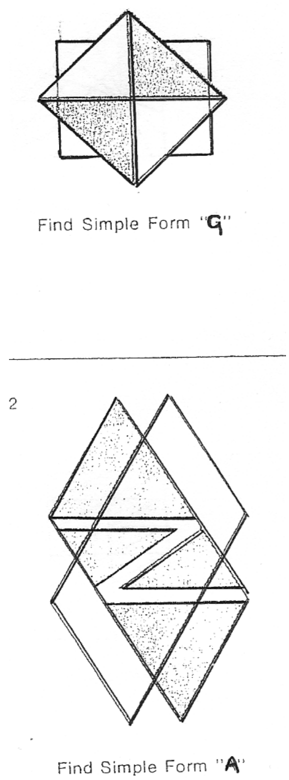

Figure 6.6: Example of complex figures within the GEFT

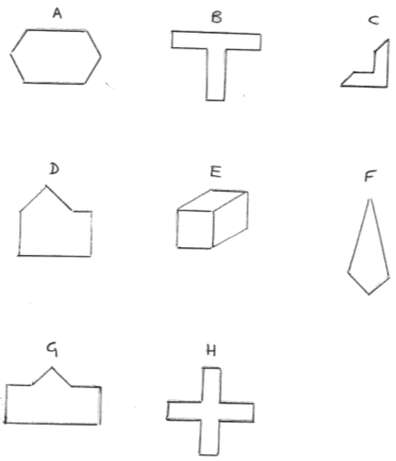

Figure 6.7: Simple figures at the back of the GEFT

The booklets were assembled such that participants could not examine any of the complex figures and its corresponding simple figure simultaneously. Each of the nine complex figures within each booklet contained an instruction to draw on to the shape one of the simple figures located at the back of the booklet. Participants then looked at the appropriate simple figure at the back of the booklet and returned to the complex figure where they were required to draw the outline of the simple figure in the same size and orientation on to the complex figure if it could be seen. Participants completed this procedure using a pencil to allow for mistakes to be erased and rectified, and each booklet of nine figures had a time limit of five minutes for completion. Possible scores on the GEFT ranged from zero (no simple figures located and copied correctly) to eighteen (all simple figures located and copied correctly). Within both studies that used the GEFT, a median split was performed on the scores of participants in order to place them into two groups of field dependent and field independent cognitive styles. Those who scored relatively few correct answers on the GEFT were classed as field dependent, and those who scored relatively many were classed as field independent.

6.6: Subjective Likeness Ratings

Subjective likeness ratings on a percentage scale were used for all four studies as a measure of the similarity of the facial composites to the original target face from which they were constructed.

Participants were given an initial briefing in which they were asked to assess each E-FIT (or EFIT-V for Study Four) for the degree of similarity to the target face, and to provide a percentage likeness rating using a scale from 1 to 100. Each participant was presented with the photographs of the original target faces and the facial composites which were made of each target face. The facial composites were assembled in the order in which they were constructed within each study (see Figure 6.8).

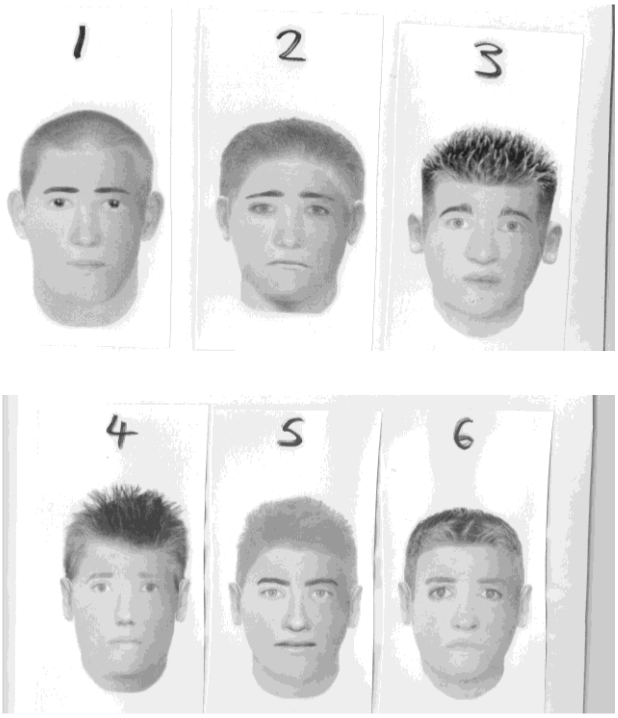

Figure 6.8: Example of layout of E-FITs for subjective likeness ratings

Participants recorded their percentage likeness ratings for each E-FIT's degree of similarity to the original target face on a prepared response sheet, and were allowed as much time as required to complete their responses. The order of the presentation of the facial composites and corresponding target faces was counterbalanced so that some raters began giving ratings on target face A, some raters began giving ratings on target face B and some began on target face C etc. This was in order to account for any effects of fatigue or other cause of changing of ratings which may have affected raters.

6.7: By-item and by-rater analyses

The subjective likeness rating data from all four studies within this thesis were analysed both by-item and by-rater. This section briefly outlines the differing inferences which can be drawn from each type of analysis, and considers the relative strengths and weaknesses of both types of analysis.

The by-item analysis of the data can be considered to be the most robust as it can be generalised to the population of people who might conceivably construct a facial composite, the eyewitnesses, as each item (composite) also represents one 'witness'. The by-item data for subjective likeness ratings refers to the overall mean rating score given to each facial composite by every rater. As such, the by-item data provide information as to whether any particular cognitive style or induced temporary processing state increased the likeness ratings given to the composite. The question of whether higher likeness ratings for individual composites leads to more frequent levels of identification of those composites in a line-up task is addressed in Study One. By-item analysis is better for objective methods of data collection such as matching tasks and line-up identification tasks. A potential weakness of by-item data analysis however, is that it lacks the power to detect small effect sizes. Thus, by-rater analysis of the data can be considered to be useful.

By-rater analysis of the data can be generalised to the population of people who might view a facial composite, and thereby make an identification based on their familiarity with the face depicted in the facial composite. By-rater data for subjective likeness ratings refers to the mean rating score given by each rater to the different experimental conditions within the studies in this thesis. This repeated-measures method of data analysis, where ratings are provided for each experimental condition aggregated by-rater, provides greater statistical power to detect what are often essentially small effect sizes within facial composite research. Additionally, although these effect sizes may be considered to be very small in an experimental context, when translated to a considerably larger real-world context they can be useful in the identification and apprehension of offenders whose composites are publicised to the general public. Within published studies on facial composite systems it is common to find that by-rater analyses are used alone, or that both types of analysis have been used as complementary measures (eg. Frowd et al., 2005b; Brace et al., 2006; Frowd et al., 2008).

This chapter has described the general methods used in the empirical research reported within this thesis. The following chapters Seven to Ten report the findings from four empirical studies which were designed to assess the role of cognitive style and temporary cognitive processing state in the production of target-accurate facial composites.

Chapter Seven: Study One

This chapter reports the first of four empirical studies which together, investigated the role of holistic and featural processing in the production of accurate facial composites. Study One was an exploratory study designed to assess whether individual differences in cognitive style specifically are associated with the ability to produce an accurate facial composite using the E-FIT system. Ellis et al. (1975) were among the first researchers to suggest that, system limitations aside, the variability they observed in the ability to construct accurate *Photo-FIT* composites might be due to differences in the field dependence/independence cognitive styles of participants. There are currently no published studies measuring the field dependence/ independence cognitive styles of individuals in relation to accurate facial composite construction.

Study One addressed the question of whether individual differences in field dependence/independence may have an impact on the quality of composites that individuals produce using E-FIT. As outlined in chapter four, the process of facial composite construction using E-FIT is arguably a very similar cognitive task to that of the embedded figures test which assesses level of FDI in individuals. Successful completion of both a facial composite and an embedded figures test requires an individual to perceptually isolate a facial feature (composite) or simple shape (embedded figures test) and ignore the distraction of the surrounding context.

Tanaka and Farah (1993) demonstrated that individual facial features viewed within the context of the whole face are more easily identified than facial features viewed in isolation, and this was a key element in the development of the E-FIT system, which allows features to be chosen within the context of a whole face shape. Although the selection of facial features within the context of a whole face shape might therefore be beneficial for some individuals, it may be the case that individuals who find it easier to perceptually isolate individual features (field independents) would produce more accurate facial composites. This is because Young, Hellawell and Hay (1987) showed that perception of a whole face image interferes with the correct identification of individual facial features, and this may be particularly detrimental to field dependent individuals who find it relatively difficult to extract detail from context.

In summary, the concept of field dependence/independence is considered to be related to the extent to which perception of the whole detracts from perception of any of its parts. A relatively field-dependent individual prefers to process visual stimuli in wholes and, as such, may have difficulty in mentally disembedding individual facial features when constructing an E-FIT because perception of the whole face may be the dominant aspect. The relatively field-independent individual is thought to be able to overcome the distraction of perception of the whole and more easily break it up into parts. Therefore field independent individuals may be able to overcome the distraction of presentation of the whole face and perceptually isolate facial features which more accurately reflect the target face being constructed.

The perceptual differences between field-dependent and field-independent individuals have been found to account for differences found in a number of areas, such as sociability and the beneficial effects of context reinstatement on memory performance (Emmett et al, 2003). The aim of Study One was to investigate whether individual differences in field-dependence/independence might account for some of the observed differences in the ability of individuals to construct an accurate facial composite. Study One used E-FIT, the facial composite construction system most widely used by the police at the present time.

7.1: Method

7.1.1: Design

A natural independent groups design was employed for this study to determine whether there were differences in the ability to produce accurate E-FITs as a function of the cognitive style of the participants. The cognitive style construct measured was field dependence/independence (Witkin et al., 1971). The study consisted of three phases, the first of which involved participants constructing an E-FIT of a previously seen target face and completing a Group Embedded

Figures Test (GEFT). Phases two and three were evaluations of the accuracy of the E-FITs produced which employed two new samples of participants.

7.1.2: Phase One –Composite Construction

Participants

Forty five university students and members of the general public (24 females, 21 males) participated in the construction phase of the study (mean age = 35.23 years, s.d. = 11.1). University students were recruited as part of their requirement for the Research Methods module research participation scheme. Forty three participants were from the same ethnic group as the target faces (Caucasian) and two participants were from a different ethnic group to the target faces. All participants were naive as to the purpose of the study, and none reported any prior knowledge or practical experience of composite construction.

Materials

Nine monochrome images of young adult Caucasian males were used as target faces. The images were obtained from the Psychological Image Collection at the University of Stirling (http://pics.stir.ac.uk) who permit their free use for non-commercial research purposes. The photographs displayed a frontal full-face pose with a neutral expression and measured 13 x 18 cm when printed for use. Thirty target faces were initially selected on the basis of having no distinguishing or outstanding features (including spectacles, piercings, distinguishing marks etc). The nine target faces used in the study were selected by a third party to ensure that the experimenter was blind to the appearance of the target faces.

All necessary materials were provided for participants to complete the GEFT (see Chapter 6, General Methods Section 6.5). The E-FIT for Windows program (Version 5.0) (Aspley, 2004) was used to construct the E-FITs, and Micrografx Picture Publisher 8 (1998) was used to make artistic enhancements to the E-FITs prior to their completion. A Compaq Presario M2000 laptop computer with a screen size of 30cm x 24 cm (1280 x 1024 pixels) was used to construct the E-FITs.

Procedure - Composite Construction and GEFT

Participants were tested individually and each produced one composite of one of nine previously selected target faces (see Chapter 6, General Methods section 6.1 for the standard composite construction procedure used for studies 1, 2 and 3). The nine target faces were pseudo-randomised such that each was viewed by five participants thus producing five composites for each target face. Participants received a standardised initial briefing which assigned them a participant number and instructed them to consult a randomised list which would inform them as to which target face they would view (the target faces were concealed in lettered A4 envelopes). They were asked to study the target face contained within the envelope they had been assigned to for one minute, with a view to making a composite construction of the face using the E-FIT program. This procedure ensured that the experimenter was blind to which target face was being viewed by the participant, as well as which faces had been selected for the study.

After viewing the target face for one minute, each participant then completed a Group Embedded Figures Test (Witkin et al, 1971) in order to establish their position on the FDI continuum, (see Chapter 6, General Methods Section 6.5).

7.1.3: Phase Two – Evaluation Stage 1 - Subjective Likeness Ratings

The facial composites (E-FITs) produced by eyewitness participants were subjected to two methods of evaluation. The first of these was a subjective likeness rating task in which participants were required to give a percentage rating to each composite when comparing it directly to the target face from which the composite was made.

Raters

Twenty one participants gave a subjective likeness rating (%) to each of the forty five E-FITS produced in the construction phase. All participants were volunteers recruited from the general population, six males and fifteen females (mean age = 28.9 years, sd = 10.9) and none had taken part in the construction phase of the study.

Materials

Each target face was presented simultaneously with print-outs of the five E-FITS which had been produced by participants in the construction phase (the target face measured 13 x 18 cm and the E-FITs measured 10 x 15 cm).

Procedure

Participants were given an initial briefing in which they were asked to assess each E-FIT for the degree of similarity to the target face, and to provide a percentage likeness rating using a scale from 1 to 100. See Chapter 6, General Methods section 6.6 for details of the subjective likeness rating procedure used in all 4 studies).

7.1.4: Phase Three – Evaluation Stage 2 – Objective Measure of Evaluation

In order to assess the validity of the subjective likeness ratings, the second method of evaluation was an objective measure in which the top 4 and bottom 4 rated E-FITs (as denoted by the subjective likeness ratings) were displayed individually alongside target-present line-ups where participants were required to pick out the target face from the line-up which most closely resembled the E-FIT. Line-ups were constructed by inserting the original target photo in a random position among foils also drawn from the PICS database (http://pics.stir.ac.uk).

Participants

Two hundred and thirty seven students and members of the general public (172 females, 59 males and nine who did not state their gender) participated in the objective evaluation phase of the study (mean age = 29.88 years, s.d.= 11.15). One hundred and sixty one participants were from the same ethnic group as the target faces, 67 participants were from a different ethnic group and nine participants chose not to give details of ethnicity. Seventy five participants completed the 4-person line up condition, and 81 participants completed the 6-person and 8-person line-ups respectively. Participants were recruited both through internal advertising at the University of Westminster and via social networking sites such as Facebook and MSN Messenger.

Materials

The online study was programmed with the help of two Psychology Department Technicians at the University of Westminster. The E-FITs which appeared in the line-ups were chosen on the basis of their overall ranking in the subjective likeness rating phase. Taking into account the fact that different target faces were to appear in each line-up, the top four and bottom four scoring E-FITs were used. Three line-up sizes were used (4-person, 6-person, 8-person) in order to assess which, if any, line-up size would best display the predicted difference in performance in terms of the number of correct identifications made of high and low rated E-FITs. The 4-person line up contained the target face plus one 'type likeness', a face of 'medium-likeness' and one face which was not considered to be a likeness the target face. The 6-person line up contained the target face plus two 'type likenesses', two medium likenesses and one face which was not considered to be a likeness to the target face. The 8-person line up contained the target face plus 3 'type-likenesses', three medium likenesses and one face which was not considered to be a likeness to the target faces. The type likenesses for the line ups were chosen in a pilot study in which 10 participants were presented with a choice of possible faces for inclusion in the line ups and asked to rate the faces on a scale ranging from 1 to 10 for similarity to the target face for each line up. The reason for using type-likenesses was that facial composites are not intended to be an exact likeness of a perpetrator's face, but ideally a 'type-likeness' which serves both to facilitate the apprehension of a suspect and also to eliminate unlikely suspects. It was decided not to pursue the question of whether type-likenesses were selected in the online study because the rationale for doing so did not fit with the overall aims of the thesis.

Procedure

Participants joined the online study via a link attached to the researcher's home page on the University of Westminster website. Full instructions were given at the beginning of the study, and informed consent was assumed by participants clicking to continue with the study itself. Participants gave information about age, gender and ethnicity before beginning the study, but were not required to do so if they did not wish to. Participants were then shown a sequence of eight screens which contained one E-FIT in a central location, and also either four, six or eight male faces in a target present line up, (see Appendices III, IV and V) and were required to select the face which they thought the E-FIT most closely resembled. After making their eight selections (each on a separate page) for the eight E-FITS presented, participants submitted their selections via a 'submit' button and were then given a full debriefing of the study in which they had participated and invited to contact the researcher should they have any queries or questions regarding the research.

7.2: Results

7.2.1: Results from evaluation phase one - Subjective likeness rating phase

In order to compare the subjective likeness ratings the facial composites received with the cognitive style of the eyewitness-participants, a median score was calculated for the overall scores participants received on the Group Embedded Figures Test. The median GEFT score was 12 and this was used to designate participants as either field dependent or field independent (cf. Emmett et al., 2003). Participants scoring 12 and above on the GEFT were classed as field-independent and those scoring 11 and below were classed as field-dependent. The overall mean GEFT score for males and females combined was 10.7 (sd-3.9). Scores on the GEFT ranged from 2 to 17 points (within a possible range of 0 to 18 points).

The data from the subjective likeness rating phase were analysed by-item (composite) and by-rater (participant). Study One addressed the question of whether individual differences in field dependence/independence may have an impact on the quality of composites that individuals produce using E-FIT. By-rater analyses were conducted first in order to assess whether there were differences in E-FIT accuracy between groups of field dependent and field independent participants.

By-rater analysis

Table 7.1 shows the mean percentage scores that individual raters gave to groups of facial composites produced by field dependent and field independent participants.

Table 7. 1: Mean subjective likeness rating scores by cognitive styles with data aggregated by-rater in Study One

Cognitive Style	Mean (SD)
Field Independent	39.24 (11.70)
Field Dependent	33.62 (11.03)

A paired samples t-test performed on the mean scores showed a significant difference by-rater (t=5.304, df =20, p < .005, two tailed) with composites produced by field independent participants as a group scoring significantly higher subjective likeness ratings than composites produced by the group of field dependent participants. Therefore, as a group, field independent participants who are relatively more able to visually disembed a simple shape from the surrounding context produced more accurate facial composites than field dependent participants using E-FIT. There was a medium effect size of 0.4 (Cohen's *d)*.

By-item analysis

A by-item analysis was conducted to assess the accuracy of individual participants in facial composite construction as rated by new participants who had not constructed a facial composite in phase one. Table 7.2 displays the mean percentage score given to individual composites produced by field dependent and field independent eyewitnesses across all raters in the subjective likeness evaluation phase.

Table 7.2: Mean subjective likeness rating scores by cognitive styles calculated by-item in Study One

Cognitive Style	Mean (SD)	N
Field Independent	39.24 (12.61)	23
Field Dependent	33.62 (10.56)	22

An independent samples t-test performed on the means showed no significant difference across composites for those made by field dependent and field independent eyewitness/participants (*t*=1.616, *df* =43, *p* = 0.113, two tailed). There was a medium effect size of 0.4 (Cohen's d). Therefore in the by-item analysis, which pertains to the individuals who produced the facial composites rather than the individuals viewing the composites, there was no difference in the accuracy ratings of the composites produced. A scattergram was employed to explore the nature of the relationship between the GEFT scores of participants who constructed the composites and the subjective likeness rating scores their composites received in the by-item analysis. The scattergram revealed no bias in the residuals across the range of GEFT scores, implying that the discrepancy between by-rater and by-item analyses was not a consequence of a sub-set of composites attracting relatively high or low scores.

Additional Analyses

An independent t-test was performed to determine whether there was a significant difference in the time taken to construct the composites between field dependent and field independent participants. Table 7.3 shows the mean construction time in minutes for field dependent and field independent participants.

Table 7.3: Mean composite construction times (minutes) for FDI cognitive styles in Study One

Cognitive Style	Mean (SD)	N
Field Independent	36.80 (8.0)	23
Field Dependent	35.60 (7.6)	22

There was no significant difference in the amount of time taken to construct a composite for field-dependent and field-independent participants (t = 0.532, df = 43, p = 0.597, two-tailed). There was a small effect size of 0.16 (Cohen's *d*).

7.2.2: Results from evaluation phase two - Objective evaluation

The data were analysed by-rater (participants who completed the online evaluation phase), which involved looking at the two groups of four E-FITs (high rated/low rated) across all three line-up conditions to see if there was a significant difference in the number of correct matches made to the target face. The minimum potential score was zero and the maximum achievable score was four. Figure 7.1 shows that the high rated E-FITs were correctly matched to the target face in nearly 3 out of 4 instances on average, whereas the low rated E-FITs were correctly matched to the target face less than half of the time on average.

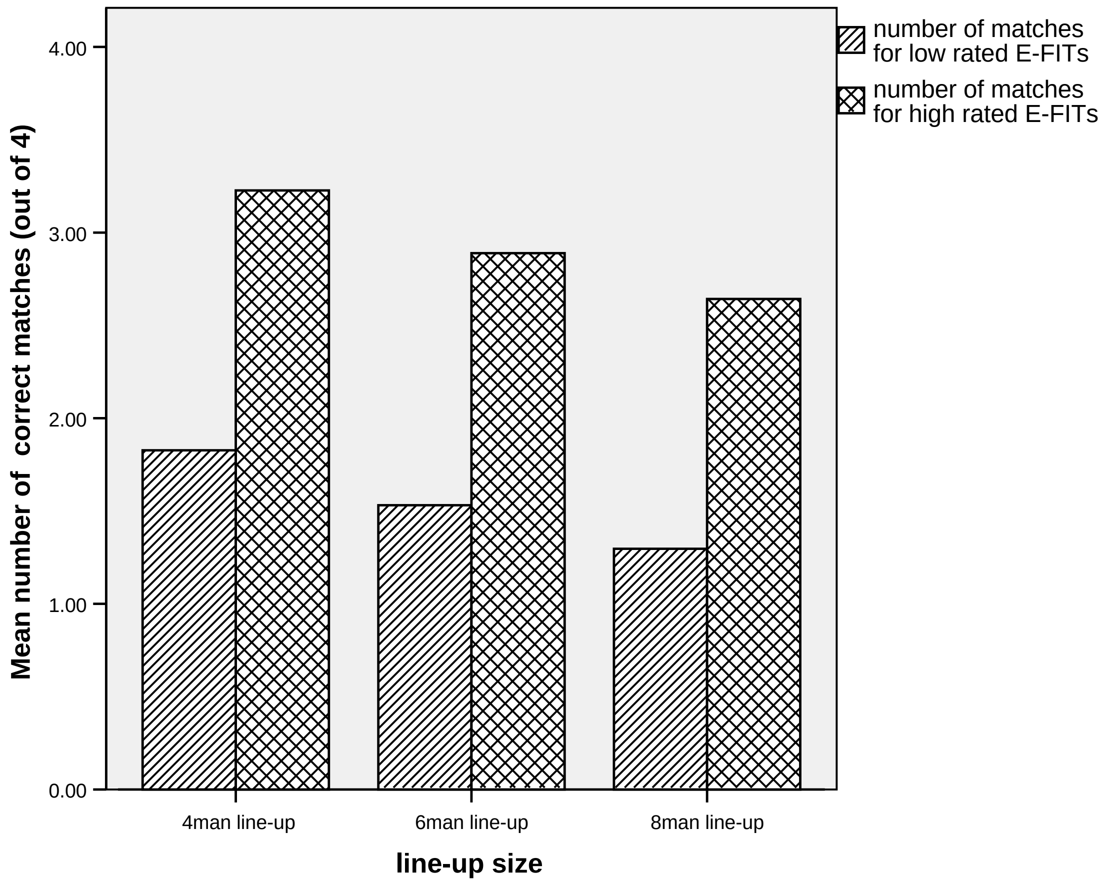

Figure 7.1: Mean number of correct matches (out of four) for low rated and high rated E-FITS over three line-up sizes in Study One.

Figure 7.2 illustrates how the number of correct matches varied as a function of the line-up size in which the target faces were presented

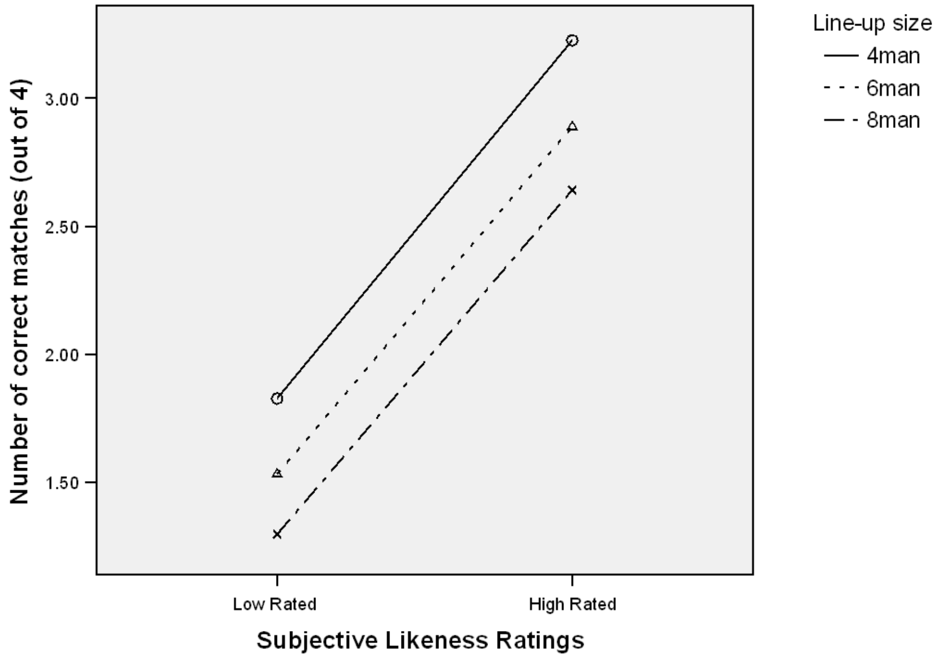

Figure 7.2: Mean number of correct matches across 3 line-up sizes for low and high rated E-FITs in Study One

A 2*3 mixed Analysis of Variance (ANOVA) was performed on the data in order to compare the performance of the two categories of E-FITs across three line-up sizes. The within subjects factor was the subjective likeness rating of the E-FITs with two levels, high rated and low rated. The between subjects factor was line-up size with three levels (4-person, 6-person, 8-person line-up).

The main effect of E-FIT rating was significant, ($F_{1,234}$ = 266.454, $p < .001$, partial $\eta^2 = .532$), with high rated E-FITs being correctly identified significantly more often than low rated E-FITs.

The main effect of line-up size was significant, ($F_{2,234}$ = 10.560, $p < .0001$, partial $\eta^2 = .083$). However, E-FIT rating was not found to interact with line-up size, ($F_{2,234}$ = .038, $p = .963$, partial $\eta^2 = .00$). Employing the Bonferroni post-hoc test, significant differences were found between the 4-person and 6-person line-up conditions (p < .05) and between the 4-person and 8-person line-up conditions (p < .005). There was no significant difference between the 6-person and 8-person line-up conditions (p = .133).

Therefore it is suggested that there is no clear advantage for facial composites produced by field independent eyewitnesses. The by-rater analysis showing higher subjective likeness ratings for the E-FITs constructed by field independent eyewitnesses was not supported by the by-item analysis, although this may be a function of a lack of power in the by-item analysis.

7.3: Discussion of Study One Results

Overall the results of Study One were inconclusive with regard to answering the question of whether individual differences in field dependence/independence might make a difference in the accuracy of EFITs produced by eyewitnesses. The by-rater analysis which assessed the performance of field dependent /independent participants as a group showed a significant difference in similarity (to target face) scores for EFITs produced by field independent participants; those who find it relatively easier to extract detail from context. However, this small advantage for field independents in the by-rater analysis was not supported when they were analysed as individuals. In the by-item analysis there were no significant differences in accuracy scores for EFITs produced by field dependent and field independent participants.

There was no significant difference in the time taken to construct a facial composite between the two cognitive style groups of participants. With regard to the ratings used to determine the similarity of the E-FITs to the original target face, a second more objective measure showed that the E-FITs given high subjective likeness ratings were correctly matched to the target face significantly more often than the E-FITs given low subjective likeness ratings. This finding confirmed the validity of the subjective likeness rating as a measure of E-FIT utility in an objective context. The objective evaluation of the high and low rated E-FITs suggested that the ratings given to the composites in the subjective likeness rating phase were accurate and indicative of performance in a practical setting. Further issues concerning the measurement of accuracy of facial composites are addressed in the general discussion.

There are a number of possible interpretations for the results found within Study One. Firstly it could be argued that there is no difference in the ability of field dependent and field independent eyewitnesses to produce a target-accurate likeness of a perpetrator's face using E-FIT. Despite modifications to facial composite systems over the last 20 years, it remains the case that accurate face reproduction is a difficult cognitive task for all witnesses. It may be that no particular style confers an advantage. Alternatively it may be that the way in which field dependence/independence is measured does not differentiate between the two styles to a precise enough degree.

The construct of field dependence/independence has been useful in differentiating groups of individuals across a range of tasks, and the by-rater analysis of Study One which showed a small advantage for field independents *as a group* in E-FIT composite construction can arguably be added to that range of tasks. However, the method of testing for FDI has an inherent limitation. The Embedded Figures Test infers field independence from success at locating simple shapes embedded within a more complex pattern, and levels of field dependence are inferred from poor performance on this task. However, low scores on the GEFT could for example be due to other possible factors such as low motivation, tiredness or misunderstanding of test instructions.

Therefore, the small effect size of a positive advantage for field independent individuals found in the by-rater analysis warrants a larger scale study using a more recently developed test which positively assesses both ends of the cognitive style continuum being measured, in order to further investigate whether differences in cognitive style may be a factor in the likeness quality of facial composites produced using E-FIT. Study Two addresses these issues by using a larger sample of participants to construct a facial composite, and by using a cognitive style test which assesses both ends of the holistic/analytic cognitive style.

Chapter Eight: Study Two

Study Two was designed to investigate further the finding that there is a small but significant increase in the by-rater analysis of subjective likeness ratings given to facial composites produced by field independents. The objective evaluation of the top and bottom rated E-FITs in Study One suggested that the ratings given to the composites in the subjective likeness rating phase were accurate and indicative of performance in an objective setting. Field independent individuals show an enhanced ability to extract detail from its surrounding context relative to field dependents. Therefore one interpretation of the finding that they produced higher rated composites is that field independents are not as distracted by perception of a whole face image as field dependent individuals. Therefore field independent individuals would find it easier to perceptually isolate and select facial features more similar to their memory.

However, a criticism of the Group Embedded Figures Test for field dependence/independence is the fact that it positively assesses only one end of the FDI continuum. Field independence is measured by higher scores on the GEFT, and field dependence is inferred from relatively lower GEFT scores. An alternative FDI test, the Extended Cognitive Styles Analysis – Holistic/Analytic Test (E-CSA-W/A) was developed by Peterson et al. (2003) and is a computerised test of cognitive style which assesses both ends of the holistic/analytic cognitive style continuum. As described in Chapter Four, field dependence/independence is a cognitive style which is considered to be subsumed within the holistic/analytic group of cognitive styles (Riding & Cheema, 1991; Kozhevnikov, 2007). The holistic cognitive style can be considered to be analogous to field dependence where visual stimuli is encoded and processed in a relatively whole picture based way. The analytic cognitive style is therefore analogous to field independence where individuals can overcome the distraction of the prevailing visual field and more easily extract detail from context.

Also described in Chapter Four, the second major cognitive style group into which most other cognitive styles can be subsumed is the verbal/imagery style.

Peterson et al. (2005b) designed a Verbal Imagery Cognitive Styles (VICS) test which is presented on computer with the E-CSA-W/A. If there is an influence of visual perceptual cognitive style (field dependence/independence and/or holist analytic styles) on facial composite construction, then there may potentially also be an influence of representational cognitive style (how stimuli are represented in memory, either in words or pictures) as denoted by the verbal/imagery style.

As described in Chapter Two, faces are generally processed in a holistic manner, but the task of producing an E-FIT is similar to that which may suit a person with an analytic cognitive style – in order to construct an E-FIT individuals are required to consider each feature of the face in isolation both for describing and choosing parts for inclusion in the facial composite – albeit within the context of a whole face. There is a verbal element to composite construction in that the witness needs to describe the target face to the E-FIT operator. It is possible that the ability to describe more accurately a face to another individual might confer an advantage for facial composite construction. Similarly there may be an imagery element in composite construction to the extent that the witness must form a mental representation of the target face in their memory in order to attempt to recreate a similar image.

Therefore, the main aims of Study Two were to investigate whether there are differences in the accuracy of composites, as measured by likeness ratings given to E-FIT facial composites as a function of the holistic/analytic and verbaliser/imager cognitive styles.

8.1: Method

8.1.1: Design

A between-subjects design was employed for this study to determine whether there were differences between cognitive styles in ability to produce more accurate (similar to target-face) E-FITs. The cognitive styles used were the holistic/analytic cognitive style as measured by the Extended Cognitive Styles Analysis – Holistic/Analytic (E-CSA-W/A) test and the verbal/imagery cognitive style as measured by the Verbal Imagery Cognitive Style (VICS) test (Peterson et al., 2005b). The study consisted of two phases, the first of which involved participants constructing an E-FIT of a previously seen face, some weeks after having been tested for their cognitive style. Sixty participants constructed an E-FIT of a white male face and thirty of these returned at a later date to construct an E-FIT of a white female face. High attrition rates of participants for Study Two accounted for a 50% response to a request to return and construct a second facial composite. The second phase was an evaluation of the accuracy of the E-FITs produced (by means of subjective-likeness ratings on a percentage scale) which employed a new sample of participants.

8.1.2: Phase One - Cognitive Style Testing

Participants

Sixty university students and members of the general public (49 females, 11 males) participated in the construction phase of the study (mean age = 24.98 years, s.d. = 8.10). University students were recruited as part of their requirement for the Research Methods module research participation scheme. The participants were not familiar with the target faces, they reported no knowledge or previous experience of using the E-FIT program, and they had not previously completed the E-CSA-W/A or the VICS prior to taking part in the study.

Materials

The study used the computerised tests for holistic/analytic and verbaliser/imager cognitive styles (Peterson et al., 2005b) and participants took around 40 minutes in total to complete both tests. The E-FIT for Windows program (Version 5.0) (Aspley, 2004) was used to construct the E-FITs and Micrografx Picture Publisher 8 (1998) was used to make artistic enhancements to the E-FITs prior to their completion. A Compaq Presario M2000 laptop computer with a screen size of 30cm x 24cm (1280 x 1024 pixels) was used to run the cognitive style tests and to construct the E-FITs.

Twelve monochrome pictures of young adult Caucasian males and six monochrome pictures of young adult Caucasian females were used as target faces. The pictures were obtained from the Psychological Image Collection at the University of Stirling (http://pics.stir.ac.uk) who permit their free use for non-commercial research purposes. The photographs displayed a frontal full-face pose with a neutral expression, and measured 13 x 18cm when printed for use. Fifty new faces which had not been used in Study One were initially selected on the basis of having no distinguishing or outstanding features. The eighteen faces used in the study were selected by a third party to ensure that the experimenter was blind to the appearance of the target faces.

Procedure - Cognitive Style Testing

At least four weeks prior to constructing a facial composite, participants completed both the VICS (Verbal/Imagery Cognitive Styles) and the E-CSA-WA (Extended Cognitive Style Analysis-Holistic/Analytic) tests. The VICS takes approximately 25-30 minutes to complete and the E-CSA-WA test takes approximately 10-15 minutes. No information about how the tests measure cognitive style was given to participants prior to testing and participants received instructions and practice at the beginning of the test. The tests were both completed within the same session with short breaks given between them. The tests were always completed in the same order; the VICS test was completed first followed by the E-CSA-W/A. It was not possible to counterbalance the administration order of the tests because they were presented within one computer program which necessitated that the VICS test was always complete first.

The VICS test was designed to measure an individual's median reaction times on judgements made about words presented on screen (verbal element) and about images presented on screen (imagery element). To test reaction times to words, participants were presented with pairs of words on screen and asked to judge whether the items are both natural (e.g. an apple and a rabbit), both man made (e.g. a kettle and a chair), or mixed (e.g. one natural item and one man made item). The verbal element of the test contained 58 word pairs. To test reaction times to images, participants were presented with pairs of pictures on screen and asked to judge whether one of the items (in real life) is bigger, smaller, or the same size as the other item on screen (the items are presented as the same size on screen). The imagery element of the test also contained 58 image pairs which corresponded to the word pairs displayed in the verbal element of the test. The order in which the imagery and verbal elements were presented was randomised. The process of creating a verbal/imagery ratio accurate to three decimal points (based on the reaction times to the verbal and imagery tasks) resulted in each participant having a score somewhere along a verbaliser/imager continuum. Verbal/imagery style ratios between .8 and 1.0 suggest little or no style preference. Scores that are closer to zero indicate a tendency towards a verbal preference and scores that are closer to 2 indicate a tendency for an imagery preference.

The E-CSA-W/A test was designed to measure an individual's median reaction time on judgements made about the similarity of two shapes, and about whether one simple shape can be found embedded within a complex shape. For the holistic element of the test, participants were presented with 2 complex shapes on screen and asked to judge whether they are exactly the same or different. For the analytic element participants were presented with one complex shape and one simple shape and asked to judge whether or not the simple shape was embedded within the complex shape. There were 20 pairs of stimuli for each task, and they were randomly presented within the test. The process of creating a holistic/analytic ratio based on reaction times resulted in each participant having a score somewhere along a holistic/analytic continuum which, similar to the VICS ratio, was accurate to three decimal points. Holistic/analytic style ratios between .97 and 1.25 suggest little or no style preference. Scores that are closer to zero indicate a tendency towards a holistic preference and scores that are closer to 2 indicate a tendency for an analytic preference. The values for the allocation of both cognitive styles were derived from test norms suggested by Peterson et al. (2005b) based on previous research.

8.1.3: Phase Two - Composite Construction

Sixty participants who had previously completed the VICS and E-CSA-W/A were tested individually and each produced one composite likeness of one of the 12 male target faces (see Chapter 6, General Methods section 6.1 for the standard composite construction procedure used for studies 1, 2, and 3). The 12 male target faces were pseudo-randomised such that each was viewed by five participants thus producing five composites for each target face. Thirty of these participants returned on a separate occasion to construct a composite of one of six female target faces (again, five composites for each target face were produced). Participants received a standardised initial briefing which assigned them a participant number and instructed them to consult a randomised list which would inform them as to which target face they would view. They were asked to study the target face for one minute, with a view to making a composite construction of the face using the E-FIT program. This procedure ensured that the experimenter was blind to which target face was being viewed by the participant, as well as which faces had been selected for the study.

8.1.4: Evaluation – Subjective Likeness Ratings

Raters

Fifty students and members of the general public who had not made an E-FIT in the construction phase (37 females, 13 males) participated in the evaluation phase of the study (mean age = 29.68 years, s.d. = 11.26). The students received course credit time towards their Research Participation requirement for the Research Methods module.

Materials

Each target face was presented simultaneously with print-outs of the five E-FITS of that target face which had been produced by participants in the construction phase (the target face measured 13 x 18 cm and the E-FITs measured 10 x 15 cm). They were arranged such that only one target face and corresponding set of E-FITs could be viewed at any one time. The E-FITs were presented in the line-up in the order in which they had been constructed.

Procedure

Participants were given an initial briefing in which they were asked to assess each E-FIT for the degree of similarity to the target face, and to provide a percentage likeness rating using a scale from 1 to 100. See Chapter 6, General Methods section 6.6 for details of the subjective likeness ratings procedure.

8.2: Results

Study Two addressed the question of whether individual differences in holistic/analytic and verbaliser/imager cognitive styles may have an impact on the quality of composites that individuals produce using E-FIT. By-rater analyses were conducted first in order to assess whether there were differences in E-FIT composite accuracy ratings across groups of holistic/analytic and verbaliser/imager participants. The data were analysed using the categories of cognitive style identified by Peterson et al. (2005b), by applying the guidelines for the interpretation of the holistic/analytic and verbaliser/imager ratios published in the E-CSA-W/A and VICS administration guides.

By-rater tests of difference

The data were analysed by-rater which involved calculating the mean scores each individual rater gave to the composites produced by holistic and analytic participants and those who fell into neither cognitive style category. A Cronbach's Alpha test was conducted on the subjective likeness ratings given by the 50 raters in the evaluation phase to assess inter-rater reliability. Cronbach's Alpha for the subjective likeness rating scale across the full sample of 50 raters was .910, indicating a high degree of agreement among raters comparing the similarity of the E-FITs to the original target face. Table 8.1 shows the mean percentage scores given to groups of holistic and analytic participants by individual raters in the subjective likeness rating task. The composites produced by participants with a holistic cognitive style received the highest overall mean ratings, and the composites produced by participants who had an analytic cognitive style or did not display a preference for either cognitive style were given similar ratings.

Table 8.1: Mean subjective likeness rating scores (%) for holistic/analytic groups aggregated by-rater in Study Two

Cognitive Style	Mean (SD)	N
Holistic	43.68 (10.31)	50
Analytic	39.44 (9.86)	50
Neither	39.03 (9.09)	50

A one-way within-subjects ANOVA was conducted on the data to examine whether there was a significant difference in the subjective likeness ratings each rater gave to composites constructed by holistic, analytic or 'neither' participants. The within-subjects factor had 3 levels (holistic style, analytic style, neither group). There was a significant effect of cognitive style group with holistic participants producing significantly higher rated E-FITs than analytic participants or those who fell into neither category, ($F_{2,98} = 8.596$, $p < .001$, partial $\eta^2 = .149$). Post-hoc paired samples t-tests, revealed a significant difference between the holistic and analytic groups ($p < .005$) (Cohen's $d = 0.4$) and between the holistic and neither groups ($p < .001$) (Cohen's $d = 0.4$). There was no significant difference between the neither and analytic groups ($p =.695$) (Cohen's $d = 0.04$). Therefore the holistic participants produced higher rated E-FITs as a group.

A further by-rater analysis was conducted calculating the mean scores each individual rater gave to the composites produced by groups of verbaliser and imager participants and those who fell into neither category. Table 8.2 shows the mean percentage scores given to groups of verbaliser and imager participants by individual raters in the subjective likeness ratings task. The mean rating scores were very similar for each group, and verbalisers received marginally the highest mean ratings.

Table 8.2: Mean subjective likeness rating scores (%) for verbaliser/imager groups aggregated by-rater in Study Two

Cognitive Style	Mean (SD)	N
Verbaliser	41.86 (10.62)	50
Imager	40.81 (9.74)	50
Neither	39.45 (8.98)	50

A one-way within-subjects ANOVA was conducted on the data to examine whether there was a significant difference in the subjective likeness ratings each rater gave to composites constructed by verbaliser, imager and 'neither' participants. The within-subjects factor had 3 levels (verbaliser style, imager style, neither group). There was no significant effect of verbaliser/imager cognitive style group on the ratings given to the E-FITs, ($F_{2,98} = 1.833$, $p = 0.165$, partial $\eta^2 = .036$).

By-item tests of difference

In addition to the by-rater analyses the data were analysed by-item which involved calculating the mean percentage score given to each individual composite across all raters. Two separate analyses were conducted; one for the holistic/analytic cognitive style and one for the verbaliser/imager cognitive style. Table 8.3 shows the mean subjective likeness rating scores given to the 60 composites of male faces produced by participants who were categorised as holistic, analytic or neither style category. The mean subjective likeness ratings show that higher accuracy ratings were given for E-FITs produced by both holistic and analytic participants in comparison to participants who did not show a cognitive style preference.

Table 8.3: Mean subjective likeness rating scores (%) by holist/analytic cognitive style groups calculated by-item in Study Two

Cognitive Style	Mean (SD)	N
Holistic	44.67 (8.97)	15
Analytic	40.08 (15.37)	13
Neither	38.97 (12.95)	32

A one-way between-subjects ANOVA was conducted on the subjective likeness rating data to see whether the by-rater effect of higher ratings for composites constructed by holistic participants was also present in the by-item analysis. The between-subjects factor had three levels (holistic style, analytic style, neither). There was no significant difference between any of the cognitive style groups and mean subjective likeness ratings ($F_{2,57}$ = 1.047, p = .358, partial η^2 = .035) when the data were analysed by-item.

Table 8.4 shows the mean subjective likeness rating scores given to the 60 composites of male faces produced by participants who were categorised as verbaliser, imager or neither style category. The mean subjective likeness ratings show that accuracy ratings across all three groups were very similar.

Table 8.4: Mean subjective likeness rating scores (%) by verbaliser/imager cognitive style groups calculated by-item in Study Two

Cognitive Style	Mean (SD)	N
Verbaliser	41.82 (15.28)	11
Imager	41.23 (13.86)	17
Neither	39.91 (11.43)	32

A one-way between-subjects ANOVA was conducted on the subjective likeness ratings data to examine whether there were any differences in accuracy ratings of the E-FITs. The between-subjects factor had three levels (verbaliser style, imager style, neither). There was no significant difference between any of the cognitive style groups and mean subjective likeness ratings when the data were analysed by-item ($F_{2,57}$ = .116, p = .891, partial η^2 = .004).

By-item correlations

Supplementary correlational analyses of the data were conducted in order to utilise the individual cognitive style ratios of participants who took part in the composite construction phase. The first analysis examined participants' individual scores on the E-CSA-W/A in relation to the mean likeness rating their composite received from all raters. Figure 8.1 displays the distribution of participants' scores on the holistic/analytic continuum in relation to the mean rating given to their E-FITs in the subjective likeness rating task.

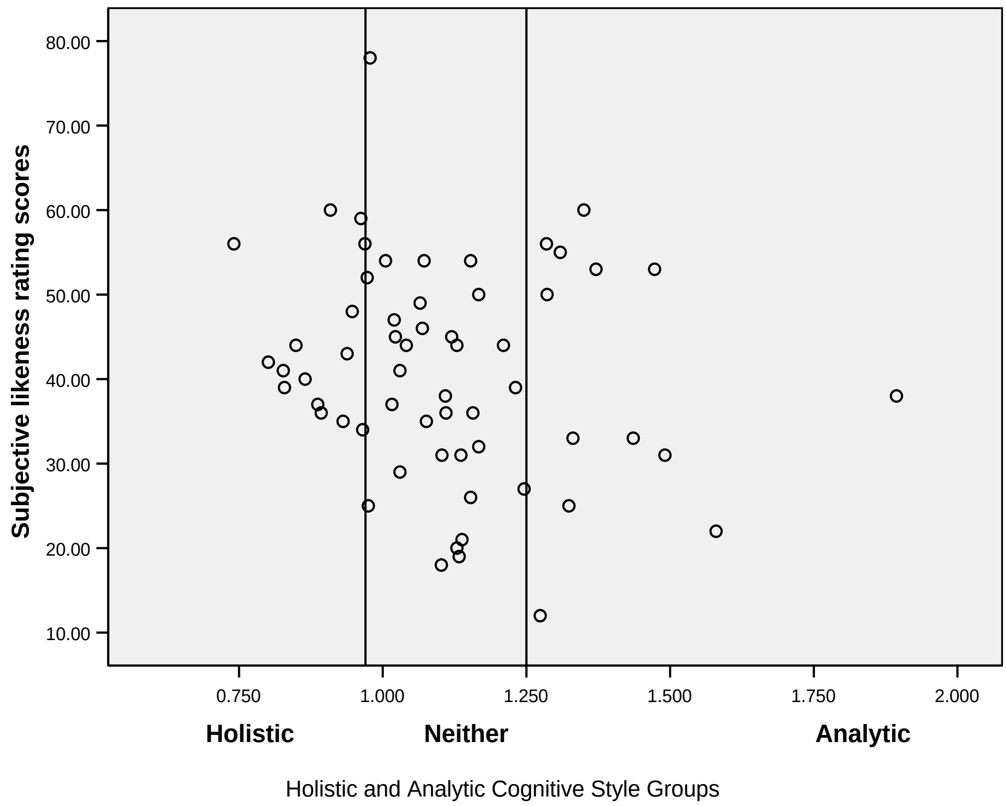

Figure 8.1 - Subjective likeness rating scores for composites produced by participants with holistic/analytic cognitive styles divided into terciles for individual style ratios in Study Two

A Pearson's correlation revealed no significant correlation between scores on the holistic/analytic continuum and the subjective likeness rating score of the E-FITs, (r = -.194, N = 60, p = .138, two tailed). The scores on the holistic/analytic scale are devised such that participants at the holistic end of the continuum achieve lower numerical scores than participants at the analytic end therefore the negative correlation coefficient indicates that subjective likeness ratings increased toward the holistic end of the continuum. Although there was a trend in the direction of higher scores for composites produced by holistic participants this did not reach statistical significance.

The second analysis examined participants' individual scores on the VICS in relation to the mean likeness rating their composite received from all raters. Figure 8.2 displays the distribution of participants' scores on the verbaliser /imager continuum in relation to the mean ratings given to their E-FITs in the subjective likeness rating task.

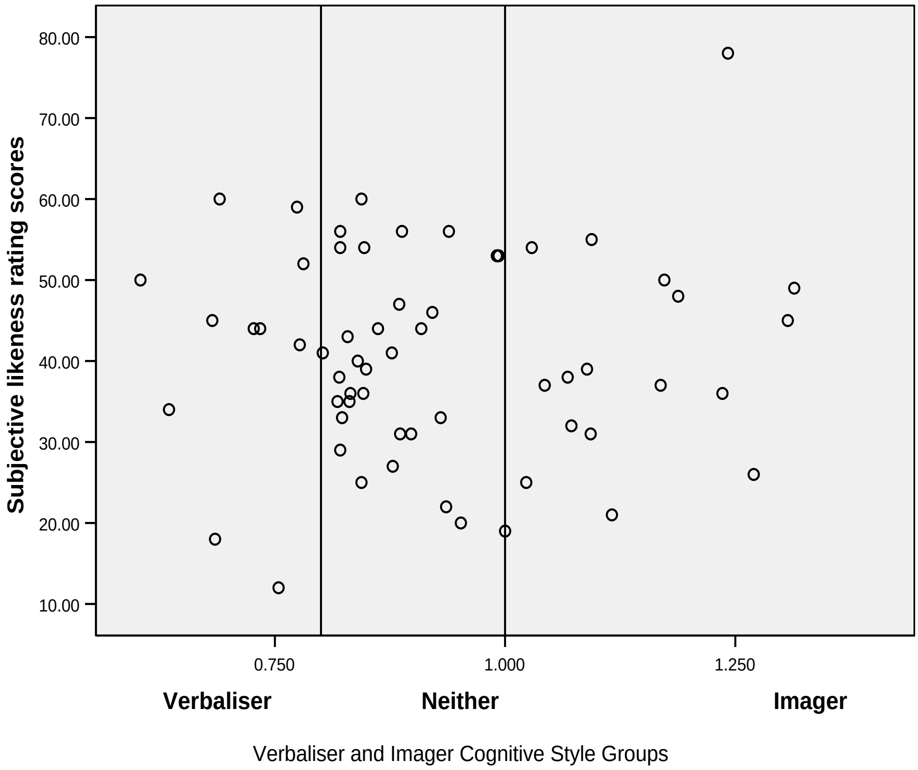

Figure 8.2: Subjective likeness rating scores for composites produced by participants with Verbaliser or Imager cognitive styles divided into terciles for individual style ratios in Study Two

A Pearson's correlation revealed no evidence for a correlation between scores on the verbaliser/imager continuum and the subjective likeness rating score of the E-FITs when cognitive style was treated as a continuous variable, (r = .042, N = 60, p = .752, two tailed). Thirty participants who produced an E-FIT of a male face returned one year later to construct another E-FIT of a female face.[2] The following analyses were based upon the sub-sample of 30 participants who constructed both a male and a female face. The strength of the association between subjective likeness rating scores given to E-FITs created on two different occasions was evaluated in order to assess any stable individual

[2] All participants who constructed an E-FIT of a male face were invited to attend one year later to construct a second E-FIT of a female target face. High rates of attrition were observed due to the fact that participants were outside the bounds of receiving course credit for participation in research.

differences in the ability to produce E-FITs which bear some similarity to the target face. Figure 8.3 shows the distribution of subjective likeness rating scores for the female faces produced by 30 participants, in relation to the subjective likeness ratings given for the male faces.

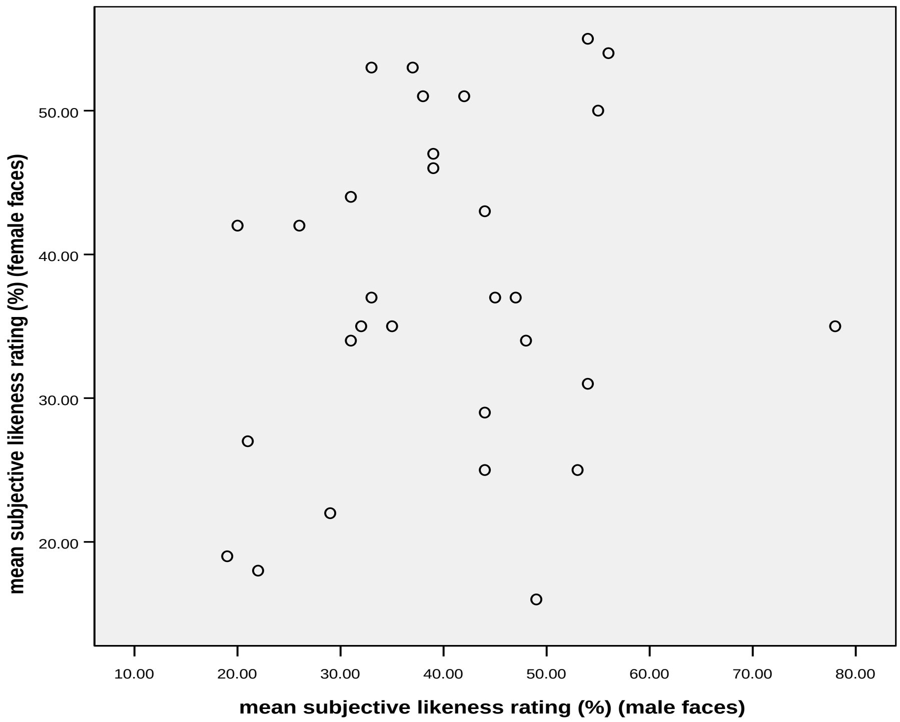

Figure 8.3: Subjective likeness rating scores for the sub-sample of participants who produced a composite of a male and a female face calculated by-item in Study Two

A Pearson's correlation revealed no evidence for a significant correlation between rating scores participants received for the male E-FIT and the female E-FIT they constructed (r = .194, N = 30, p = .303, two-tailed). A paired t-test was also conducted on the subjective likeness rating scores for the male and female E-FITs. The mean scores obtained for the male and female target faces were very similar (male mean = 37.57, s.d = 11.54; female mean = 39.93, s.d. = 13.08). The paired t-test revealed that there was no significant difference between the mean scores obtained by the E-FITs of male faces and female faces (t = .827, df = 29, p = .415, two tailed) (Cohen's d = 0.19). This would suggest that overall there was no practice effect acquired by participants who produced two E-FITs. An additional paired t-test showed no significant difference between the mean scores obtained by the E-FITs of male faces between the group of thirty participants who returned at a later date to construct a second facial composite and the group of thirty participants who did not (t = .424, df = 58, p = .673, two tailed) (Cohen's d = 0.11). This suggests that there were no large differences in motivation or conscientiousness between the group of participants who returned to construct a second E-FIT and those who did not, at least in terms of the subjective likeness ratings that their E-FITs received.

Individual performance over two attempts was averaged to increase the reliability of the measure of individual performance. Figure 8.4 displays the distribution of participants' scores on the holistic/analytic continuum in relation to the mean ratings given to the male and female E-FITs they produced.

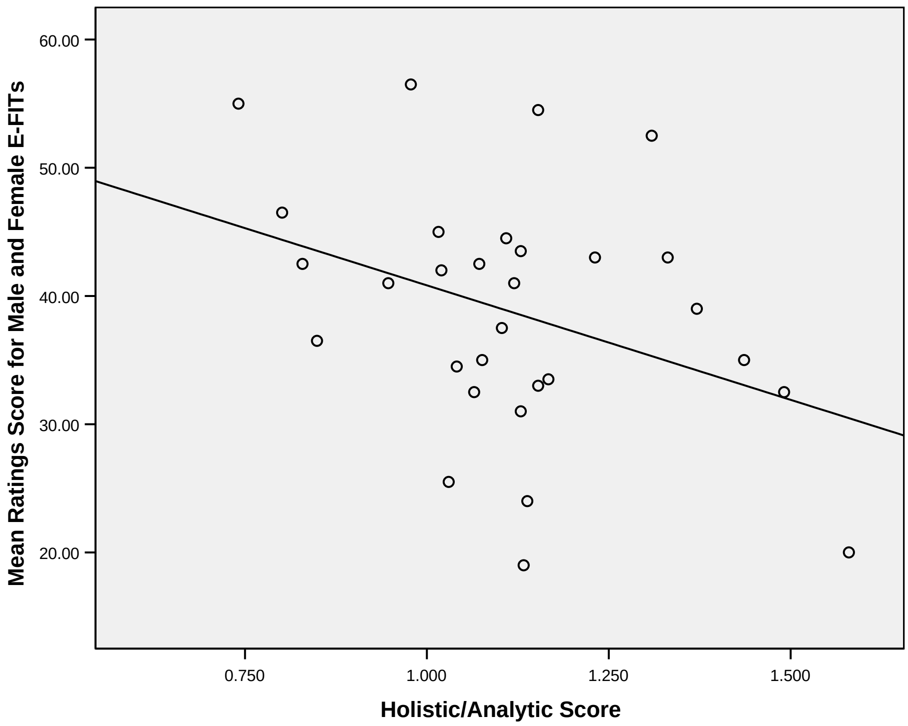

Figure 8.4: Mean subjective likeness rating scores of male and female composites combined and holistic/analytic cognitive style scores in Study Two

A Pearson's correlation revealed some evidence for a correlation between scores on the holistic/analytic continuum and the mean subjective likeness rating score achieved over two E-FIT attempts, (r = -.365, N = 30, p = .047), which suggests that individuals with a holistic cognitive style score achieved higher mean subjective likeness ratings when their scores were averaged over two E-FIT attempts. A further Pearson's correlation revealed no evidence for a correlation between scores on the verbaliser/imager continuum and the mean subjective likeness rating score achieved over two E-FIT attempts $r = .062$, N = 30, $p = .746$.

<u>Further Analyses</u>

The E-CSA-W/A test for holistic/analytic cognitive style produces a median reaction time score for both the holistic and the analytic parts of the test. The task for the analytic part of the test is similar to the group embedded figures test for field dependence/independence, as the participant is required to state whether a simple figure on the screen can be located within a complex figure displayed on the screen simultaneously. The task for the holistic part of the test is not assessed within the group embedded figures test. The data from the two halves of the E-CSA-W/A were broken down into separate components to represent each component in the GEFT. This was in order to check whether the differences found in Studies One and Two for cognitive style could be attributed to the way in which cognitive style is measured by each test.

Two correlations were conducted comparing median reaction time scores for each of the holistic and analytic elements of the E-CSA-W/A with the mean subjective likeness rating scores. There was no correlation between the median reaction time for the analytic half of the test and likeness ratings, ($r = -.077$, $N = 52$, $p = .588$, two tailed), and no correlation between the median reaction time for the holistic half of the test and likeness ratings, ($r = -.163$, $N = 52$, $p = .248$, two tailed).[3] If the difference between the studies had been due to the measurement properties of the two scales used, then the same relationship would be expected between the holistic half of the test and the likeness ratings as was observed for the GEFT in Study One. This analysis does not provide evidence that the two halves of the E-CSA-W/A function differentially, or that scores on this measure of cognitive style are associated with likeness ratings.

The significant correlation for holistic cognitive style and subjective likeness ratings was found in the mean score given to two facial composites that the

[3] N = 52 in the correlational analyses reported above as some reaction time data from one testing laboratory were lost due to computer error.

sub-sample of 30 participants produced. Therefore two further correlations were conducted. The only correlation to approach significance was found when the mean score for male and female E-FITs combined was correlated with median reaction time for the holistic half of the test, (r = -.354, N = 28, p = .064, two tailed). The correlation of the mean subjective likeness rating score for male and female E-FITs and median reaction time for the analytic half of the test was not significant (r = -.098, N = 28, p = .618). The negative correlations suggest that faster reaction times are correlated with higher ratings.

8.3: Discussion of Study Two results

Study Two was designed to investigate further the finding in Study One that the subjective likeness ratings given to facial composites produced by field independent individuals were significantly higher than the ratings given to composites produced by field dependent individuals. A computerised test of holistic/analytic and verbaliser/imager cognitive styles was administered to participants who then went on to produce a facial composite of a previously seen unfamiliar face using E-FIT. The results of Study Two showed some differences and correlations which suggested that the holistic group of participants received significantly higher ratings for the facial composites they produced than the analytic group and those who fell into neither category. However for the verbaliser/imager cognitive style dimension there was no evidence of differences in performance between the cognitive style categories.

By-rater there was no increase in likeness ratings associated with being either a verbaliser or an imager when participants were separated into cognitive style groups. Additional by-item analyses also showed that there was no correlation between likeness ratings and individual scores, based on individual verbaliser/imager style ratios. The task of producing a facial composite using the E-FIT system is one that requires both verbal and imagery elements in combination. With this in mind, the present results could be interpreted as suggesting that being either a verbaliser or an imager does not confer an advantage in facial composite construction because the individual advantage of being *either* style might cancel the other out. However, if this were the case it might be expected that being either a verbaliser or an imager could be more advantageous than being in the 'neither' style category. The fact that being *neither* style was no more detrimental to composite construction than being *either* style suggests that there is no individual advantage of being either a verbaliser or an imager for the purpose of accurate facial composite construction.

The results were however, mixed with regard to any association between the holistic/analytic cognitive styles and the production of accurate facial composites. By-rater tests of difference showed that composites produced by holistic participants as a group received significantly higher likeness ratings than composites produced by analytic participants or those who fell into neither category. By-item tests of difference revealed no significant differences between the two style categories and likeness ratings, although the pattern of mean subjective likeness scores was in the same direction as the by-rater analysis, as both the holistic and the analytic groups received higher likeness ratings than the neither style group.

By-item analyses based on individual holistic/analytic style ratios revealed no significant correlations between cognitive style and likeness ratings. However, in a sub-set of thirty participants who constructed two facial composites (one male face and a year later one female face) there was a significant correlation between cognitive style and their averaged likeness ratings for both composites combined; participants with a holistic style received significantly higher combined ratings for two composites than participants with an analytic style or those with neither cognitive style. It should be noted that there was no significant correlation between the ratings given to the individual male face and female face composites made by each participant. This may indicate that a single measure of E-FIT performance is weakly indicative of the general performance level of an individual, so an average over two E-FIT attempts could be a more consistent measure. Alternatively it may be that the reliability of the effect of cognitive style is masked by other individual variables such as personality, motivation or general ability, as the effect of cognitive style when detected is small.

Further analyses on the sub-set of thirty participants who constructed two facial composites revealed that there were no significant differences between the rating scores given to E-FITs of male faces and E-FITs of female faces; this demonstrates that there was no practice effect found after completion of a second facial composite. Additionally, this suggests that there was no detriment to constructing a female face within E-FIT, where a detriment might be expected due to the fact that there are considerably fewer exemplars of female features in the female database within E-FIT than there are of male features within the male database.

To summarise, holistic individuals produced E-FITs which received higher subjective likeness ratings than those produced by analytic individuals (who have a preference for part based processing) and this would seem to contradict the finding from Study One that field independents (part based processors) did produce higher rated E-FITs. However, the contradictory nature of the findings from both Studies One and Two are based on the assumption that the cognitive style of field independence/independence is entirely analogous to the holistic/analytic cognitive style. Although the two cognitive style constructs share many similarities, and the field dependence/independence style is said to be subsumed within the holistic/analytic family of styles, there are differences between the two style constructs which are addressed further in the general discussion.

Overall, there were differences in the results of Studies One and Two in terms of the cognitive styles which appeared to be beneficial for facial composite construction, and in terms of by-item and by-rater analyses. The inconsistent nature of the results from Studies One and Two suggests that perhaps the way in which information is processed and represented in memory has little relation to a face recall task which requires several cognitive elements. Regardless of which cognitive style an individual prefers (or does not) the task of constructing a facial composite using E-FIT requires both holistic processing in terms of remembering the target face and viewing the whole face on screen, *and* featural processing (selecting, resizing and repositioning features).

There are no published studies examining the relationship between cognitive style and facial composite construction, and this may be because if there is any effect of cognitive style on composite construction performance it is too small to be of practical benefit. According to Kozhevnikov (2007) individual differences in cognitive style do exist, but their effects are often overwhelmed by other factors such as general abilities and the cognitive constraints that all human minds have in common. An interesting avenue of research which has produced larger effect sizes in eyewitness face *recognition* performance is the manipulation of cognitive processing *state* (rather than style) first reported by Macrae and Lewis (2002). Study Three applied the manipulation of cognitive processing state to the task of constructing a facial composite using E-FIT to investigate whether this may be beneficial in terms of producing facial composites with a higher degree of similarity to the target face.

Chapter Nine: Study Three

Study Three was designed to investigate whether manipulating the cognitive processing state (rather than style) of an individual might have an effect on face recall and subsequently on the accuracy of the facial composites they produce. Previous research has found that manipulating the cognitive processing state of individuals specifically through completion of a Navon task, (Navon, 1977) has an effect on face *recognition* performance (Macrae & Lewis, 2002; Weston et al., 2008). The process of building a facial composite using E-FIT draws more on recognition the nearer the witness gets to the end of the process. In the initial stages, the witness must select individual features within the context of the whole face. Once all the initial features have been selected, the witness must then decide whether the composite face they have constructed matches their memory of the target face they previously viewed. Thus, cognitive processing states influenced by the Navon task might be predicted to affect the accuracy of facial composites that individuals produce by affecting the cognitive processes which are utilised for face recognition.

There are currently no published studies investigating the effect of inducing cognitive processing orientation using the Navon task on subsequent composite construction performance, therefore it is unclear which type of Navon task might be beneficial for eyewitnesses. Following on from the finding of Study One that field independent or part based processing individuals produced higher rated E-FITs in the by-rater analysis, it might be predicted that inducing a featural processing strategy by using a local Navon task during composite construction would produce higher rated composites. However, following the finding from Study Two that holistic or global based processing individuals produced higher rated E-FITs in the by-item analysis when the rating scores for two E-FITs were averaged, it might instead be predicted that inducing a holistic processing strategy by using a global Navon task during composite construction would produce higher rated composites.

In addition to a global or local Navon task being used during facial composite construction, holistic and featural processing was introduced at the face encoding stage to investigate if this would have an effect on composite accuracy. Wells and Hryciw (1984) encouraged participants to encode faces on either a featural level or a holistic level. Results showed that those who had been encouraged to encode faces in a featural way made higher rated facial composites, and those who had been encouraged to encode faces in a holistic way made more correct identifications in a line-up task. This lends support to the idea that if there is a match between cognitive processing state at encoding and processing in a subsequent composite construction or identification task there will better performance by eyewitnesses. Therefore it may be the case that the way in which faces are initially encoded has an impact on the effect of the Navon task used during composite construction, depending on whether there is a match between encoding and Navon task. The current study investigated whether there would be a featural face encoding advantage using E-FIT, and also whether there would be an interaction between holistic/featural encoding and the Navon task that participants completed during E-FIT construction.

In summary, the main aim of Study Three was to assess the effect of introducing a local or global Navon task into the composite construction process. The Navon task has consistently been found to affect face recognition performance, with the global Navon facilitating face recognition performance and the local Navon task hindering face recognition performance relative to a no Navon task control group. However, the question remains of whether the Navon effect can be generalised to facial composite construction, and if so what direction that effect would take when the process of constructing an E-FIT is one which has both featural and holistic elements.

9.1: Method

9.1.1: Design

A 2 (holistic/featural encoding) x 3 (global, local, control) independent groups design was employed in which participants were assigned to either the holistic or the featural encoding group and were then subsequently assigned to complete global Navon tasks or local Navon tasks during composite construction, or were assigned to a control group which had no Navon intervention during composite construction. The study consisted of two phases, the first of which involved participants constructing an E-FIT of a previously seen unfamiliar face while completing either a global or local Navon task intermittently throughout the composite construction process. The second phase was an evaluation of the accuracy of the E-FITs produced (by means of subjective likeness ratings on a percentage scale) which employed a new sample of participants.

Participants (Construction Phase)

Seventy two University students and members of the general public participated in the construction phase of the study (13 males, 59 females, mean age = 22.11 years, s.d = 5.54). University students were recruited via the Research Participation Scheme and received course credit for participation. The participants were not familiar with the target faces, they had no knowledge or previous experience of using the E-FIT program, and they had not previously completed a Navon task prior to taking part in the study.

Materials

Two sets of questions were devised which were designed to induce either featural or holistic encoding of the target faces (see Appendices VI and VII). The featural encoding questions required participants to rate individual features of the face out of 10, and the holistic questions required participants to give personality ratings (out of 10) to the target face.

Twelve monochrome pictures of young adult Caucasian males were used as target faces. The pictures were obtained from the Psychological Image Collection at the University of Stirling (http://pics.stir.ac.uk) and displayed a frontal full-face pose with a neutral expression (measuring 13 x 18 cm when printed for use).

The E-FIT for Windows program (Version 5.0) (Aspley, 2004) was used to construct the E-FITs and the Micrografx Picture Publisher 8 (1998) program was used to make artistic enhancements to the E-FITs prior to their completion. A Compaq Presario M2000 laptop computer with a screen size of 30 x 24 cm (1280 x 1024 pixels) was used both to construct the E-FITs and to run the Navon task.

Navon task materials were provided on request by Professor Tim Perfect from the University of Plymouth. The Navon task materials used for this study were global precedence white Navon letters on a black background (see section 6.4), and they measured 5 x 8 cm when displayed on the computer screen. The same Navon materials have previously been used by Perfect et al., (2008).

9.1.2: Phase One – Composite Construction

Seventy two participants were tested individually and each produced one composite likeness of one of the twelve male target faces. The target faces were pseudo-randomised such that each was viewed by six participants thus producing six composites for each target face. The target faces were also pseudo-randomised such that each face was viewed once within each of six cells of the design. Participants were assigned to a particular condition based on the order in which they arrived to participate in the study. Participants received a standardised initial briefing which assigned them a participant number and instructed them to consult a list which would inform them as to which target face they would view. Participants then selected the envelope which contained the target letter corresponding to their participant number, and then proceeded to take the picture of the face from the envelope.

The participants were asked a series of questions relating to the face, and these were either global processing questions such as 'can you rate this face out of 10 for honesty', or local processing questions such as 'can you rate the nose out of 10 for attractiveness'. The interaction was timed so that each participant was exposed to the picture of the face for no longer than one minute. Participants then replaced the picture in the envelope, and placed all envelopes back in their original alphabetical order.

A short maths task followed in which participants were required to answer a series of maths questions which required them to make subtractions of seven from a list of numbers (duration 2 minutes) after which time participants were stopped. After the 2 minute intervention participants gave a free-recall description of the face they had seen, this was followed by cued recall from the experimenter. Following on from this, the participants were introduced to the E-

FIT system and its features. The process of making an E-FIT with the system was described to them.

Those participants in the control condition then constructed a facial composite in the standard way as described for Studies One and Two (see Chapter 6, General Method Section 6.1). The participants in the global processing condition were given a global Navon task for one minute before commencing selection of their first feature. Participants in the local processing condition were given a local Navon task, also for one minute, before commencing selection of their first feature (see Chapter 6, General Method section 6.2). Once participants had decided upon the feature they wished to select, they were then given another Navon task before proceeding on to the next feature. The process of including the Navon tasks in between feature selection added approximately seven to eight minutes to the whole construction process as this was repeated for face shape, hair, brows, eyes, nose, ears, lips and artistic enhancement (if required by the participant). Once the composite was completed to the participant's satisfaction, it was then saved and the participant thanked and debriefed as to the purpose of the study.

9.1.3: Phase Two – Evaluation – Subjective Likeness Ratings

Raters

Forty students and members of the general public who had not made an E-FIT in the construction phase (14 males, 26 females) participated in the evaluation phase of the study (mean age = 23.2 years, s.d. = 7.23). The students received half an hour course credit time towards their Research Participation requirement for the Research Methods module.

Materials

Each target face was presented simultaneously with print-outs of the six E-FITs of that target face which had been produced by participants in the construction phase (the target face measured 13 x 18 cm and the E-FITs measured 10 x 15 cm). They were arranged such that only one target face and corresponding set of E-FITs could be viewed at any one time. The E-FITs were presented in a line-up which represented the order in which they appeared in the original randomised list in the construction phase. Participants were provided with a prepared response sheet which provided spaces for each single evaluation of E-FIT compared to the original target face.

Procedure

Subjective likeness ratings were elicited in accordance with the general methodology used for all Studies (see Chapter 6, General Method, Section 3.6).

9.2: Results

The aim of Study Three was to investigate whether manipulating cognitive processing state might have an effect on the accuracy of the facial composites that individuals produce. Participants were assigned to one of six conditions and completed a Navon task (or no Navon task in the control group) both prior to and during construction of a facial composite using E-FIT. The data were analysed by-item (composite) and by-rater (participant).

A Cronbach's Alpha test was conducted on the subjective likeness ratings given by the 40 raters in the evaluation phase to assess inter-rater reliability. Cronbach's Alpha for the subjective likeness rating scale across the full sample of 40 raters was .940, indicating a high degree of agreement among raters comparing the similarity of the E-FITs to the original target face.

By-rater analysis

The by-rater analysis showed that featural encoding of faces resulted in facial composites which received higher rating scores when averaged across all experimental conditions. Figure 9.1 shows the mean subjective likeness rating scores by condition, aggregated by-rater.

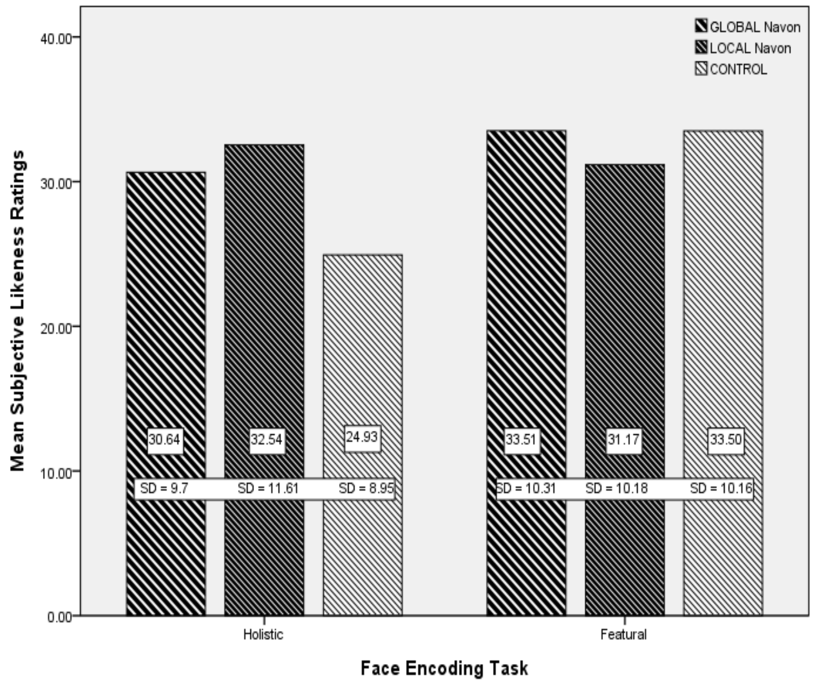

Figure 9.1: Mean subjective likeness rating scores for six experimental conditions aggregated by-rater in Study Three

A 2 (face encoding) x 3 (Navon task) within-subjects ANOVA was conducted on the by-rater data to see if there were any differences in the ratings given to the composites when applied to the 6 cells of the design. The main effect of face encoding was significant, with participants in the featural face encoding group producing significantly higher rated E-FITs than participants in the holistic face encoding group ($F_{1,39}$ = 54.728, p <.001, partial η^2 = .584) . The main effect of Navon task was also significant, ($F_{2,78}$ = 7.425, p < .001, partial η^2 = .160), with participants who completed either Navon task producing higher rated E-FITs than participants in the control condition. However these main effects need to be interpreted in light of the significant interaction between encoding condition and Navon task condition, ($F_{2,78}$ = 16.755, p < .001, partial η^2 = .301), where the holistic encoding of faces and local Navon task produced higher rated E-FITs, as did the featural encoding of faces and global Navon task.

Figure 9.2 shows the interaction between face encoding and Navon task for the by-rater analysis.

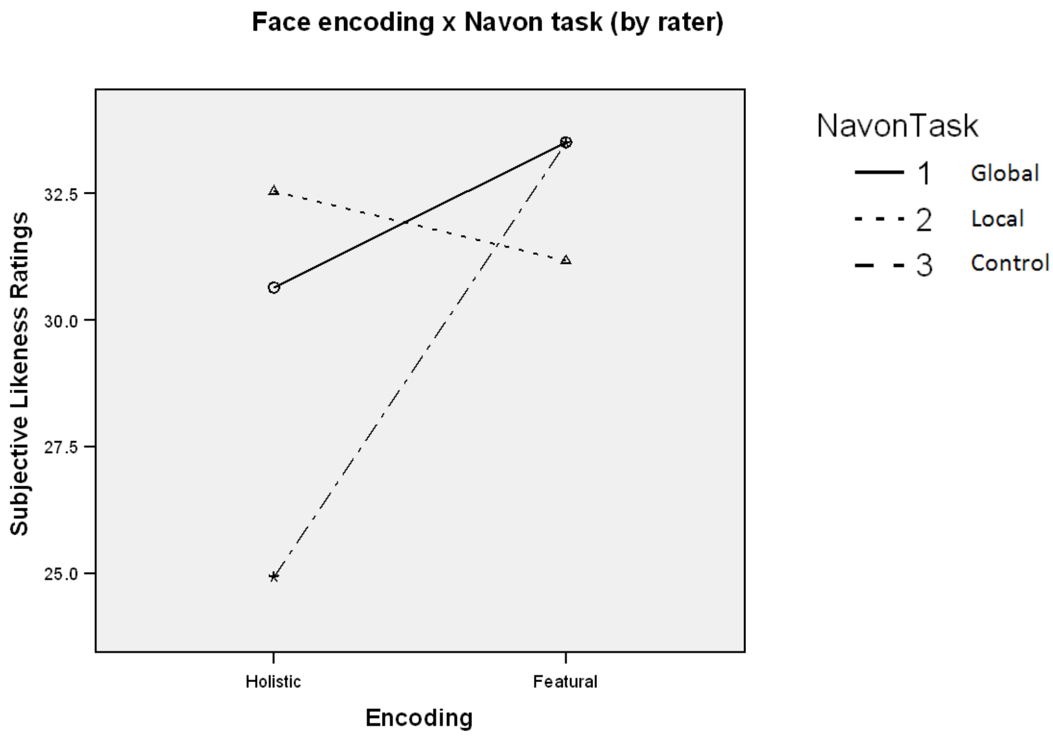

Figure 9.2: The interaction between face encoding and Navon task groups aggregated by-rater in Study Three

A one-way ANOVA looking at the simple main effect of Navon task in the holistic face encoding condition showed that the effect of Navon task was significant ($F_{2,78} = 27.499$, $p < .001$, partial $\eta^2 = .414$). Post hoc paired t-tests revealed that there was a significant difference in subjective likeness ratings, with the global Navon group receiving higher ratings for their composites than the control group, ($t = 5.544$, $df = 39$, $p < .001$, two tailed) (Cohen's $d = 0.61$). The local Navon group also received higher ratings for their composites than the control group, ($t = 6.736$, $df = 39$, $p < .001$, two tailed) (Cohen's $d = 0.73$). There was no significant difference in likeness ratings between the two Navon conditions, thus showing that following holistic encoding, completing either type of Navon task led to composites which received significantly higher ratings than the control group.

By contrast, a one-way ANOVA looking at the simple main effect of Navon task following featural face encoding found that there was no main effect of Navon task, ($F_{2,78} = 2.124$, $p = .126$, partial $\eta^2 = .052$).

Finally, looking at the no-Navon control group over both encoding conditions it was found that the featural face encoding condition led to significantly higher ratings for composites than the holistic encoding condition (t = 7.296, df = 39, p < .05, two tailed) (Cohen's d = 0.8).

By-item analysis

Table 9.1 shows the mean subjective likeness rating scores given to the 72 E-FITs produced by participants who were in one of six conditions; holistic/featural encoding x global Navon/local Navon/control group. E-FITs constructed by participants in the featural face encoding condition generally received higher ratings than E-FITs in the holistic encoding condition.

Table 9.1: Mean subjective likeness rating scores (%) by face encoding x Navon group conditions calculated by-item in Study Three

Encoding Task	**Navon task**	**Mean (SD)**	**N**
Holistic	**Global**	30.64 (10.87)	12
	Local	32.54 (12.92)	12
	Control	28.20 (11.94)	12
Featural	**Global**	33.50 (15.17)	12
	Local	31.17 (11.47)	12
	Control	33.80 (14.15)	12

A 2 (encoding condition) x 3 (Navon task) between-subjects (ANOVA) was conducted to see if there were any differences in the subjective likeness ratings of the E-FITs which were produced by participants in the six conditions. There was no significant main effect of Navon task, $F_{2,66}$ = .051, p = .950, partial η^2 = .002. There was also no significant main effect of encoding task, $F_{1,66}$ = .595, p = .443, partial η^2 = .009 and no significant interaction between encoding task and Navon task, $F_{2,66}$ = .438, p = .647, partial η^2 = .013.

9.3: Discussion of Study Three results

Study Three was designed to investigate whether the Navon effect, which has previously been observed in face recognition studies, could be extended to facial composite construction using E-FIT. In addition the way in which the target faces were encoded was manipulated. This was intended to determine whether any effect of Navon task was specific to encoding strategy, in light of previous research which suggests that if there is a match between cognitive processing strategy at encoding and retrieval this will enhance memory performance on a facial composite construction task (Wells & Hyrciw, 1984).

The results of the by-rater analysis revealed a significant main effect of face encoding condition with featural encoding leading to significantly higher rated composites than holistic encoding, as evidenced by the simple effect of encoding showing a featural advantage for the control group. There was also a significant main effect of Navon task. A significant interaction between Navon task and encoding condition and follow-up tests revealed there to be an advantage for either Navon task group over the control group which was restricted to those encouraged to encode the face holistically. Finally, inspection of the means for the Navon x encoding indicated that the combinations of featural encoding and global Navon task, and holistic encoding and local Navon task, tended to lead to higher rated facial composites. In the by item analysis there were no significant main effects or interactions.

There was a simple effect of encoding (by-rater), where featural encoding of the target face led to higher ratings of accuracy for the facial composites produced within the control group. Wells and Hyrciw (1984) demonstrated a featural encoding advantage for composites made using Photo-FIT, and Frowd et al. (2008) demonstrated a featural encoding advantage for composites made using PRO-fit. Frowd et al. (2012) also showed a featural encoding advantage for EvoFIT, and the present study is the first to demonstrate a featural encoding advantage for E-FIT. The Photo-FIT system relies on featural processing at the composite construction stage, as witnesses choose features from a book where they are viewed in isolation from the whole face, and the chosen features are then assembled to make a face image. The E-FIT system was designed to capitalise on the idea that faces are processed and remembered in a holistic manner and not as individual sets of features, and features can be chosen within the context of a whole face image. Given that a whole face image is present for witnesses during E-FIT construction it might be expected that holistic processing at the encoding stage would lead to higher rated E-FITs. The advantage for featural encoding in the present study suggests that for witnesses, featural processing contributes to the quality of composites produced using E-FIT.

Inspection of the means showed that there was a tendency for featural encoding/global Navon task and holistic encoding/local Navon task to lead to higher rated facial composites. This is in contrast to the finding of Weston et al. (2008) who demonstrated that administration of the Navon task was beneficial for face recognition, but this was context dependent on the original encoding task employed by participants. In line with transfer appropriate processing theory, Weston et al. (2008) found that the featural face encoding group were better at face recognition following completion of a local Navon task, and the holistic face encoding group were better at face recognition following completion of a global Navon task compared to a control group who did not complete a Navon task. However, the contrasting finding from the present study that the featural encoding/global Navon and holistic encoding/local Navon conditions received higher ratings, which does not support transfer appropriate processing theory, could potentially be attributed to differences in the operationalisation of the dependent variable (face recognition vs. recall) and the individual cognitive styles of the participants. These issues are addressed in the general discussion.

It was noteworthy that both the global and local Navon task groups tended to make higher rated E-FITs than the control group, at least for those participants encouraged to encode the target face holistically. This finding is difficult to interpret within a cognitive processing framework, as it would be predicted that one style of cognitive processing would be more advantageous than the other, as is the case for featural vs. holistic encoding of faces. However, Weston et al. (2008) also found an advantage for either Navon task over a control group albeit regardless of the face encoding task which had been employed, and suggested there may be an effect of Navon task as yet undiscovered. A possible alternative interpretation for the advantage of both types of Navon task over the control group is the idea that participants in the Navon task group may have been aware that they were in an experimental condition, and might have responded to demand characteristics. There is also the issue of the frequent breaks taken by the Navon task groups during facial composite construction. Participants in the Navon conditions had to pause periodically during composite construction in order to complete a Navon task in between choosing individual facial features.

In addition, it is unclear what effect verbalisation might have on the utility of the Navon task during facial composite construction. In studies of face recognition the Navon task is used once, and this is followed by face recognition being measured (Weston et al., 2008; Lewis et al., 2009). However in the present study the Navon task is used intermittently between the selection of each individual feature. Verbalisation may have caused a shift to a featural state of cognitive processing (Fallshore & Schooler, 1995) and, as such, may have negated the specific effect of the global Navon task, as participants were required to choose descriptors of each feature they were choosing within E-FIT after completion of the global Navon task. Another possible interpretation within the present study is the idea that some participants may have been differentially affected by the Navon task depending on their stable cognitive style rather than on encoding condition; this issue is explored further in Study Four and in the general discussion.

In summary, it would appear that the Navon task does have an impact on the subjective likeness ratings that facial composites receive, and this impact appears to be context dependent on the way in which faces are originally encoded. If faces are originally encoded in a holistic manner, then the administration of either Navon task leads to composites which are given higher ratings than those given to a control group. However, if faces are originally encoded in a featural manner, then the Navon task does not confer an advantage over the control group in terms of the accuracy of facial composites produced.

Several issues arise from the above results with regard to the effect of incorporating the Navon task within the facial composite construction process. This is because composite construction using E-FIT can be considered to be a predominantly featural process as demonstrated by the advantage in Study Three for featural encoding in the absence of any Navon intervention, and by the advantage in Study One for E-FITs produced by field independent (feature based) participants. In addition the results of Study Three are mixed for the

Navon task, with an advantage for administration of a Navon task being found only in a holistic face encoding context, and not in a featural face encoding context.

Firstly, previous research has demonstrated an advantage for using the Navon task in face recognition but not the recall of faces in a featural fashion as required by witnesses constructing a facial composite using E-FIT. It may be the case that the utility of a Navon task for eyewitness memory may be better demonstrated within a composite construction system based on the recognition of whole faces. Secondly, it could be that the effectiveness of the Navon task is context dependent not on the way in which faces are originally encoded, but may interact with the natural propensity of an individual for holistic or featural processing; namely their cognitive style. Study Four addresses these issues by using EFIT-V to construct facial composites, which is a whole-face recognition based composite system, and by investigating whether there is an interaction between cognitive style and Navon task in the construction of target accurate facial composites.

Chapter Ten: Study Four

Study Four was designed to investigate whether manipulating the encoding of faces and the cognitive processing state of an individual using the Navon task might have an effect on face recall and subsequently on the accuracy of facial composite they produce using the EFIT-V composite construction system. Additionally, level of field dependence/independence among participants was measured in order to assess whether FDI could be a mediating factor when assessing the utility of the Navon task in facial composite construction.

Emmett et al. (2003) demonstrated that the cognitive style of individuals was a mediating factor which influenced their susceptibility to cued recall within a cognitive interview, and that cognitive style could potentially account for previous differential research findings with regard to the efficacy of the cognitive interview as a memory tool. Their finding that field independent individuals respond particularly well to cued recall where they are essentially being asked to scan their memory for specific details, may have implications for the way in which eyewitnesses interact with facial composite systems. Whilst field independent individuals might be predicted to produce more accurate composites using the feature based E-FIT system, a finding supported by the results from Study One, with EFIT-V which is based on whole-face recognition, field dependent individuals (who habitually use a more holistic processing style) might be predicted to produce more accurate composites.

Darling, Martin, Hellmann and Memon (2009) addressed the question of whether propensity to either global or local processing relates to face identification. They found that individuals who are relatively more global processors, as defined by reaction time performance on both local and global Navon tasks, were better at face identification in a line-up task than individuals who are more local processors. This differs from the Macrae and Lewis (2002) study where cognitive processing was manipulated, as cognitive processing propensity was simply measured by Darling et al. (2009). However, the results from both studies are complementary in that holistic/global processing, whether occurring naturally within individuals or induced through the global Navon task leads to superior whole face recognition. EFIT-V is a system based entirely on the recognition of whole faces, and it is the task of the witness to choose the face which is most similar to their memory of the target face. This selection is then used to generate a new set of nine faces for the witness to choose from (Gibson et al., 2003).

It could be argued therefore, that following a fairly extensive literature demonstrating that processing global Navon letters leads to greater face recognition, those in the global Navon task group will have a higher success rate at recognising the face most similar to the target face within the EFIT-V system. However, Weston et al. (2008) found that the local Navon task led to greater recognition of faces if it was preceded by featural face encoding. A further study by Weston et al. (2008) found that the administration of either Navon task led to greater face recognition, regardless of how faces were originally encoded. Study Three of this thesis also found a positive effect of either type of Navon task on composite construction performance, and this raises a number of issues which are addressed within the present study.

Therefore the main aims of Study Four are as follows. Firstly to investigate whether the positive effect of Navon task could be applied to a face *recognition* based composite construction system, and if so in what direction. Secondly, to investigate the possibility that the cognitive style of the participant could be a mediating factor in the way the Navon task might affect performance. Thirdly, to assess the impact of face encoding strategy and the interaction between face encoding condition and Navon task on subsequent facial composite construction performance using EFIT-V.

10.1: Method

10.1.1: Design

A 2*3*2 independent groups design was employed. The first factor was face encoding with 2 levels (holistic encoding/featural encoding), the second factor was Navon task with 3 levels (global Navon task/local Navon task/control group), and the third factor was FDI with 2 levels (field dependence/field independence). The study consisted of two phases, the first of which was a construction phase which involved participants constructing an EFIT-V (see

Chapter 6, General Methods section – 6.6) of a previously seen unfamiliar face after completing either a global or local Navon task (or no task in the control group). Participants within the three Navon conditions were also either assigned to a holistic or featural face encoding task while viewing the target face from which they constructed an EFIT-V. Furthermore, participants in the construction phase completed a Group Embedded Figures Test after EFIT-V construction. Phase two was an evaluation of the accuracy of the EFIT-Vs produced (by means of subjective-likeness ratings on a percentage scale) which employed a new sample of participants.

Participants

Seventy-two students and members of the general public (53 females, 19 males) participated in the construction phase of the study (mean age = 21.44 years, s.d. = 4.7). University students were recruited as part of their requirement for the Research Methods module research participation scheme and each received course credit. The participants were not familiar with the target faces, they reported no knowledge or previous experience of using the EFIT-V program, and they had not previously completed a Group Embedded Figures Test prior to taking part in the study.

Materials

Three sets of materials used in Study 3 were used again for this study. Firstly, the questions used during face encoding designed to encourage featural or holistic encoding of the target faces were repeated. Secondly, the twelve monochrome pictures of young adult Caucasian male faces were used in the same order as for Study 3. Finally, the global precedence white Navon letters on a black background provided by Professor Tim Perfect from the University of Plymouth were again used.

The EFIT-V for Windows program (Version 4.020) (Visionmetric Ltd, 2010) was used to construct the facial composites and the Micrografx Picture Publisher 8 (1998) program was used to make artistic enhancements to the EFIT-Vs if required by the participant. A Compaq Presario M2000 laptop computer with a screen size of 30 x 24 cm (1280 x 1024 pixels) was used both to construct the

EFIT-Vs and to run the Navon task. The Group Embedded Figures Test (GEFT) (Witkin et al, 1971) was completed by all participants (see Chapter 6, General Methods section 6.5).

10.1.2: Phase One – Composite Construction

Participants were tested individually and each produced one EFIT-V of one of 12 previously selected target faces. The target faces, face encoding conditions and Navon conditions were pseudo-randomised in the same way as for Study 3. After participants had viewed the target face for one minute, they completed a short maths-based distracter task for two minutes duration. On completion of the distracter task participants in the global or local Navon conditions completed a Navon task which lasted for five minutes. The Navon letters were presented sequentially in a PowerPoint presentation which presented each Navon letter on screen for 1 second. [4] On completion of the Navon task participants were shown the EFIT-V system and how it works, and proceeded to construct an EFIT-V of the target face they viewed from memory. The participants in the control group constructed an EFIT-V of the target face they had viewed from memory, but did not complete a Navon task. Following on from completion of the facial composites, all participants completed a Group Embedded Figures Test.

10.1.3: Phase Two – Evaluation – Subjective Likeness Ratings

Participants

Fifty students and members of the general public who had not made an EFIT-V (34 females, 16 males) participated in the evaluation phase (mean age = 20.4 years, s.d. = 3.59). Materials and procedure were the same as for Study 3 (see Chapter 6, General Methods section 6.6).

[4] The Navon task was not presented intermittently during composite construction as it was for Study 3 in the construction of E-FITs, as there are no natural breaks during construction of an EFIT-V facial composite: during E-FIT construction there is a natural break which occurs following the selection of each individual facial feature.

10.2: Results

The aim of Study Four was to investigate whether manipulating cognitive processing state might have an effect on the accuracy of facial composites that individuals produce using EFIT-V. The same procedure was used as for Study Three, participants were assigned to one of six conditions and completed a Navon task (or were in a control group) immediately before constructing a facial composite of a previously seen unfamiliar face using EFIT-V. New participants gave subjective likeness ratings on the facial composites that were produced. A Cronbach's Alpha test was conducted on the subjective likeness ratings given by the 50 raters in the evaluation phase to assess inter-rater reliability. Cronbach's Alpha for the subjective likeness rating scale across the full sample of 50 raters was .953, indicating a high degree of agreement among raters comparing the similarity of the E-FITs to the original target face.

By-rater analyses

Figure 10.1 shows the mean subjective likeness rating scores split by field dependence/independence when applied to each experimental condition. The by-rater analysis showed that featural encoding of faces generally resulted in facial composites which received higher rating scores.

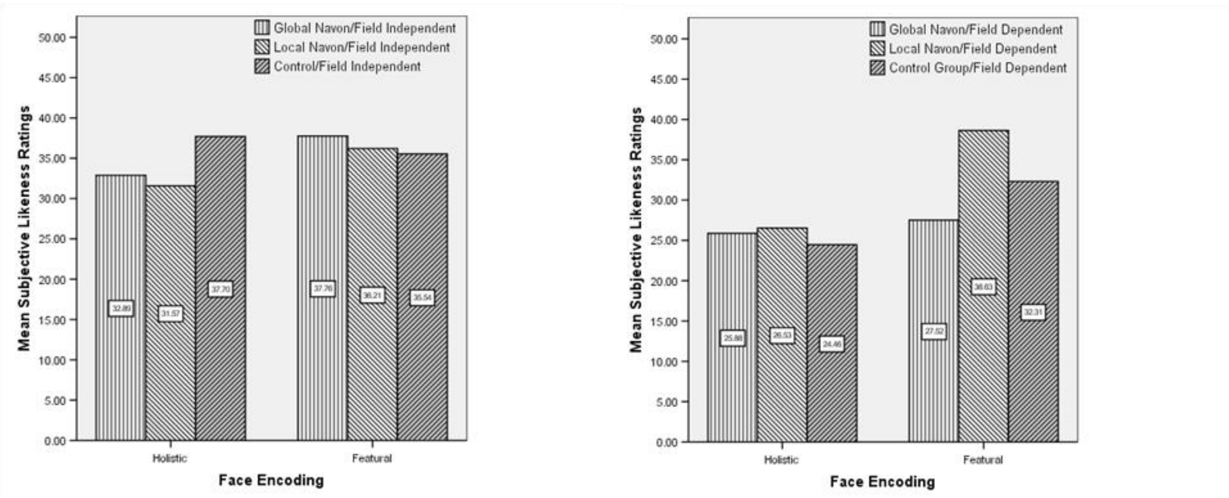

Figure 10.1 – Mean subjective likeness rating scores for six experimental conditions for field independence and field dependence aggregated by-rater in Study Four

A 2 x 3 x 2 (encoding x Navon task x FDI) within-subjects ANOVA was conducted on the twelve experimental conditions to see if there were any differences in the conditions when they were rated as groups in the subjective likeness ratings phase. There was a significant main effect of encoding, ($F_{1,49} = 41.538$, $p < .001$, partial $\eta^2 = .459$) with the featural encoding condition achieving significantly higher ratings. The overall mean score for composites produced after featural encoding was 34.67 compared to 29.84 for composites produced after holistic encoding. There was also a significant main effect of Navon task, ($F_{2,98} = 4.399$, $p < .05$, partial $\eta^2 = .082$), with the local Navon task producing higher mean scores than the control group who in turn produced higher mean scores than the global Navon group. The overall mean score for composites produced following a local Navon task was 33.23, compared to 32.5 for the control group, and 31.01 for the global Navon task. There was a significant main effect of FDI, ($F_{1,49} = 59.311$, $p < .001$, partial $\eta^2 = .548$), with composites produced by field independent participants achieving significantly greater mean likeness ratings overall (35.28 field independents, 29.22 field dependents).

There was a significant interaction between encoding and Navon task, ($F_{2,98} = 5.637$, $p < .05$, partial $\eta^2 = .103$), also a significant interaction between encoding and FDI, ($F_{1,49} = 9.697$, $p < .005$, partial $\eta^2 = .165$), and a significant interaction between Navon task and FDI, ($F_{2,98} = 9.647$, $p < .001$, partial $\eta^2 = .164$). There was a significant 3-way interaction between encoding, Navon task and FDI, ($F_{2,98} = 6.951$, $p < .05$, partial $\eta^2 = .124$).

Figure 10.2 shows the interactions between encoding condition, Navon task for field independent and field dependent participants.

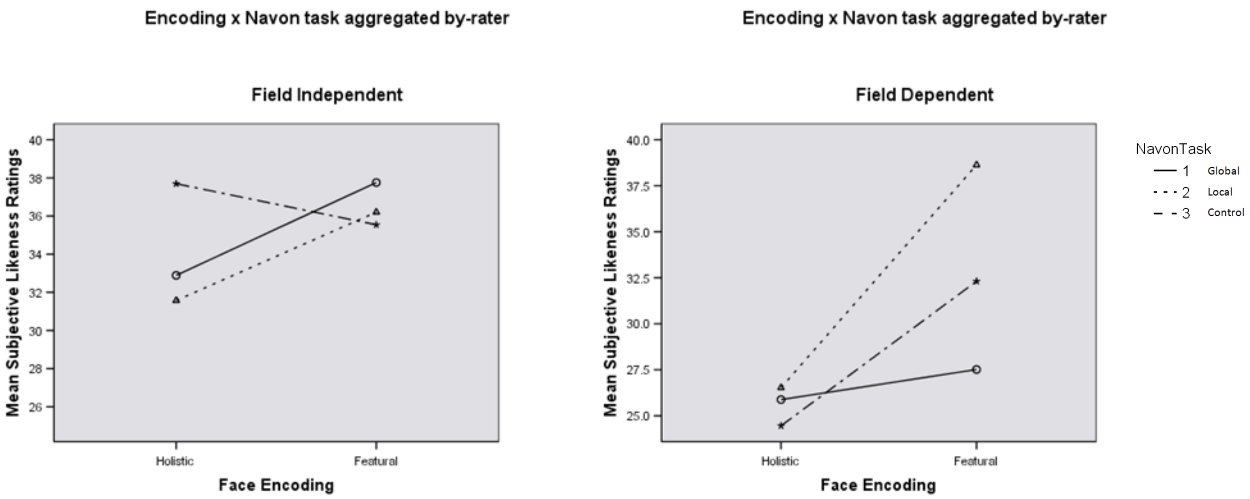

Figure 10.2 – showing the interactions between the Navon task and face encoding for field independent and field dependent participants aggregated by-rater in Study Four

In order to further investigate the interactive effect of the Navon task on both encoding and FDI, a series of one-way within-subjects ANOVAs were conducted examining the effect of Navon task for each of the four combinations of holistic/featural encoding, and field dependence/independence.

The first one-way within-subjects ANOVA was conducted on the subjective likeness rating scores of the composites produced by participants following holistic encoding who were field independent. There was a significant effect of Navon task, ($F_{2,98} = 8.513$, $p < .005$, partial $\eta^2 = .148$).

Paired samples t-tests showed that there was a significant difference between the control group and the global Navon task group, ($t = -3.024$, $df = 49$, $p < .005$, two-tailed) (Cohen's $d = 0.36$), and a significant difference between the control group and the local Navon task group ($t = -3.632$, $df = 49$, $p < .005$, two-tailed) (Cohen's $d = 0.46$). There was no significant difference between the global and local Navon task groups ($t = .942$, $df = 49$, $p = .351$, two-tailed) (Cohen's $d = 0.09$). Therefore field independent participants under holistic encoding conditions who were in the control group produced significantly higher rated composites than either of the Navon task groups.

The second one-way ANOVA was conducted on the subjective likeness rating scores of the composites produced by participants following featural encoding who were field independent. There was no significant effect of Navon task, ($F_{2,98}$ = .790, *p* =.457, partial η^2 = .016). Therefore, field independent participants in the featural encoding condition produced similarly rated facial composites across all Navon task conditions.

The results from the two one-way ANOVAs reported above show that the main effect of Navon task can be interpreted through the interaction with field independence and the way in which faces are originally encoded. For field independent participants there was an effect of Navon task only when faces are encoded in a holistic manner, in which instance the administration of neither Navon task was advantageous to facial composite construction, and the control group produced composites with higher subjective likeness ratings.

A further one-way ANOVA was conducted on the subjective likeness rating scores of the composites produced by participants following holistic encoding who were field dependent. There was no significant effect of Navon task, ($F_{2,98}$ = .683, *p* =.507, partial η^2 = .014). Field dependent participants in the holistic encoding condition produced similarly rated facial composites across all levels of the Navon task.

A further one-way ANOVA was conducted on the subjective likeness ratings given to the composites produced by participants following featural encoding who were field dependent. There was a significant effect of Navon task, ($F_{2,98}$ = 16.547, $p < .005$, partial η^2 = .252).

Paired samples t-tests showed that there was a significant difference between the control group and the global Navon task group, (*t* = 3.023, *df* = 49, $p < .005$, two-tailed) (Cohen's *d* = 0.34), and a significant difference between the control group and the local Navon task group (*t* = 2.767, *df* = 49, $p < .005$, two-tailed) (Cohen's *d* = 0.4). There was also a significant difference between the global and local Navon task groups (*t* = 5.912, *df* = 49, $p < .05$, two-tailed) (Cohen's *d* = 0.7). Therefore, field dependent participants under featural encoding conditions who were in the local Navon task group produced significantly higher rated composites than field dependent participants in the control group. Those in the control group produced significantly higher rated composites than participants in the global Navon task group.

To summarise, the conditions where the Navon task had an effect on facial compositing performance were field independent participants in the holistic encoding condition and field dependent participants in the featural encoding condition. For field independent participants in the holistic encoding condition there was no beneficial effect of Navon task, the control group in the holistic encoding condition gained significantly greater likeness ratings than either of the Navon task groups. The effect of Navon task was of much greater benefit in the featural encoding condition for participants who are field dependent. Following featural encoding, the field dependent participants gained significantly higher likeness ratings after completing the local Navon task than the control group, and global Navon task participants gained the lowest ratings of the three groups. Therefore, the greatest effect of Navon task overall was found for field dependent participants in the featural encoding condition who completed a local Navon task. [5]

By-item analyses

A further analysis was conducted in order to investigate the main effects of encoding and Navon within six experimental conditions when the subjective likeness ratings were adjusted for differences associated with field dependence/independence. A 2 (encoding condition) x 3 (Navon task) between-subjects Analysis of Covariance (ANCOVA) was conducted on the by-item data. The dependent variable was the subjective likeness ratings the composites received and the covariate was the scores that individuals received on completion of the Group Embedded Figures Test. Table 10.1 shows the mean subjective likeness rating scores given to individual composites across all raters for each of the six cells of the design taking cognitive style into account.

[5] Cell sizes were uneven in the by-rater analyses for Study Four, so a median split for FDI was applied within the 12 experimental conditions and the above analyses were repeated. The findings from the analyses with even cell sizes were identical except for the holistic face encoding condition. In the complementary analyses field independents in the global Navon group produced higher ratings than those in the local Navon group – but both Navon groups still produced composites with lower ratings than the control group, as in the analysis reported above.

Table 10.1 Mean subjective likeness rating scores (%) for face encoding x Navon group conditions adjusted for the differences associated with field dependence/independence calculated by-item in Study Four

Encoding	Navon Task	Mean (SD)	N
Holistic	**Global**	30.82 (10.44)	12
	Local	28.63 (10.86)	12
	Control	31.08 (10.62)	12
Featural	**Global**	32.64 (9.37)	12
	Local	36.86 (11.73)	12
	Control	33.66 (12.08)	12

A 2*3 (encoding, Navon task,) between-subjects ANCOVA with GEFT scores as the covariate revealed no significant main effect of encoding task, ($F_{1,65}$ = 1.833, p = .180, partial η^2 = .027). There was also no significant main effect of Navon task, ($F_{2,65}$ = .107, p = .899, partial η^2 = .003). The covariate, field dependence/independence, had a significant effect ($F_{1,65}$ = 4.047, $p < 0.05$, partial η^2 = .059), with composites produced by relatively field independent participants receiving significantly higher likeness ratings. There was no significant interaction between encoding and Navon condition, ($F_{2,65}$ = .436, p = .648, partial η^2 = .013).

A scattergram was employed to explore the nature of the relationship between the GEFT scores of the participants who constructed the facial composites and the subjective likeness ratings given to the composites (by-item). The scattergram revealed no bias in the residuals across the range of GEFT scores, implying that the discrepancy between the by-item and by-rater analyses was not a consequence of a sub-set of composites attracting relatively high or low scores.

Further Analyses

The same target faces were used in the same order for Studies Three and Four along with holistic and featural encoding of the target faces which gave an opportunity to compare the likeness ratings given to both the E-FIT and EFIT-V systems. Although different raters were used for Studies Three and Four, both sets of ratings had a high level of inter-rater reliability. A 2 (holistic encoding/featural encoding) x 2 (E-FIT/EFIT-V) ANOVA was performed on the ratings given to the composites produced by the control group in both studies. Figure 10.3 shows the mean subjective likeness ratings given to the control groups following holistic or featural encoding of the target face in Studies Three and Four.

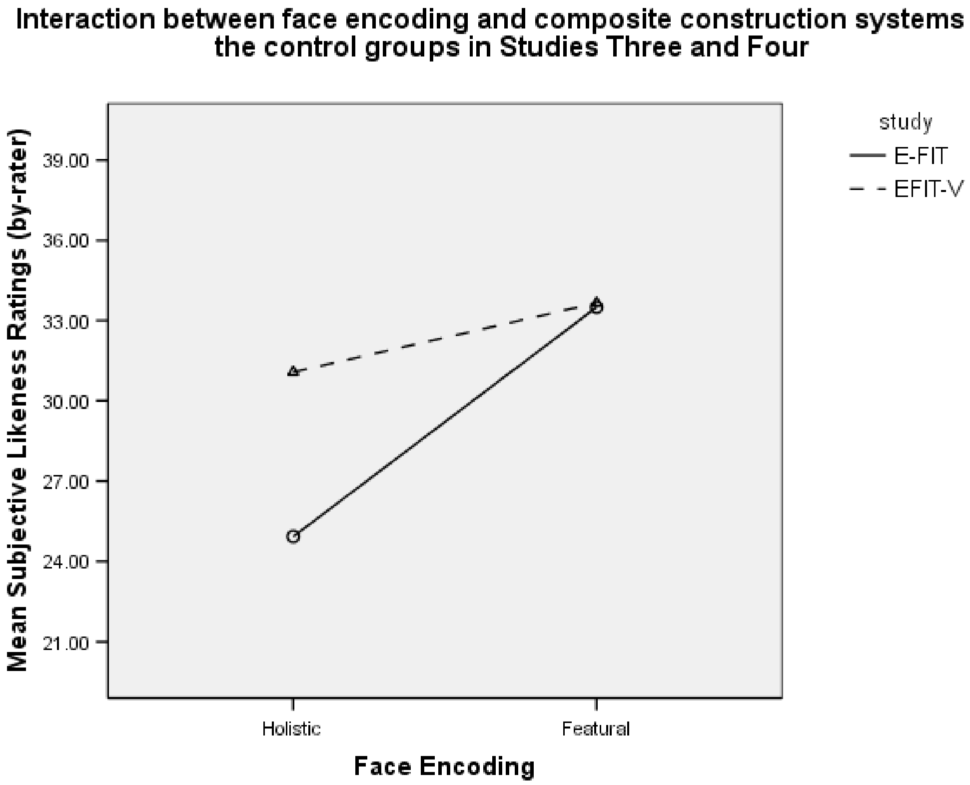

Figure 10.3 – mean subjective likeness rating scores for holistic and featural face encoding for the control conditions in Studies Three and Four

A by-rater analysis revealed a significant main effect of encoding with featural encoding of the target face being advantageous for both the E-FIT and EFIT-V composite systems, ($F_{1,88} = 35.144$, $p < 0.05$, partial $\eta^2 = .285$). There was also a significant interaction between encoding and composite system ($F_{1,88} = 10.142$, $p < 0.05$, partial $\eta^2 = .103$) but no main effect of composite system, ($F_{1,88} = 2.050$, $p = .156$, partial $\eta^2 = .023$).

A post-hoc independent t-test showed there was a significant difference in the holistic encoding conditions between E-FIT and EFIT-V, (t = 2.639, df = 88, p < 0.05,) (Cohen's d = 0.56) with composites produced using the EFIT-V system given significantly higher ratings than composites produced using the E-FIT system following holistic encoding. Therefore, even though featural encoding was better for both systems, invoking holistic encoding was not so detrimental if the EFIT-V system was used to construct a composite.

10.3: Discussion of Study Four results

Study Four was designed to investigate whether the Navon effect, which has previously been observed in face recognition studies, could be generalised to facial composite construction using EFIT-V which is a face recognition based system. In addition the field dependence/independence cognitive style of the participants was measured, as previous research has suggested that an inherent predisposition to either local or global processing can have an influence on face recognition (Darling et al., 2009). Finally, holistic and featural face encoding was included as a factor as the results from Study Three and previous research suggests that encoding may influence the direction of the Navon effect on face recognition.

The by-rater analyses revealed a significant main effect of encoding, with featural encoding of faces generally leading to higher rated facial composites. This result is surprising given the previous results of Wells and Hyrciw (1984) to explain their finding that holistic encoding of faces led to better face recognition and featural encoding led to better Photo-FIT facial composite construction. This is because EFIT-V is a recognition-based composite construction system, so it might have been expected that holistic encoding of faces would lead to more accurate and higher rated facial composites. However, facial composite construction using EFIT-V is still a task of recall in addition to recognition in that the faces viewed on the computer screen must be matched with the representation of the perpetrator's face in the witness's memory. The advantage for featural encoding using a face-recognition based system supports the finding that featural encoding led to better sorted facial composites using EvoFIT (Frowd et al., 2007b).

The significance of a featural cognitive style for facial composite construction is also highlighted by the finding that field independent participants produced higher rated composites as a group than field dependent participants. In complementary by-item analyses field dependence/independence was the only significant main effect, strongly in favour of field independent individuals. This is again a striking finding that was not in the expected direction, given that the disembedding of features is not required in the construction of an EFIT-V. If the interpretation of the results from Study One that field independent individuals are better at constructing facial composites because they are better able to disembed individual features from their surrounding context is correct, then it might be expected that this disembedding advantage would be dissipated when constructing a composite using just whole faces as with EFIT-V.

The main effect of Navon task was such that participants in the local Navon task condition produced higher rated composites overall. However, this was mitigated by both the face encoding condition and FDI cognitive style of the participants. When the effect of the Navon task was investigated further it was found that the local Navon task was particularly beneficial for field dependent individuals in the featural face encoding condition. This pattern of results supports the findings of Weston et al (2008) who found that the beneficial effect of the Navon task on face recognition was dependent on the context in which the faces had been originally encoded.

However, the by-rater analyses from Study Four also revealed that there was no effect of Navon task in the holistic face encoding condition for field dependent individuals. When a median split for FDI was not performed within each condition there was a detrimental effect of Navon task for field independent individuals following holistic face encoding, with the control group producing significantly higher rated composites than the global Navon task group. The local Navon task group did not perform significantly better than the global Navon task group or significantly worse than the control group.

Overall the findings from Study Four suggest that the field dependence/ independence cognitive style of an individual may be a mediating factor in the effectiveness of the Navon task in aiding accurate facial composite construction, even using a whole face recognition system such as EFIT-V. From the perspective of potential real world applicability, if the way in which faces are encoded is removed from the analysis, as face encoding cannot be manipulated after a witnessed event, then results show that for field dependent individuals there was a significant benefit of completing a local Navon task prior to composite construction compared to the group who completed a global Navon task. This could arguably be attributed to the idea that the local Navon task is orienting field dependent individuals to a more featural or piecemeal mode of cognitive processing which appears to be beneficial to composite construction for individuals who are naturally field independent. However, even though field independent individuals in the control group received higher ratings for their composites, there was no statistically significant difference in composite accuracy ratings for field independent individuals who completed a local Navon task. There was a statistically significant difference in ratings between the control group and the global Navon task group for field independent individuals, which indicates that the global Navon task may be detrimental to composite construction performance. In summary, the local Navon task would arguably be non-detrimental to field independent individuals for facial composite construction and beneficial for field dependent individuals. The following chapter is a general discussion which summarises the results from the previous four empirical chapters, and provides a general overview of the research presented within this thesis.

Chapter Eleven: General Discussion

This chapter presents a summary of results from the empirical research presented in Chapters Seven to Ten, and how they relate to facial composite construction using E-FIT and EFIT-V. This is followed by a discussion of the methodological issues inherent within the four research studies presented in this thesis, and consideration of the practical issues which arise from the results of the research along with suggestions for future research.

11.1: Summary of results from Study One

The aim of Study One was to assess the role that the cognitive style of field dependence/independence might have in the production of target-accurate facial composites using E-FIT. Previous research on the efficacy of facial composite systems identified individual differences in the ability to construct accurate facial composites. It was suggested that field dependence/ independence might be a source of those perceived individual differences (Ellis et al., 1975). This is because level of field dependence relates to the ability to perceptually isolate shapes from their surrounding context. This is a cognitive task which can arguably be considered to be analogous to the ability to perceptually isolate individual facial features when choosing them within the context of a whole face, which is the way in which an E-FIT is constructed.

Participants completed a Group Embedded Figures Test for field dependence /independence and constructed a composite of a previously viewed unfamiliar male face from memory using E-FIT. The E-FITs were compared to the original target face by new participants and rated for similarity to the target face on a percentage scale. The results showed that there was some evidence that field independent individuals, those who are relatively more able to perceptually isolate simple shapes from a surrounding context as denoted by the Group Embedded Figures Test, constructed E-FITs which received significantly higher likeness ratings than E-FITs constructed by field dependent individuals. One interpretation of this finding is that field independent individuals are not as distracted by perception of a whole face image when comparing each feature of the face with their memory of that feature. Field independents might therefore find it easier to select a feature within E-FIT which is more likely to be an accurate representation of that feature. This would be consistent with findings that perception of a whole face image interferes with the perception of parts of a face (e.g., Young et al., 1987).

However, several issues arose from Study One relating to the main findings. Firstly the advantage of field independence was only observed when consideration was given to ratings based on groups of participants in the by-rater analysis. When considered individually in the by-item analysis the advantage for field independence was not observed. Secondly the small effect size observed and the uni-directional nature of assessing field independence using the GEFT warranted caution in the interpretation of the findings from Study One. The GEFT positively assesses for field independence and as such, could possibly be measuring some other aspect of individual differences such as conscientiousness or motivation. Therefore a further investigation into the possible effect of a natural predisposition to whole-based or part-based processing on facial composite construction was carried out. Study Two addressed some of these issues.

11.2: Summary of results from Study Two

The aim of Study Two was to assess the role that holistic/analytic and verbaliser/imager cognitive styles might have in the production of target-accurate facial composites using E-FIT. A computerised test of holistic/analytic (2003) and verbaliser/imager (2005) cognitive styles was developed by Peterson, Deary and Austin which positively assesses both ends of the above cognitive style constructs. The E-CSA-W/A (holistic/analytic style) and VICS (verbaliser/imager style) tests were administered to participants who later completed a facial composite of a previously seen unfamiliar male face using E-FIT. Subjective likeness ratings were used to assess the similarity of the E-FITS to the target faces in the same way as for Study One.

Results showed that there was some evidence that individuals with a holistic cognitive style, those who have a relative tendency to process visual information in wholes rather than parts, constructed E-FITs which received significantly higher likeness ratings than E-FITs constructed by analytic individuals. Similar to the result from Study One, the advantage for holistic cognitive style was only observed for ratings based on groups of participants in the by-rater analysis. When considered individually in the by-item analysis the advantage for holistic cognitive style was not observed. However thirty participants returned some time later to construct a second facial composite. The subjective likeness ratings for the two E-FITs constructed by each participant were averaged to give a 'mean performance' score over two E-FIT construction attempts. Within this sub-set of thirty participants a by-item advantage for composites produced by those with a holistic cognitive style was found. There were no differences in the ratings given to E-FITs produced by individuals with either a verbaliser or imager cognitive style.

The advantage for E-FITs produced by individuals with a holistic cognitive style is difficult to interpret if the holistic cognitive style is considered to be entirely analogous to a field dependent whole-based processing style. This is because the results from Study Two seem to go in a reverse direction to the results from Study One. The discrepancy in the results from Studies One and Two is further addressed in section 11.5.2. However, the finding that an advantage can be observed for either a part-based processing style in Study One, or a whole-based processing style in Study Two highlights the possibility that construction of a facial composite using E-FIT does not rely solely on either type of processing. Construction of a facial composite is a process that includes elements of both holistic and featural processing. If the field independence /independence and holistic/analytic styles are considered to be similar constructs with different words to describe those constructs (Zhang & Sternberg, 2005) then the following interpretation can be applied.

It is possible that the field independent participants in Study One were advantaged by making better initial choices of facial features in order to construct the face. It is equally possible that holistic participants in Study Two were advantaged more toward the end of the process of composite construction, when the full face was viewed and artistic enhancements or other small changes could be made to the face. The majority of participants who constructed a facial composite tended to have an opinion at the end of the process on whether the likeness they had created was a good representation of the face they were attempting to construct or not. Of those participants who felt that the likeness was not a good representation, only some felt confident to make overall changes to the face which they felt would make a difference. Others commented that it was not a good likeness, but felt that they were unsure what to do to improve the likeness. The advantage for holistic individuals therefore might potentially be in the match they are able to make between the whole target face as represented in memory, and the complete composite image as represented on screen near the end of the composite construction process.

In summary the results from Studies One and Two suggest that the stable cognitive style of an individual can make a significant difference to the quality of facial composites produced using E-FIT. However the results were inconclusive with regard to the direction in which the effect can most clearly be observed. The results from Study One indicate that a stable tendency to process visual information in a part based manner might be advantageous, but it is unclear whether this relates to how a face is initially encoded or stored in memory, or to the composite construction process itself. The results from Study Two indicate that a stable tendency to process visual information in a holistic manner might be advantageous, but again it is unclear which stage of face perception or recall that this may be related to. Overall, the cognitive style of an individual has been found to have an effect on the accuracy of the facial composites they produce. However, the previous research highlighted in Chapter Five has demonstrated that the way in which people process information can be manipulated to suit the cognitive task to be performed (Macrae & Lewis, 2002). From an applied perspective, system variables which can be manipulated both experimentally and in an applied context may contribute to the understanding of how cognitive processes affect facial composite construction. The small effect sizes observed in Studies One and Two led to an investigation of the possible role of the manipulation of temporary cognitive processing state, where research has demonstrated much larger effect sizes with respect to face recognition.

11.3: Summary of results from Study Three

The aim of Study Three was to investigate the role of manipulating cognitive processing state in the production of target-accurate facial composites using E-FIT. Previous research has indicated that completing a global Navon task prior to a face recognition test can have a positive effect on face recognition relative to a control group (Macrae & Lewis, 2002). The positive effect of the global Navon task on face recognition was attributed to a switch within participants to a global mode of processing. Global processing is believed to be consistent with the holistic manner in which faces are generally processed, and was originally suggested to support the transfer-inappropriate processing account for the verbal overshadowing effect. Participants in Study Three completed either a global or local Navon task periodically throughout facial composite construction with E-FIT. The way in which faces were encoded was also manipulated in order to investigate if there was an interaction between cognitive processes at the encoding and retrieval stages of face recall and composite construction. Face encoding was manipulated by asking participants to give personality ratings to induce holistic encoding or ratings of individual features to induce featural encoding for the target face they were viewing prior to constructing a facial composite.

Results indicated that featural encoding of the target faces prior to composite construction led to higher ratings for the E-FITs which participants produced across all Navon conditions. When the data were analysed by-rater there was a significant advantage for composites produced by the participants who had completed either a local or global Navon task over a control group who had not completed a Navon task. The significant interaction between encoding and Navon task meant that the advantage for composites produced following either Navon task was only observed following holistic encoding of the target face. In the by-item analysis the advantage for either Navon task over the control group was not observed.

The by-rater finding of an advantage for either type of Navon task following holistic processing of the target face over the control group is difficult to interpret within a cognitive processing framework where it might be considered that one style of cognitive processing would be advantageous for face recall and composite construction. This is however, consistent with the findings of Weston et al. (2008) who found a face recognition advantage for either type of Navon task over a control group. One possible interpretation for this finding is that in a similar way to the contrasting findings of Studies One and Two, facial composite construction does not rely solely on either holistic or featural processing but rather is a combination of both types. Therefore orienting participants to either style of processing throughout the composite construction process may have been advantageous. Other possible interpretations of the findings from Study Three are discussed in section 11.5.

11.4: Summary of results from Study Four

The aim of Study Four was to investigate the role of manipulating cognitive processing state in the production of target accurate facial composites using EFIT-V, which is a whole-face recognition-based composite construction system. The majority of published research using the Navon task has investigated its effect on face recognition (Macrae & Lewis, 2002; Weston et al. 2008; Lewis et al, 2009). This research has generally found that completion of a global Navon task leads to better face recognition relative to a control group. Face encoding was also manipulated in the same way as for Study Three. In addition the field dependence/independence cognitive style of the participants was measured in order to assess how stable trait also affects the construction of facial composites using EFIT-V.

The results from both the by-item and the by-rater analyses showed that field independent participants produced facial composites which received higher ratings than those produced by field dependent participants. This finding is consistent with the finding from Study One where a small advantage for field independents was found in the by-rater analysis. Several possible interpretations arise from this finding, which are discussed in section 11.5.

The results from the by-rater analysis also showed that across all experimental conditions, featural encoding of the target faces led to EFIT-V composites which received higher likeness ratings. This finding is consistent with the finding from Study Three that featural encoding of faces was advantageous for E-FIT construction. It is noteworthy that this should be the case also with EFIT-V construction, because featural encoding would be a processing strategy arguably at odds with the holistic way in which EFIT-V works as a system. The advantage for featural encoding was driven by the large effect featural encoding had on field dependent participants in particular. For field independent participants, the way in which faces were initially encoded was not a factor in their performance.

There was a significant main effect of Navon task but this was moderated by both the encoding conditions and the field dependence/independence cognitive style of the participants. The greatest effect of Navon task was observed in the featural encoding condition for field dependent participants, where the featural Navon task was beneficial for facial composite construction. The advantage for the featural Navon task was significantly greater than the advantage for the control group who in turn received significantly higher ratings than the global Navon task group. This effect was maintained when a median split for FDI was applied to the data to produce even cell sizes in each encoding/Navon condition. The only other effect of Navon task was observed in the holistic encoding condition for field independent participants. Here the control group produced EFIT-Vs with higher ratings than either Navon task group. In a by-item analysis field dependence/independence was the only factor to maintain significance, highly in favour of field independent participants overall. Section 11.5 is an overview discussion which shall examine the theoretical and practical issues arising from the results of the four empirical studies within this thesis.

11.5: Discussion of combined results from all empirical studies

The main aim of the research presented within this thesis was to investigate the role of holistic and featural processing in the construction of facial composites from two perspectives; (1) stable individual cognitive style and (2) the manipulation of cognitive processing sets. Two main questions therefore arise from these aims. Firstly, is there a relationship between stable cognitive style and accurate facial composite construction? Secondly, can cognitive processing be manipulated during or prior to facial composite construction to have a positive effect on the accuracy of facial composite that an individual produces?

11.5.1: Cognitive Style – Field dependence/independence

Three separate methodologies of cognitive style were used in the research presented in this thesis. In Studies One and Four participants were measured for their level of field dependence/independence (Witkin et al., 1971) and in Study Two participants were measured for their level of holistic/analytic (Peterson et al., 2003) and verbaliser/imager (Peterson et al., 2005b) cognitive styles. These styles shall be considered individually within the following discussion, beginning with field dependence/independence.

The results from both Study One using E-FIT and Study Four using EFIT-V indicated that people who have a field independent cognitive style have a significant advantage in facial composite construction. Field independent individuals received higher ratings for composites produced with both the E-FIT and EFIT-V systems. Field independent individuals are considered to be better able to disembed and locate simple shapes from their surrounding distracting context than field dependent individuals. Previous research has also demonstrated that faces are generally processed as whole entities and they fuse into a perceptual gestalt from which it is difficult to extract individual featural information (Maurer et al, 2002). Therefore, these results seem to suggest that there may be an advantage when recalling faces from memory, of being able to perceptually disembed individual facial features.

If the above account is correct, then the disembedding advantage may surface in one of two ways. Firstly, this could be in the initial description of the to-be-recalled face. Alternatively this may be when viewing a whole face on screen and deciding which facial feature is the best representation of the face held in memory. Facial composite construction using E-FIT requires the eyewitness to describe first the individual features of the target face. These descriptors are then transferred into the E-FIT program in order to potentially reduce the number of examples of each feature an eyewitness must view in order to select one which they consider to be a sufficiently good representation of the feature in their memory. It is possible therefore that field independent individuals are better able to disembed and therefore describe individual facial features accurately to an E-FIT operator. It could be this initial accuracy advantage in describing features which leads to more accurate facial composites. However, the interpretation of an advantage at the featural description stage of E-FIT composite construction for field independent individuals is not applicable to the composite construction process using EFIT-V given that the process of composite construction is different to that of E-FIT.

Facial composite construction using EFIT-V does not require the eyewitness to provide a description of each individual facial feature. Eyewitnesses must give global information at the beginning of the EFIT-V construction process such as sex, age, hairstyle and ethnicity of the target face. This global information is then used to begin the composite construction process. This would be particularly advantageous for witnesses who cannot remember specific details of the face they are attempting to construct. The advantage for field independent individuals within the context of constructing an EFIT-V composite therefore, may be seen in their ability to isolate individual facial features within the context of the whole faces displayed on screen, and not at the description stage of construction as may be the case with E-FIT construction. There is the option within the EFIT-V program to manipulate individual facial features as there is within the E-FIT program. A witness may choose to work on a single feature rather than whole faces, and change the size and/or shape of that feature and the position of the feature on the face.

The results of Studies One and Four are therefore consistent with the interpretation that there is a perceptual advantage for field independent individuals based on their relatively superior ability to disembed simple shapes (features) from their surrounding context (the whole face). This would be consistent with the matching superiority hypothesis (Wells & Hyrciw, 1984) where it was found that featural encoding led to more accurate facial composites using Identikit. This is because there would be a match for field independent people in the way in which they encoded faces, and the way in which they interact with facial composite construction systems. This interpretation would indicate that there is a strong featural encoding element to facial composite construction even using a whole face system such as EFIT-V. An advantage for featural encoding has been found using IdentiKit, PRO-fit and

EvoFIT (Frowd et al., 2012), and the research within this thesis suggests there is also a strong featural encoding element in E-FIT and EFIT-V construction.

An informal observation consistent with a featural element to EFIT-V construction is that many participants who used the EFIT-V system asked which part of the face they should be giving precedence to when deciding which of the nine faces on screen was most familiar to them, even though they had been instructed to consider the faces as wholes. Study Four used EFIT-V version 4.020 which was almost entirely whole face based from the beginning of the composite construction process. A newer version of EFIT-V (Version 5) developed in 2011 has incorporated an option of choosing the size and shape of individual features at the beginning of the construction process as it was recognised by the developers of EFIT-V that some individuals prefer to narrow down the choice of available features if any of them could be remembered sufficiently well. These developments are compatible with the present findings.

An alternative interpretation for the advantage of field independence is that field independent individuals may have a tendency to process faces and encode them in the first instance in a more featural manner as a general rule. Although it is generally proposed that faces are encoded and stored as a gestalt, there is evidence that some individuals may behave differently from this general rule. For instance, Olsson and Juslin (1999) reported that 25% of participants who had viewed a video clip reported a spontaneous natural featural encoding strategy of the face they had viewed. The differences in self-reported face encoding strategies were supported by differential performance on a subsequent face identification task. Similarly, Martin and Macrae (2010) found that individuals who display a weak global precedence are faster at identifying inverted faces. Two groups of individuals with relatively weak and strong global precedence were formed on the basis of reaction times in responding to consistent and inconsistent local and global Navon stimuli. The findings of Martin and Macrae (2010) suggested that there may be systematic individual differences in the manner in which faces are processed. In turn these differences in processing visual information at a global level impact face recognition performance. Individuals who habitually process information in a relatively more holistic manner are more susceptible to the face inversion effect, which is purported to interrupt holistic processing of faces. If field independent individuals process and represent faces in a relatively more featural way then they would arguably be more accurate at disembedding and identifying similar facial features when viewed within the context of a whole face.

A significant main effect for field independence on ratings given to composites was found in Study Four where EFIT-V was used to construct the composites. One possible interpretation of this is that the advantage for field independent individuals resides in their relative ability to disembed features when viewing a whole face, and not so much at the description stage as suggested for E-FIT construction. This is because there is no detailed description stage using EFIT-V.

Additionally, the way in which faces were encoded also had an effect on composite ratings, with featural encoding being advantageous for composite construction. A featural encoding advantage in facial composite construction is a consistent finding across several composite sytems (Wells and Hyrciw, 1984; Frowd et al., 2007b). Study Four of this thesis was the only study to combine holistic or featural encoding with a measure of field dependence/independence and results showed that for field independent individuals there was no advantage conveyed by the manipulation designed to induce a featural face encoding strategy. It may be the case that if field independent individuals are naturally predisposed to featural encoding, then no added advantage can be gained from inducing a featural type of encoding that is already being utilised. For field dependent individuals, those who have a tendency to process visual information in wholes, featural encoding of faces was beneficial. This was particularly so when used in combination with the local Navon task. This finding supports the interpretation that being field independent is advantageous for facial composite construction because field independent individuals may have a tendency to encode faces in a more featural manner as a matter of course.

There was some evidence that the main effect of an advantage for featural encoding was mediated by the field dependence/independence cognitive style of participants in Study Four, and this interaction in turn affected the way in which the Navon task influenced composite construction. Field independent participants in the featural face encoding condition showed no evidence of being affected by either Navon task or by not completing a Navon task. There was also no differential effect for the Navon task for field dependent participants following holistic face encoding. These findings have implications for research in which a Navon effect has not been found such as the study by Lawson (2007). It might be the case that the proportion of field dependent and independent participants within a sample of participants may mediate the Navon effect in the same way as field dependence/independence was suggested to mediate the effect of cued recall within studies investigating the efficacy of the cognitive interview (Emmett et al., 2003).

However, for field dependents who encoded the target faces in a featural manner there was a positive effect of completing a local Navon task. It is possible that encouraging field dependent individuals to encode faces in a featural manner encourages them to engage in a featural form of cognitive processing. A featural form of cognitive processing would be similar to that employed naturally by field independents, and therefore would potentially enable them to construct more accurate facial composites. Consistent with this interpretation, it was found that featural encoding of faces made no difference to the performance of field independent individuals in Study Four. These results suggest that using a featural processing style when faces are encoded is advantageous for subsequent facial composite construction, and that if a person does not encode faces in a relatively featural way naturally, then drawing their attention to facial features at the face encoding stage can lead to more accurate facial composites.

In terms of real world application, the way in which faces are encoded by witnesses cannot be manipulated after the perpetrator of a crime has been viewed. Therefore, although it is of theoretical interest to assess how encoding interacts with field independence and the Navon task, it is not a factor which can be of practical benefit. The way in which field dependence/independence interacts with the Navon task when face encoding is removed from the analysis is discussed in the section below on the effects of the Navon task on facial composite construction.

11.5.2: Holistic/analytic cognitive style

The aim of Study Two was to investigate whether individual differences in holistic/analytic cognitive style had an effect on the accuracy of facial composites that individuals produce. A by-rater advantage was found for composites produced by holistic individuals, those who habitually tend to process visual information in wholes rather than in parts. As previously suggested, this result is difficult to interpret if the holistic/analytic cognitive style is considered to be entirely analogous to the style of field dependence/ independence, as holistic individuals would be categorised as field dependent. If the field dependence/independence and the holistic/analytic styles are the same constructs using different names, then the results of Study Two run counter to the results found in Study One where an advantage for field independent cognitive style was found. The differential findings from Studies One and Two therefore support the side of the ongoing debate within cognitive style research which suggests that these style constructs are not the same constructs with different names, but different constructs (Zhang & Sternberg, 2005).

It may be the case that a field dependent or independent cognitive style is not entirely the same as a holistic or analytic style even though both are concerned with parts and wholes in visual perception. Zhang (2004) suggested that the construct of field dependence/independence represents perceptual ability, but not a broad cognitive style. Two issues arise from this point, the first being what it is that differs between the field dependence/independence and holistic/ analytic cognitive styles, if field dependence is argued to be subsumed within the holistic/analytic family of styles. The second issue, if the styles are different, is why having a holistic cognitive style should be beneficial for facial composite construction.

The tests for field dependence/independence and holistic/analytic cognitive styles are different even though they are both believed to assess the ability to locate and isolate a simple shape within a more complex shape. Although the E-CSA-W/A tests positively for both holistic and analytic ends of the continuum, there is a 50% chance of getting any of the answers within the test correct. This is because the test involves giving 'yes' or 'no' answers to the question of whether a simple shape can be located within a complex shape when both are displayed on screen. In the GEFT, the simple shape must be drawn directly onto the complex shape, thereby ensuring that the test taker has successfully located and isolated the simple shape from its surrounding context. The GEFT is therefore potentially measuring a different aspect of visual perception than the E-CSA-W/A. The major difference between the two tests however is that the GEFT positively measures only field independence. Field dependence is inferred from relatively lower scores on the test. This has been an area of criticism of the GEFT, with some researchers claiming that the GEFT is a test of general intelligence or ability (Sternberg & Grigorenko, 1997). However, the results from the study by Emmett et al. (2003) which showed that there are circumstances in which field dependent individuals perform better than field independent individuals would seem to suggest that this is not the case.

With regard to the issue of what may differ between the field dependence/independence and holistic/analytic cognitive styles apart from the way in which they are assessed, it has been suggested by Miller (1987) that the styles are related to different aspects of visual perception. Accordingly, field dependence/independence is related to selective attention and holistic/analytic style is related to pattern recognition. The premise that field dependence /independence is related to selective attention fits the interpretation from Study Four that the benefit for field independence was related to attending to the features of the face at encoding, which field independent individuals are more naturally predisposed to do. This interpretation was supported by the finding that inducing featural encoding of the faces was advantageous for only the field dependent individuals, those who would not have spontaneously processed the face in a featural manner.

By this account selective attention to the features of the face at the encoding stage of memory may lead to composites with higher likeness ratings. This could apply whether due to the natural propensity of a field independent cognitive style or to the inducement of featural processing. In summary the field dependence/independence and holistic/analytic styles may have differences in whether they relate to attention or pattern recognition. These differences in turn may be advantageous in either the way in which the target face is originally encoded or to the composite construction process. A field independent cognitive style may be advantageous if it leads to faces being encoded in a featural way. A holistic cognitive style may be advantageous near the end of the composite construction process.

The holistic/analytic style is related to pattern recognition and not to selective attention (Miller, 1987). It may be the case therefore that the advantage for a holistic cognitive style is conferred later on in the E-FIT composite construction process. At the end composite construction the witness must decide whether the composite they can see on the computer screen is a good match to the target face held in their memory. At this stage, the cognitive task changes from one of predominantly recall to one of predominantly recognition. Research on face processing and recognition suggests that faces are generally processed in a holistic way (Young et al., 1987) so therefore individuals with a holistic cognitive style may be better at recognising whole faces than those with an analytic style. The ability to recognise the composite of the whole face as presented on computer screen may also be advantageous in that individuals with a holistic style may be better able to judge when an optimum likeness of the target face has been reached. They may therefore stop trying to alter aspects of the facial composite through artistic enhancement and/or feature modification.

Facial composites are not supposed to represent an absolute likeness of the perpetrator's face as this would be almost impossible to achieve, instead they are a representation of a type-likeness of that face. A facial composite that would constitute a good type-likeness of a person's face would be one that has similar *global* features such as a similar hair style, colour and length, a similar face shape and similar skin tone. An individual with a holistic cognitive style would potentially pay more attention to the global aspects of a face at both the encoding and the recognition phases of composite construction. The match between global pattern recognition processes at encoding and recognition might confer an advantage in facial composite construction for holistic individuals in terms of producing an accurate type-likeness.

A further interpretation for the advantage of a holistic cognitive style in Study Two is that the holistic participants may have had an advantage from utilising two forms of processing (both holistic and featural) at the face encoding stage. Face encoding was not manipulated in Study Two and therefore participants encoded the target face they viewed in a spontaneous manner. Although it might be expected that holistic individuals would encode faces in general in a relatively holistic manner, participants were aware that they were going to construct a facial composite of the target face they were viewing. There is evidence to suggest that when participants know that a face perception task is to follow, that they switch encoding strategy to a more featural style (Laughery et al., 1986).

Facial composite construction is a cognitive task that encompasses both featural and holistic processing, so therefore, those individuals who utilised both types of processing at the face encoding stage may have an advantage over individuals who utilised only one type of processing. This advantage may be conferred in two ways. Firstly the advantage may be in the ability to judge the accuracy of individual facial features for inclusion in the facial composite. Secondly the advantage may be in the judgement of the relative accuracy of the whole facial composite at the end of the construction process. Frowd et al., (2012) found that inducing both types of cognitive processing prior to EvoFIT construction led to higher rated composites than inducing either featural or holistic processing in isolation.

Following this reasoning, analytic individuals who viewed the target faces would therefore mainly be utilising a featural processing strategy. This is because they have a natural featural cognitive style and would have adopted or maintained a featural encoding strategy in the knowledge that a face perception task was to follow. Therefore analytic individuals would potentially be exposed to only one type of processing of the target face at the encoding stage, and be disadvantaged by this during facial composite construction. Holistic individuals however, would have naturally employed both a holistic encoding strategy, and arguably switched to a featural encoding strategy due to the demands of the task and the knowledge that they would later be constructing a facial composite of the target face they were viewing. An advantage of utilising both holistic and featural cognitive processing at the face encoding stage is consistent with the finding in Study Four that within the control group, field independent individuals who were encouraged to encode the target face in a holistic manner produced higher rated composites than those who encoded the face in a featural manner. Similarly, field dependent individuals who encoded the target face in a featural manner produced higher rated composites than those who encoded the face in a holistic manner.

11.5.3: Verbaliser/imager cognitive style

A second aim of Study Two was to investigate whether individual differences in verbaliser/imager cognitive style had an effect on the accuracy of facial composites that individuals produce. Results showed that individual scores on the verbaliser/imager continuum as denoted by Peterson et al. (2005b) for the VICS test were not related to the accuracy of facial composites that participants produced. According to Riding and Cheema (1991) the verbaliser/imager cognitive style construct is the second major construct alongside the holistic/analytic in which all other styles identified in the literature can be subsumed.

The verbaliser/imager cognitive style test developed by Peterson et al. (2005a) showed a high degree of test-retest reliability. However the form of the verbaliser/imager test was criticised by Massa and Mayer (2005) who suggested that it does not measure an individual's primary mode of representing information in memory. This is because no questions are asked about how information is being processed by the test taker. Other tests such as the VVIQ (Richardson, 1977) ask questions directly about how information is being processed. At least two possible interpretations arise from the result in Study Two that the verbaliser//imager cognitive style did not affect the accuracy of facial composites produced using E-FIT. Firstly, it may be the case that holding a vivid mental image of a target face in memory is not sufficient alone to confer an advantage in composite construction, as the construction process is one that requires both imagery processes *and* the ability to describe the image to a composite system operator. Secondly, it may be the case that the VICS test for verbaliser/imager cognitive style is not a sensitive enough measure to determine any differences in cognitive style within individuals. Despite the criticism that the VICS has received, in a survey of members of the European Learning Styles Information Network who are active researchers in the field of cognitive style the VICS is the test most often used to assess verbaliser/imager style (Peterson, Rayner & Armstrong, 2009).

11.5.4: Summary

This section presents a summary of the evidence relating to the first aim of the research reported in this thesis. Is there a relationship between stable cognitive style and accurate facial composite construction? The role of holistic and featural processing in facial composite construction from the perspective of stable individual cognitive style appears to be mediated by the way in which faces are encoded. Overall the biggest effect of stable cognitive style was found for the construct of field dependence/independence. There is evidence from Studies One and Four that a field independent or part-based cognitive processing style, as measured by the GEFT, is advantageous for facial composite construction regardless of whether the composite system used is predominantly featural as with E-FIT or predominantly holistic as with EFIT-V. The results from Study Two indicated that a holistic or whole-face based cognitive style as measured by the E-CSA-W/A is advantageous for facial composite construction using E-FIT.

Taken together, these results suggest that there are at least two possible interpretations for these apparently conflicting results. Firstly it might be the case that even though field dependence is subsumed within the holistic/analytic family of cognitive styles, it is a separate construct. Secondly, if the style constructs can be considered to be essentially the same, then it may be the case that an advantage for part-based or whole-face based processing is conferred at either the face encoding stage or the composite construction stage. Field dependence/independence was found to interact with the manner in which the target face was encoded, suggesting that it is possibly the utilisation of both holistic and featural cognitive processes which is particularly useful for subsequent facial composite construction. The following section considers the effect of manipulating cognitive processing using the Navon task on facial composite construction using both E-FIT and EFIT-V.

11.6: Manipulating cognitive processing using the Navon Task

The research presented in Studies Three and Four was designed to address the issue of whether manipulating cognitive processing using the Navon task could have a positive effect on the accuracy of facial composites produced using E-FIT and EFIT-V. The results of Study Three indicated that there is a differential effect of Navon task on accurate facial composite construction using E-FIT, and this was dependent on the way in which the target faces had been encoded. When the target face was encoded in a featural manner there was no effect of Navon task on the accuracy of facial composites produced. When the target face was encoded in a holistic manner, then there was an advantage for both the local and global Navon task groups over the control group.

The advantage for the global Navon task group following holistic encoding of the face lends support to the matching superiority hypothesis that similar cognitive processes at encoding and retrieval will aid memory performance. This finding is in line with that of Weston et al. (2008) who found that if a target face was encoded in a holistic manner, the global Navon task led to greater face recognition, which may have been advantageous for participants towards the end of the E-FIT construction process. However, the advantage for the local Navon task group following holistic encoding lends support to the hypothesis that the utilisation of both featural and holistic cognitive processing is advantageous for composite construction. If the matching superiority hypothesis is to be accepted, then the featural encoding condition should have resulted in an advantage for facial composite construction following the local Navon task, but this was not the case. It may be however, that the effect of featural processing at encoding may mask the effects of the Navon task. The by-rater results from Study Three showed an effect size of .584 (partial eta squared) for face encoding and a smaller effect size of .160 for the Navon task. However, this smaller effect size for the Navon task was collapsed across all face encoding conditions. The effect size for Navon task rose to .414 within the holistic face encoding condition. Similarly, in Study Four, the by-rater results showed an effect size of .459 for face encoding and .082 for Navon task collapsed across all experimental conditions. Therefore it is unclear whether matching superiority, or the utilisation of both types of cognitive processing, can best account for the finding that both Navon task groups produced more accurate facial composites following holistic encoding of the target face.

An alternative explanation for the advantage of either Navon task over the control group in Study Three is that there could potentially have been a Hawthorne effect (Adair, 1984) on the performance of the participants who constructed the facial composites. Participants who completed a Navon task were unaware of whether the Navon task condition they were assigned to was designed to assist or to hinder their facial composite construction performance. They would arguably have assumed that the condition they were assigned to would be designed to give them a performance advantage, and this may have influenced their motivation to do well in the subsequent facial composite construction task. Alternatively, it may have been the case that the frequent pauses that participants encountered during construction of an E-FIT, where they were required to pause following the final selection of each feature and complete a Navon task before moving on to the next feature, could have been advantageous in terms of the accuracy of the composite they produced. This is because the act of having to pause between the selection of each feature made composite construction slower and more deliberate, and may have therefore emphasised concentration on each feature.

In Study Four, the target faces were encoded in the same way as for Study Three, the Navon task was used prior to composite construction, and the field dependence/independence cognitive style of the participants was measured. A different pattern of Navon effect was found in Study Four to that which was found in Study Three. Firstly if the participant was encouraged to encode the target face according to their cognitive style type, then there was no effect of Navon task on facial composite accuracy. There was no Navon effect in the holistic face encoding condition for field dependent individuals, or in the featural face encoding condition for the field independent individuals. An interesting finding was that under holistic encoding conditions, the field independent participants in the control group produced higher rated composites than either

Navon task group. This is the opposite finding to Study Three where holistic face encoding coupled with either Navon task led to better performance than the control group. This raises the question of why the Navon task should have such a contrasting effect under similar encoding conditions. In Study Four, it might be that as both types of cognitive processing were already potentially utilised during face encoding, the field independent participants were encouraged to process the faces holistically, this meant that those participants were already operating at their optimal performance level in terms of facial composite construction. Therefore the Navon task would have conferred no added advantage to the memory of the participants. The cognitive style of the participants in Study Three was not measured, so the effect of cognitive style on Navon performance using E-FIT could not be assessed.

Continuing the interpretation that a dual cognitive processing strategy is best for subsequent face recall, it follows that if optimal face recall performance is achieved when both holistic and featural processing is utilised at the face encoding stage, then field dependent participants in the featural encoding condition should also have produced higher rated facial composites, regardless of which Navon task intervention was used. Within this experimental group of participants in Study Four it was found that those who completed a local Navon task produced significantly higher rated composites than the control group, and that both groups produced significantly higher rated composites than the global Navon group. This finding is not entirely inconsistent with the dual cognitive processing advantage interpretation and additionally highlights the importance of featural cognitive processing at all stages of facial composite construction.

This finding is therefore different to Frowd et al. (2012) who found that encouraging holistic processing immediately prior to constructing a facial composite using EvoFIT was advantageous for facial composite construction. However, holistic processing in the Frowd et al. study was induced by asking for personality judgements of the target face, and it is probable that the cognitive processing invoked by the global Navon task is different to the processing invoked by making personality judgements about a face. For example, Weston et al. (2008) found that inducing cognitive processing using a Navon task led to a different pattern of face recognition results than inducing cognitive processing by asking participants to make judgements about personality or facial features. There are also differences between the experimental conditions of Studies Three and Four which may account for the finding that the Navon task appears to have differing effects between studies.

Firstly, two different facial composite construction systems were used within the studies, Study Three used E-FIT which is predominantly featural, and Study Four used EFIT-V which is predominantly holistic. The significant main effect of featural encoding in both Studies Three and Four highlights the importance of featural processing irrespective of which facial composite system is used to construct a facial composite. However, the difference in subjective likeness ratings for the holistic and featural encoding groups was greater when the E-FIT system was used (Cohen's $d = 0.8$). With EFIT-V, although featural encoding led to composites with higher ratings, the difference between the featural and holistic encoding groups was far less (Cohen's $d = 0.2$) suggesting that any relatively detrimental effect of holistic encoding may have been ameliorated by using a holistic composite construction system.

Another difference between Studies Three and Four which could account for the differential effects of the Navon task is the number of times that the Navon task was used during composite construction. Previous research has indicated that the Navon effect may be short-lived. For example Weston et al. (2008) found that the effects of inducing Navon processing on face identification were strongest in the first trial. If individual performance was averaged over several trials of face recognition following a single cognitive processing manipulation, then the Navon effect disappeared. In Study Three of this thesis, E-FIT was used, and a Navon task was completed by participants before the selection of each feature in order to try to maintain any effect of Navon processing for individual feature selection. In Study Four, EFIT-V was used, and the Navon task was administered once immediately prior to the beginning of the composite construction procedure. The Navon task was not administered during EFIT-V composite construction because there are no natural breaks in the construction process as there are with the E-FIT system. Additionally, the process of constructing a facial composite using EFIT-V is generally much shorter than that of constructing an E-FIT. It is possible, therefore, that the cognitive processing effect induced by the Navon task in either or both Studies Three and Four was too short-lived to have an effect on the accuracy of composites that participants produced. The effect of Navon processing was potentially stronger for those who made a composite using E-FIT, as the Navon task was administered repeatedly throughout the process. However this is not possible to ascertain as there is a lack of consensus in the literature on the duration of the Navon effect. However, the observation of a Navon effect in Study Four where the Navon task was used only once suggests one of two possible interpretations. Firstly the Navon effect may not be as short-lived as first suggested. Secondly, the Navon effect may be strongest early in the EFIT-V composite construction process. It is early on in the process of EFIT-V construction that the faces displayed on screen vary to the greatest degree.

Alternatively, the effects of inducing cognitive processing using the Navon task may have been mediated by the cognitive style of the participants. Consistent with this interpretation, the result from Study Four reported in this thesis suggests that the field dependence/independence cognitive style of participants had a large effect on the accuracy of facial composites (.548, partial eta squared). The way in which the target faces were encoded also had a large effect on the accuracy of facial composites produced (.459, partial eta squared), and interacted with the cognitive style of participants. If the way in which the target faces were encoded is removed from the analysis of the Study Four data, as encoding cannot be manipulated in a real-world setting, a tentative interpretation can be made with regard to the effect of the Navon task on facial composite construction using EFIT-V.

Field independent participants in the control group performed best overall within the field independent group of participants, but not significantly better than the local Navon group. Field dependent participants in the local Navon group performed best overall within the field dependent group of participants, but not significantly better than the control group. Within all participants it was those in the global Navon group who produced facial composites with significantly lower likeness ratings than the best performing group within each condition. As global precedence Navon stimuli were used throughout the experiments within this thesis, it might be that the local Navon task groups were encouraged to adopt a controlled or analytic processing mode, as they were responding to the non-dominant aspect of the Navon stimuli. As an automatic processing strategy has been found to lead to greater accuracy in face recognition tasks (Dunning & Stern, 1994), it may be that controlled or analytic processing may lead to more accurate facial composites.

The idea that controlled processing could lead to more accurate facial composites, could account for the Study Three finding that either Navon task led to facial composites which received higher likeness ratings. It is possible that making eyewitness-participants pause between the selection of different facial features led to a more considered selection of each feature, in turn leading to a more accurate facial composite. The interpretation of an advantage for analytic processing in facial composite construction is further supported by the finding that field independent individuals produce the highest rated facial composites regardless of whether the composite system used for composite construction is mainly featural or holistic. It is also supported by the finding that controlled encoding of a face through use of a featural encoding strategy also generally leads to facial composites with higher ratings.

Even with the newest holistic face recognition based composite construction systems such as EFIT-V where whole faces are viewed on the screen, the task is still one of face recall. This is because the composite faces viewed on screen are never going to be an exact match of the target face viewed by the witness. Therefore the witness is always going to be undergoing a process where they are attempting to combine recognition processes with recall processes. The witness will always be trying to make a match between the face they are viewing on screen, and the representation of that face they hold in their memory.

11.6.1: Summary

This section presents a summary of the evidence relating to the second aim of the research reported in this thesis. Can cognitive processing be manipulated (using a Navon task) during or prior to facial composite construction to have a positive effect on the accuracy of facial composites? All published research to date has investigated the effect of the Navon task on face or part-face recognition. Studies Three and Four presented in this thesis are the first to investigate the effect of the Navon task on the accuracy of facial composites that individuals produce using E-FIT and EFIT-V. Facial composite construction contains elements of both face recall and recognition and elements of both holistic and featural cognitive processing. E-FIT is a predominantly featural system where the eyewitness must recall and describe individual facial features for the purpose of composite construction. EFIT-V is a predominantly recognition-based system where the eyewitness views whole faces and does not have to manipulate individual features. The different emphasis on featural and holistic processing within E-FIT and EFIT-V may account for the different effects of Navon task observed between Studies Three and Four. The Navon task may also produce differential effects for face recall and face recognition.

Using the E-FIT system in Study Three there was an advantage for either Navon task over a control group which interacted with the way in which faces were encoded. However, Study Four using EFIT-V demonstrated that the effect of the Navon task also interacted with the field dependence/independence cognitive style of the individual in addition to interacting with the way faces were encoded. There are several further possible reasons for the differential effects observed between composite construction systems. Firstly there were differences in the way in which the Navon task was administered. This may have implications for the length of time the induced Navon effect lasts, although the observed effect of Navon task in Study Four suggests that it can be useful when administered once prior to composite construction. An interesting interpretation for the finding that either Navon task was advantageous in Study Three is the idea that the Navon task may have induced controlled processing during the composite construction process. Automatic processing has been found to be beneficial for face recognition, and it may be the case that controlled processing is beneficial for E-FIT composite construction. Future research on cognitive style and manipulating cognitive processing using the Navon task prior to or during the composite construction process may help to elucidate some of the issues arising from the research reported in this thesis. Research manipulating cognitive processing by manipulating the type of interview that participants receive prior to composite construction has produced some promising results (Frowd et al., 2008). It has been demonstrated that orienting participants to a featural processing style using a cognitive interview for feature description, and then inducing a holistic style by asking personality questions leads to more recognisable facial composites using both a featural and a holistic system (Frowd et al., 2008; Frowd et al., 2012).

11.7: Further research

The findings presented herein have raised a number of possible avenues for future investigation. Several questions remain unanswered with regard to the nature and origin of the Navon effect in both face recognition tasks and in composite construction tasks. For example, it is unclear whether the Navon effect can be attributed to a shift in cognitive processing style or to a switch between automatic and controlled processing, or to other factors such as an increase in motivation or arousal. Future research using the Navon task within facial composite construction could use comparisons with other tasks thought to affect cognitive processing style such as verbalisation or making holistic and featural judgements about faces.

In practical terms, further research on the length of time for which the Navon effect lasts could be informative in determining the extent to which it is useful for facial composite construction and how often it may need to be administered during the composite construction process. Future research could also use both global and local precedence Navon stimuli during facial composite construction to assess whether it is cognitive processing or automatic versus controlled processing which is being manipulated by the Navon task.

Relating to cognitive style, Studies One and Four tested participants for field dependence/independence and Study Two tested participants for their holistic/analytic cognitive style. Future research could test participants for both cognitive styles in order to assess how they interact, and their impact on composite construction performance. Similarly, it would be interesting to conduct a study directly comparing a featural system such as E-FIT with a holistic system such as EFIT-V in order to assess the relative impact of cognitive style on composites produced by the two systems.

Individual differences in cognitive style may have an impact on which composite construction system would most successfully be employed by an individual eyewitness: it may be the case that individuals who habitually process visual information in a generally more featural way would produce a more accurate facial composite using the E-FIT system, where a face is constructed feature by feature, and that individuals who are relatively more holistic would produce more accurate composites using the EFIT-V system. Although Study Four assessed the impact of field dependence/independence on the Navon effect, it may be useful to investigate the possible impact of the holistic/analytic cognitive style on the effect of the Navon task.

The words used by participants in Studies One to Four to describe the target face they viewed were not analysed for differences. However some participants tended to provide predominantly featural descriptions while others provided predominantly holistic descriptions. Future research could investigate whether there is a correlation between cognitive style and the type of description of the face that eyewitnesses provide. This in turn could be an easy way of determining either the cognitive style of an eyewitness, or the way in which they encoded the target face. Knowledge of either a natural featural style or of featural encoding of a face may help to determine who might make a better composite. In the case of a single witness, it may help determine the likelihood of any composite produced being accurate, in conjunction with other factors known to affect composite quality.

To summarise, the present research has demonstrated that cognitive style and cognitive processing can have an effect on the accuracy of facial composites that individuals produce. The manipulation of cognitive processing is a promising avenue for research in that it offers the possibility of developing a short intervention which can be used with eyewitnesses to improve their facial compositing performance. The following section describes some methodological issues inherent within the current research.

11.8: Methodological Issues

There are limitations and issues of ecological validity within the research reported in this thesis which are common in the field of facial composite research. Firstly it would not be ethical to induce the type or degree of stress in participants that is often involved with real eyewitnesses. An associated issue is the potentially lower level of motivation applied to the task of facial composite construction by participants in an experimental setting. A positive aspect concerning motivation is the fact that participants found the task of producing a facial composite both interesting and engaging. Several participants requested feedback on how their composites had been rated in comparison to others following the subjective likeness ratings phases, indicating a positive engagement with the composite construction process.

A further issue with the current research is the fact that there was no realistic time delay between viewing the target face and constructing the facial composite. In reality, there would be a gap of at least two days between a person witnessing an event and constructing a facial composite for the police, and sometimes a gap of much longer than two days (Frowd et al., 2005a). There was a gap between viewing the face and composite construction in the present research of around ten minutes, and during that time a maths calculation task was employed with participants which prevented rehearsal of the target face in short term memory. However, the lack of a realistic time delay might have implications for the way in which faces are remembered, in that after a delay of a couple of days the memory of a face becomes more global, and more general in nature.

The reasons for not employing a significant time delay within the current research are two-fold. Firstly there is a high attrition rate among student participants which would have been exacerbated by requiring them to attend on two occasions several days apart. Secondly, in some studies they already had to attend twice, once for cognitive style testing, and later to construct a composite. To have exposed participants to a target face at the same time as cognitive style testing may have unintentionally influenced the way in which the target face was encoded, therefore to avoid participants having to attend on three separate occasions the target face was viewed and the composite constructed within the same session. An advantage of this approach was that it negated the need to conduct a cognitive interview with participants in order to reinstate the context in which the target face was viewed, as viewing took place in the same research room in which the facial composite was subsequently constructed.

In all Studies apart from Study One, subjective likeness ratings were used as the only method of evaluating the composites. The most ecologically valid method for evaluating composites is acknowledged to be the naming of those composites by a participant who is familiar with the person that the composite is supposed to depict, following construction of that composite by a person who is unfamiliar with the target face used for composite construction. However, the number of participants used in the construction and evaluation phases of each study, meant that in practical terms, to find target faces which were unfamiliar to all the composite constructors but familiar to all the evaluators would have been extremely difficult. The issue of finding suitable target faces for each study, coupled with the floor effects that naming data produce in some published studies meant that naming was not considered to be a viable measure of the individual differences investigated within this thesis.

Another common method of evaluating composites in the literature is the matching task, where participants are given a series of facial composites and are required to match them to the target face they are supposed to depict from an array of target faces on display. Matching tasks were not used in the present research because they may not indicate how good or how accurate a facial composite is. If some facial composites are matched by at least one person, and other facial composites are matched by several people, this does not give an indication of which composite may work best in a practical context, as only one positive match is required by the police in order to apprehend a perpetrator of a crime.

Ranking tasks, where composites are placed in order of accuracy or similarity to the target face, may exaggerate slight differences between composites. This is something that should not occur with subjective likeness ratings, because raters could potentially give the same rating score to composites that are perceived to be very similar in terms of accuracy. In the studies reported in this thesis subjective likeness ratings were considered to be the best choice of evaluation for investigating differences in composite accuracy as a result of cognitive style or induced cognitive processing state. This is because the continuous data on a percentage scale generated by the likeness ratings task allows for analyses which would best show up the differences between experimental groups.

Even though subjective likeness ratings were considered to be the best individual method of evaluating the facial composites produced in all four studies reported in this thesis, there are limitations with using this method. Most notably, published research on facial composite construction systems and techniques typically uses several methods of evaluation in conjunction. This is because differing inferences can be drawn from different methods of evaluation, and also because the usefulness of any system or technique of composite construction can be confirmed by several converging methods of evaluation.

A second problem which arises from the sole use of subjective likeness ratings is the possibility that the person doing the evaluating is basing their rating on the accuracy of the global features of the facial composite they are evaluating. This is because there is the tendency to process unfamiliar faces by their external features such as hair and face shape (Young et al., 1985) and to process familiar faces by the relatively more internal features such as the eyes and mouth (Ellis, 1986). It may be the case therefore, that the subjective likeness rating task is not the best measure of the accuracy of the facial features of the composites produced in the current research. Given that the results of Studies One to Four emphasise the importance of featural encoding, field independent featural cognitive style, and the featural, local Navon task it would be useful to assess whether the internal features of the facial composites produced were enhanced. This could be achieved by conducting a further analysis where the external regions of the facial composites produced are masked, and identification of the composites is attempted on inspection of the internal face region alone. Both by-item and by-rater analyses were conducted on the subjective likeness rating data as the two types of analysis allow different inferences to be drawn about the utility of the facial composites produced.

The by-rater analysis of subjective likeness ratings gives an indication of the proportion of the population who are likely to make a connection between the composite and someone they know in the event that they recognise the person depicted in the facial composite. By-rater analyses were carried out in addition to by-item analyses because experimental power is increased and therefore effect sizes tend to be larger within by-rater analyses. For example, the by-rater analysis of the Study Three data had observed power of .933 and showed a main effect of Navon task (partial $\eta^2 = .160$), whereas the by-item analysis of the same data set had observed power of .057 and did not detect a main effect of Navon task (partial $\eta^2 = .002$).

The possibility remains that the non-significant by-item analyses in Studies One, Three and Four might be due to insufficient statistical power to detect modest population effect sizes. This might be particularly so in Study Four where the inclusion of three factors resulted in there being only twelve participants in each experimental condition, even though 72 facial composites were produced. Two possible solutions to this issue would be to reduce the number of experimental conditions or preferably, to increase the number of facial composites constructed. An increase in participant composite-constructors would be desirable because by-item analyses (which pertain to the composites produced and not the people who evaluate them), would be more informative about the effect of experimental conditions on the individuals who constructed the facial composites. It takes only one observer/evaluator to make a positive identification of a facial composite, but that composite must be at least an accurate type-likeness of the perpetrator in the first instance, for positive identification to occur. Therefore, the population of particular interest in facial composite research remains the population of constructors, that which lends itself to by-item analyses.

The correlation between subjective likeness ratings and the objective measure of evaluation used in Study One demonstrates that subjective likeness ratings are a valid measure of the likely usefulness of facial composites in an applied setting. One potential problem with a rating task is that the target photos are present during ratings, and participants may therefore carry out a feature by feature match and ignore the holistic aspects of the facial composites. However, all evaluations in an applied real world setting are essentially subjective, and it may be that witnesses viewing facial composites in the media adopt a feature matching strategy.

A potential limitation within the research was that all of the target faces used were Caucasian, whereas the participants comprised all ethnicities. Given the diverse nature of the potential participant pool it was not possible to control for race of participant, or to match the ethnicity of the participant to the target faces. However, the data from all four studies were checked and there was no evidence for an own-race effect, either within the participants who constructed a facial composite or in the participants who acted as raters, in the rating scores given to the facial composites. According to the contact hypothesis of the own-race bias in face recognition it is probable that in an increasingly multi-ethnic society, the-own race effect is likely to be greatly diminished. This is due to the fact that individuals within that society would likely have contact with others from diverse ethnic backgrounds on a daily basis (Meissner & Brigham, 2001).

Finally, the Group Embedded Figures Test used to measure field dependence/independence has been criticised because it only measures one end of the FDI continuum. Relatively high scores indicate that an individual is field independent, and field dependence is inferred from attaining a low score. However, other possible reasons for a low score on the GEFT include lack of motivation, or boredom for example, and may not therefore necessarily be measuring an individual's visual processing preference. The largest potential limitation however is the application of a median split to GEFT scores within the literature (Emmett et al., 2003). Applying a median split to GEFT scores means that in some studies a score around the middle of the scale would mean a participant being classed as field independent and in another study the same score might mean the same participant being classed as field dependent. While it is acknowledged that the same individual being classed as field independent or field dependent in differing studies based on a median split of the data is a potential confound, it is the case that the GEFT still distinguishes an individual's position on the GEFT scale in *relation* to the other individuals in that particular sample. Therefore, the GEFT is still a useful tool with which to distinguish individuals based on their test scores, and the differences observed with scores on the GEFT do translate to differences observed in performance in forensic settings (Emmett et al., 2003; Studies One and Four, this thesis).

11.9: Practical applications

Psychological research on facial composite systems has focused on the technical development of the systems and the way in which witnesses interact with them. The experiments presented herein have contributed to the body of knowledge concerning the underlying cognitive processes of eyewitnesses when recalling a face for the purpose of facial composite construction. One potential way of determining which of several eyewitnesses might make a relatively more accurate facial composite is to test the eyewitnesses for their level of field dependence/independence using the GEFT. The GEFT is an easy test to administer and score, and takes only ten minutes to complete. The eyewitness with the highest score on the GEFT could then potentially be the witness most likely to produce an accurate facial composite if other factors are equal, such as length of time the witnesses had to view the face of the perpetrator, and lighting conditions for example (Devlin, 1976). However, administration of the GEFT to witnesses would require extra training for composite system operators, and could lead to a situation where some witnesses might complete the GEFT but then perhaps not be required to construct a facial composite if they received a low score on the GEFT relative to other witnesses.

Considering the results from Study Two, it might also be possible to determine who may construct a more accurate facial composite by administering the E-CSA-W/A test for holistic/analytic cognitive style to eyewitnesses. Although the E-CSA-W/A is also an easy test to administer and score, there would remain the same issue as with getting witnesses to complete the GEFT. It would be a waste of a witness's time to get them to complete a test which might eliminate them from constructing a facial composite depending on the score they attained.

A natural featural encoding strategy may be beneficial for facial composite construction. Therefore there may be a simpler way of determining which eyewitness may construct a more accurate facial composite than testing for their cognitive style. This could be just to ask eyewitnesses whether they deliberately noticed the features of the perpetrator. Previous research has suggested that self reported encoding strategy has an impact on subsequent face recognition (Olsson & Juslin, 1999). Therefore if an eyewitness reported intentionally encoding a face for future reference, they might construct a relatively more accurate facial composite than an eyewitness who did not intentionally encode a face.

Alternatively the Navon task could be administered to all eyewitnesses as part of the composite construction process. The results from Studies Three and Four indicate that overall, completion of a local Navon task prior to composite construction may lead to more accurate facial composites, particularly if the witness is field dependent. For field independent witnesses there may be no significant advantage to be gained by completing a local Navon task, but no significant detriment either. All composite construction systems are computerised, therefore it could be easy to administer the featural Navon task prior to composite construction because the Navon task is run on a computer. Further research is needed on cognitive style and cognitive processing before such changes could be recommended for use in police practice.

11.10: Conclusion

Overall, the results of the research presented within this thesis suggest that there may be a strong featural encoding element to facial composite construction. If witnesses were encouraged to encode faces in a featural way they generally produced composites of higher rated accuracy than those witnesses who encoded faces in a holistic way. If witnesses were naturally predisposed to process in a featural manner, if they were field independent, they generally produced composites of higher rated accuracy than field dependent witnesses. The way in which faces are encoded, and the cognitive style of witnesses cannot be manipulated. However knowledge of how these variables affect eyewitness performance in producing a facial composite may assist police in the decision of whether to publish a composite that a witness has constructed. It may be possible to influence the cognitive processing strategy of witnesses using a short Navon task prior to facial composite construction. It may be that completion of a local Navon task could be advantageous to composite construction. It is not clear whether this advantage may reside in invoking a featural cognitive processing style or in invoking a controlled rather than automatic processing style. Although automatic processing has been found to be advantageous for face recognition, these findings suggest that it may be controlled featural processing that is advantageous for facial composite construction. Further research is required to further our understanding of this phenomenon before making strong recommendations about the value of introducing a Navon task into face composite protocols.

<u>References</u>

Adair, J. G. (1984). The Hawthorne effect: A reconsideration of the methodological artefact. *Journal of Applied Psychology,* **69 (2),** 334-345.

Allport, G. W. (1937). cited in Kozhevnikov, M. (2007). Cognitive Styles in the Context of Modern Psychology: Toward an Integrated Framework of Cognitive Style. *Psychological Bulletin,* **133 (3),** 464-481.

Aspley Limited (1993). E-FIT for Windows (Version 5.1)

Berman, G. L. & Cutler, B. L. (1998). The influence of processing instructions at encoding and retrieval on face recognition accuracy. *Psychology, Crime and Law,* **4**, 89-106.

Bloom, L. C. & Mudd, S. A. (1991). Depth of processing approach to face recognition: A test of two theories. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **17,** 556-565.

Bower, G. H. & Karlin, M. B. (1974). Depth of processing pictures of faces and recognition memory. *Journal of Experimental Psychology,* **103 (4),** 751-757.

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

Brown, C. & Lloyd-Jones, T. J. (2002). Verbal overshadowing in a Multiple Face Presentation Paradigm: Effects of Description Instruction. *Applied Cognitive Psychology,* **16,** 873-885.

Bruce, V. (1982). Changing faces: Visual and non-visual coding processes in face recognition. *British Journal of Psychology,* **73,** 105-116.

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

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

Christie, D., Davies, G., Shepherd, J. & Ellis, H. (1981). Evaluating a New

Computer-Based System for Face Recall. *Law and Human Behavior,* **5,** 209-218.

Christie, D. & Ellis, H. (1981). Photo-FIT Constructions Versus Verbal Descriptions of Faces. *Journal of Applied Psychology,* **66,** 358-363.

Coin, C. & Tiberghien, G. (1997). Encoding Activity and Face Recognition, *Memory,* **5 (5),** 545-568.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior.* **11 (6),** 671-684.

Darling, S., Martin, D., Hellmann, J,7 Memon, A. (2009). Some witnesses are better than others. *Personality and Individual Differences,* **47,** 369-373.

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.

Davies, G. M., Ellis, H. & Christie, D. (1981). Exploring New Strategies for Facial Recall. *Medicine, Science and Law,* **21,** 137-145.

Davies, G. M., Shepherd, J. W. & Ellis, H. D. (1978). Remembering Faces: Acknowledging our Limitations. *Journal of Forensic Science,* **18,** 19-24.

Davies, J. & Graff, M. (2006). Wholist-analytic cognitive style: A matter of reflection. *Personality and Individual Differences,* **41,** 989-997.

Davies, G. & Oldman, H. (1999). The Impact of Character Attribution on Composite Production: A Real World Effect? *Current Psychology: Developmental, Learning, Personality, Social.* **18 (1),** 128-139.

Davies, G., van der Willik, P. & Morrison, L. J. (2000). Facial Composite Production: A Comparison of Mechanical and Computer-Driven Systems. *Journal of Applied Psychology,* **85,** 119-124.

Devlin, P. (1976). *Report to the Secretary of State for the Home Department on the departmental committee on evidence of identification in criminal cases.* London: HMSO.

Diamond R., & Carey, S. (1986). Why Faces Are and Are Not Special: An Effect of Expertise. *Journal of Experimental Psychology: General.* **115 (2),** 107-117.

Dodson, C. S., Johnson, M. K. & Schooler, J. W. (1997). The verbal overshadowing effect: why descriptions impair face recognition. *Memory and Cognition,* **25,** 129-139.

Dunning, D. & Stern, L. S. (1994). Distinguishing accurate from inaccurate eyewitness identifications via inquiries about decision processes. *Journal of Personality and Social Psychology,* **67 (5),** 818-835.

Ellis, H. (1986). Face Recall: A Psychological Perspective. *Human Learning,* **5,** 189-196.

Ellis, H., Davies, G. M., & Shepherd, J. W (1978). A Critical Examination of the Photo-FIT System for Recalling Faces. *Ergonomics,* **21,** 297-307.

Ellis, H., Shepherd, J. W., & Davies, G. M. (1975). An investigation of the use of the Photo-fit technique for recalling faces. *British Journal of Psychology,* **66,** 29-37.

Ellis, H., 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,** 431-439.

Emmett, D., Clifford, B. R. & Gwyer, P. (2003). An investigation of the interaction between cognitive style and context reinstatement on the memory performance of eyewitnesses. *Personality and Individual Differences,* **34,** 1495-1508.

Endler, N. S. (2000). The Interface between Personality and Cognition. *European Journal of Personality,* **14,** 377-389.

Farah, M. J., Tanaka, J. & Drain, M. (1995). What Causes the Face Inversion Effect? *Journal of Experimental Psychology: Human Perception and Performance,* **21,** 628-634

Fallshore, M. & Schooler, J. W. (1995). Verbal vulnerability of perceptual expertise. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **21 (6),** 1608-1623.

Finger, K. & Pezdek, K. (1999). The Effect of the Cognitive Interview on Face Identification Accuracy: Release From Verbal Overshadowing. *Journal of Applied Psychology,* **84,** 340-348.

Frowd, C. D., Bruce, V., McIntyre, A., & Hancock, P. (2007a). The relative importance of external and internal features of facial composites. *British Journal of Psychology,* **98 (1),** 61-77.

Frowd, C. D., Bruce, V., Ness, H., Bowie, L., Paterson, J., Thomson-Bogner, C., McIntyre, A. & Hancock, P. J. B. (2007b). Parallel approaches to composite production: interfaces that behave contrary to expectation. *Ergonomics,* **50 (4),** 562-585.

Frowd, C. D., Bruce, V., Smith, A. J., & 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.

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

Frowd, C. D., Carson, D., Ness, H., Richardson, J., Morrison, L., McLanaghan, S. & Hancock, P. (2005b). A forensically valid comparison of Facial Composite Systems. *Psychology, Crime and Law,* **11,** 33-52.

Frowd, C. D., Hancock, P. & Carson, D. (2004). EvoFIT: A Holistic, Evolutionary Facial Imaging Technique for Creating Composites. *ACM Transactions on Applied Perception (TAP),* **1 (1),** 19-39.

Frowd, C. D., Nelson, L., Skelton, F., Noyce, R., Atkins, R., Heard, P., Morgan, D., Fields, S, Henry, J., Mcintyre, A., & Hancock, P. J. B. (2012). Interviewing Techniques for Darwinian Facial-Composite Systems. *Applied Cognitive Psychology,* Wiley Online Library.com, DOI: 10.1002 /acp.2829

Gibling, F. & Bennett, P. (1994). Artistic enhancement in the production of photo-fit Likenesses: An examination of its effectiveness in leading to suspect identification. *Psychology, Crime and Law*, **1 (1),** 93-100.

Gibson, S. J., Solomon, C. J. & Pallares-Bejarano, A. (2003). Synthesis of photographic quality facial composites using evolutionary algorithms. In R. Harvey and J. A. Bangham (Eds.) Proceedings of the British Machine Vision Conference (pp. 221-230).

Goldstein, K. M. & Blackman, S. (1978). *Cognitive Style: Five approaches and relevant Research.* New York: Wiley.

Goodenough, D. R. (1976). The Role of Individual Differences in Field Dependence as a Factor in Learning and Memory, *Psychological Bulletin,* **83 (4),** 675-694.

Gordon, R. (1949). cited in Laughery, K. R. & Fowler, R. H. (1980). Sketch Artist and

Identi-kit Procedures for Recalling Faces. *Journal of Applied Psychology,* **65,** 307-316.

Griffiths, R. & Sheen, R. (1992). Disembedded Figures in the Landscape: A Reappraisal of L2 Research on Field Dependence/Independence. *Applied Linguistics,* **13, (2),** 133-148.

Hancock, P. J. B., Bruce, V. & Burton, M. A. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences,* **4 (9),** 330-337.

Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces, *Perception,* **23 (1),** 65-74.

Homa, D., Haver, B. & Schwartz, T. (1976). Perceptibility of schematic face stimuli: Evidence for a perceptual Gestalt. *Memory and Cognition,* **4,** 176-185.

Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience,* **3,** 759-762.

King, D. (1971). The use of Photo-FIT 1970-1971: A progress report. *Police Research Bulletin.*

Kozhevnikov, M. (2007). Cognitive Styles in the Context of Modern Psychology: Toward an Integrated Framework of Cognitive Style. *Psychological Bulletin,* **133 (3)** 464-481.

Laughery, K. R., Duval, C., & Wogalter, M. S. (1986). cited in Frowd, C. D., Bruce, V., Ness, H., Bowie, L., Paterson, J., Thomson-Bogner, C., McIntyre, A., & Hancock, P. J. B. (2007). Parallel approaches to composite construction: interfaces that behave contrary to expectation. *Ergonomics,* **50 (4),** 562-585.

Laughery, K. R. & Fowler, R. H. (1980). Sketch Artist and Identi-Kit Procedures for Recalling Faces. *Journal of Applied Psychology,* **65,** 307-316.

Lewis, M. B., Mills, C., Hills, P. J., & Weston, N. (2009). Navon Letters affect Face Learning and Face Retrieval, *Experimental Psychology,* **56 (4),** 258-264.

Macrae, C. N. & Lewis, H. L. (2002). Do I Know You? Processing Orientation and Face Recognition. *Psychological Science,* **13,** 194-196.

Martin, D., & Macrae, C. N. (2010). Processing style and person recognition: Exploring the face inversion effect. *Visual Cognition,* **18 (2),** 161-170.

Massa, L. J. & Mayer, R. E. (2005). Three obstacles to validating the Verbal-Imager Subtest of the Cognitive Styles Analysis. *Personality and Individual Differences,*

39, 845-848.

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

McKone, E. (2004). Isolating the Special Component of Face Recognition: Peripheral Identification and a Mooney Face. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **30 (1),** 181-197.

McKone, E., Martini, P. & Nakayama, K. (2001). Categorical Perception of Face Identity in Noise Isolates Configural Processing. *Journal of Experimental Psychology: Human Perception and Performance,* **27 (3),** 573-599.

McQuiston-Surrett, D., & Topp, L. D. (2008). Examining the Accuracy of Facial Descriptions vs. Composites as a Function of the Own-Race Bias. *Experimental Psychology,* **55 (3),** 195-202.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces. *Psychology, Public Policy and Law,* **7 (1),** 3-3

Miller, A. (1987). Cognitive styles: An integrated model. *Educational Psychology,* **7,** 251 - 268.

Milne, E., & Szczerbinski, M. (2009). Global and local perceptual style, field-independence, and central coherence: An attempt at concept validation, *Advances in Cognitive Psychology,* **5,** 1-26.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior,* **16 (5),** 519-533.

Navon, D. (1977). Forest before Trees: The Precedence of Global Features in Visual Perception. *Cognitive Psychology,* **9 (3),** 353-383.

Navon, D. (2003). What does a compound letter tell the psychologist's mind? *Acta Psychologica,* **114,** 273-309.

Olsson, N. & Juslin, P. (1999). Can Self-Reported Encoding Strategy and Recognition Skill Be Diagnostic of Performance in Eyewitness Identifications? *Journal of Applied Psychology,* **84,** 42-49.

Parmentier, F., & Andres, P. (2003). Switch cost asymmetry: Effects of Task Dominance. cited in Perfect, T. J., Weston, N. J., Dennis, I., & Snell, A. (2008). The effects of precedence on Navon-induced processing bias in face recognition. *The Quarterly Journal of Experimental Psychology,* **61 (10),** 1479-1486.

Perfect, T. J., Dennis, I. & Snell, A. (2007). The effects of local and global processing orientation on eyewitness performance. *Memory,* **15 (7),** 784-798.

Perfect, T. J., Weston, N. J., Dennis, I., & Snell, A. (2008). The effects of precedence on Navon-induced processing bias in face recognition. *The Quarterly Journal of Experimental Psychology,* **61 (10),** 1479-1486.

Peterson, E. R., Deary, I. J. & Austin, E. J. (2003). The reliability of Riding's Cognitive Style Analysis test. *Personality and Individual Differences,* **34,** 881-891.

Peterson, E. R., Deary, I. J. & Austin, E. J. (2005a). Are intelligence and personality related to verbal-imagery and wholistic-analytic cognitive styles? *Personality and Individual Differences,* **39,** 201-213.

Peterson, E. R., Deary, I. J. & Austin, E. J. (2005b). A new measure of Verbal-Imagery Cognitive Style: VICS. *Personality and Individual Differences,* **38,** 1269-1281.

Peterson, E. R., Deary, I. J. (2006). Examining wholistic-analytic style using preferences in early information processing. *Personality and Individual Differences,* **41,** 3-14.

Peterson, E.R., Rayner, S. G., & Armstrong, S.J. (2009). Researching the psychology of cognitive style and learning style: Is there really a future? *Learning and Individual Differences,* **19 (4),** 518-523.

Rakover, S. S. (2002). Featural vs. configurational information for faces: A conceptual and empirical analysis. *British Journal of Psychology,* **93,** 1-30.

Rayner, S. (2011). Researching style: Epistemology, paradigm shifts and research interest groups. *Learning and Individual Differences,* **21 (3),** 255-262.

Richardson, A. (1977). The meaning and measurement of memory imagery. *British Journal of Psychology,* **68 (1),** 29-43.

Riding, R. (1991). cited in Peterson, E. R., Deary, I. J., & Austin, E. J. (2003). The reliability of Riding's Cognitive Style Analysis test. *Personality and Individual Differences,* **34,** 881-891.

Riding, R. & Cheema, I. (1991). Cognitive Styles-an overview and integration. *Educational Psychology,* **11,** 193-215.

Riding, R. & Douglas, G. (1993). The effect of cognitive style and mode of presentation on learning performance. *British Journal of Educational Psychology,* **63 (2),** 297-307.

Riding, R. & Pearson, F. (1994). The Relationship between Cognitive Style and Intelligence. *Educational Psychology,* **14,** 413-425.

Riding, R. & Sadler-Smith, E. (1992). Type of Instructional Material, Cognitive Style and Learning Performance. *Educational Studies,* **18 (3),** 323-340.

Riding, R. J. & Wigley, S. (1997). The relationship between cognitive style and personality in further education students. *Personality and Individual Differences,* **23,** 379-389.

Roediger, H. L. (1990). Implicit memory: Retention without remembering. *American Psychologist,* **45 (9),** 1043-1056.

Ryan, R. S. & Schooler, J. W. (1998). Whom do words hurt? Individual Differences in Susceptibility to Verbal Overshadowing. *Applied Cognitive Psychology,* **12,** 105-125.

Sadler-Smith, E. (2001). The relationship between learning style and cognitive style. *Personality and Individual Differences,* **30,** 609-616.

Schooler, J. W. & Engstler-Schooler, T.Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology,* **22 (1),** 36-71.

Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology,* **75,** 221-242.

Shepherd, J. W., Ellis, H. D.,McMurran, M. & Davies, G. M. (1978). Effect of character attribution on Photo-FIT construction of a face. *European Journal of Social psychology,* **8,** 263-268.

Smith, S. M. & Rothkopf, E. Z. (1984). Contextual Enrichment and Distribution of Practice in the Classroom, *Cognition and Instruction,* **1 (3),** 341-358.

Solomon, C. J, Gibson, S. J. & Pallares-Bejanaro, A. (2002). EigenFIT – the generation of photographic quality facial composites. Presented at: Advancing Biometric Technologies: One Day British Machine Vision Association Symposium at the Royal

Statistical Society.

Sternberg, R. J. & Grigorenko, E. L. (1997). Are Cognitive Styles Still in Style? *American Psychologist,* **52,** 700-712.

Stroop, R. (1938). Factors affecting speed in serial verbal reactions. *Psychological Monographs: General and Applied,* **50 (5),** 38-48.

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

Tanaka, J. W. & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory and Cognition,* **25 (5),** 583-592.

Tiedemann, J. (1989). Measures of Cognitive Styles: A Critical Review. *Educational Psychologist,* **24,** 261-275.

Tulving, E. & Thomson,D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review,* **80,** 352-373.

Valentine, T. (1988). Upside-down faces: a review of the effect of inversion upon face recognition. *British Journal of Psychology*, **79 (4),** 471–491.

Valentine, T., Davis, J. P., Thorner, K., Solomon, C. J., & Gibson, S. J. (2010). Evolving and Combining Facial Composites: Between-Witness and Within-Witness Morphs Compared. *Journal of Experimental Psychology: Applied.* **16 (1),** 72-86.

Visionmetric Ltd. (2004). EFIT-V for Windows (Version 4.020).

Wells, G. L. & Hasel, L. E. (2007). Facial Composite Production by Eyewitnesses. *Current Directions in Psychological Science,* **16 (1),** 6-10.

Wells, G. L. & Hyrciw, B. (1984). Memory for faces: Encoding and retrieval operations. *Memory and Cognition,* **12,** 338-344.

Wells, G. L. & Turtle, J. W. (1988). What is the best way to encode faces? In Gruneberg, M. M., Morris, P. E. & Sykes, R. N. (Eds). Practical aspects of memory: Current research and issues, Vol. 1: Memory in everyday life., (pp. 163-168). Oxford, England: John Wiley & Sons.

Weston, N. J., Perfect, T. J., Schooler, J. W. & Dennis, I. (2008). Navon processing and verbalisation: A holistic/featural distinction. *European Journal of Cognitive Psychology,* **20 (3),** 587-611.

Weston, N. J. & Perfect, T. J. (2005). Effects of processing bias on the recognition of composite face halves. *Psychonomic Bulletin and Review,* **12,** 1038-1042.

Wickham, L. H. V. & Lander, K. (2008). The effect of verbal description and processing type on face identification. *European Journal of Cognitive Psychology,* **20 (3),** 577-586.

Wickham, L. H. V. & Swift, H. (2006). Articulatory Suppression Attenuates the Verbal Overshadowing Effect: A Role for Verbal Encoding in Face Identification. *Applied Cognitive Psychology,* **20,** 157-169.

Winograd, E. (1976). Recognition memory for faces following nine different judgements. *Bulletin of the Psychonomic Society,* **8 (6),** 419-421.

Winograd, E. (1981). Elaboration and distinctiveness in memory for faces. *Journal of Experimental Psychology: Human Learning and Memory,* **7 (3),** 181-190.

Witkin, H. A. (1950). Individual differences in ease of perception of embedded figures. *Journal of Personality,* **19 (1),** 1-15.

Witkin, H. A. & Goodenough, D. R. (1977) cited in Endler, N. S. (2000). The Interface between Personality and Cognition. *European Journal of Personality,* **14,** 377-389.

Witkin, H. A., Oltman, P. K., Raskin, E. & Karp, S. A. (1971). A Manual for the Embedded Figures Tests. Palo Alto: Consulting Psychologists Press.

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

Young, A. W., Hellawell, D. & Hay, D. C. (1987). Configurational information in face perception. *Perception,* **16,** 747-759.

Yovel, G., Paller, K. A. & Levy, J. (2005). A whole face is more than the sum of its halves: Interactive processing in face perception. *Visual Cognition,* **12 (2),** 337-352.

Zhang, L. (2004). Field-dependence/independence: cognitive style or perceptual ability? – validating against thinking styles and academic achievement. *Personality and Individual Differences,* **37,** 1295-1311.

Zhang, L. & Sternberg, R. J. (2005). A Threefold Model of Intellectual Styles. *Educational Psychology Review,* **17,** 1-53.

Appendix I: Example of facial composites constructed using E-FIT

Appendix II: Example of an EFIT-V facial composite

Appendix III: Four-person line-up from evaluation phase two of Study One

Question 1. Please select one image of the possible 4 photographed faces that best represents the computerised large image...

Next question

Appendix IV: Six-person line-up from evaluation phase two of Study One

Question 1. Please select one image of the possible 6 photographed faces that best represents the computerised large image...

Next question

Appendix V: Eight-person line-up from evaluation phase two of Study One

Question 1. Please select one image of the possible 8 photographed faces that best represents the computerised large image...

Next question

Appendix VI: Global/holistic face encoding questions for Studies 3/4

Global Processing

On a scale of 1 to 10 can you rate this face for?

Honesty

Intelligence

Trustworthiness

Friendliness

Happiness

Sincerity

Kindness

Attractiveness

Adventurousness

Seriousness

Sociability

Charm

Astuteness

Reliability

Appendix VII: Local/featural face encoding questions for Studies 3/4

Local Processing

On a scale of 1 to 10 can you rate the following?

How current or modern is the hairstyle

How attractive is the face shape

How wrinkled does the forehead look

How trustworthy are the eyes

How attractive are the ears

How attractive is the nose

How kind do the lips look

How striking is the chin