‘He looks like your typical average guy’

Eliciting person descriptions during frontline interviews

Gary Dalton

August 2021

This thesis is submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy of The University of Portsmouth

Abstract

The overarching aims of the current PhD thesis were to explore how person descriptions are obtained from witnesses during frontline police interviews, and to examine the effectiveness of information elicitation techniques in facilitating the reporting of person description information. **Study 1** used body worn video footage to examine how person descriptions are elicited in real-world frontline police interviews. The results showed that over half of the person descriptions obtained by frontline police officers were inappropriate, with leading questions being the most commonly used question type. Appropriate questions (i.e. open questions) led to more person description information being provided (cf. inappropriate questions). **Study 2** examined cross-race differences in the reporting of person description information using a mock witness paradigm. The results showed that white participants provided more person descriptors than black participants, but there was no own-race bias for the number of descriptors provided. However, an own-race bias was obtained for accuracy of person descriptors. **Study 3** tested the effectiveness of using self-generated cues to facilitate the reporting of person descriptions in a single perpetrator event **(Experiment 1)** and an event involving multiple perpetrators **(Experiment 2)**. The results of Experiments 1 and 2 showed that participants who used person description cues provided a higher number of person descriptors than participants who used self-generated cues or a free recall. However, participants who used self-generated cues were more accurate when providing person descriptors than those participants who used person description cues. The set of studies that form the current PhD thesis have contributed to the limited research that has examined how person descriptions are obtained and reported, specifically during the early stages of a criminal investigation. Throughout the PhD thesis, the theoretical implications for person recall and description are discussed, alongside avenues for future research and the practical implications for obtaining person description information.

Table of contents

Abbreviations

ANOVA: Analysis of variance

BWV: Body Worn Video

CI: Cognitive Interview

ICC: Intra-Class Correlation

PDC: Person description cues

PDF: Person description form

MANOVA: Multivariate analysis of variance

MRC: Mental reinstatement of context

SAI: Self-Administered Interview

SGC: Self-Generating Cues

Declaration

Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

Word count: 33,554

List of Tables and Figures

Tables

Figures

Dedication

Dedicated to Angela-Maria Lowth. I hope I have made you proud.

Acknowledgements

When a PhD takes longer than 9 years to complete, it comes as no surprise that I owe my extended thanks to many people. I would like to start by thanking my supervisors, Professor Becky Milne and Professor Lorraine Hope. Becky, it is difficult to do justice with these few words, but thank you for everything. You gave me the opportunity to start my academic journey and the belief you have shown in me will not be forgotten. You have provided me with numerous opportunities to be involved in research that has a real impact in the policing realm. You are the reason I have continued my journey into academia. You have not only been an excellent supervisor, but also a great friend. I look forward to collaborating with you on many more projects in the future. Lorraine, your support as a second supervisor cannot be underestimated. Your research speaks for itself, but your patience and guidance has strengthened my work and challenged me to think more critically and of course with theory in mind.

Next, I need to thank Dr Alex McIntyre, Professor Amina Memon and Professor Aldert Vrij. Alex, you gave me the passion to study and explore forensic psychology, and I will never forget the support you provided to me during my undergraduate degree. Amina, you gave me my first job, and an opportunity to see how applied research impacts the real-world. You supported my early career with patience and kindness. Aldert, you gave me opportunities that I would class as life-changing. Thanks to you, I have been able to collaborate on projects with several external agencies and be part of publications in journals with real impact. There is no question that the opportunities you gave me went a long way to securing my first job as a lecturer.

During this PhD journey, I have been fortunate enough to explore one of my passions – travelling. Aldert, thank you for giving me the opportunity to collaborate with Professor

Galit Nahari in Israel. Zarah and I, have so many good memories of these trips, sharing time with many new and old friends - Ameen, Sharon, and Ami to name a few. I am also grateful to Becky and the Home Team Behavioural Sciences Centre (HTBSC) for funding my internship to Singapore. To all those at the HTBSC, I thank you for your kindness and friendship. Finally, I thank the International Centre for Research in Forensic Psychology for funding my trip to South Africa to work with Professor Colin Tredoux at the University of Cape Town. Cape Town and the people of South Africa are simply wonderful. In particular, I would like to thank Katie whose friendship and sheer kindness allowed me to visit places I could only dream of.

To my colleagues in both the Department of Psychology and in the Institute of Criminal Justice Studies, I would like to offer my thanks. Whilst there are so many staff that have been a part of my journey, I want to start by thanking the Admin and Technician teams. These are the unsung heroes of the University (without them nothing would function). Their smiles and laughs on a 'tough' day cannot be overlooked. Val, Tina, Beth, Rhi, Francis, Nikki, Amy, Dave, Paul and Paul, I thank you.

I would next like to thank my family, who have provided love, financial support, and 'Dalton' wit. I know I am extremely lucky to have a family that value education and whilst perhaps not always appreciated by myself – they have pushed me to succeed.

Finally, and for me most importantly, I would like to thank my friends (many of whom have already been mentioned). Friends also referred to as 'colleagues' are like a second family, and without these people, this dissertation just would not have been possible. Zarah, some might argue that this PhD is only possible because of you – and being honest they may have a point. You have pushed me and believed in me to the point that this is now possible. Thank you for being such a great friend. Gemma, you were the first person I met at

Portsmouth, and your continued friendship – including now supporting Blackburn Rovers – means more than you can know. Through you, I have met Pete and now Kev, creating the 'Thursday beer club', which is always a welcome distraction. To Helene, Jordan, Brandon, Marco, Feni, Ali, Renan, Megan, Eva, Amy, Andie, Sam, Cristina, Maria, Eglantine, Jamie, Sean, Dom, Jerome, Mathilde, Linda, et al., whilst many of you have now left Portsmouth, your friendship and support have meant that these last 9 years have flown by with laughter and amazing experiences. Raul, I am not sure anyone will appreciate how difficult I must be to live with, but thank you for the endless support including beers on a tough day. We have shared some amazing experiences and visited some great places. Archie and Clarke (dogs) thank you for adding to my stress but then showing me that its true – a dog is a man's best friend.

Dissemination

Publications and Book Chapters

Study 1 has been published in Applied Cognitive Psychology, and the findings have been discussed in more broad terms (i.e. frontline communication) in two book chapters (i) Police psychology: New trends in forensic psychological science, and (ii) Encyclopedia of Security and Emergency Management.

Dalton, G., Milne, R., Hope, L., Vernham, Z., & Nunan, J. (2021). 'He was just your typical average guy' Examining how person descriptions are elicited by frontline police officers. *Applied Cognitive Psychology, 35*(2), 517-525. doi: 10.1002/acp.3778.

Dalton, G., & Milne, R. (2020). Interviewing witnesses and victims of crime. Chapter in P. Marques and M. Paulino. (Eds.), *Police psychology: New trends in forensic psychological science.* Elsevier Academic Press.

Milne, B., Shawyer, A., **Dalton, G.,** May, B. L., Nunan, J. H., & Bull, R. (2019). Interviewing adults: communication at the front line. In L. R. Shapiro, & M-H. Maras (Eds.), *Encyclopedia of Security and Emergency Management* Springer. doi: 10.1007/978-3-319-69891-5_160-1.

Conference proceedings

Study 1 was presented at the International Investigative Interviewing Research Group (iIIRG) in June 2016; the Society of Applied Research Memory and Cognition (SARMAC) Conference in January 2017; and at the European Association of Psychology and Law (EAPL) Conference in June 2018.

Dalton, G., Milne, R., Hope, L., & Vernham, Z. (2018). 'He was just your typical average guy': Examining how person descriptions are elicited by frontline police officers. *European Association of Psychology and Law (EAPL) Conference, June 2018*

Dalton, G., Milne, R., Hope, L., & Vernham, Z. (2017). Person description information: An examination of frontline communication. *Society of Applied Research Memory and Cognition (SARMAC) Conference, January 2017.*

Dalton, G., Milne, R., Hope, L., & Vernham, Z. (2016). Body worn video cameras and Investigative Interviewing: Transferring knowledge to the field. *International Investigative Interviewing Research Group (iIIRG), June 2016.*

Study 2 was presented at the European Association of Psychology and Law (EAPL) Conference in July 2019.

Dalton, G., Tredoux, C., Hope, L., & Milne, R. (2019). Person description information: Is there an own race bias? *European Association of Psychology and Law (EAPL) Conference, July 2019.*

Study 3 (Experiments 1 and 2) were presented at the International Investigative Interviewing Research Group (iIIRG) in July 2018.

Dalton, G., Hope, L., Milne, R., & Vernham, Z. (2018). The utility of a self-generated cue mnemonic instruction to facilitate enhanced recall of person descriptions *International Investigative Interviewing Research Group (iIIRG), July 2018.*

Chapter 1

General Introduction

An important application of witness testimony is in the initial apprehension of the perpetrator (Fisher, 1995). Moreover, the successful progress of a criminal investigation depends on eliciting reliable and detailed information about what happened and who was involved (Milne & Bull, 2006). Person descriptions represent an important element of policing when investigating a crime, particularly when locating a perpetrator(s) in the immediate aftermath of an incident (e.g. a terror attack). When providing an account of what they saw, witnesses are typically expected to be able to report everything they can remember about the incident, including a detailed person description about the perpetrator(s). However, people are not that good at providing detailed descriptions in the first instance. Moreover, police interviewing techniques are inadequate (especially on the frontline) and fail to support the reporting of more detailed information about the perpetrator. Although psychological research has contributed to the development of established interviewing protocols that support the retrieval of information from witnesses (e.g. the Cognitive Interview [CI]; Fisher & Geiselman, 1992), there is only a small body of research that has examined interviewing by police officers in frontline contexts. The programme of research conducted for the current thesis aimed to examine the role of person descriptions within the initial account and what this means for the investigative process. The thesis also aimed to test whether components of existing interview techniques (e.g. Self-Administered Interview [SAI], Self-Generated Cues [SGC]) can be adapted to improve the diagnostic value of person descriptions.

After viewing a crime, witnesses are likely to be asked by frontline police officers to describe the perpetrator. Often, in the course of these interactions, witnesses are prompted by the police officer to consider the perpetrator's ethnicity, age, weight, height, accent, etc. (Association of Chief Police Officers, 2016), but beyond these prompts, there is no specific protocol for obtaining a person description. For the purpose of the current thesis, *person description information* is the term used to describe individuating information obtained from

a witness about the characteristics of a person-of-interest. Whilst a witness is expected to provide a full description of the perpetrator, the use of ineffective communication practices by frontline police officers can restrict reporting (i.e. a police officer may employ inappropriate interview techniques that fail to support the retrieval of information from a witness's memory). This often leads to the resulting descriptions lacking diagnostic value and consisting of only a few references to the physical appearance (e.g. gender, ethnicity, height) and clothing of the perpetrator (Douglass et al., 2013; Gabbert & Brown, 2015).

There is now a wealth of evidence-based techniques that support the reporting of detailed and reliable information from witnesses (e.g. the CI; Geiselman et al., 1984, and the SAI; Gabbert et al., 2009). However, there is no specific technique aimed solely at eliciting person description information, and few techniques have been designed for frontline police officers interviewing at the scene of an incident. It is important to note here that frontline interviews are not the same as 'formal interviews' (where the majority of academic research has been conducted). Frontline interviews typically fall short of the behaviours, techniques and cautions that we expect from formal interview settings (e.g. Dando et al., 2009). This is compounded by the fact that frontline officers have many activities to engage in. For example, frontline police officers need to prioritise managing witnesses at the scene (e.g. separating fights, consoling victims). The management of witnesses can therefore restrict comprehensive interviews from taking place for days, or even weeks, after the event (Smith & Milne, 2018). Very little research has been carried out to examine frontline interviewing, partly because, to date, no formal records of the conversations between police officers and witnesses have been retained (i.e. a police officer might make pocketbook notes about what the witness said, but not record the questions asked to generate the response). Hence, the current thesis involved examining real-world frontline police interviews to explore how

frontline police officers elicit person descriptions and the extent to which witnesses report person descriptions.

Advances in digital recording technology are transforming modern policing (Jennings et al., 2014). Cameras mounted on police uniforms (referred to as Body Worn Video [BWV]) capture and record activity, and create a permanent digital recording of police encounters with members of the public. BWV has been credited with creating greater transparency, efficiency, and effectiveness of police conduct all over the world (e.g. America, the United Kingdom, Australia; Drover & Ariel, 2015). Early research observed that the use of BWV is associated with a reduced number of complaints against the police by both witnesses and fellow officers (Katz et al., 2014; Ready & Young, 2015), and a reduced use of force by police officers during interactions with members of the public (Ariel et al., 2015). More recent research has started to use BWV as a tool to examine the communication skills of frontline police officers and how information is obtained from witnesses. To date, results suggest that frontline police officers often fail to deploy strategies to build rapport with witnesses, and use inappropriate questions (e.g. closed or leading questions) to gather information (Dalton et al., 2021; Gabbert et al., 2016).

Following the introduction of BWV, it is now possible to examine in full, the interactions that take place between frontline police officers and witnesses at the scene of an incident. Capitalising on this technological advance, the research presented in the current thesis aimed to (i) examine *for the first time* whether frontline police officers use effective interviewing practices and support memory retrieval when eliciting person descriptions; (ii) explore the phenomenon of own-race bias by witnesses reporting person descriptions; and (iii) test targeted retrieval techniques with the goal of improving the diagnostic value of person descriptions.

Obtaining reliable person descriptions about single or multiple perpetrators are necessary across a wide range of crime types. Person descriptions can be of persons known (e.g. domestic incident where the perpetrator knows the witness, but leaves the scene) or unknown (e.g. an assault by a stranger) to the witness. Person descriptions are often required from witnesses who have seen complex events, such as terrorist attacks, physical assaults involving multiple perpetrators, or sexual assaults. The reporting of person descriptions from witnesses about complex events are further influenced by a variety of factors, including the internal state of the witness (e.g. intoxication, substance misuse, stress; Deffenbacher et al., 2004; Yullie & Tollestrup, 1990). Person descriptions can also be affected by the demographics (e.g. sex, age, race) of both the witness and the perpetrator (Ellis et al., 1975; Searcy et al., 2001; Sporer, 1992).

With increased globalisation, we now live in multi-cultural communities; thus, the possibility for witnesses to fall victim to a crime where the perpetrator is of a different-race is more likely. Research has already shown that witnesses are poor at providing person descriptions (even when the witness and the perpetrator are of the same race; Sporer, 2007); therefore, it is important to explore how reliable witnesses are at providing the police with descriptions when the race of the witness and suspect is different. Whilst there is a wealth of research examining own-race bias (people are more accurate at remembering faces from their own-race compared to faces of a different-race; Malpass & Kravitz, 1969; Meissner & Brigham, 2001) in recognition memory (Meissner & Brigham, 2001; Sporer, 2001), very little is known about how people describe faces from their own-race compared to faces of another race. Research has demonstrated that individuals attend to features deemed relevant to faces of their own-race and further attempt to apply this encoding scheme inappropriately when examining faces of another race (Ellis et al., 1975; Fallshore & Schooler, 1995; Shepherd & Deregowski, 1981).

Research into the elicitation of person description information is limited with very few studies being conducted and findings being inconsistent. As a result, numerous academics in the field have called for more empirical research to be conducted into the factors that impact the reporting of person description information (e.g. Lindsay, 2012; Meissner et al., 2007). Across four studies, the current thesis examined how real-world person descriptions are elicited by frontline police officers (Study 1; see Chapter 2), explored whether there is an own-race bias when reporting person descriptions (Study 2; see Chapter 3), and tested the use of a cognitive mnemonic to facilitate the recall of person description information for both single perpetrator (Study 3, Experiment 1; see Chapter 4) and multiple perpetrator (Study 3, Experiment 2; see Chapter 4) events. The purpose of this first chapter is to introduce the importance of person descriptions to the criminal investigation and review the theoretical literature on memory processes for reporting witness events. It will start by discussing the constraints of memory when providing person descriptions.

Remembering person details: A challenge for memory

Theories of memory typically distinguish between working memory, which holds memories for a brief period, and long-term memory, which maintains memories for longer periods. The latter comprises semantic memory and episodic memory, with semantic memory referring to general world knowledge (e.g. facts) and episodic memory referring to the recollection of unique personal experiences (Tulving, 1972, 2002). According to Tulving (1985), episodic memory is a specialised category of memory relating to the conscious recall of personally experienced events. In this sense, episodic memory allows us to become witnesses to the past and give testimony about it (Mahr & Csibra, 2020).

However, memory is fallible (Frenda et al., 2011). Fallible not just because we forget information, but fallible because we also “remember” things that did not happen. This can be a simple case of misremembering (e.g. that we had orange juice at lunch, when in fact we had

apple juice), or more seriously, misremembering entire events that never happened. We do this not because memory is fundamentally flawed but because it is reconstructive. That is, our memory for events is not a verbatim playback of what happened (Loftus, 1979). Rather, it is a reconstruction based on the retrieval of stored remnants of the original experience that have persisted in memory, along with schema-driven information that serves to make the memory coherent and fill in the blanks (Ahn et al., 1992; Brewer & Nakamura, 1984).

Memory is malleable. That is, once created, a memory is not necessarily a fixed stable entity (Wixted et al., 2018). When a memory is reactivated, whether it be to discuss an event with friends or to provide an account of an event when questioned by the police, the mere reactivating of the memory alters the memory. Post-event information (e.g. the misinformation effect; Pickrell et al., 2016) can also reduce the accuracy of a witness's memory. The misinformation effect has been well-established in the academic literature and refers to the tendency for post-event misleading information to reduce one's memory accuracy for the original event. For example, participants in a memory experiment can be led to believe that they saw a 'stop' sign when they actually saw a 'yield' sign (Loftus et al., 1978) or that they saw broken glass even though broken glass was not present (Loftus & Palmer, 1974). Despite the inherent reconstructive nature of memory, memory often serves as the centrepiece of the criminal investigation (Howe et al., 2017). That is, witnesses attempt to provide person descriptions about the perpetrator(s) from memory, and therefore police officers should follow best-practice interview techniques in order to support memory retrieval from witnesses.

Witnesses are poor at providing person descriptions and this may be because of a mismatch between the process of recognising a face and the process of recalling the features of the face. Faces are encoded in memory and recognised using a holistic process (Tanaka & Sengco, 1997). Thus, each feature of a face is strongly linked with other facial features.

However, the activity of providing a description of a perpetrator consists of a sequential process (Kask & Bull, 2009), which may be challenging given the holistic representation stored in memory. For example, descriptions of faces tend to emphasise the individual features of the face (e.g. thin lips, bushy eyebrows; Wells & Turtle, 1987). Consequently, these conflicting processes (i.e. holistic vs. sequential) may explain why people are poor at describing faces.

Reporting of person descriptions

Despite the importance of person descriptions in the investigative process, research has shown that descriptions provided by witnesses are often limited in terms of the quantity and quality of descriptors provided and, as a consequence, can frequently implicate multiple people within the vicinity of the crime (Demarchi & Py, 2009). Kebbell and Milne (1998) surveyed 159 police officers and found that 76% of police officers agreed with the statement that witnesses 'rarely' or 'never' provide as many person details as they would like. Brown et al. (2008) found an even higher percentage of police officers (90% of 72 police officers) agreed with the statement that witnesses 'rarely' or 'never' provide as many person details as they would like.

Archival data from real-world witness statements support the view held by police officers. For example, Kuehn (1974) analysed 100 witness statements based on the accounts of a single witness. The statements were taken from a random sample of criminal cases (e.g. homicides, rapes, and armed robberies). Results showed that witnesses provided on average 7.2 physical details about the perpetrator. A number of other archive-based studies have found that witnesses typically provide between 3.9 and 9.7 physical characteristic details when describing a single perpetrator (see Table 1). Some research suggests that witnesses tend to provide more physical details than descriptions of clothing (Sporer, 1996; van Koppen & Lochun, 1997). For physical descriptors, Demarchi and Py (2009) reported that witnesses

tended to report more general features such as gender, height and ethnicity compared to more specific details pertaining to facial features such as the eyes, ears and mouth. Interestingly, Kuehn (1974) found that fewer than 25% of the witnesses reported facial features. In other studies (e.g. Demarchi, 2003; Fahsing et al., 2004; Sporer, 1996; van Koppen & Luchun, 1997), the majority of facial descriptors referred to the upper half of the face, in particular the hair and eye colour, but other facial characteristics (e.g. chin, nose, mouth) were rarely mentioned.

Table 1

Frequency and accuracy of descriptions from archive studies (adapted from Demarchi & Py, 2009)

Studies	**Number of witness statements**	**Mean number of details reported**	**Accuracy**
Kuehn (1974)	100	7.2	-
Lindsay, Martin, & Webber (1994)	205	3.9	-
Tollestrup, Turtle, & Yuille (1994)	164	6.9	-
Sporer (1996)	172	9.7	-
van Koppen & Lochun (1997)	2,299	8.4	76
Demarchi (2003)	216	8.6	78
Fashing, Ask, & Granhag (2004)	250	9.4	87

In terms of accuracy of person descriptions in real-world settings, Yuille and Cutshall (1986) found an average accuracy rate of 75.6%. Similar results have been found in other studies (see Table 1). When examining the contents of person descriptions in more detail, the descriptions about the perpetrators gender and ethnicity were almost always accurate,

whereas facial and other characteristics were less accurate across studies (see Table 2). However, there are limitations to drawing conclusions from such data. For example, van Koppen and Lochun (1997) compared the person descriptions given by witnesses to the descriptions of the convicted offenders as recorded by police officers at the time of the arrest. However, this does not take into account that the appearance of the perpetrator may have changed over time, nor does it confirm whether the perpetrator was in fact the guilty party. It should be noted that the person descriptions obtained in the previous studies (see Table 2) were elicited using standard police interview protocols, and thus techniques to support retrieval (e.g. the CI; Geiselman et al., 1984) were not used to obtain the descriptions.

In summary, it seems witnesses have a general impression of the perpetrator, but do not report discrete features (e.g. nose, mouth, etc.). Thus, witnesses' verbal descriptions of criminals are typically too few in number and too vague to identify a perpetrator.

Table 2

Occurrence and accuracy of descriptors from archive studies (Demarchi & Py, 2009)

Attributes	*van Koppen & Lochun (1997) (N = 2,299)*		*Demarchi (2003) (N = 216)*		*Fahsing et al. (2004) (N = 250)*	
	Occurrence (%)	*Accuracy (0%)*	*Occurrence (%)*	*Accuracy (0%)*	*Occurrence (%)*	*Accuracy (0%)*
Gender	95	100	100	100	100	100
Height	70	97	83	58	91	78
Ethnicity	64	80	85	97	50	92
Build	48	51	65	74	84	90
Age	55	98	86	43	62	60
Hair	-	-	-	-	18	97
Hair colour	36	82	43	77	-	-
Hair length	-	-	51	84	-	-
Hair type	34	33	0	-	-	-
Face shape	12	100	0	-	10	85
Eye colour	5	75	11	78	4	100
Nose	3	37	2	-	-	-
Facial hair	10	3	0	-	2	100
Mouth	2	39	1	-	2	20
Accent	31	32	-	-	47	99

Laboratory studies examining human memory (e.g. verbal recall, construction of facial composites, identification parades) have found that witnesses provided on average 10 physical details about the perpetrator (see Table 3). Whilst the average number of person descriptions is slightly higher in laboratory studies than real-world studies, participants in laboratory studies, and under good quality encoding conditions, can still provide poor person descriptions. For example, Sporer (2007) found participants only reported on average 4.46 person details even when they gave their descriptions immediately after viewing the target in

a filmed event. In terms of accuracy, laboratory studies have reported an average accuracy rate of 83.8% (see Table 3).

In terms of the content of person descriptions in the laboratory, Lindsay et al. (1994) found that height was reported by 86% of participants, follow by build (51%), gender (46%), age (45%), ethnicity (43%) and weight (22%). In terms of facial features, upper features (e.g. hair and eyes) were reported more than lower features (e.g. chin, lips, mouth; Ellis et al., 1980; Laughery et al., 1986). Lindsay et al. (1994) also compared the frequency of report of various features by witnesses to real-world crimes (using newspaper accounts) versus staged crimes, and found that data obtained in the laboratory and field differ in meaningful ways. Although there is little evidence that real-world descriptions are better than those obtained in research settings, these data suggest that different features are described at different rates between real-world and laboratory studies. For example, in person descriptions obtained in the field, weight was rarely mentioned, and gender nearly always mentioned, but this was not the case in the laboratory.

Table 3

Frequency and accuracy of descriptions from laboratory studies (adapted from Demarchi & Py, 2009.

Studies	Mean number of details reported	Accuracy
Wells, Rydell, & Seelau (1993)	5.1	-
Lindsay, Martin, & Weber (1994)	7.4	-
Geiselman et al. (2000)	5.3	87
Tunnicliff & Clark (2000)	6.3	-
Meissner, Brigham, & Kelly (2001)	9.0	90
Brown & Lloyd-Jones (2003)	15.9	71
Pozzulo & Warren (2003)	9.9	87

An issue with research on person descriptions for eyewitness events is the lack of an agreed definition. Meissner et al. (2008) found that the methods of eliciting person descriptions vary between studies in terms of the types of physical characteristics they include. For example, some researchers focused specifically on asking for facial features and hair (e.g. Brown & Lloyd-Jones, 2002, 2003), whereas others included assessments of the whole body such as height, build, weight, posture and clothing (e.g. van Koppen & Lochun, 1997). This variability in how descriptors are obtained from witnesses has added complexity when researchers have attempted to compare measures of description quality across different studies (Meissner et al., 2008).

In summary, whilst person descriptions in laboratory studies appear to be more detailed than those reported by real witnesses, they remain insufficient in terms of the diagnostic value needed to help police officers to identify the perpetrator(s).

Why do person descriptions lack diagnostic value?

Diagnostic value refers to the quantity and quality of the person descriptions provided by witnesses (e.g. a person description with high diagnostic value will include numerous references to the physical appearance and clothing of the perpetrator). To have 'diagnostic' value, the person description needs to enable an investigator to discriminate accurately in order to identify the correct perpetrator. Incomplete person descriptions lack diagnostic value and are problematic for the investigative process as they provide few clues for the police to find the perpetrator. There are numerous reasons why person descriptions lack diagnostic value, such as strategic monitoring of memory, insufficient instructions provided to witnesses, and inadequate interviewing techniques.

Strategic monitoring of memory

One reason why person descriptions lack diagnostic value is because they lack fine-grain details about the appearance of the perpetrator. Witnesses do not provide more detail because they are attempting to regulate or balance the informativeness and accuracy of the information that they provide (Sauer & Hope, 2016). One method for balancing accuracy and informativeness is to regulate the level of detail (grain size) they report (Goldsmith et al., 2002; Sauer & Hope, 2016). Fine-grain responses are more informative but less likely to be accurate compared with coarse-grain details. When a premium is placed on accurate reporting (e.g. when providing person descriptions to the police), it is likely that witnesses use a higher threshold for accurate reporting and may screen out information that is likely to be incorrect.

Thus, witnesses may be strategically regulating their memory to enhance accuracy (Koriat & Goldsmith, 1994, 1996).

Witnesses need to adopt appropriate strategies to enhance their memory for the task at hand (Bjorklund et al., 2009). Witnesses control how much information they report based on how confident they are about the accuracy of their recollection. In addition, witnesses regulate their responses by adjusting the precision of the reported information (control over grain size; Ackerman & Goldsmith, 2008). According to the satisficing model of the minimum-confidence criterion (Goldsmith et al., 2002), the interviewee will start by assessing fine-grain answers which they will volunteer (i.e. report) if it is likely to be correct (e.g. he had piercing blue eyes), otherwise they will provide coarse-grain answers (e.g. he had dark eyes) in order to preserve accuracy. In the immediate aftermath of a crime when searches are often initiated to locate a perpetrator, more informative (fine-grain) details are preferable. However, if an interviewee is unsure of their recollection, fine-grain details are less likely to be correct (cf. coarse-grain details; Koriat & Goldsmith, 1996; Weber & Brewer, 2008). When a premium is placed on accurate reporting (e.g. when providing person descriptions for the police), it is likely that witnesses attempt to enhance the accuracy of what they say by screening out information that is likely to be incorrect. Given that witnesses are unlikely to have extensively encoded the face, this strategic regulation may result in minimal detail about the perpetrators face being provided by the witness.

McCallum et al. (2016), during an eyewitness experiment, found that participants reported more fine-grain than coarse-grain information. Observing a similar prioritisation of fine-grain details, Brewer et al. (2018) suggested that while witnesses to a crime may be extremely confident about some coarse-grain details they recalled, they may also believe the police would find such details uninformative (e.g. police already know the information, or details are too general). The Dual-Criterion Model suggests that informativeness mediates

reporting, in that even if coarse-grain details are more likely to be accurate (though imprecise; Evans & Fisher, 2011), they may be withheld if assessed by the interviewee as not sufficiently informative. This potential assumption of police knowledge that witnesses hold is problematic when gathering person description information from witnesses because in the immediate search for the perpetrator, coarse-grain details could be invaluable to narrowing down the search (e.g. knowing someone is wearing a dark rather than a light jacket could inform a CCTV search for a target fleeing the scene).

The ability of a witness to regulate the quality and quantity of their reporting is impacted by the type of retrieval technique employed. Koriat and Goldsmith (1996) examined how participants responded to a general knowledge test when using free-report and forced-report conditions. Participants first gave responses to each question and had to rate how confident they were that the answer was correct. Later, participants could decide whether to answer a question or not. During this stage, participants were informed that whilst there was no penalty for not answering a question, they would receive monetary incentives for accurate responses and incur penalties for incorrect responses. The results revealed that when free to choose whether to report an answer or not, participants answered questions that they were certain of, with participants who received high penalties for incorrect answers exercising a stricter criterion at reporting (cf. participants with moderate penalties). Participants in the free-report condition were also able to monitor the accuracy of their responses by volunteering answers that exceeded a pre-set criterion of confidence that the answer was correct (*satisficing model;* Ackerman & Goldsmith, 2008). Finally, participants were more effective at monitoring their accuracy when they were able to control whether to volunteer or withhold an answer (*control of report option*). Furthermore, Scoboria and Fisico (2013) found that interviewees were able to effectively monitor the accuracy of their responses using the phrase ‘I don’t know’.

These findings have a number of important implications for frontline police interviews. Frontline police officers should utilise free-report formats as these are more likely to facilitate effective monitoring of accuracy, due to witnesses being in control of what information they volunteer. In addition, officers should not pressure or demand a witness to report information. Instead, officers should encourage witnesses not to guess information that they are unsure of, in order to promote accurate reporting (Memon et al., 2010). Thus, one of the objectives of the current thesis was to examine how person descriptions are elicited with a focus on the reporting format used by frontline police officers.

Instructions given to witnesses

Person descriptions have been shown to be dependent on the nature of the instructions given to witnesses (Demarchi & Py, 2009). Poor instructions can limit or caution a witness against providing detailed and accurate person descriptions. For instance, Meissner et al. (2001) found that providing a warning instruction 'be sure to report only those details that you are confident of, and do not attempt to guess any particular feature' significantly increased the overall description accuracy (cf. standard and forced recall). However, it should be noted that even when the warning instruction was used, the accuracy of person descriptions was still poor. Thus, the question of how person descriptions can be improved remains.

Inadequate interviewing techniques to support the provision of detailed person descriptions

Securing detailed and accurate accounts about what actually happened from witnesses is critical for criminal investigations (e.g. locating the perpetrator who may still be at large; Hope & Gabbert, 2019). In the United Kingdom, when collecting person descriptions, frontline police officers typically use a suspect description form (see Appendix E) that is completed by the police officer taking the witness's account (Centrex, 2004). This form is

designed to prompt the reporting of critical information from witnesses via a series of specific closed questions (e.g. prompts on height, build, and clothing). However, this form can encourage police officers to ask witnesses a series of short-answer closed questions, alongside other poor practices (e.g. asking witness-incompatible questions, interrupting witnesses; Fisher et al., 1987).

The use of ineffective interviewing practices and failure to support the retrieval of information from memory can restrict the reporting of information by witnesses (Evans et al., 2010; Vrij et al., 2014). This may explain why witnesses provide poor person descriptions, as police officers fail to provide adequate retrieval support for the task and fail to make clear what the task actually is (Fisher, 1995; Wright & Alison, 2004), or police officers adopt inadequate questioning strategies (Oxburgh et al., 2010). The current guidance on interviewing witnesses (Ministry of Justice, 2011) recognises that witnesses tend not to realise that the interviewer requires detailed descriptions of the perpetrator(s), and as a result, person descriptions tend to be short and incomplete.

There is limited field research examining police interviewing at the scene of an incident, and therefore little is known about the nature of interactions between frontline police officers and witnesses. The research that is available demonstrates that police officers have an overreliance on closed questions and leadings questions (Myklebust & Alison, 2000; Wright & Alison, 2004), which is problematic because (i) closed questions limit the amount of information reported by the witness (Oxburgh et al., 2010), and (ii) leading questions have a detrimental effect on the accuracy of witness memory (Loftus & Palmer, 1974; Eisen et al., 2002). In contrast, open-ended questions have been found to elicit more accurate and detailed information compared to closed questions (Myklebust & Bjorklund, 2006; Oxburgh et al., 2010). Hence, witnesses may be able to remember information about the appearance of the perpetrator(s) but are being asked inappropriate questions that inhibit the reporting of person

descriptions. It is not surprising that frontline police officers have been shown to rely on closed or leading questions as they themselves have reported feeling inadequately trained, ill equipped and under pressure when conducting frontline interviews (Dando et al., 2008). However, research has found that even trained interviewers often fail to implement best practices when interviewing witnesses (Soukara et al., 2009).

Frontline police officers need to ensure they are following interview guidelines that have been informed by best-practice. For example, to actively listen to witnesses (and to not interrupt witnesses when they provide their account of what happened; Beune et al., 2011). Fisher (1995) argues that interruptions discourage witnesses from providing extensive, detailed responses. If witnesses are constantly interrupted then the witness may interpret this as an indication that they only have a limited time to respond to the police officer, so may start to shorten their responses accordingly. This can further reduce the quantity of the person descriptions reported by the witness, which are already limited.

To avoid contaminating a witnesses memory of an event, police officers should be warned against providing witnesses with feedback during the interview as it can inflate their confidence (though this has mainly been documented in recognition tasks as opposed to recall tasks; Semmler et al., 2004). For example, when a witness is given explicit feedback, or even subtle cues, regarding their account, they are likely to adjust their confidence to reflect their belief that they were accurate (confirming feedback) or inaccurate (disconfirming feedback). Witness confidence can impact upon how the witness is perceived by the police and in court, with higher confidence being associated with honesty (Wells et al., 1979; Whitley & Greenberg, 1986). Frontline police officers therefore need to be aware that providing feedback can distort memory (Garven et al., 2000), and ultimately lead to inaccurate person descriptions.

Finally, witnesses are rarely warned against discussing their experiences with other witnesses and, furthermore, witnesses are sometimes interviewed in front of each other at the scene of a crime (Hope & Gabbert, 2019). Failing to separate witnesses can have significant implications for the criminal investigation, including the reliability of witness accounts. For example, co-witnesses can affect how one witness remembers the event and this can lead to memory conformity (Gabbert & Hope, 2013). Memory conformity refers to the formation of a memory about the incident that is acquired by a witness following the event (e.g. during discussions with other people; Gabbert et al., 2003). In other circumstances, one witness may report details learned from another in order to avoid the perceived social costs of disagreeing (referred to as social conformity; Cialdini & Goldstein, 2004). The problem with memory conformity and social conformity is that witnesses may go on to report details that were learnt through discussions with a co-witness because they have forgotten the source of the information (Zaragoza & Lane, 1994) or because they believe the new information to be accurate (Deutsch & Gerard, 1955). In dynamic incidents where there are multiple witnesses and importantly multiple police officers, it is paramount that witnesses are separated to avoid contamination of memory (Gabbert et al., 2006; Ito et al., 2019).

The failure to use best practices in interviewing raises serious concerns about the ability of police officers to obtain complete and accurate information (including descriptions of the perpetrator) needed to conduct criminal investigations successfully and efficiently. Thus, person descriptions provided by witnesses are not inherently unreliable, but poor interviewing techniques can negatively affect the nature and content of witness accounts.

Using mnemonics to improve person descriptions

Despite person descriptions being a significant source of information in the identification of potential perpetrators, few studies have attempted to increase the quantity

and accuracy of descriptions. That being said, there are a number of evidence-based mnemonics that support the reporting of detailed and reliable information by witnesses, and some of these are already embedded in best-practice police interviewing (e.g. context reinstatement forms part of the CI and the SAI).

Cognitive Interview (CI)

Interview techniques that use cognitive mnemonics to provide retrieval support have been shown to be useful for increasing the completeness of information across the whole interview (e.g. the CI; Fisher & Geiselman, 1992). The CI proposes ways of enhancing retrieval of episodic information from memory and centres around two main principles: (i) a memory trace consists of several elements of related information, and (ii) there are several possible ways of retrieving an encoded event, so information that cannot be retrieved via one pathway need not be lost, but may be retrieved using an alternative pathway. The CI is embedded in best-practice police interviewing and incorporates several distinct mnemonics. One of these is the mental reinstatement of context (MRC) instruction. The key premise of MRC is that recall is facilitated when there is an overlap between the condition present at encoding and at retrieval (*encoding-specificity principle*; Thomson & Tulving, 1970; Tulving & Thompson, 1973). The MRC instruction therefore encourages a witness to recreate the psychological (e.g. their emotional state and their thoughts) and physical environment (e.g. their surroundings) that existed at the time of the to-be-remembered event via a set of pre-defined instructions (Memon et al., 2010). Research has consistently reported that the administration of the MRC instruction has elicited more correct information than a free recall condition (Dando et al., 2009; Davis et al., 2005; Memon & Bruce, 1985).

Geiselman et al. (1984) found that during the CI, information gained for person descriptions was often weak. However, there has been some success with certain aspects of

the CI, such as the *report all* and *context reinstatement* instructions (Boon & Noon, 1994; Clifford & George, 1996), but even then, results were still poor for person description recall. Removing inappropriate instructions (forming a modified version of the CI) has also been found to increase completeness. In studies that focused specifically on whether the CI improves perpetrator descriptions, more positive results were found. For instance, Finger and Pezdek (1999) removed the *change order* and *change perspective* instructions and reported that the modified CI was more effective than a standard interview. Py and Demarchi (2006) modified the CI and added a *holistic processing instruction* (Craik & Lockhart, 1972), which aimed to take into account the holistic character (e.g. not only the person's appearance but also other characteristics, such as intelligence and likeability). Through combining these techniques, the modified CI increased the quantity and quality of person descriptions (cf. standard police interview).

Although the CI has substantial empirical support and far-reaching application, Satin and Fisher (2019) identified limitations with the former studies including the omission of important components of the CI such as building rapport and instructing the eyewitness not to guess. Consequently, two conspicuous gaps in the literature have been identified. First, whether the CI can elicit a more complete, accurate description of the perpetrator (cf. standard police interview; Brown et al., 2008; Gabbert & Brown, 2015). Second, if so, whether the description provided by witnesses is useful to the police in identifying the perpetrator. To address these gaps in the literature, Satin and Fisher (2019) restricted person descriptions to include only descriptors of the physical appearance. College students (N = 67) witnessed a simulated robbery and were then interviewed using either the CI or a standard police interview to elicit a description of the robber. The results showed that the CI elicited almost three times as many descriptors, and at comparable levels of accuracy (cf. standard police interview). These findings suggest that the CI can be used to successfully elicit person

descriptions about a perpetrator. However, it is unlikely that frontline police officers would have the time or resources to carry out the CI (Dando et al., 2009) and therefore alternative retrieval techniques need to be considered.

Self-Administered interview (SAI)

There are evidence-based interview tools available to support the retrieval and reporting of detailed information that police officers can administer at the scene of an incident (e.g. the SAI; Gabbert et al., 2009). The SAI is based upon the same underlying theoretical principles as the CI and uses some of the same memory-enhancing components (e.g. report everything and context reinstatement) that are designed to elicit a written free-recall statement. In addition, the SAI aims to provide a high-quality statement by discouraging witnesses from guessing and by using non-leading questions and cues (Hope et al., 2011). Studies using the SAI have reported an increase in the quantity and accuracy of person description details (Gabbert et al., 2009; Krix et al., 2016), as well as overall details (for a review see Horry et al., 2021). It should be noted that research suggests that the SAI should be administered as soon as possible (to preserve witness memory), as recall accuracy decreases and post-event misinformation susceptibility increases when the SAI is administered more than 24 hours after an event (Paterson et al., 2015).

Self-Generated Cues (SGC)

Previous research on eyewitness memory has focused on interviewer-led cues to generate information (e.g. mental reinstatement of context; Dando et al., 2009). However, interviewers cannot know what the most salient information to the interviewee is. One solution is to ask the interviewee to provide the information themselves using the SGC mnemonic. SGC are salient details from a witnessed event that can be actively generated by the individual themselves to facilitate their memory of a target event (cf. interviewer-

generated cues and no cues; Kontogianni et al., 2018; Wheeler & Gabbert, 2017). When witnesses recall episodic information, stored traces are activated and these subsequently prompt related details. Thus, 'spreading activation' through an associative network (*activation theory*; Anderson, 1983). With every attempt at remembering details, the stronger the memory trace becomes. Research has shown that the stronger a memory, the more likely it is that it will be recalled at a later date and that it will activate associated memories (Anderson, 1983). Similarly, Anderson and Conway (1993) showed that, when asked to list event details using a free recall, participants first listed 'distinctive details', before then listing other details that were highly associated with those distinctive details. Thus, self-generation of distinctive cues can trigger related memories by tapping a common theme (Anderson & Conway, 1993). The use of SGC should trigger the retrieval of related event details while excluding unrelated details, thus satisfying both the encoding-specificity principle (Tulving & Thomson, 1973), and the principle of cue overload (Nairne, 2002).

For retrieval cues to effectively support recall, a number of key qualities have been suggested: (i) constructability, so that cues generated at encoding can be reliably reproduced at recall; (ii) there needs to be consistency between encoding and retrieval within a given context; (iii) strong associations are needed with the target and these need to be easily associated with newly learned information; and (iv) bidirectionality of association (i.e. the cue recalling target information and target information recalling the cue; Wheeler & Gabbert, 2017). In addition, retrieval cues need to be distinctive and discriminable so that it is possible to distinguish cues from one another and so that it is possible to differentiate the target memories associated with each (Tullis & Benjamin, 2015). If retrieval cues are not distinct or discriminable then the cues are likely to become associated with more information which reduces the effectiveness of the cue in promoting the recall of target information. Watkins and Watkins (1975) refer to this as *cue overload* which is problematic as it leads to lower and

less accurate recall. This is because the cue contains too many associative links (*the fan effect*; Anderson, 1983).

Recent research has focused on using SGC to examine whether they facilitate the retrieval of target memories in eyewitness contexts. Wheeler and Gabbert (2017) defined SGC as: *any detail salient to the individual themselves which facilitates more complete retrieval of a target.* Salient information is more easily recalled than non-salient information; thus, the SGC instruction allows the witness to run with cues that come to their mind rather than cues implemented by the interviewer.

Kontogianni et al. (2018) found that using SGC led to more correct details from mock witnesses (cf. other-generated cues and no cues). In terms of perpetrator descriptions, Wheeler et al. (2017) showed participants a live event, before they were asked to carry out a free recall. Participants were provided with three self-generated cue techniques: a timeline, keywords, and a concept map. They found that the use of SGC improved free recall (i.e. provided more details), without reducing accuracy. In addition, they found that participants in the keyword condition reported the most person description information (cf. a timeline and a concept map). Thus, it appears that self-generated cue techniques can aid retrieval of person description details which are important to an investigation. To this end, one objective of the current thesis was to further test whether the use of SGC increased the reporting of person description information.

Thesis Outline

This thesis presents four empirical studies that examine how person descriptions are elicited in real-world cases (Study 1) and within laboratory settings (Studies 2 and 3). Within the laboratory studies, I also explored whether an own-race bias exists when mock-witnesses describe a perpetrator, and tested the effectiveness of using a self-generated cue mnemonic

for improving the quantity and quality of person description during single and multiple perpetrator events. Beyond estimator factors (such as poor encoding conditions), there are several reasons why person descriptions may lack diagnostic value. However, for the purposes of the current research, I focused on strategic monitoring of memory and inadequate interviewing techniques.

Study 1 (see Chapter 2) explores how person descriptions are elicited by frontline police officers in the United Kingdom. The study involved the use of real-world body worn video footage to examine the questioning techniques employed by frontline officers in order to elicit person descriptions from witnesses.

Study 2 (see Chapter 3) examined whether witnesses demonstrated an own-race bias when reporting person descriptions. During this laboratory experiment, the person description component of the SAI was extended to include prompts and cues to facial features, physical characteristics and clothing. The study examined interviewees' reporting of person description information for own-race (cf. different-race) during a staged crime event in which a single perpetrator committed a theft. The study draws upon the research knowledge base of metacognitive monitoring and precision over reporting.

Study 3 (see Chapter 4) comprised two laboratory experiments that tested a theory-driven mnemonic (SGC) to facilitate participant reporting of person description information. In Experiment 1, participants witnessed a staged theft of a single perpetrator and in Experiment 2, participants witnessed a staged theft of multiple perpetrators, with participants in both experiments being tasked with providing a description of the perpetrator(s). Based on previous research on effective cued-recall retrieval for person descriptions, the study tested whether the use of self-generated cues (cf. free recall and person description cues) increased the quantity and quality of person description information reported by the participants.

In the general discussion (see Chapter 5), I begin by providing an overview of the key findings. I consider both the theoretical and applied implications of the results on memory retrieval and how person descriptions can be better elicited. Finally, I provide recommendations for enhancing the retrieval of person descriptions from witnesses, as well as discuss limitations of the current research, before offering suggestions for future research.

Thesis Format: PhD by compilation

The format of the current PhD thesis follows a compilation style, starting with a general introductory chapter, followed by individual chapters presenting each study in a journal article format (i.e. abstract, introduction, method, results and discussion). As a result, there is some repetition throughout the thesis because each chapter has been written so that it is independent of all other chapters, and more digestible for the reader. The current PhD thesis ends with a general discussion chapter that summarises the findings of each study and explains how the studies together have advanced the research knowledge base in the field of frontline policing.

This thesis is structured as a series of published or publishable articles in line with Christianson et al. (2015) in which PhD students are encouraged to focus on ensuring their PhD studies are ready for publication in order to strengthen their CV alongside the completion of their research. The choice to undertake a compilation thesis was made in conjunction with the supervisory team and the graduate school at the University of Portsmouth. It was deemed most appropriate for the chosen subject area of the PhD thesis due to the research consisting of a small number of separate, stand alone, studies suitable for journal articles. Consequently, each chapter has been reviewed by the supervisory team and peer reviewed by external reviewers, if published. In addition, this style of PhD thesis is perceived to work best for students who are supervised by experienced researchers with a

strong publication record (see work of Professor Milne and Professor Hope for extensive research outputs).

Chapter 2

‘He was just your typical average guy’ Examining how person descriptions are elicited by frontline police officers

Published as:

Dalton, G., Milne, R., Hope, L., Vernham, Z., & Nunan, J. (2021). ‘He was just your typical average guy’ Examining how person descriptions are elicited by frontline police officers. *Applied Cognitive Psychology, 35*(2), 517-525. doi: 10.1002/acp.3778.

Foreword

Chapter 2 presents a field study for which I coded body worn video footage to examine how frontline police officers elicit person descriptions from witnesses. My supervisors and I identified a gap in the literature recognising the need to explore frontline policing, specifically, the questions frontline officers ask to elicit person information from witnesses. Until the current study, there was no research on the topic of frontline questioning due to police officers only manually recording the responses that a witness provides and not the questions that the officers ask (i.e. the handwritten statement). The design of the study was developed through discussions with a senior police inspector who facilitated access to body worn video footage at a UK police station. This process required me to become a police volunteer and entailed six months of police vetting, confidentiality agreements, and ethical approval from my host University. I was responsible for designing the study (which was discussed with the supervisory team). I was then required to set up the study, access the data, view the data, code and transcribe the data, analyse the data, and prepare the study for publication. Dr Vernham assisted with statistical analyses and Dr Nunan completed the inter-rater reliability. All authors read drafts of the manuscript prior to journal submission.

Abstract

Person descriptions often lack the level of detail necessary to assist in the apprehension of a perpetrator. To date, it is not clear how person descriptions are obtained by frontline police officers. Worldwide, many police forces now use body worn video (BWV), which provides a unique opportunity to examine how frontline police officers gather person descriptions from witnesses. I examined how person descriptions (N = 207) were elicited by frontline police officers, with a particular focus on the types of questions used. BWV of 81 interactions involving 45 frontline police officers and 141 witnesses were analysed. Person descriptions were obtained using inappropriate questions 50.54% of the time, with leading questions being the most commonly used (44.84%). Appropriate questions (i.e. open questions) led to more information being provided (cf. inappropriate questions), including more fine- and coarse-grain details. Implications for the training of frontline police officers are discussed.

Keywords

Person descriptions, Investigative interviewing, Eyewitness memory, Body worn video, Police

Introduction

Frontline police officers attending the scene of a crime are usually tasked with obtaining descriptions about persons of interest (i.e. perpetrators) as quickly and as accurately as possible. Such descriptions are passed on to other law enforcement teams for the identification of potential perpetrators in the vicinity of the crime scene. Person descriptions are both prescriptive (i.e. look for someone who matches the description) and proscriptive (i.e. ignore those who do not fit the description). According to police officers, person descriptions are a significant source of information in the identification of potential perpetrators (Brown et al., 2008). Verbal descriptions of perpetrators are therefore a critical component of the preliminary investigation of a crime (Demarchi & Py, 2009; Meissner et al., 2007; Milne & Bull, 2006).

Despite the importance of person descriptions in the investigative process, research has shown that descriptions provided by witnesses or victims[1] often lack diagnostic value and can frequently implicate multiple people within the vicinity of the crime (Demarchi & Py, 2009). Indeed, Brown et al. (2008) found that 80.6% of police officers agreed with the statement that 'witnesses rarely provide as many person details as they would like' (p.537). Given that the purpose of a person description is to help someone recognise that an unknown individual walking down the street is the potential perpetrator, this lack of detail may hamper effective investigations and the apprehension of perpetrators. The aim of the current study was to explore the nature of real-world person descriptions and to examine the types of questions frontline police officers ask at the scene of the incident to obtain such descriptions.

Perpetrator descriptions provided by witnesses

[1] For the remainder of this manuscript, the term witness will be used to describe not only those who observe the crime but also those who fall victim to the crime.

Kuehn (1974) analysed 100 witness statements and found that witnesses provided accurate recall for actions and objects, but made numerous errors when reporting information about people. Similar findings have been reported across other studies (e.g. Fahsing et al., 2004; Sporer 1996). Research has also examined the frequency of different types of person descriptors provided by witnesses. For example, it has been shown that witnesses provide more physical details than descriptions of clothing (Sporer, 1996; van Koppen & Lochun, 1997). For physical descriptors, Demarchi and Py (2009) reported that witnesses tended to report more general features such as gender, height and ethnicity compared to more specific details pertaining to facial features such as the eyes, ears and mouth. Furthermore, Kuehn (1974) found that fewer than 25% of the witnesses reported facial features. A number of archive-based studies have found that witnesses typically provide between 3.9 and 9.7 physical characteristic details when describing a single perpetrator (Lindsay et al., 1994; Sporer, 1996).

Lindsay et al. (1994) compared the frequency of report of various features by witnesses to real-world crimes (using newspaper accounts) versus staged crimes and found that data obtained in the laboratory and field differs in meaningful ways. Although there is little evidence that real-world descriptions are better than those obtained in research settings, these data suggest that different features are described at different rates between real-world and laboratory descriptions. For example, in person descriptions obtained in the field, weight was rarely mentioned, and gender nearly always mentioned but this was not the case in research settings. In a survey by Brown et al. (2008), police officers reported that descriptions of perpetrators tended to consist of general characteristics (e.g. sex, age, race, height, hair colour and hair length) and less often contained information concerning facial features (e.g. eyes, nose and mouth). One reason why person descriptions lack diagnostic value may be that witnesses are reluctant to report more fine-grain details or it does not occur

to them that such information might be relevant. With regards to person descriptions, fine-grain detail is specific and detailed information, such as ‘light blue’ eyes or ‘long, dark brown, curly’ hair.

Witnesses tend to regulate the information they provide (Goldsmith et al., 2002; Sauer & Hope, 2016). In other words, when people remember an event, they do not necessarily report everything they know, but instead strategically regulate what they report (Koriat & Goldsmith, 1996). When a premium is placed on accurate reporting (e.g. when providing person descriptions for the police), it is likely that witnesses attempt to enhance the accuracy of what they say by screening out information that is likely to be incorrect. In other words, witnesses may be strategically regulating their memory to enhance accuracy (Koriat & Goldsmith, 1994, 1996). Given that witnesses are unlikely to have extensively encoded the face, this strategic regulation will result in minimal detail about the face being provided about the perpetrator.

In addition to fine-grain details, witnesses can also provide coarse-grain details (i.e. details that lack specific information, such as “he had dark eyes”). McCallum et al. (2016) found that participants reported more fine-grain than coarse-grain information when providing accounts of witnessed events. Observing a similar prioritisation of fine-grained details, Brewer et al. (2018) suggested that while witnesses to a crime may be extremely confident about some coarse-grain details they recalled, they may also believe the police would find such details uninformative (e.g. police already know the information or details are too general). Coarse-grain details are more likely to be accurate, but imprecise (Evans & Fisher, 2011), and since such details are not thought to be informative by witnesses, witnesses do not spontaneously report them. This potential assumption of police knowledge that witnesses hold is problematic when gathering person description information from witnesses because in the immediate search for the perpetrator, coarse-grain details could be

invaluable to narrowing down the search (e.g. knowing someone is wearing a dark rather than a light jacket could inform a CCTV search for a target fleeing the scene).

Another reason why witnesses provide poor person descriptions may be that police officers fail to provide adequate retrieval support for the task (Fisher, 1995; Wright & Alison, 2004) or adopt inadequate questioning strategies (Oxburgh et al., 2010). When collecting person descriptions, frontline police officers typically use a suspect description form (Centrex, 2004). This form is designed to prompt the reporting of critical information from witnesses via a series of specific closed questions (e.g. questions on height, build, and clothing). However, it could be argued that the suspect description form is problematic because it is not normative for witnesses to provide descriptions of people at the required level of detail, and they may not necessarily have the language to do it with precision.

Unfortunately, there is limited field research examining police interviewing at the scene of an incident. Thus, little is known about the interactions occurring between witnesses and frontline police officers. There are many competing priorities in the frontline contexts, including securing the scene and maintaining public order. When interviewing witnesses in frontline contexts, at a basic level, is it important for frontline police officers to actively listen to witnesses (and to not interrupt witnesses when they provide their account of what happened; Beune et al., 2011; Fisher, 1995), to avoid providing feedback (Semmler et al., 2004), and to manage multiple witnesses (Hope & Gabbert, 2019). In dynamic incidents where there are multiple witnesses and multiple police officers, it is paramount that the above steps are taken to avoid contamination of memory (Gabbert et al., 2006; Ito et al., 2019).

The few studies that have been conducted in this context have noted shortcomings in the way in which police officers interview witnesses (Brown et al., 2008), including an over reliance on the use of closed questions (Myklebust & Alison, 2000; Wright & Alison, 2004)

and the use of leading/suggestive questions (Wright & Alison, 2004). This is problematic as closed questions limit the amount of information likely to be reported (Oxburgh et al., 2010) and leading/suggestive questions have a detrimental effect on the accuracy of witness memory (Loftus & Palmer, 1974; Eisen et al., 2002). The issue for frontline police officers gathering person descriptions is that research has shown that if a witness encounters misinformation prior to a detailed statement being taken then this can increase the chance of misinformation being recalled in subsequent retrieval attempts (Frenda et al., 2011).

Research has shown that asking open-ended questions elicits more accurate and detailed information compared to specific closed questions (Myklebust & Bjorklund, 2006; Oxburgh et al., 2010). Therefore, it is possible that witnesses can remember information about the appearance of the perpetrator, but are being asked unhelpful or other inappropriate questions that inhibit the retrieval or reporting of this critical person description information. It is worth bearing in mind that frontline police officers typically have the least amount of investigative interview training, but it is these officers who are responsible for carrying out the majority of interviews at the initial stages of the investigation (Dando et al., 2009). Indeed, frontline police officers have reported feeling inadequately trained, ill-equipped and under pressure when conducting frontline interviews (Dando et al., 2008). Thus, person descriptions provided by witnesses are not inherently unreliable, but poor interviewing techniques can negatively affect the nature and content of witness accounts.

Capturing person descriptions via body worn video

Advances in digital recording technology are transforming modern policing (Jennings et al., 2014). Cameras mounted on police uniforms (referred to as Body Worn Video [BWV]) have been credited with creating greater transparency, efficiency, and effectiveness of police conduct all over the world (e.g. America, the United Kingdom, and Australia; Drover &

Ariel, 2015). Early research observed that the use of BWV is associated with a reduced number of complaints against the police by both witnesses and fellow officers (Katz et al., 2014), and reduced use of force by police officers during interactions with members of the public (Ariel et al., 2015). More recent research has started to use BWV as a tool to examine the communication skills of frontline police officers and how information is elicited from witnesses. To date, results suggest that frontline police officers often fail to deploy strategies to build rapport with witnesses, and use inappropriate questions (e.g. closed or leading questions) to gather information (Gabbert et al., 2016).

The current study

The aim of this exploratory study was to examine BWV of real-life interviews taken from frontline police officers interactions with witnesses where a description of the perpetrator(s) is elicited. With this footage, and for the first time since the introduction of BWV, I had a unique opportunity to examine the interactions occurring between frontline police officers and witnesses, and to evaluate the types of questions frontline police officers ask witnesses when gathering information about the perpetrator(s). To date, it has not been possible to examine such interactions as, previously, police officers recorded the responses provided by the witnesses in a written statement but did not record the specific questions used to elicit such information. Based on previous research suggesting that (i) police officers ask inappropriate questions (Wright & Alison, 2004); (ii) witnesses provide more fine-grain details than coarse-grain details (e.g. McCallum et al., 2016); and (iii) witnesses are poor at describing faces (e.g. Demarchi & Py, 2009), I expected that open questions would lead to the provision of more information by witnesses than other question types and that, overall, appropriate question types (e.g. open, specific closed, appropriate yes/no) would lead to the provision of more information than inappropriate question types (e.g. leading, forced choice,

multiple). I also expected that descriptions provided by witnesses would include more fine-grain details than coarse-grain details.

Method

Data

Ethical approval was obtained from the Faculty of Humanities and Social Sciences ethics committee at the University of Portsmouth (see Appendix C). The BWV footage available for analysis was provided by a UK police force, spanned a 20-month period, and included 2,095 recordings drawn from officer interactions at the frontline with witnesses. From this corpus of recordings, 95 were identified as footage that contained interactions regarding the appearance of potential perpetrators. This footage comprised 81 separate incidents including allegations of assault (N = 31), domestic incidents (N = 17), theft (N = 16), and sexual offences (N = 10). Interviews were conducted by a total of 45 different frontline police officers who were captured on BWV interviewing witnesses to gather a verbal description of alleged perpetrators. Across the 81 incidents, police officers interviewed 141 witnesses (both children and adults[2]), which led to the description of 207 potential perpetrators (note: at a scene there can be multiple interactions by police with witnesses who all saw the same perpetrator; hence, the 207 potential perpetrators do not represent the total number of perpetrators, as one perpetrator can be independently described multiple times).

The criteria for selecting an incident involving a person description from the available footage was that the footage should involve an interaction between police and witnesses at a frontline incident captured on BWV, where a police officer asked for a description of a

[2] It is not possible to say with confidence how many child witnesses were involved in the body worn video interactions, because police officers do not ask for age details during the frontline interview and I am reluctant to rely on the subjective judgements of coders to determine the age of the witness. However, if appearance alone might be considered an indicator, then the number of child witnesses were very low within the sample with the youngest witness potentially being in their teenage years.

perpetrator. BWV footage was viewed in a police station using a computer programme called Digital Evidence Management Software (DEMS). Footage that was marked evidential (note: footage marked non-evidential is automatically deleted after 31 days) was viewed and footage that contained a description of a perpetrator was tagged. The segment of the body worn footage containing person descriptions was then transcribed verbatim and coded using a coding scheme developed by the research team.

Coding

Both coders were blind to the true purpose of the current study. A Coding Manual was created for coders to follow. As part of coding training, one transcript was then selected at random and was coded by each coder independently to make sure that (i) each coder fully understood the coding scheme and (ii) the coding was well calibrated across coders. This training exercise confirmed that coders were able to follow the manual appropriately.

Question type

Based on previous work by Shepherd and Griffiths (2013), questions posed by frontline police officers to witnesses were categorised as either appropriate or inappropriate (see Table 4). To establish inter-rater reliability, a selection of 24 recordings consisting of 173 questions were coded by an independent researcher for question type. Cohen's Kappa showed there was strong agreement (see Altman, 1999) between the two coders, $_K$=.96, 95% CI [.93, .99], $p < .001$.

Interruptions

An interruption was coded when the frontline police officer(s) spoke over the witness (i.e. an interruption was not coded if the police officer was spoken over by the witness, or if the witness was spoken over by another witness).

Witness separation

The number of witnesses versus the number of police officers was recorded at the scene. If there was an opportunity for witnesses to be interviewed individually by frontline police officers but this did not occur it was coded as a witness separation error.

Providing feedback

Feedback was coded when the frontline police officer(s) (in)validated the description of the perpetrator by the witness (e.g. “Yes you’re right, he is tall”).

Table 4

Type of appropriate vs. inappropriate question

Question type	**Operational definitions** (Shepherd & Griffiths, 2013) **and examples**
Appropriate	
Open	An unstructured question in which the answer cannot be answered with a yes/no response; the answer is not suggested; and requires developed thought. For example: *'Tell me, what did he look like?'*
Specific closed	Specifying what precise information is required, allowing witness to generate a response. Typically begin with WH. For example: *'You said he had long hair, what colour was his hair?'*
Appropriate yes/no	A question where the expected answer is either a yes or a no. Typically only one answer is acceptable (yes or no). Used at the conclusion of a topic where open and probing questions have been exhausted. Appropriateness is based on the context. For example: *'Did the man have any other distinctive features other than the ones you described?*
Inappropriate	
Leading	Prompting or encouraging the witness to a desired or assumed response. For example: *'He is 5'8, yeah?'*
Multiple	Several questions are asked at once, without giving the witness a chance to respond to the first question. For example: '*What was he wearing and how tall was he?'*
Forced choice	Specifying the precise information that is required in which the witness chooses a response option that indicates a definitive option. These questions eliminate the response of "don't know". For example: *'Was he black or white?'*
Inappropriate yes/no	A question where the expected answer is either a yes or a no. Typically only one answer is acceptable (yes or no). Used at the wrong point in the interview. Inappropriateness is based on the context. For example: *'Did the man have a tattoo?'*

Grain size

Following Lindsay et al. (1994), the person descriptions provided by witnesses were coded to determine what descriptors witnesses commonly provided when they were asked by the police to describe the potential perpetrator. To assess the level of detail provided in descriptions, the grain size of reported information was coded as either fine- or coarse- grain. Specifically, similar to coding reported by Weber and Brewer (2008), and, Sauer and Hope (2016), responses including numerical values were coded as fine-grain if they included up to three possible values (e.g. aged between 25 and 27 years). However, a response was coded as coarse-grain if it included four or more possible values (e.g. aged 20 to 30 years). For categorical data, pre-determined parameters defined when responses were considered fine-grain (e.g. white British male) and when responses met the permitted range for coarse-grain (e.g. medium build). Vague responses (i.e. 'about my height') were not coded. To establish inter-rater reliability, a selection of 24 recordings consisting of 229 grain size details were coded by an independent researcher. Cohen's Kappa showed there was strong agreement between the two coders, $_K$=.89, 95% CI [.82, .95], $p < .001$.

Results

Across 95 recordings (81 separate incidents), frontline police officers asked 556 questions to elicit person descriptions from witnesses. There was variation in the number of questions asked per recorded incident, ranging from 1 to 22 (M = 5.85, SD = 4.50). As shown in Table 5, the most commonly asked question type was Specific Closed (e.g. what was the colour of the jacket?) and the least commonly asked question was Appropriate Yes/No. Of all the questions asked, 49.46% (N_{total} = 275) were classified as appropriate and 50.54% (N_{total} = 281) were classified as inappropriate. Of the 281 inappropriate questions asked, 44.84%

(N_{total} = 126) were leading (see Table 5). An example of a leading question asked by an officer was: “All IC1 [Native White British], all white, all white skinned?”.

Table 5

Frequency for the question types asked by frontline police officers

Type of question	Frequency (%) (Total questions = 556)	Categorisation
Specific closed	174 (31%)	Appropriate
Leading	126 (23%)	Inappropriate
Open	76 (14%)	Appropriate
Multiple	65 (12%)	Inappropriate
Forced choice	47 (8%)	Inappropriate
Inappropriate yes/no	43 (7%)	Inappropriate
Appropriate yes/no	25 (5%)	Appropriate

When gathering person descriptions from witnesses, frontline police officers interrupted witnesses 182 times (M = 1.92, SD = 1.79). There was again sizeable variation in the number of times officers interrupted, ranging from 0 to 9. An example of an interruption taken from the footage is:

Witness: I can tell you what the man in the back of the car looked like as well…

[Interrupted by police officer]

Police officer: Is he white?

Frontline police officers did not separate witnesses before asking for a person description in 28 out of the 95 recordings, with between two and five witnesses being interviewed collectively. For 14 of these recorded incidents involving multiple witnesses,

there were several frontline police officers in attendance at the scene, which means there was potentially an opportunity to interview witnesses individually, but this did not occur. In addition, in 7 of the 95 recordings frontline police officers provided feedback on the accuracy of the descriptions provided by the witnesses. An example of a police officer providing feedback to a witness is:

Witness: They sounded like they were from London, they had a London accent

Police officer: You're spot on there

Description details provided by witnesses

Across the 95 recordings, witnesses provided a total of 700 details about the 207 potential perpetrators (M = 7.37, SD = 5.09). The number of details provided about a potential perpetrator ranged from 0 to 21 details (see Table 6 for a list of the person description details provided). The most frequent detail provided by witnesses was gender of the perpetrator (N = 197, 95.17%). The least frequent details provided about the perpetrator were details pertaining to weight and jewellery (N = 1, 0.48%). No witnesses provided details about the head shape, nose, mouth, ears or eyebrows of the perpetrator.

Table 6

Frequency of descriptor category

Descriptor category	Frequency (*N* = 207)	Percentage
Gender	197	95.17%
Clothing	115	55.56%
Race	94	45.41%
Age	69	33.33%
Height	42	20.29%
Hair colour	42	20.29%
Build	40	19.32%
Hair length	33	15.94%
Accent	21	10.14%
Facial hair	10	4.83%
Hair style	9	4.35%
Glasses	6	2.90%
Tattoo	5	2.42%
Face	5	2.42%
Eyes	4	1.93%
Complexion	2	0.97%
Teeth	2	0.97%
Weight	1	0.48%
Jewellery	1	0.48%
Mouth/Nose/Head shape/Ears/Eyebrows	0	0.00%

Reporting of fine- and coarse-grain details

The 556 questions asked by frontline officers elicited significantly more fine-grain details (M = 1.56, SD = 2.13, 95% CI [1.40, 1.74]) than coarse-grain details (M = 0.48, SD = 0.92, 95% CI [0.41, 0.56]), $t(555) = 11.97$, $p < .001$, $d = 0.66$ (see Table 7). The number of

fine-grain details across all questions ranged from 0 to 22 details, whereas the coarse-grain details ranged from 0 to 7 details.

Further analyses were conducted using a one-way MANOVA with Question Type (open vs. specific closed vs. appropriate yes/no vs. leading vs. multiple vs. forced choice vs. inappropriate yes/no) as the between-subjects factor and total number of details, number of fine-grain details and number of coarse-grain details as the three dependent variables. There was a significant multivariate main effect of Question Type, Wilks' $\lambda = .82$, $F(12,1096) = 9.44$, $p < .001$, $\eta_p^2 = .09$. Significant main effects were obtained for Question Type in terms of total details, $F(6,549) = 18.06$, $p < .001$, $\eta_p^2 = .17$, fine-grain details, $F(6,549) = 16.47$, $p < .001$, $\eta_p^2 = .15$, and coarse-grain details, $F(6,549) = 5.11$, $p < .001$, $\eta_p^2 = .05$. Post-hoc tests using a Bonferroni correction showed that open questions led to more total details and fine-grain details than each of the other question types (all *p*-values <.001; see Table 7). Open questions also led to more coarse-grain details than specific closed questions ($p = .006$), leading questions ($p < .001$), and forced choice questions ($p = .003$). All other comparisons were non-significant (*p*-values ranged from 0.56 to 1.00).

Table 7

Means and standard deviations for the number of fine- and coarse-grain details reported by question type

Question type	Categorisation	Fine grain *M* (*SD*)	Coarse grain *M* (*SD*)
Open	Appropriate	3.63 (3.58)	0.91 (1.50)
Specific closed	Appropriate	1.33 (1.50)	0.46 (0.76)
Leading	Inappropriate	1.10 (1.43)	0.26 (0.57)
Multiple	Inappropriate	1.22 (1.85)	0.67 (1.09)
Forced choice	Inappropriate	1.13 (1.39)	0.27 (0.54)
Inappropriate yes/no	Inappropriate	1.19 (1.58)	0.44 (0.91)
Appropriate yes/no	Appropriate	1.56 (2.18)	0.40 (0.58)

A between-subjects MANOVA was conducted with Categorisation (appropriate vs. inappropriate) as the between-subjects factor and total number of details, number of fine-grain details and number of coarse-grain details as the three dependent variables. There was a significant multivariate main effect of Categorisation, Wilks' $\lambda = .96$, $F(2,553) = 12.32$, $p < .001$, $\eta_p^2 = .04$. A significant univariate main effect was found with witnesses who were interviewed using appropriate question types providing more details than those witnesses who were interviewed using inappropriate question types, $F(1,554) = 24.58$, $p < .001$, $\eta_p^2 = .04$, $d = 0.37$ (see Table 8). Additionally, witnesses who were interviewed using appropriate question types provided more fine-grain details than those witnesses who were interviewed using inappropriate question types, $F(1,554) = 22.40$, $p < .001$, $\eta_p^2 = .04$, $d = 0.40$. Finally, witnesses who were interviewed with appropriate question types provided more coarse-grain

details than those witnesses who were interviewed using inappropriate question types, $F(1,554) = 6.04$, $p = .014$, $\eta_p^2 = .01$, $d = 0.21$ (see Table 8).

Table 8

Means and standard deviations for the number of fine- and coarse-grain details reported by question category

Question categorisation	Total details *M* (*SD*)	Fine grain *M* (*SD*)	Coarse grain *M* (*SD*)
Appropriate	2.56 *(3.03)*	1.99 *(2.53)*	0.58 (1.03)
Inappropriate	1.53 *(2.50)*	1.15 *(1.55)*	0.39 *(0.79)*

Discussion

The aim of this exploratory study was to examine the types of questions frontline police officers ask to obtain person description information from witnesses. The current study found that the use of appropriate questions resulted in witnesses reporting a higher overall number of person description details about the perpetrator than when inappropriate questions were asked, and that this was particularly the case when (appropriate) open-ended questions were asked. In the BWV data, over half of all the questions witnesses were asked by police officers were inappropriate. Of the inappropriate questions, leading questions were the most commonly used. Leading questions introduce information to witnesses that may not be true (Bowles & Sharman, 2014); hence, it is worrying to see such a high number of leading questions in frontline interactions.

Frontline officers used appropriate questions in just under half of all questions asked with specific closed questions being the most commonly used – despite the fact that the use of open-ended questions allows for unlimited free recall responses from witnesses, which

produce higher accuracy than closed questions (Fisher et al., 2000). The current study also found that witnesses provided more fine-grain details than coarse-grain details within their person descriptions. Both fine- and coarse- grain details were more likely to be elicited when frontline police officers asked appropriate questions compared to inappropriate questions. Person descriptions typically lack coarse-grain details (e.g. 'taller than X'). However, in the immediate search for a perpetrator, reporting of coarse-grain details could help narrow down the potential number of persons of interest and directly inform the allocation of resources (Brewer et al., 2018; McCallum et al., 2016). Appropriate questions lead to more coarse-grain details than inappropriate questions, and so appropriate questions need to be utilised more by frontline officers as such questions increase the likelihood of otherwise withheld information being elicited. Whilst there have been advances in effective interview practices, the benefits of asking appropriate questions have not always translated into frontline interviewing.

In the current study and consistent with laboratory research, an average of 7.4 details about the perpetrator were provided by witnesses. The current results fit with archival data showing witnesses providing between 3.9 and 9.7 physical characteristics (Lindsay et al., 1994; Sporer, 1996). However, inconsistent with laboratory research, the current study found that 'weight' was rarely mentioned about the perpetrator whereas 'gender' was almost always mentioned. This finding suggests that there are differences between the types of descriptors provided in the field compared to the laboratory, emphasising the need for more research to be conducted in the field if we are to obtain a better representation of person descriptions gathered by frontline police officers. The current study also revealed that facial features were rarely mentioned in the descriptions provided which is consistent with the findings of previous research (e.g. Demarchi & Py, 2009; Fahsing et al., 2004; Kuehn, 1974). Lastly, the current findings are consistent with the observation made by police officers in Brown et al. (2008) that descriptions tend to consist of general characteristics (e.g. sex, age and race) and

contain little information concerning facial features. One reason for this low reporting of facial feature details could be due to the types of questions police officers ask. For example, Brown et al. (2008) found that police officers reported that they would follow up the witness's free recall with specific probes, but that these probes directed the witness to think about further physical characteristics and clothing of the perpetrator as opposed to detailed descriptions of facial features. However, it should be noted that providing facial features is a difficult task, so it is not surprising to find that witnesses are poor at describing faces.

A further interesting observation made in the current study was that frontline police officers often failed to follow best-practice for interviewing witnesses. For example, some frontline police officers provided feedback to witnesses about their descriptions, which is problematic because it can inflate a witness's confidence (Semmler et al., 2004). When witnesses are given explicit feedback, or even subtle cues, regarding their account, they are likely to adjust their confidence to reflect their belief that they were accurate (confirming feedback) or inaccurate (disconfirming feedback). Frontline police officers therefore need to be aware that providing feedback can distort memory (Garven et al., 2000). It was also observed that frontline police officers interrupted witnesses when they were attempting to recall a person description. If witnesses are constantly interrupted then the witness may interpret this as if they have limited time to talk and may start to shorten their responses (Fisher, 1995). Thus, person descriptions may be less complete or informative because interruptions by police officers lead to witnesses holding back information.

Finally, it was noted that police officers did not separate witnesses before getting a description in 28 of the 95 recordings. In dynamic incidents where there are multiple witnesses and importantly multiple police officers, it is paramount that witnesses are separated to avoid contamination of memory (Gabbert et al., 2006; Ito et al., 2019). In 14 of the 28 recorded incidents where multiple witnesses were questioned together, there were

multiple police officers in attendance at the scene which meant the resources were available to interview the witnesses individually, but this was not prioritised. Even when there are not multiple officers in attendance, there are interview tools available to support the retrieval and reporting of detailed information at the scene or shortly afterwards (e.g. the Self-Administered Interview [SAI©]; Gabbert et al., 2009; see also College of Policing, 2019). Whilst officers on the frontline are often in resolution mode (e.g. breaking up a domestic disturbance), it is still vital that best practice is adhered to (i.e. they remain in investigative mode, e.g. questioning witnesses).

Overall, the current research shows that data obtained in the field can complement experimental data from the laboratory in a number of ways (e.g. number of details elicited from witnesses are similar regardless of whether a real or a mock crime is witnessed). However, there are a number of limitations associated with the current data. First, the absence of ground truth regarding the actual appearance of the perpetrator means I could not determine the accuracy of the details elicited from witnesses. Second, potential confounds such as crime type, individual differences between police interviewers (see Hudson et al., 2018), police interview training or previous experience of witness questioning were not measured. Police training could impact the types of questions asked and a witness that has been interviewed previously may be aware of the types of details they need to provide to inform a police investigation. Third, the study relied on the BWV footage accurately capturing the interaction between frontline police officers and witnesses. In the UK, it is the responsibility of the individual police officer to decide whether to turn on the BWV and thus it is possible that some relevant interactions were not recorded and hence not part of the available sample. These limitations highlight the need to avoid drawing casual inferences from field data (Wright, 2006). Nevertheless, these data from the field contribute to our existing knowledge of frontline policing, interviewing and person description information,

whilst also identifying priorities for future research (e.g. using BWV footage to explore the ability of vulnerable witnesses, such as those who are intoxicated or experiencing mental illness, to provide person descriptions and/or a police officer's ability to interview these vulnerable groups).

Conclusion

Body worn video footage provides a unique insight into the interactions at the frontline of policing, including the types of questions officers ask to obtain person descriptions from witnesses. Coding BWV footage from a sample of frontline officers in the U.K., the current study revealed that the frontline officers frequently asked inappropriate questions to elicit person descriptions, with leading questions a common occurrence. However, a comparative analysis of person descriptions showed that more person description information was obtained when appropriate questions were asked than when inappropriate questions were asked. Asking appropriate questions also led to more fine- and coarse- grain details being elicited, with open questions providing the highest number of details overall.

The current study demonstrates the need for continued collaboration between academics and practitioners if we are to contribute to the research knowledge base and maintain an evidence based approach to interviewing practice. Specifically, academics need to take advantage of the available BWV footage and use it as a tool to examine, not only person descriptions, but also frontline interactions more broadly. The focus for the field should be to ensure that practitioners know about the limitations of memory, are trained in the use of effective questioning techniques for obtaining detailed and reliable information and are aware of existing tools for managing frontline contexts involving multiple witnesses.

Chapter 3

Are there cross-race differences when providing person descriptions?

Foreword

Chapter 3 presents a laboratory study that was conducted in South Africa to investigate whether cross-race differences existed in the reporting of person descriptions. When conducting a literature review (see Chapter 1), cross-race differences with person descriptions was an area identified as needing further research. With that in mind and when analysing the body worn video footage as part of Study 1 (see Chapter 2), I discovered that person descriptions were less diagnostic when the witness and perpetrator were of a different race (e.g. one witness reported that "they were black lairy teenagers – best I can do mate"). My current circumstances meant that I was able to apply for an Internship through the International Centre for Research in Forensic Psychology (Department of Psychology, University of Portsmouth). I contacted Professor Tredoux at the University of Cape Town, South Africa, to explore any research opportunities and determine whether it would be possible for me to complete a research study within his institution. Together we designed a research study on cross-race differences and person descriptions, and I went out to South Africa to conduct the study for eight weeks. South Africa was chosen due to its diverse population which ensured that there was enough power for the study design. I was responsible for designing the study (which was discussed with the supervisory team). I was then required to set up the study, recruit research assistants (with the assistance of Dr Kempen and Dr Nortje), create study materials, code and analyse the data, and prepare the study for publication. Dr Vernham assisted with statistical analyses and completed the inter-rater reliability. All authors are currently in the process of reading the manuscript prior to journal submission.

Abstract

Globalisation has led to an increase in the number of multi-cultural communities, increasing the likelihood of victims or witnesses encountering perpetrator(s) of a different race or ethnic background to themselves. The accuracy of eyewitness testimony can be affected by a number of factors, including the race of the perpetrator. The current research examined cross-race differences in the reporting of person description information. Participants (N = 160; 86 black participants, 74 white participants) acted as mock-witnesses and viewed a simulated theft event involving either a white or black perpetrator. Participants were then required to provide a description of the perpetrator. Detail quantity, grain size and accuracy were coded. Overall, white participants provided more person descriptors than black participants, but there was no own-race bias. However, an own-race bias was obtained for accuracy of person descriptors. In terms of grain size, white participants provided more coarse-grain details than black participants and an own-race bias was found for white participants, but not black participants. In terms of accuracy of grain size, an own-race bias was present in the accuracy of fine-grain details, but not coarse-grain details. Implications for frontline policing are discussed.

Keywords: Memory, eyewitness memory, investigative interviewing, own-race bias, cross-race effect, person description.

Introduction

Person descriptions form a critical component of the preliminary investigation of a crime (Demarchi & Py, 2009; Milne & Bull, 2006). To be informative for an investigator, person descriptions need to have what is often referred to as 'diagnostic value' (i.e. include sufficiently detailed information to accurately identify a specific individual). To have 'diagnostic' value, the person description needs to enable an investigator to discriminate accurately in order to identify the correct perpetrator. As such, diagnostic value is reflected in the quantity and quality of the information provided in the person description. For instance, a person description with high diagnostic value will include numerous accurate references to the physical appearance and clothing of the perpetrator. Vague, imprecise or incomplete person descriptions lack diagnostic value and are problematic for the investigative process as they provide few clues for the police to find the perpetrator. Despite the importance of person descriptions in the investigative process, research has shown that descriptions provided by witnesses often lack diagnostic value and can frequently implicate multiple people within the vicinity of the crime (Demarchi & Py, 2009).

Cross-race differences can influence the quality and content of person descriptions provided by witnesses (Meissner et al., 2007). Although research has investigated recognition memory for own-race faces versus different-race faces (for a review see Meissner & Brigham, 2001; Sporer, 2001a), few studies have examined whether participants differ in the way they provide person descriptions from memory for target individuals from their own-race as opposed to target individuals from another race (Sporer, 2001b). The cross-race effect for memory recognition has been shown to impact on criminal investigations (e.g. identification line-ups; Kassin et al., 2001), so it is also reasonable to suggest that the cross-race effect will be a factor in the early stages of the criminal investigation (e.g. when obtaining a person

description from a witness and searching for the perpetrator in the immediate aftermath of an incident).

The few studies examining witness descriptions of own-race versus different-race faces have generally concluded that individuals attend to features they deem relevant to own-race faces and then inappropriately apply this same encoding scheme when assessing different-race faces (Ellis et al., 1975; Shepherd & Deregowski, 1981). For instance, in an early, small-sample study, Ellis et al. (1975) found that white Scottish participants were more likely than black African participants to report details concerning hair colour and texture, as well as the colour of the eyes. Sporer (2001a) suggested that this may be due to white participants having a greater familiarity with other white faces, and thus they have learned to pay greater attention to these specific features and use them as cues to distinguish members of their own-race. Ellis et al. (1975) also found that black African participants were more likely than white Scottish participants to report descriptions of the chin, ears, eyebrows, size of eyes, style of hair (e.g. combed to the right) and face shape. When nose and nostrils were collapsed to form one category, black African participants were more likely to provide these details (cf. white Scottish participants). In addition, the study revealed that black African participants used a larger number of facial characteristics compared with white Scottish participants when providing descriptions of both black faces and white faces. It is important to note that the black African participants in Ellis et al.'s (1975) study were not native speakers of English so it is possible that some of the reported differences were a function of language and the specific vocabulary for facial characteristics (Sporer, 2001a). Unfortunately, Ellis et al. (1975) did not assess the facial descriptions for accuracy, though they did note that white Scottish participants tended to provide facial descriptions that were limited (e.g. he has black skin).

Fallshore and Schooler (1995) examined white North American participants' ability to describe and identify black African American faces and white faces. As expected, they reported a cross-race effect for recognition memory, in that white North American participants were better able to identify white faces compared to black African American faces. However, they found no cross-race effect in terms of accuracy for the descriptions of white faces and black African American faces. Fallshore and Scholar (1995) suggested that the differences between cross-race face recognition and cross-race description may be down to differences in the reliance on configural processing (recognition, i.e. processing not just the shapes of individual features but also processing the relations among the features of the face) versus featural processing (description, i.e. processing individual features of the face; Maurer et al., 2002; Rhodes et al., 1989).

In terms of individual person descriptors, research examining cross-race differences has found an own-race bias when estimating height (Chen & Geiselman, 1993; Lee & Geiselman, 1994). Chen and Geiselman (1993) found that white and Asian participants underestimated the height of an Asian perpetrator compared with a white perpetrator of equal height. Similarly, Lee and Geiselman (1994) compared white and Asian participants but also had an additional comparison group of Hispanic participants. They found that in terms of height, Hispanic participants gave the highest (over)estimates of all groups whereas Asian participants gave the lowest estimates of all groups.

When reporting information, cooperative witnesses generally aim to be accurate and informative (Horry et al., 2016). However, under conditions of uncertainty, the demands of accuracy and informativeness compete and witnesses may have to reach a compromise between the two (Koriat & Goldsmith, 1996). One method for balancing accuracy and informativeness is to regulate the level of detail (grain size) they report (Goldsmith et al., 2002; Sauer & Hope, 2016). Fine-grain responses are more informative but less likely to be

accurate compared with coarse-grain responses (Brewer et al., 2018). When a premium is placed on accurate reporting (e.g. when providing person descriptions to the police), it is likely that witnesses use a higher threshold for accurate reporting and may withhold information that is likely to be incorrect. Thus, witnesses may be strategically regulating their memory to enhance accuracy (Koriat & Goldsmith, 1994, 1996). If cross-race differences exist when providing descriptions, it might be because witnesses regulate the information they provide about other-races. Witnesses are already unlikely to have extensively encoded the face or person details (regardless of race), so adding an extra variable in the form of race could increase metacognitive monitoring and result in fewer person description details being provided compared to when there is no cross-race difference (i.e. the witness and perpetrator are of the same race).

The aim of the current study was to expand on the limited research on cross-race differences in person descriptions. Previous research has shown that (regardless of race) the reporting of person descriptors is low (on average between 5.3 and 15.9; e.g. Geiselman et al., 2000; Pozzulo & Warren, 2003; Wells et al., 1993). In order to explore how good person descriptions could be, if witnesses are supported properly, participants in the current study were provided with a tool to help them to recall person description information (i.e. person description form; PDF). Whilst this was an exploratory study, it was predicted that participants from the same race as the perpetrator (i.e. own-race) would provide more person descriptors than participants from a different-race to the perpetrator (Hypothesis 1); participants from the same race as the perpetrator would provide more accurate person descriptions compared to participants from a different-race to the perpetrator (Hypothesis 2); participants from the same race as the perpetrator would provide more fine-grain and coarse-grain details than participants from a different-race to the perpetrator (Hypothesis 3); and

participants from the same race as the perpetrator would provide more accurate fine-grain and coarse-grain details than participants from a different-race to the perpetrator (Hypothesis 4).

Method

Participants and Design

A power analysis revealed that a minimum of 20 participants would be sufficient in each group (i.e. a total of 80 participants) for the study to have good statistical power (.849) and a large effect size ($\eta_p^2 = .138$) at the .05 significance level. To avoid discrimination and allow all students the same opportunity to participate (students gained course credits for participation), participation in the study was open to all students of all races, and participants were asked to self-identify their race in a demographics questionnaire. A total of 242 participants from South Africa were recruited through the Department of Psychology's participation pool (awarded 1 course credit for participation). 82 participants who did not self-identify their race as either "black" or "white" were excluded from the data analyses (e.g. participants who self-identified as "coloured" or "mixed race"). The reported analysis is based on the data for the remaining 160 participants (122 female, 38 male) with a mean age of 19.80 years (SD = 2.07; 86 were black [68 female], M_{age} = 19.95 years, *SD* = 2.39; 74 were white [54 female], M_{age} = 19.62 years, *SD* = 1.60). Taking into account the race of the participants, participants were pseudo-randomly allocated to a 2 (Participant Race: black participant vs. white participant) x 2 (Perpetrator Race: own-race vs. different-race) between-subjects design, with 36 white participants viewing an own-race perpetrator, 42 black participants viewing an own-race perpetrator, 38 white participants viewing a perpetrator of a different-race, and 44 black participants viewing a perpetrator of a different-race.

Materials

Stimulus event

The stimulus event was a short film lasting 40 seconds. Four short recordings were made each depicting a single male perpetrator committing a theft from a library. To reduce stimulus specificity, each recording showed a different perpetrator committing the same crime and behaving in an identical manner (two black perpetrators and two white perpetrators; see Figure 1). 17 white participants and 22 black participants saw Video 1 (white perpetrator), 19 white participants and 22 black participants saw Video 2 (white perpetrator), 19 white participants and 21 black participants saw Video 3 (black perpetrator), and 19 white participants and 21 black participants saw Video 4 (black perpetrator). The event starts with a scene from a library and more specifically a desk that has been left unattended. On the desk is a laptop, a mobile phone and a wallet. The perpetrator then enters the room. He looks around before approaching the desk and taking a seat. He sits for a short period of time and takes his own mobile phone from his pocket. After a short time and, checking that no-one else was around, he takes the wallet from the desk and removes a number of bank notes. The perpetrator then stands up, again checks that no-one is looking, before also taking the mobile phone from the desk. The event ends with the perpetrator walking away with the stolen money and the stolen mobile phone. There was no audio, and each video was matched in terms of number of possible person descriptors that could be provided by participants. That is, all perpetrators were dressed similarly and wore similarly accessories (e.g. all perpetrators wore a watch and a lanyard).

Figure 1

Screenshot of perpetrators used in the stimulus event

Person description form (PDF) instructions

To maximise the potential number of person descriptors reported within the current study and magnify any cross-race differences, mock-witnesses were provided with a tool to help them to recall person description information. Instructions in the person description form (PDF) were adapted from the Self-Administered Interview (SAI; Gabbert et al., 2009). Mock-witnesses were instructed to provide a description of the perpetrator they saw in their short film. The instructions contained a full list of descriptors (e.g. race, gender, height) and mock-witnesses were asked to provide details relating to each of these listed descriptors, but only if they had a memory for them. Mock-witnesses were instructed not to guess, and were encouraged to close their eyes (to support memory retrieval). The instructions then provided a series of non-leading prompts, and mock-witnesses were asked to report any new details,

including new details about the perpetrators face, clothing, and any distinguishing features (See Appendix A).

Procedure

Ethical approval was obtained from the Department of Psychology Ethics Committee at the University of Cape Town in South Africa (see Appendix C). Participants arrived at the laboratory at pre-arranged times and were informed that the study would be about eyewitness memory. Participants first read an information sheet and signed a consent form indicating their agreement to take part. Participants were then asked to complete a short demographic survey which asked participants to self-identify their gender and ethnicity, as well as provide their age. Participants were then pseudo-randomly allocated to one of four short films. All participants were given the following instructions prior to watching the short film: *'During this study you will watch a short film. The film has no audio. However, please pay close attention to the film, as you will be asked questions about what you saw later on.'*

After witnessing the event, all participants completed a 10-minute distractor task (word search). Participants in all conditions were told to complete the PDF which had the following instructions *'Please write down a detailed physical description of the person you saw in the video. Try to describe the person in enough detail so that someone else could use your description to pick them out of a crowd?'* Participants were also instructed not to guess and were encouraged to close their eyes or stare/ look at a blank wall or the floor to help focus and concentrate on the task. Participants were not asked any questions by the researcher. Once participants had completed their description of the perpetrator they were thanked for their time and fully debriefed. Participation in the study took approximately 20 minutes.

Coding

A coding scheme was devised to categorise each person descriptor. The person descriptors reported by participants were coded according to previous research (see Lindsay et al., 1994). Categories included face items (e.g. face, mouth, nose), exterior face items (e.g. hair length, hair style, hair colour), body (e.g. height, weight), clothing (e.g. blue jeans, white all-star converse, blue jumper), and accessory descriptors (e.g. blue lanyard, brown leather watch). Race, age and gender were coded separately.

Similar to previous research (see Pozzulo & Warren, 2003), age, height and weight were given an acceptable range (age ± 3 years; height ± 2 inches; weight ± 2kg). Any descriptor that exceeded the acceptable range (e.g. 'aged between 20 and 30 years') was considered an error and coded as incorrect. Details that were subjective or vague (e.g. 'he was neither tall nor short') were not coded for accuracy. To assess overall inter-rater reliability, 32 PDFs were randomly selected and coded independently by a second rater blind to the experimental conditions. Inter-rater reliability for overall accuracy of person descriptors was excellent, ICC = .90, 95% CI [.74, .96].

To assess the quality of elicited person descriptors, the grain size of reported information was coded, in terms of fine-grain and coarse-grain. Specifically, coding for descriptions with numerical values, and similar to Weber and Brewer (2008) and Sauer and Hope (2016), were coded as fine-grain if they included up to three possible values (e.g. aged between 18 and 21 years). However, a response was referred to as coarse-grain if it included four or more possible values (e.g. aged 18 to 25 years). For categorical data, pre-determined parameters defined when responses were considered fine-grain (e.g. bushy eyebrows). Vague and subjective responses (i.e. 'about my height') were not coded. To assess overall inter-rater reliability for grain size, 32 PDFs were randomly selected and coded independently by a second rater who was blind to the experimental conditions. Inter-rater reliability for accuracy

of fine-grain details was excellent, ICC = .91, 95% CI [.78, 96] and inter-rater reliability for accuracy of coarse-grain details was very good, ICC = .84, 95% CI [.66, .93].

Each participant received a score for (i) the total number of person descriptors provided, (ii) the proportion (accuracy) of person descriptors provided (i.e. correct descriptors divided by total [correct + incorrect] number of descriptors), (iii) the total number of fine-grain details and the total number of coarse-grain details, and (iv) the proportion (accuracy) of fine-grain details (i.e. correct fine-grain details divided by total number of fine-grain details) versus coarse-grain details (i.e. correct coarse-grain details divided by total number of coarse-grain details) provided. These newly computed scores then formed the dependent variables in the current study.

Results

Manipulation check

To ensure that the findings reported below were not an outcome of the specific video viewed by participants, a one-way ANOVA was conducted with Video (1 vs. 2 vs. 3 vs. 4) as the only between-subjects factor and the total number of person descriptors provided by participants as the dependent variable. Levene's Test for Homogeneity of Variance was met (p =.992), therefore equal variances can be assumed. The ANOVA revealed no differences between the four different videos in terms of the number of person descriptors provided (p = .950). Therefore, Video was not used as a covariate in the future analyses.

Reporting of person descriptors

A 2 (Participant Race: black participant vs. white participant) x 2 (Perpetrator Race: own-race vs. different-race) between-subjects ANOVA was conducted with the total number of person descriptors provided by participants as the dependent variable. Levene's Test for Homogeneity of Variance was met (p =.353), therefore equal variances can be assumed. The

ANOVA revealed a main effect for Participant Race, with white participants providing more person descriptors than black participants, $F(1,156) = 12.29$, $p = .001$, $\eta_p^2 = .07$ (see Table 9). There was no difference between the number of person descriptors provided by participants who viewed a perpetrator from their own-race and those who viewed a perpetrator from a different-race, $F(1,156) = .36$, $p = .548$, $\eta_p^2 < .01$. The Participant Race X Perpetrator Race interaction effect was not significant, $F(1,156) = .03$, $p = .874$, $\eta_p^2 < .01$.

Accuracy of person descriptors

A 2 (Participant Race: black participant vs. white participant) x 2 (Perpetrator Race: own-race vs. different-race) between-subjects ANOVA was conducted with the accuracy of the person descriptors provided by participants as the dependent variable. Levene's Test for Homogeneity of Variance was met ($p = .085$), therefore equal variances can be assumed. The ANOVA revealed a main effect for Participant Race, with higher accuracy rates for person descriptions being provided by white participants than black participants, $F(1,156) = 6.12$, $p = .014$, $\eta_p^2 = .04$ (see Table 9). The ANOVA also found a main effect for Perpetrator Race, with participants who saw a perpetrator from their own-race scoring a higher accuracy for person descriptors than participants who saw a perpetrator from a different-race, $F(1,156) = 5.35$, $p = .022$, $\eta_p^2 = .03$ (see Table 9). The Participant Race X Perpetrator Race interaction effect was not significant, $F(1,156) = .10$, $p = .752$, $\eta_p^2 < .01$.

Table 9

Descriptive statistics for the number and accuracy of person descriptors reported

Dependent variable	**Participant Race**		**Perpetrator race**	
	Black participants Mean *(SD), 95% CI*	**White Participants** Mean *(SD), 95% CI*	**Own-race** Mean *(SD), 95% CI*	**Different-race** Mean *(SD), 95% CI*
Total number of person descriptors	19.51 (3.96), *18.68-20.36*	21.77 (4.13), *20.78-22.74*	20.76 (4.01), *19.87-21.63*	20.37 (4.36), *19.44-21.34*
Overall accuracy of person descriptors	0.81 (0.11), *0.78-0.83*	0.85 (0.09), *0.83-0.87*	0.84 (0.09), *0.83-0.86*	0.81 (0.11), *0.78-0.83*
Total number of fine-grain details	16.64 (4.06), *15.67-17.36*	17.92 (4.68), *16.88-19.00*	17.71 (4.39), *16.68-18.66*	16.78 (4.36), *15.74-17.54*
Overall accuracy of fine-grain details	0.84 (0.11), *0.81-0.86*	0.87 (0.09), *0.84-0.89*	0.87 (0.09), *0.85-0.89*	0.83 (0.12), 0.80-0.86
Total number of coarse-grain details	3.08 (1.43) *2.77-3.39*	3.82 (1.61), *3.45-4.20*	3.32 (1.31), *3.06-3.65*	3.52 (1.76), *3.20-3.96*
Overall accuracy of coarse-grain details	0.69 (0.30), *0.62-0.75*	0.76 (0.24), *0.70-0.81*	0.72 (0.27), *0.66-0.76*	0.72 (0.28), 0.66-0.78

Reporting of fine-grain and coarse-grain details

Pearson correlation analyses were conducted to explore whether there was a relationship between participant reporting of fine-grain and coarse-grain details. It was found

that there was a significant negative correlation between the total number of fine-grain details and total number of coarse-grain details, $r(N = 160) = -.23$, $p = .004$. Hence, the assumption for conducting a MANOVA was met.

A 2 (Participant Race: black participant vs. white participant) x 2 (Perpetrator Race: own-race vs. different-race) between-subjects MANOVA was conducted with total number of fine-grain details and total number of coarse-grain details as the two dependent variables. Significant multivariate main effects were obtained for Participant Race, Wilks' $\lambda = .90$, $F(2, 155) = 8.51$, $p < .001$, $\eta_p^2 = .10$, as well as the Participant Race X Perpetrator Race interaction effect, Wilks' $\lambda = .94$, $F(2, 155) = 5.42$, $p = .005$, $\eta_p^2 = .07$. There was no multivariate main effect for Perpetrator Race, Wilks' $\lambda = .98$, $F(2, 155) = 1.36$, $p = .259$, $\eta_p^2 = .02$; however, since Hypothesis 3 was specifically about perpetrator race it was deemed appropriate to also analyse the univariate between-subjects effects for this factor. Levene's Test for Homogeneity of Variance was met for both dependent variables (p-values ranged from .236 to .905), therefore equal variances can be assumed. A univariate main effect was found for Participant Race with white participants providing more coarse-grain details than black participants, $F(1, 156) = 9.57$, $p = .002$, $\eta_p^2 = .06$ (see Table 9). No such univariate main effect was found between white participants and black participants in terms of fine-grain details, $F(1,156) = 3.63$, $p = .058$, $\eta_p^2 = .02$. No univariate main effects were found for Perpetrator Race. Participants who saw a perpetrator from their own-race did not differ from participants who saw a perpetrator from a different-race in terms of total number of fine-grain details, $F(1,156) = 2.15$, $p = .145$, $\eta_p^2 = .01$.

Additionally, participants who saw a perpetrator from their own-race did not differ from participants who saw a perpetrator from a different-race in terms of total number of coarse-grain details, $F(1,156) = 1.23$, $p = .270$, $\eta_p^2 = .01$. In terms of the univariate main effects for the Participant Race X Perpetrator Race interaction, there was no interaction effect

for the total number of fine-grain details, $F(1,156) = 2.32$, $p = .129$, $\eta_p^2 = .02$; however, there was a significant interaction effect for the total number of coarse-grain details, $F(1, 156) = 10.32$, $p = .002$, $\eta_p^2 = .06$. Simple main effects analyses were conducted to interpret the interaction effect. The simple main effects analyses showed that white participants provided more coarse-grain details when describing a perpetrator from a different-race ($M = 4.32$, $SD = 1.77$, 95% CI [3.84, 4.79]) than when describing a perpetrator from their own-race ($M = 3.31$, $SD = 1.24$, 95% CI [2.82, 3.79]), $F(1, 156) = 8.68$, $p = .004$, $\eta_p^2 = .05$. Black participants, however, did not differ in their reporting of coarse-grain details when describing a perpetrator of their own-race ($M = 3.33$, $SD = 1.39$, 95% CI [3.00, 3.80]) versus a perpetrator of a different-race ($M = 2.84$, $SD = 1.45$, 95% CI [2.51, 3.33]), $F(1, 156) = 2.40$, $p = .124$, $\eta_p^2 = .02$.

Accuracy of fine-grain and coarse-grain details

Pearson correlation analyses were conducted to explore whether there was a relationship between participant accuracy of fine-grain and coarse-grain details. It was found that there was a positive correlation between participant accuracy of fine-grain details and participant accuracy of coarse-grain details, $r(N = 158) = .19$, $p = .017$. Hence, the assumption for conducting a MANOVA was met.

A 2 (Participant Race: black participant vs. white participant) x 2 (Perpetrator Race: own-race vs. different-race) between-subjects MANOVA was conducted with accuracy of fine-grain details and accuracy of coarse-grain details as the two dependent variables. A significant multivariate main effect was obtained for Perpetrator Race, Wilks' $\lambda = .96$, $F(2, 153) = 3.13$, $p = .046$, $\eta_p^2 = .04$. There were no multivariate main effects for Participant Race, Wilks' $\lambda = .97$, $F(2, 153) = 2.81$, $p = .063$, $\eta_p^2 = .04$, or for the Participant Race X Perpetrator Race interaction effect, Wilks' $\lambda = .98$, $F(2, 153) = 1.38$, $p = .256$, $\eta_p^2 = .02$. Levene's Test

for Homogeneity of Variance was met for both dependent variables (*p*-values ranged from .064 to .246), therefore equal variances can be assumed. A significant univariate main effect was found for Perpetrator Race with participants being more accurate at providing fine-grain details when the perpetrator was of their own-race compared to when the perpetrator was of a different-race, $F(1, 154) = 6.11$, $p = .015$, $\eta_p^2 = .04$ (see Table 9). No such univariate main effect was found in terms of accuracy of course-grain details between participants viewing a perpetrator of their own-race and participants viewing a perpetrator of a different-race, $F(1, 154) < .01$, $p = .984$, $\eta_p^2 < .01$.

Exploratory analyses: Types of descriptors

The types of descriptors reported were examined to assess whether there were any differences between black participants and white participants in the types of person descriptors they provided and whether this was dependent upon the race of the perpetrator (see Table 10). To account for numerous analyses, Bonferroni correction was applied to all pairwise comparisons.

Two 2 (Participant Race: black participant vs. white participant) x 2 (Perpetrator Race: own-race vs. different-race) between-subjects ANOVAs were conducted with either race or age as the dependent variable. No significant main effects were obtained for Participant Race, Perpetrator Race, or the Participant Race X Perpetrator Race interaction effect (*p*-values for race ranged from .145 to .988; *p*-values for age ranged from .054 to .725).

A 2 (Participant Race: black participant vs. white participant) x 2 (Perpetrator Race: own-race vs. different-race) between-subjects MANOVA was conducted with height and weight as the two dependent variables. A significant multivariate main effect was obtained for Participant Race, Wilks' $\lambda = .96$, $F(2, 155) = 3.45$, $p = .035$, $\eta_p^2 = .04$. There were no multivariate main effects for Perpetrator Race ($p = .750$) or for the Participant Race X

Perpetrator Race interaction effect (p = .479). A significant univariate main effect was found for Participant Race with white participants describing the weight/ build of the perpetrator more than black participants, $F(1, 156) = 5.62$, p = .019, η_p^2 = .04 (see Table 10). No such univariate main effect was found for Participant Race in terms of describing the height of the perpetrator (p = .145).

A 2 (Participant Race: black participant vs. white participant) x 2 (Perpetrator Race: own-race vs. different-race) between-subjects MANOVA was conducted with eyes, ears, mouth, nose, eyebrows, face, complexion, hair colour, hair length, hair style, and facial hair as the dependent variables. Significant multivariate main effects were obtained for Participant Race, Wilks' λ = .88, $F(11, 146) = 1.89$, p = .045, η_p^2 = .13, Perpetrator Race, Wilks' λ = .80, $F(11, 146) = 3.30$, $p < .001$, η_p^2 = .20, as well as the Participant Race X Perpetrator Race interaction effect, Wilks' λ = .80, $F(11, 146) = 3.32$, $p < .001$, η_p^2 = .20.

In terms of Participant Race, univariate main effects were obtained for the face, $F(1, 156) = 4.49$, p = .036, η_p^2 = .03, and hair colour, $F(1, 156) = 8.05$, p = .005, η_p^2 = .05 (see Table 10). White participants mentioned the face of the perpetrator more than black participants. Likewise, white participants mentioned the hair colour of the perpetrator more than black participants. No other univariate main effects were found for Participant Race in terms of the differing facial features (p-values ranged from .058 to .953).

In terms of Perpetrator Race, univariate main effects were found for the eyes, $F(1, 156) = 6.32$, p = .013, η_p^2 = .04, complexion, $F(1, 156) = 9.49$, p = .002, η_p^2 = .06, hair colour, $F(1, 156) = 8.05$, p = .005, η_p^2 = .05, and hair style, $F(1, 156) = 7.51$, p = .007, η_p^2 = .05 (see Table 10). Participants mentioned the eyes more often when the perpetrator was of a different-race than when the perpetrator was of their own-race. Likewise, participants mentioned the hair colour more often when the perpetrator was of a different-race than when

the perpetrator was of their own-race. Participants mentioned the complexion of the perpetrator more often when the perpetrator was of their own-race than when the perpetrator was of a different-race. Likewise, participants mentioned the hair style more often when the perpetrator was of their own-race than when the perpetrator was of a different-race. No other univariate main effects were found for Perpetrator Race in terms of the differing facial features (*p*-values ranged from .171 to .996).

Finally, there were significant univariate main effects for the Participant Race X Perpetrator Race interaction for hair colour, $F(1, 156) = 16.96$, $p < .001$, $\eta_p^2 = .10$ and hair style, $F(1, 156) = 4.58$, $p = .034$, $\eta_p^2 = .03$. No other univariate main effects for the Participant Race X Perpetrator Race interaction were obtained (*p*-values ranged from .074 to .747).

Simple main effect analyses were conducted to interpret the significant interaction effects. The simple main effects analyses showed that black participants mentioned hair colour more often when the perpetrator was from a different-race ($M = 1.00$, $SD < .01$, 95% CI [.92, 1.08]), than of their own-race, ($M = .71$, $SD = .46$, 95% CI [.64, .80]), $F(1, 156) = 26.15$, $p < .001$, $\eta_p^2 = .14$. White participants, however, did not differ in their reporting of hair colour when describing a perpetrator of their own-race versus a perpetrator of different-race ($p = .384$). The simple main effect analyses also showed that white participants mentioned hair style more often when the perpetrator was from their own-race ($M = .69$, $SD = .47$, 95% CI [.53, .86]) than of a different-race ($M = .32$, $SD = .47$, 95% CI [.16, 1.47]), $F(1, 156) = 11.09$, $p = .001$, $\eta_p^2 = .07$. Black participants, however, did not differ in their reporting of hair style when describing a perpetrator of their own-race versus a perpetrator of different-race ($p = .660$).

An additional two 2 (Participant Race: black participant vs. white participant) x 2 (Perpetrator Race: own-race vs. different-race) between-subjects MANOVAs were conducted with clothing (trousers, jumper/ t-shirt, shoes) forming the dependent variables of the first MANOVA and accessories (jewellery, lanyard, mobile phone) forming the dependent variables of the second MANOVA. Across both MANOVAs, there were no multivariate main effects for Participant Race, Perpetrator Race or the Participant Race X Perpetrator Race interaction (all *p*-values ranged from .191 to .862). There were also no univariate main effects across the two MANOVAs (all *p*-values ranged from .052 to .886).

Table 10

Descriptive statistics for the types of person descriptors reported

Dependent variable	**Participant Race**		**Perpetrator race**	
	Black participants Mean *(SD), 95% CI*	**White Participants** Mean *(SD), 95% CI*	**Own-race** Mean *(SD), 95% CI*	**Different-race** Mean *(SD), 95% CI*
Race	0.99 (0.24), *0.94-1.04*	0.99 (0.20), *0.94-.1.03*	1.00 (0.28), *0.94-1.06*	0.98 (0.16), *0.94-1.00*
Age	0.83 (0.38), *0.74-0.90*	0.91 (0.30), *0.83-0.97*	0.81 (0.40), *0.72-0.89*	0.91 (0.28), *0.85-0.97*
Weight/build	0.90 (0.31), *0.83-0.95*	0.99 (0.12), *0.96-1.00*	0.95 (0.22), *0.89-0.99*	0.93 (0.26), *0.86-0.98*
Height	0.86 (0.35), *0.78-0.93*	0.93 (0.25), *0.87-0.99*	0.91 (0.29), *0.85-0.97*	0.88 (0.33), *0.79-0.95*
Eyes	0.65 (0.48), *0.54-0.75*	0.77 (0.42), *0.68-0.86*	0.62 (0.49), *0.51-0.72*	0.79 (0.41), *0.70-0.89*
Ears	0.09 (0.29), *0.03-0.16*	0.19 (0.39), *0.10-0.28*	0.14 (.35), *0.06-0.22*	0.13 (0.34). *0.06-0.21*

Mouth	0.37 (0.49), *0.28-0.49*	0.46 (0.50), *0.35-0.57*	0.44 (0.50), 0.33-0.54	0.39 (0.49), *0.29-.050*
Nose	0.41 (0.49), *0.30-0.52*	0.55 (0.50), *0.44-0.67*	0.46 (0.50), *0.35-0.58*	0.49 (0.50), *0.38-0.60*
Eyebrows	0.08 (0.28), *0.03-.014*	0.12 (0.33), *0.05-0.20*	0.09 (0.29), 0.03-0.15	0.11 (0.32), *0.05-0.18*
Face	0.62 (0.71), *0.47-0.77*	0.85 (0.72), *0.66-1.00*	0.64 (0.66), *0.60-0.80*	0.80 (0.76), *0.64-0.98*
Complexion	0.73 (0.45), *0.64-0.82*	0.77 (0.42), *0.67-0.86*	0.86 (0.35), *0.78-0.94*	0.65 (0.48), *0.54-0.76*
Hair Colour	0.86 (0.35) *0.78-0.93*	0.97 (0.16), *0.93-.1.00*	0.85 (0.36), 0.77-0.92	0.98 (0.16), *0.94-1.00*
Hair length	0.60 (0.49), *0.50-0.71*	0.70 (0.46), *0.59-0.81*	0.63 (0.49), *0.51-0.73*	0.67 (0.47), *0.56-0.77*
Hair style	0.50 (0.50), *0.39-0.60*	0.50 (0.50), *0.39-0.61*	0.60 (0.49), *0.50-0.72*	0.40 (0.49), *0.29-0.51*
Facial hair	0.76 (0.43), *0.66-0.84*	0.84 (0.37), *0.75-0.92*	0.82 (0.39), *0.74-0.90*	0.77 (0.43), *0.68-0.86*
Jumpers/t-shirt	2.35 (0.61), *2.22-2.48*	2.46 (0.66), *2.29-2.61*	2.36 (0.60), *2.23-2.49*	2.44 (0.67), *2.29-2.57*
Trousers	1.86 (0.38), *1.77-1.94*	1.85 (0.36), *1.77-1.93*	1.82 (0.42), *1.73-1.91*	1.89 (0.32) 1.82-1.96
Shoes	1.87 (1.40), *1.59-2.16*	1.84 (1.48), *1.49-2.15*	1.88 (1.44), *1.57-2.19*	1.83 (1.44) *1.51-2.13*
Jewellery	0.80 (0.85), *0.62-0.99*	0.91 (1.00), *0.69-1.14*	0.96 (0.97), *0.75-1.18*	0.74 (0.86), *0.56-0.93*
Lanyard	0.42 (0.89), *0.24-0.60*	0.73 (1.13), *0.48-0.99*	0.67 (1.15), 0.43-0.93	0.46 (0.86), *0.28-0.65*
Mobile phone	0.38 (0.79), *0.23-0.55*	0.46 (0.92), *0.26-0.69*	0.45 (0.91), *0.26-0.65*	0.39 (0.80), *0.22-0.58*

Discussion

The current study explored potential cross-race differences between white participants and black participants when they provided person descriptions of a perpetrator. Although white participants provided more person descriptors than black participants, there was no own-race bias and therefore Hypothesis 1 was not supported. Previous studies (e.g. Ellis et al., 1975; Fallshore & Schooler, 1995) have focused on examining facial features only: To the knowledge of the researcher, the current study is the first to examine the overall frequency of descriptors (including details such as clothing, height, weight, etc.). When the current study explored individual person descriptors (e.g. eyes, hair colour, trousers, lanyard, etc.), participants were more likely to mention complexion and hair style for a perpetrator of their own-race compared to a perpetrator of a different-race. Interestingly, participants were more likely to report the eyes and hair colour of a different-race compared to their own-race. Unlike Chen and Geiselman (1993) and Lee and Geiselman (1994), the current study did not obtain an own-race bias for height. However, the perpetrators used in the videos of the current study were of different heights and therefore height was not controlled for like it was in previous studies. In line with Ellis et al. (1975), white participants mentioned hair colour more often than black participants regardless of the race of the perpetrator. In direct contrast to Ellis et al. (1975), white participants provided descriptors about the face of the perpetrator more than black participants, irrespective of the race of the perpetrator. The current study also found that white participants provided details about weight more than black participants. It appears apparent that there are differences across races for individual types of person descriptors but there is a need for more research to be conducted.

As predicted in Hypothesis 2, participants who saw a perpetrator from their own-race provided more accurate person descriptions than participants who saw a perpetrator from a different-race. Regardless of the race of the perpetrator, white participants were more

accurate when providing person descriptions, than black participants. Previous research on facial recognition has demonstrated that expertise, in part, comes from meaningful experiences with faces across the lifespan (Brigham & Malpass, 1985; Bukach et al., 2012). Thus, researchers have concluded that the power to recognise faces with high accuracy comes from experience in discriminating between individuals from within a homogeneous population of highly similar faces (i.e. faces of one's own-race; Chiroro & Valentine, 1995; Tham et al., 2017). It is therefore possible that expertise with one's own-race also crosses over to being more accurate with person recall and description.

Unfortunately, very few studies have looked at accuracy for descriptions rather than recognition. For example, Ellis et al. (1975) did not report accuracy in their study. The one study that has examined accuracy is not supported by the current study: Fallshore and Schooler (1995) found no cross-race effect in terms of accuracy for the descriptions of white American and black African American faces. However, the study by Fallshore and Schooler (1995) used a different paradigm to measure accuracy (i.e. communication accuracy paradigm) compared to the current study which used description analysis (where the number of correct and incorrect feature were counted). Therefore, the opposing findings may be due to the differing ways in which person descriptions were measured. The contradictory findings reinforce the need for future research to be conducted into cross-race descriptions.

The current study was the first of its kind to explore whether participants regulate the reporting of person descriptions depending on the race (own versus different) of the perpetrator. In contrast to Hypothesis 3, there was no overall own-race bias for the number of fine-grain details or coarse-grain details reported by participants. However, white participants provided more coarse-grain details about perpetrators from a different-race than perpetrators from their own-race. White participants also provided more coarse-grain details than black participants overall. This pattern of results might suggest that white participants were

strategically regulating their memory when providing coarse-grain details, particularly for individuals of a different-race. In partial support of Hypothesis 4, participants (regardless of race) were more accurate at providing fine-grain details when describing a perpetrator of their own-race compared to a perpetrator of a different-race. Hence an own-race bias appears to be present in the accuracy of fine-grain details, but not coarse-grain details. Since this was the first study to look at own-race bias for fine- and coarse-grain details, more research is needed to explore the strategic reporting of person descriptions when mock-witnesses describe faces of either their own-race or a different-race.

Practical implications

In a multi-cultural society the likelihood of a witness and perpetrator being of a different race increases. Whilst a wealth of research has found evidence for a cross-race effect for recognition (see Meissner & Brigham, 2001 for a review), the current study shows that it remains unclear whether a cross-race effect for recall (e.g. when providing person description) exists. Person descriptions are often elicited by frontline police officers using inappropriate questioning techniques that do not adhere to best-practice (e.g. separating witnesses, not providing feedback), resulting in poor person descriptions being reported by witnesses (Dalton et al., 2021). In addition, frontline police officers need to be aware that the race of the perpetrator and/ or witnesses may affect the reporting of person description information.

Previous research has found that the reporting of person descriptors is low (on average between 5.3 and 15.9; e.g. Geiselman et al., 2000; Pozzulo & Warren, 2003; Wells et al., 1993). Therefore, to maximise the potential number of person descriptors reported within the current study and magnify any cross-race differences, participants were provided with a tool to help them to recall person description information (i.e. a person description form). As

a result, participants in the current study reported, on average, 20.6 person descriptors, illustrating the importance of using non-leading tools to support witnesses during the recall of their person descriptions (e.g. during frontline communication). Future research should continue to explore alternative ways of facilitating witness reporting of person description information.

Limitations

One limitation associated with the current study is that it did not make any assessment of 'contact'. Even with the consensus that cross-race differences exist (see Meissner & Brigham, 2001), there is no generally agreed upon reason for why cross race-differences occur. The best-known explanation for cross-race differences is the 'contact hypothesis' (e.g. Brigham & Malpass, 1985; Valentine et al., 1995). According to the contact hypothesis, people become experts at discriminating faces from their own-race because they have more contact with people of their own-race compared to people of a different-race. Whilst some studies support the contact hypothesis, others have disputed it (see Slone et al., 2000). Future studies should ask mock-witnesses to self-report their amount of contact with individuals from different-races compared to their own-race (creating comparison groups that can be examined, e.g. low-contact with different-races versus high contact with different-races; see Chiroro & Valentine, 1995) or use groups that are predicted to have different amounts of interracial contact (see Ng & Lindsay, 1994).

A second limitation of the current study involves the format in which person descriptions were obtained. In the current study, mock-witnesses provided written reports for their descriptions of the perpetrator. Sauerland and Sporer (2011) in a mock-crime scenario found spoken accounts provided a higher quantity and accuracy of person descriptors than written accounts. This suggests that a written account may not be as effective as one obtained

orally. Additionally, research has shown that witnesses who are asked to provide written accounts may be disadvantaged, especially if they are unaccustomed to writing, and embarrassed by their handwriting or spelling difficulties (McPhee et al., 2014). Furthermore, witnesses may be less able to communicate as effectively through writing as they would orally (Hope et al., 2011). Research has shown that written statements are more cognitively-demanding than their oral counterparts (Kellogg, 2007). However, Sauerland et al. (2014) found that written accounts yielded more quantity and detail (but not accuracy) than oral accounts. Future research should examine different reporting methods to explore whether orally obtained descriptions contain more person descriptors than other reporting formats (e.g. written statements), and whether cross-race differences are heighted or reduced depending on reporting mode.

A third limitation of the current study is that it was carried out in a laboratory under optimal conditions with mock-witnesses. For example, participants were only instructed to provide a person description and did not have to report what happened or provide information about the scene; participants likely experienced no heightened emotions about the event (e.g. fear); and there were no potential areas for memory contamination (e.g. participants could not discuss the event with each other). Whilst difficult, real-world data (e.g. through Body Worn Video footage) would be useful when examining the number of descriptors (accuracy would not be possible as ground truth cannot be established) and how these vary depending on the race of the witness and/ or perpetrator.

Conclusion

Although the expectations about a cross-race effect were not fully met, the results of the current study demonstrate that witness race may have an impact on the reporting and accuracy of person descriptions, with a partial own-race bias. Due to the complexities of this

area of research and the conflicting findings of the current study, as well as previous research, future research needs to be conducted into the area of own-race bias with a particular focus on the elicitation of person description information.

Chapter 4

Interview techniques that facilitate recall and reporting of person descriptions: An examination of the use of self-generated cues

Foreword

Chapter 4 presents two experiments that aimed to increase the quantity and accuracy of person descriptions elicited from witnesses. When conducting a literature review (see Chapter 1), it was found that one reason person descriptions lack detail was because of inadequate interview techniques. This was supported by the research findings in Study 1 (see Chapter 2) that showed that police officers failed to follow best practice guidelines in their interviews with witnesses. Whilst Chapter 3 showed that using an enhanced person description form could increase the number of person descriptors, there was a promising separate line of research being conducted that suggested a new interview technique in which the interviewee provides the information themselves (using self-generated cues). The current set of experiments aimed to examine whether self-generated cues could increase the total number and accuracy of person descriptors. One of the constraints of previous research was that it focused on single perpetrator events. Study 1 (see Chapter 2) showed that a large proportion of crimes involved multiple perpetrators and thus the current chapter involved the design of two experiments to investigate single perpetrator events and multiple perpetrator events. I was responsible for designing the study (which was discussed with the supervisory team). I was then required to set up the study, create study materials, code and analyse the data, and prepare the study for publication. Dr Vernham assisted with statistical analyses. Mr May completed the inter-rater reliability. All authors are currently in the process of reading the manuscript prior to journal submission.

Abstract

Obtaining detailed and accurate person descriptions is important for the apprehension of perpetrators, particularly in the immediate aftermath of an incident, when guilty parties may still be in close proximity. Person descriptions often lack the level of detail necessary to assist in the apprehension of the perpetrators. In the present research (Study 3), two experiments were conducted to examine whether the use of self-generated cues (SGC) could enhance the reporting of person descriptions during a single perpetrator event (Experiment 1) and a multiple perpetrator event (Experiment 2). In both experiments, participants witnessed a mock-crime and provided a description of the perpetrator(s) using a mnemonic – SGC, person description cues (PDC) or a free recall (control). In both Experiment 1 and Experiment 2, mock-witnesses who provided a description using PDC provided more person descriptors than mock-witnesses who used SGC or a free recall. However, mock-witnesses were more accurate using SGC than PDC. In terms of grain size, both experiments showed that using PDC led to more fine-grain details being provided by mock-witnesses than using SGC or a free recall. However, again, mock-witnesses were more accurate in the reporting of fine-grain details when using SGC than PDC. The findings for coarse-grain details were inconsistent and dependent on whether mock-witnesses witnessed a single perpetrator or multiple perpetrators. Further findings and implications of the research are discussed.

Keywords: Eyewitness memory, Person description, Investigative interviewing, Mnemonics, Self-generated cues**,** Memory

Introduction

The successful progress of a criminal investigation depends on eliciting reliable and detailed information about what happened and who was involved (Milne & Bull, 2006). Obtaining person descriptions is important for police officers investigating any crime, particularly when locating a perpetrator(s) in the immediate aftermath of an incident (e.g. a physical assault). When providing an account of what they saw, witnesses are typically expected to be able to report everything they can remember about the incident, including a detailed person description of those involved.

Laboratory studies have found that witnesses typically provide between 5.3 and 15.9 physical details when describing a single perpetrator (e.g. Geiselman et al., 2000; Pozzulo & Warren, 2003; Wells et al., 1993), with an accuracy rate of approximately 83.8% (e.g. Geiselman et al., 2000; Meissner et al., 2001; Pozzulo & Warren, 2003). Whilst the average number of person descriptors is slightly higher in laboratory studies than in real-world studies (3.9 to 9.4 physical details), participants in laboratory studies often provide poor person descriptions even under good quality encoding conditions. For example, Sporer (2007) found participants only reported on average 4.46 person details even when they gave their descriptions immediately after viewing the target in a filmed event, despite the potential for a richer description given the stimuli. In terms of the content of person descriptions in the laboratory, Lindsay et al. (1994) found that some reference to height was included by 86% of participants, followed by references to build (51%), gender (46%), age (45%), ethnicity (43%) and weight (22%). With regard to facial features, upper features (e.g. hair, eyes) were mentioned more than lower features (e.g. chin, lips, mouth; Ellis et al., 1980; Laughery et al., 1986).

On the whole, to be informative for an investigator, person descriptions need to have what might be referred to as ‘diagnostic value’ (i.e. include sufficiently detailed information

to accurately identify a specific individual). To have 'diagnostic' value, the person description needs to enable an investigator to discriminate accurately in order to identify the actual perpetrator. As such, diagnostic value is reflected in the quantity and quality of the information provided in the person description. For instance, a person description with high diagnostic value will include detailed and accurate information pertaining to the physical appearance and clothing of the perpetrator. Vague, imprecise, incomplete, or inaccurate person descriptions lack diagnostic value and are problematic for the investigative process as they provide few clues for the police to find the perpetrator. Beyond estimator factors (such as poor encoding conditions), there are several reasons why person descriptions may lack diagnostic value. For the purposes of the current research, the focus will be on strategic monitoring of memory and inadequate interviewing techniques.

Strategic monitoring of memory

Witnesses adopt appropriate strategies when providing information from their memory (e.g. reporting a description of the perpetrator; Bjorklund et al., 2009) and control how much information they report based on how confident they are about the accuracy of their recollection and the demands of the situation (Koriat & Goldsmith, 1996). One way witnesses regulate their responses is by adjusting the precision of the reported information (control over grain size; Ackerman & Goldsmith, 2008). According to the satisficing model of the minimum-confidence criterion (Goldsmith et al., 2002), the interviewee starts by assessing a fine-grain answer which they will volunteer (i.e. report) if they believe it to be correct (e.g. he had piercing blue eyes). Otherwise, the interviewee will provide a coarse-grain answer (e.g. he had light eyes) in order to preserve accuracy. In the immediate aftermath of a crime when searches are often initiated to locate a perpetrator, more informative (fine-grain) details are preferable – and more likely to be diagnostic. However, if an interviewee is unsure of their recollection, fine-grain details are less likely to be reported

(cf. coarse-grain details; Koriat & Goldsmith, 1996; Weber & Brewer, 2008). When a premium is placed on accurate reporting (e.g. when providing person descriptions for the police), it is likely that witnesses attempt to enhance the accuracy of what they say by screening out information that is likely to be incorrect. Given that witnesses are unlikely to have extensively encoded the face, this strategic regulation may result in minimal detail about the perpetrator's face being provided by the witness.

The Dual-Criterion Model (Ackerman & Goldsmith, 2008) proposes that informativeness mediates reporting, in that even if coarse-grain details are more likely to be accurate (though imprecise; Evans & Fisher, 2011), they may be withheld if assessed by the interviewee as not sufficiently informative. In a mock-witness paradigm, McCallum et al. (2016) found that participants reported more fine-grain details than coarse-grain details about the mock-crime event, with a further study finding that mock-witnesses perceived fine-grain details to be more informative than coarse-grain details (McCallum et al., 2019). Observing a similar prioritisation of fine-grain details in mock-witness accounts, Brewer et al. (2018) suggested that while witnesses to a crime may be extremely confident about some coarse-grain details they remember, they may also believe the police would find such details uninformative (e.g. police already know the information or details are too general) and so tend not to report coarse-grain details spontaneously. This potential assumption by witnesses with respect to the informativeness of their knowledge is problematic as in the immediate search for the perpetrator, coarse-grain details could actually be invaluable for narrowing down the search. For instance, knowing someone is wearing a dark rather than a light jacket could inform a CCTV search for a target fleeing the scene.

The ability of a witness to regulate the quality and quantity of their reporting is impacted by the type of reporting instructions employed. Koriat and Goldsmith (1996) examined how participants responded to a general knowledge test when using free-report and

forced-report conditions. When free to choose whether to report an answer or not, participants answered questions that they were certain of, with participants who received high penalties for incorrect answers exercising a stricter criterion at reporting (cf. participants with moderate penalties). Participants in the free-report condition were also able to monitor the accuracy of their responses by volunteering answers that exceeded a pre-set criterion of confidence that the answer was correct (*satisficing model;* Ackerman & Goldsmith, 2008). Participants were more effective at monitoring their accuracy when they were able to control whether to volunteer or withhold an answer (*control of report option*). In order to enhance memory reporting of person description information, witnesses should be instructed not to guess responses and should be encouraged to give free recall accounts (thus, witnesses should not be interviewed using inappropriate or leading questions).

Using mnemonics to improve person descriptions

Several evidence-based techniques use mnemonics to support the reporting of detailed and reliable information by witnesses (e.g. the Cognitive Interview [CI]; Fisher & Geiselman, 1992; Geiselman et al., 1984). The CI is embedded in best-practice police interviewing and incorporates several distinct mnemonics. One of these is the mental reinstatement of context (MRC) instruction. The key premise of MRC is that recall is facilitated when there is an overlap between the conditions present at encoding and the conditions present at retrieval (*encoding-specificity principle*; Thomson & Tulving, 1970; Tulving & Thomson, 1973). The MRC instruction therefore encourages a witness to recreate the psychological (e.g. their emotional state, their thoughts) and physical environment (e.g. surroundings) that existed at the time of the to-be-remembered event via a set of pre-defined instructions. The administration of the MRC instruction has consistently been found to elicit more correct information compared to a free recall instruction (Dando et al., 2009; Davis et al., 2005; Memon & Bruce, 1985).

It is important to highlight that whilst the encoding-retrieval match appears to aid memory, it is the quality of cues that moderates the extent to which retrieval improves (Nairne, 2002). Cues effectively facilitate retrieval when they are distinctive in addition to satisfying the encoding-retrieval match (Tullis & Benjamin, 2015; Watkins & Watkins, 1975). When retrieval cues are associated with multiple details, they are not distinctive and they reduce the effectiveness of promoting the recall of target information. A distinctive cue uniquely matches a memory to the exclusion of other related memories (*principle of cue overload*; Nairne, 2002). Therefore, to be effective, cues need to be encoded within the context of the witnessed event (encoding specificity principle), and to offer diagnostic information identifying a single target to the exclusion of others, rather than matching multiple related targets (i.e. matching but not distinctive; Goh & Lu, 2012; Nairne, 2002). Previous research on eyewitness testimony has focused on interviewer-led cues to generate information (e.g. context reinstatement; Dando et al., 2009). However, interviewers cannot know what the most salient information to the interviewee is. Therefore, one solution is to ask interviewees themselves to provide such information using the self-generated cue mnemonic (Wheeler & Gabbert, 2017).

Self-generated cues (SGC) can be defined as salient details that are actively generated by the witnesses themselves and facilitate retrieval of a target memory (e.g. what the suspect looked like; Kontogianni et al., 2018; Wheeler & Gabbert, 2017). When witnesses recall episodic information, stored traces are activated and these subsequently prompt related details. Thus, this 'spreading activation' through an associative network (*activation theory*; Anderson, 1983) leads to the strengthening of memory traces across repeated retrieval attempts. Research has shown that the stronger a memory, the more likely it is that it will be recalled at a later date and that it will activate associated memories (Anderson, 1983). Similarly, Anderson and Conway (1993) showed that, when asked to list event details in free

recall, participants first listed ‘distinctive details’, before then listing other details that were highly associated with those distinctive details. Thus, self-generation of distinctive cues can trigger related memories by tapping a common theme (Anderson & Conway, 1993). The use of SGC should trigger the retrieval of related event details while excluding unrelated details, thus satisfying both the encoding-specificity principle (Tulving & Thomson, 1973), and the principle of cue overload (Nairne, 2002).

Previous research has focused on using SGC to examine whether the use of this mnemonic facilitates retrieval of target memories in eyewitness contexts. For example, Kontogianni et al. (2018) found that the use of SGC led to the reporting of more correct details from mock-witnesses (cf. other-generated cues and no cues). In terms of perpetrator descriptions, Wheeler et al. (2017) found (across two studies) that the use of SGC techniques increased reporting of person description details, with no cost to accuracy (cf. other-generated cues and free recall). To date, however, research has not examined whether the use of an SGC mnemonic can specifically improve the reporting of person descriptions by witnesses for both single and multiple perpetrator events. Most previous research (e.g. Kuehn, 1974; Meissner et al., 2001) has only addressed the reporting of person descriptions for a single perpetrator. This is problematic, as in the real-world, crimes are often committed by multiple perpetrators (McGloin & Piquero, 2009; van Mastrigt & Farrington, 2009). The limited previous research has shown that increasing the number of perpetrators leads to poorer person descriptions (Sporer, 1996; van Koppen & Lochun, 1997). Perhaps the use of SGC could increase the reporting of person descriptions for multiple perpetrators as using SGC would result in more unique associated memories about the event and this should facilitate higher rates of correct recall.

Current Study

Two experiments were conducted to test the effectiveness of using SGC in comparison to person description cues (PDC) and a free recall (control) with single perpetrator (Experiment 1) and multiple perpetrator (Experiment 2) events. Frontline police officers often use a short-answer format (i.e. checklist) to elicit person descriptions from witnesses. Therefore, to reflect real-life practice, the current study manipulated the use of PDC as part of the experimental design. As previously stated, the use of SGC has not been specifically examined in the context of person descriptions. Given that the use of SGC should prompt the most salient memories for mock-witnesses (cf. other generated cues) and the fact that SGC draws upon activation theory (i.e. mock witnesses start with the most salient memory and then report the next most salient memory and so on; Kontogianni et al., 2018), it was predicted that the use of SGC would result in a higher number of person descriptors than when using PDC or a free recall in both single perpetrator events (Hypothesis 1a) and multiple perpetrator events (Hypothesis 2a). Based on previous research using mock-witness paradigms (Kontogianni et al., 2018; Wheeler et al., 2017), it was predicted that the accuracy of reported person descriptions would be higher when using SGC, than when using PDC or a free recall, for both single perpetrator events (Hypothesis 1b) and multi-perpetrator events (Hypothesis 2b). It was also predicted that the use of SGC would activate unique associated memories and lead to the reporting of more fine-grain details and coarse-grain details in both single perpetrator events (Hypothesis 1c) and multiple perpetrator events (Hypothesis 2c), than the use of PDC or a free recall. Finally, it was predicted that the use of SGC would lead to a higher accuracy of fine-grain details and coarse-grain details in both single perpetrator events (Hypothesis 1d) and multiple perpetrator events (Hypothesis 2d), than the use of PDC or a free recall.

Experiment 1

Method

Participants and design

A power analysis revealed that a minimum of 25 participants would be sufficient in each group for the study to have good statistical power (.873) and a large effect size (η_p^2 = .138) at the .05 significance level.

A total of 103 participants were recruited through the Department of Psychology's participant pool. Participants were pseudo-randomly allocated to one of three conditions (Mnemonic Type: self-generated cues [SGC], person description cues [PDC], free recall) in a between-subjects design. Data were excluded for one participant who did not follow the instructions. The reported analysis is based on the data for the remaining 102 participants (M_{age} = 20.02 years, *SD* = 4.59, 79 females), with 34 participants allocated per group cell.

Materials

Stimulus event

The stimulus event was a short, recorded crime event lasting 58 seconds (see Figure 2 for crime timeline). The event showed a theft by a single male perpetrator from a building reception area. The event starts with the male perpetrator entering the reception area. He looks around before jumping over the reception desk and rifling through a woman's handbag that has been left unattended behind the desk. The perpetrator steals some items from the bag, before jumping back over the reception desk. At the end of the event the perpetrator exits the building with the stolen items. There was no audio.

Figure 2

Screenshots showing timeline of criminal event (Study 3, Experiment 1)

Self-Generated Cues Instructions

Instructions in the SGC condition were adapted from Gabbert et al. (2014) and Kontogianni et al. (2018). Participants were instructed to *'without thinking too hard, write down the first six things that you remember seeing or thinking about the person you saw in the film'*. They were then instructed to focus on each of these things, one at a time, and were encouraged to close their eyes or stare/look at a blank wall or the floor to help focus and concentrate on remembering the perpetrator (see Appendix B).

Person Description Cues Instructions

Participants in the PDC condition were provided with, and had to select from, a list of descriptors that frontline police officers often use to gather person descriptions. Similar to a checklist, participants were required to consider a set of specified descriptors including the perpetrators ethnicity, height, apparent age, sex and whether they had any marks, scars, or peculiarities. For build, hair colour, eyes, complexion and facial hair, participants were provided with a list of options to choose from (e.g. Eyes - Blue; Brown; Grey; Cast; Staring; Other). Participants then had to answer a yes/no response question to indicate whether the perpetrator wore glasses. Finally, participants were asked to comment on the perpetrators clothing. This section was multiple-choice, and participants had to select items of clothing that were provided (e.g. hat, jacket, dress, belt, shoes, etc.). Participants could then provide additional details about each item of clothing that they selected (see Appendix B).

Procedure

Ethical approval was obtained from the Faculty of Science and Health ethics committee at the University of Portsmouth (see Appendix C). Individuals arrived at the laboratory at pre-arranged times and completed consent procedures which informed them that the study would be about eyewitness memory. Participants were then asked to complete a short demographic survey which asked participants to self-identify their gender and ethnicity,

as well as provide their age. Participants were then randomly allocated to one of three conditions (SGC; PDC; free recall). All participants were given the following instructions prior to watching the short film: *During this study you will watch a short film. The film has no audio. However, please pay close attention to the film, as you will be asked questions about what you saw later on.*

After witnessing the event, all participants completed an 8-minute filler task (searching for items in a complex scene). Participants in all conditions were then instructed to *'provide a detailed person description of the person you saw in the event. Please remember that you are the only person who has seen the individual. Imagine that the success of this case hinges solely on the description you provide. I have no information about the person in the event other than what you write down. It is therefore important to write down as many details as possible when giving the description and take as much time as you require. A description can refer to physical descriptors of an individual but also items of clothing. Try to describe the person in enough detail so that someone else could use your description to pick them out of a crowd. Please do not guess about any detail that you are unsure of'*. To promote accurate reporting, participants were instructed not to guess (Memon et al., 2010). Depending on condition, participants also received instructions to use SGC or PDC. Participants in the free recall (control) condition did not receive any further instructions. Once participants had completed their description of the perpetrator they were thanked for their time and fully debriefed. Participation in the study took approximately 20 minutes.

Coding

A coding scheme was devised to categorise each person descriptor. The person descriptors reported by participants were coded according to previous research (see Lindsay et al., 1994). Categories included face items (e.g. face, mouth, nose), exterior face items (e.g. hair length, hair style, hair colour), body (e.g. height, weight), clothing (e.g. grey jogging

bottoms, white Lonsdale trainers, black Gillet), and accessory descriptors (e.g. brown leather watch). Race, age and gender were coded individually.

Similar to previous research (e.g. Pozzulo & Warren, 2003), age, height and weight were coded within an acceptable range (age ± 3 years; height ± 2 inches; weight ± 2kg). Any descriptor that exceeded the acceptable range (e.g. 'aged between 20 and 30 years') was considered an error and coded as incorrect. Details that were subjective or vague (e.g. 'he was neither tall nor short') were not coded for accuracy. To assess overall inter-rater reliability for the accuracy of person descriptors, 21 person descriptions were randomly selected and coded independently by a second rater who was blind to the experimental conditions. The inter-rater reliability was excellent, ICC = .95, 95% CI [.82, .98].

To assess the precision of elicited person descriptors, the grain size of reported information was coded, in terms of fine-grain and/or coarse-grain. Specifically, coding for descriptions with numerical values, and similar to Weber and Brewer (2008), and, Sauer and Hope (2016), responses including numbers were coded as fine-grain if they included up to three possible values (e.g. aged between 25 and 28 years). However, a response was referred to as coarse-grain if it included four or more possible values (e.g. aged 25 to 35 years). For categorical data, pre-determined parameters defined when responses were considered fine-grain (e.g. bushy eyebrows). Vague and subjective responses (i.e. 'about my height') were not coded. To assess overall inter-rater reliability for grain size, 21 person descriptions were randomly selected and coded independently by a second rater who was blind to the experimental conditions. Inter-rater reliability for accuracy of fine-grain details was excellent, ICC = .94, 95% CI [.87, .98] and inter-rater reliability for accuracy of coarse-grain details was acceptable, ICC = .72, 95% CI [.33, .89].

Each participant received a score for the total number of descriptors provided, a score for the proportion of correct descriptors provided (correct descriptors divided by total number

of descriptors), a score for total number of fine-grain details and coarse-grain details, and a score for the proportion of fine-grain details (correct fine-grain details divided by total number of fine-grain details) versus coarse-grain details (correct coarse-grain details divided by total number of coarse-grain details) provided. These newly computed scores then formed the dependent variables in the current study.

Results and Discussion

For the hypothesis-testing analyses conducted below, the between-subjects factor in each of the one-way ANOVAs was always Mnemonic Type (self-generated cues vs. person description cues vs. free recall). Unless otherwise stated, the assumption for Levene's Test for Homogeneity of Variance was met (*p*-values ranged from .168 to .805), therefore equal variances can be assumed.

Reporting of person descriptors

A one-way between-subjects ANOVA with total number of person descriptors reported as the dependent variable revealed a significant main effect for Mnemonic Type, $F(2, 99) = 28.76$, $p < .001$, $\eta_p^2 = .37$. Post-hoc tests using Tukey HSD showed that the use of person description cues led to the reporting of significantly more person descriptors than the use of self-generated cues, $p < .001$, and a free recall, $p < .001$ (see Table 11). There was no difference between using self-generated cues and using a free recall, $p = .904$, in terms of the number of person descriptors reported.

Accuracy of person descriptors

A one-way between-subjects ANOVA with accuracy of person descriptors provided by participants as the dependent variable revealed a significant main effect for Mnemonic Type, $F(2, 99) = 25.92$, $p < .001$, $\eta_p^2 = .34$. Post-hoc tests using Tukey HSD showed that the use of self-generated cues led to a higher accuracy for person descriptors than the use of person description cues, $p < .001$. Free recall reporting also resulted in a higher accuracy for

person descriptors than the use of person description cues, $p < .001$ (see Table 11). There was no difference between the use of self-generated cues and a free recall, $p = .672$, in terms of the accuracy of person descriptors reported.

Reporting of fine-grain and coarse-grain details

A one-way between-subjects ANOVA with total number of fine-grain details as the dependent variable revealed a significant main effect for Mnemonic Type, $F(2, 99) = 20.73$, $p < .001$, $\eta_p^2 = .30$. Post-hoc tests using Tukey HSD showed that using person description cues led to significantly more fine-grain details than using self-generated cues, $p < .001$, and a free recall, $p < .001$ (see Table 11). There was no difference between using self-generated cues and using a free recall, $p = .856$, in terms of the number of fine-grain details reported.

A one-way between-subjects ANOVA with total number of coarse-grain details as the dependent variable revealed a significant main effect for Mnemonic Type, $F(2, 99) = 3.96$, $p = .022$, $\eta_p^2 = .07$. Post-hoc tests using Tukey HSD showed that the use of person description cues led to significantly more coarse-grain details than using a free recall, $p = .028$. There were no differences between the use of self-generated cues and the use of person description cues, $p = .075$, or between the use of self-generated cues and a free recall, $p = .912$ (see Table 11).

Accuracy of fine-grain and coarse-grain details

A one-way between-subjects ANOVA with the accuracy of fine-grain details reported by participants as the dependent variable revealed a significant main effect for Mnemonic Type, $F(2, 99) = 27.02$, $p < .001$, $\eta_p^2 = .35$. Post-hoc tests using Tukey HSD showed that the use of self-generated cues led to a higher accuracy for fine-grain details than the use of person description cues, $p < .001$. Free recall reporting also resulted in a higher accuracy for fine-grain details than using person description cues, $p < .001$ (see Table 11). There was no

difference between using self-generated cues and using a free recall, $p = .740$, in terms of the accuracy of fine-grain details.

A one-way between-subjects ANOVA with the accuracy of coarse-grain details reported by participants as the dependent variable was conducted. Levene's Test for Homogeneity of Variance was not met ($p = .007$). Nevertheless, analyses of variance are robust to issues of normality (Lantz, 2013). The ANOVA revealed that there was no significant differences between using self-generated cues, person description cues, and a free recall, $F(2, 99) = 2.74$, $p = .069$, $\eta_p^2 = .05$ (see Table 11).

Table 11

Descriptive statistics for the number and accuracy of person descriptors reported (Study 3, Experiment 1)

Dependent Variable	**Type of instruction**		
	Self-generated cues	**Person description cues**	**Free recall**
	Mean *(SD), 95% CI*	Mean *(SD), 95% CI*	Mean *(SD), 95% CI*
Total number of person descriptors	15.88 (3.45), *13.86-16.19*	21.15 (3.92). *19.82-22.48*	15.47 (3.69), *14.24-16.75*
Overall accuracy of person descriptors	0.86 (0.10), *0.82-0.89*	0.71 (0.11), *0.67-0.75*	0.88 (0.92) *0.85-0.91*
Total number of fine-grain details	13.38 (3.38), *12.26-14.55*	18.53 (4.06), *17.21-20.00*	13.85 (3.45), *12.61-14.95*
Overall accuracy of fine-grain details	0.88 (0.09), *0.85-0.91*	0.74 (0.11), *0.70-0.78*	0.90 (0.08), *0.87-0.93*
Total number of coarse-grain details	1.71 (1.17), 1.32-2.14	2.50 (1.78), *1.93-3.14*	1.56 (1.44), *1.09-2.05*
Overall accuracy of coarse-grain details	0.64 (0.44), *0.49-0.79*	0.41 (0.34), 0.31-0.53	0.58 (0.46), *0.41-0.73*

Summary of Experiment 1

The aim of Experiment 1 was to investigate the use of self-generated cues (SGC) in comparison to using person description cues (PDC) or a free recall when mock-witnesses viewed a single perpetrator event. The current experiment is the first to evaluate the use of SGC with only person descriptors. In contrast to Hypothesis 1a, mock-witnesses who used PDC reported more person descriptors than mock-witnesses who used SGC or a free recall, but no difference was found in terms of the number of person descriptors when using SGC compared to a free recall. Unlike the current experiment, Wheeler et al. (2017) found a benefit to using SGC (cf. free recall) to obtain person-details in a mock-crime scenario. In line with Hypothesis 1b, mock-witnesses who used SGC were more accurate with the person descriptors they reported compared to mock-witnesses who used PDC. In contrast to Hypothesis 1b, there was no difference between the accuracy of person descriptors for SGC and free recall. This is line with Kontogianni et al. (2018) who found that mock-witnesses who used SGC reported more correct details than mock-witness in other conditions. However, it is important to note that the study by Kontogianni et al. (2018) measured all types of details (e.g. person, action, setting details) and not just person descriptors; hence, the comparison of using SGC versus a free recall has not previously been examined in the context of the current experiment which is solely person descriptions. In refute of Hypothesis 1c, the use of PDC led to the reporting of more fine-grain details than the use of SGC or a free recall. Also in contrast to Hypothesis 1c, there were no differences in the reporting of coarse-grain details when using SGC, PDC, or a free recall. When mock-witnesses view a single perpetrator event, the use of PDC leads to more person descriptors being reported than the use of SGC, but to the detriment of accuracy. Of most importance, the use of SGC led to more accurate person descriptors being provided than the use of PDC, which is essential to ensuring the apprehension of the correct perpetrator during frontline policing.

In partial support of Hypothesis 1d, the use of SGC led to a higher accuracy rate for fine-grain details about the perpetrator than the use of PDC, but the use of a free recall also led to a higher accuracy rate for fine-grain details than the use of PDC. No difference was found when using SGC versus a free call in terms of accuracy of fine-grain details reported. Based on the previous literature (e.g. Kontogianni et al., 2018), it comes as no surprise that the spontaneous reporting of person descriptions, such as through the use of SGC and a free recall, outperformed (i.e. led to a higher accuracy of fine-grain details) the forced-choice style of reporting, such as the use of PDC. In refute of Hypothesis 1d, there were no differences between the use of SGC, PDC and a free recall in terms of the accuracy of coarse-grain details reported about the perpetrator. However, it is important to note that the overall number of coarse-grain details reported in the current experiment was extremely low. This is in line with previous research that mock-witnesses hold back on the reporting of coarse-grain details (Brewer et al., 2018).

Experiment 2

Method

Participants and design

As in Experiment 1, a power analysis revealed that a minimum of 25 participants would be sufficient in each group for the study to have good statistical power (.873) and a large effect size ($\eta_p^2 = .138$) at the .05 significance level.

A total of 101 participants were recruited through the Department of Psychology's participant pool and through advertisements online (e.g. Facebook). The study involved a one-way between-subjects design with participants pseudo-randomly allocated to one of three conditions (Mnemonic Type: self-generated cues [SGC], person description cues [PDC], free recall). Data were excluded for eleven participants who did not follow the instructions ($N = 5$) or contained too many missing data ($N = 6$). The reported analysis is based on the data for the

remaining 90 participants (M_{age} = 22.90 years, *SD* = 4.91, 68 females), with 30 participants allocated per group cell.

Materials

Stimulus event

The stimulus event was a recorded crime event, involving multiple perpetrators which lasted 1 minute 58 seconds (see Figure 3 for crime timeline). The event showed a theft from a parked car in a residential street. At the start of the event three male targets who were very distinct from one another (e.g. different clothing, different accessories, different appearance, etc.) can be seen sitting on a garden wall (in front of a terraced property). The men jump off the wall and walk down a street of parked cars. The men can be seen scanning the area and trying a number of car doors before finding one car that has been left unlocked. One of the group opens the passenger side door and a second man walks around the car and opens the back-passenger door. The third man keeps watch. The men then leave with some stolen items (one man can be seen putting something in his pocket, another man is seen leaving with a laptop which he places under his jacket). There was no audio.

The instructions for participants in Experiment 2 (e.g. SGC instructions, PDC instructions) were the same as in Experiment 1. Participants were asked to provide descriptions of each perpetrator, starting with the person most clear in their memory before moving on to report a description of the next perpetrator and so on.

Figure 3

Screenshots showing timeline of criminal event (Study 3 Experiment 2).

Procedure and coding

The procedure and coding were the same as in Experiment 1. However, as this was a multiple perpetrator experiment compared to a single perpetrator experiment, participants were required to provide three person descriptions instead of the one person description (as was the case in Experiment 1). All participants in Experiment 2 viewed the same three perpetrators. Participants were first instructed to describe the perpetrator that they felt they had the best memory of. A person description detail was scored as accurate if it was present in the stimulus event and described correctly. Details that were subjective or vague were not coded. Subsequently, once participants provided a description of the first perpetrator, they were then instructed to provide a description of the next perpetrator, and so on. Each perpetrator was distinct (see Figure 3). For example, one perpetrator wore a green hoodie, one wore a blue denim jacket, and, one wore a red hoodie. In addition, one perpetrator had glasses, the other two perpetrators did not have glasses. Participation in the study took approximately 30 minutes.

To assess overall inter-rater reliability for accuracy of person descriptors in the multiple perpetrator event, 18 person descriptions were randomly selected and coded independently by a second rater blind to experimental conditions. The inter-rater reliability was excellent, ICC = .99, 95% CI [.98, 1.00]. To assess overall inter-rater reliability for grain size in the multiple perpetrator event, 18 person descriptions were randomly selected and coded independently by a second rater blind to experimental conditions. Inter-rater reliability for accuracy of fine-grain details was excellent, ICC = .99, 95% CI [.97, 1.00] and inter-rater reliability for accuracy of coarse-grain details was also excellent, ICC = .97, 95% CI [.92, .99].

Results and Discussion

For the hypothesis-testing analyses conducted below, the between-subjects factor in each of the one-way ANOVAs was always Mnemonic Type (self-generated cues vs. person generated cues vs. free recall). Unless otherwise stated, the assumption for Levene's Test for Homogeneity of Variance was met (*p*-values ranged from .104 to .803), therefore equal variances can be assumed.

Reporting of person descriptors

A one-way between-subjects ANOVA with total number of person descriptors as the dependent variable revealed a significant main effect for Mnemonic Type, $F(2, 87) = 18.08$, $p < .001$, $\eta_p^2 = .29$. Post-hoc tests using Tukey HSD showed that the use of person description cues led to significantly more person descriptors than the use of self-generated cues, $p = .033$, and a free recall, $p < .001$ (see Table 12). The use of self-generated cues led to significantly more person descriptors than using a free recall, $p = .003$.

Accuracy of person descriptors

A one-way between-subjects ANOVA with accuracy of person descriptors provided by participants as the dependent variable revealed a significant main effect for Mnemonic Type, $F(2, 87) = 7.02$, $p = .001$, $\eta_p^2 = .14$. Post-hoc tests using Tukey HSD showed that the use of self-generated cues led to a higher accuracy rate for person descriptors than the use of person generated cues, $p = .002$. Free recall led to a higher accuracy of person descriptors than the use of person generated cues, $p = .015$ (see Table 12). There was no difference in the accuracy of person descriptors when using self-generated cues compared to a free recall, $p = .782$.

Reporting of fine-grain and coarse-grain details

A one-way between-subjects ANOVA with total number of fine-grain details as the dependent variable revealed a significant main effect for Mnemonic Type, $F(2, 87) = 16.13$, $p < .001$, $\eta_p^2 = .27$. Post-hoc tests using Tukey HSD showed that the use of person generated

cues led to significantly more fine-grain details than the use of self-generated cues, $p = .034$, and a free recall, $p < .001$ (see Table 12). The use of self-generated cues led to significantly more fine-grain details than using a free recall, $p = .007$.

A one-way between-subjects ANOVA with total number of coarse-grain details as the dependent variable was conducted. Levene's Test for Homogeneity of Variance was not met ($p = .009$). Nevertheless, analyses of variance are robust to issues of normality (Lantz, 2013). The ANOVA revealed a significant main effect for Mnemonic Type, $F(2, 87) = 4.95$, $p = .009$, $\eta_p^2 = .10$. Post-hoc tests using Tukey HSD showed that the use of person generated cues led to significantly more coarse-grain details than using a free recall, $p = .015$. The use of self-generated cues led to significantly more coarse-grain details than using a free recall, $p = .030$ (see Table 12). There was no difference between the use of self-generated cues and the use of person description cues in terms of the number of coarse-grain details reported, $p = .965$.

Accuracy of fine-grain and coarse-grain details

A one-way between-subjects ANOVA with the accuracy of fine-grain details reported by participants as the dependent variable revealed a significant main effect for Mnemonic Type, $F(2, 87) = 8.53$, $p < .001$, $\eta_p^2 = .16$. Post-hoc tests using Tukey HSD showed that the use of self-generated cues led to a higher accuracy for fine-grain details than the use of person description cues, $p < .001$. Free recall reporting also resulted in a higher accuracy for fine-grain details than the use of person description cues, $p = .045$ (see Table 12). There was no difference when using self-generated cues versus a free recall in terms of the accuracy of fine-grain details, $p = .218$.

A one-way between-subjects ANOVA with the accuracy of coarse-grain details reported by participants as the dependent variable revealed a significant main effect for Mnemonic Type, $F(2, 87) = 6.34$, $p = .003$, $\eta_p^2 = .13$. Post-hoc tests using Tukey HSD

showed that using a free recall led to a higher accuracy for coarse-grain details than using self-generated cues, $p = .005$, and using person generated cues, $p = .012$ (see Table 12). There was no difference between the use of self-generated cues and the use of person description cues in terms of accuracy of coarse-grain details, $p = .960$.

Table 12

Descriptive statistics for the number and accuracy of person descriptors reported (Study 3, Experiment 2)

Dependent Variable	**Type of instruction**		
	Self-generated cues	**Person description cues**	**Free recall**
	Mean *(SD), 95% CI*	Mean *(SD), 95% CI*	Mean *(SD), 95% CI*
Total number of person descriptors	38.13 (11.43), *34.41-42.41*	45.27 (11.42), *41.15-49.13*	28.50 (9.57), 25.18-32.05
Overall accuracy of person descriptors	0.85 (0.09), *0.81-0.88*	0.76 (0.09), *0.73-0.79*	0.83 (0.11), *0.79-0.87*
Total number of fine-grain details	32.73 (10.11), *29.14-36.26*	39.27 (10.98), *35.21-43.20*	24.70 (8.62), *21.62-27.65*
Overall accuracy of fine-grain details	0.87 (0.09), *0.84-0.90*	0.76 (0.11), *0.72-0.80*	0.83 (0.11), *0.78-0.87*
Total number of coarse-grain details	5.40 (3.00), *4.31-6.41*	5.60 (3.74), *4.26-7.00*	3.37 (2.19), *2.56-4.14*
Overall accuracy of coarse-grain details	0.65 (0.32), *0.53-0.76*	0.67 (0.32), *0.54-0.78*	0.90 (0.27), *0.78-0.99*

Summary of Experiment 2

The aim of Experiment 2 was to investigate the use of self-generated cues (SGC) in comparison to using person description cues (PDC) and a free recall when mock-witnesses viewed a multiple perpetrator event. In contrast to Hypothesis 2a, mock-witnesses who used

PDC reported more person descriptors than mock-witnesses who used SGC or a free recall. This goes against recent research that showed a benefit of using SGC for person-details in a mock-crime scenario (Kontogianni et al., 2018). However, in line with Hypothesis 2a, the use of SGC led to more person descriptors being reported than using a free recall. A potential reason for the increase in person descriptions when using SGC compared to a free recall could be the initial identification of six person details that the mock-witnesses were asked to provide prior to recall. Research suggests that an initial list of details could activate stronger memories (Anderson, 1983); thus, facilitating further recall for person descriptions.

In line with Hypothesis 2b, mock-witnesses who used SGC were more accurate with the person descriptors they reported compared to mock-witnesses who used PDC. In contrast to Hypothesis 2b, there was no difference between the accuracy of person descriptors reported when using SGC compared to a free recall. These findings are similar to Experiment 1 and in line with Kontogianni et al. (2018). In refute of Hypothesis 2c, the use of PDC led to the reporting of more fine-grain details than the use of SGC and a free recall. However, in support of Hypothesis 2c, the use of SGC led to more fine-grain details and more coarse-grain details than the use of a free recall. There was no difference between using PDC and using SGC in terms of coarse-grain details (refuting Hypothesis 2c).

In partial support of Hypothesis 2d, the use of SGC led to a higher accuracy rate for fine-grain details about the perpetrators than the use of PDC, but the use of a free recall also led to a higher accuracy rate for fine-grain details than the use of PDC. No difference was found when using SGC versus a free call in terms of accuracy of fine-grain details reported for multiple perpetrators. In refute of Hypothesis 2d, the use of a free recall led to a higher accuracy rate for coarse-grain details than the use of PDC and the use of SGC, but there was no difference in accuracy of coarse-grain details when using PDC versus using SGC. However, as with Experiment 1, it is important to note that the overall number of coarse-

grain details reported in Experiment 2 was extremely low, so the findings associated with coarse-grain accuracy need to be interpreted with caution.

When mock-witnesses view a multiple perpetrator event, the use of PDC leads to more person descriptors being reported than the use of SGC, but to the detriment of accuracy. As with Experiment 1, the use of SGC led to more accurate person descriptors being provided than the use of PDC, which is essential to ensuring the apprehension of the correct perpetrators during frontline policing.

General Discussion

The current two-part experiment tested the effectiveness of using self-generated cues (SGC) in comparison to using person description cues (PDC) and a free recall when mock-witnesses viewed a single perpetrator event (Experiment 1) or a multiple perpetrator event (Experiment 2).

In both Experiment 1 and Experiment 2, the use of SGC did not elicit as many person descriptors as using PDC. Whilst the use of PDC produced a more exhaustive recall of person descriptors across both single perpetrator and multiple perpetrator events, the accuracy of such descriptors was poor, with the use of SGC leading to a higher accuracy rate of person descriptors than the use of PDC in both single perpetrator events and multiple perpetrator events. One of the primary aims of eliciting a person description on the frontline is the quick apprehension of the perpetrator(s) in the immediate vicinity of the incident. Hence, it is essential that the person descriptions are accurate in order to apprehend the correct perpetrator(s). The reason for the low accuracy rates when using PDC could be because the use of PDC forces witnesses to consider particular features (e.g. eye colour), even when they have no memory for these features. Feature checklists that are similar to using PDC have been shown to produce more incorrect person descriptors than a free recall task (see Wogalter, 1991, 1996), which is problematic given that the use of PDC is similar to frontline

policing in the real-world. Further research is needed to examine how person descriptions are best elicited using mnemonics, such as using SGC.

In the current study, the comparison of using SGC versus a free recall showed that the use of SGC led to more person descriptors and grain size details than a free recall, but only for the multiple perpetrator event (Experiment 2), and not during the single-perpetrator event (Experiment 1). Experiment 2 was a more difficult task than Experiment 1 as mock-witnesses had to encode more information in Experiment 2 than in Experiment 1 (i.e. three perpetrators instead of one perpetrator). It is therefore possible that the person cues were more distinctive in the multiple perpetrator event (Experiment 2), which resulted in mock-witnesses being able to offer more diagnostic information about an individual target than when viewing the single perpetrator event (Experiment 1). Furthermore, it may well be that the use of a retrieval technique, such as using self-generated cues, is of greater benefit when completing a more challenging task (e.g. in the case of providing person descriptors about multiple perpetrators compared to a single perpetrator). Research suggests that whilst increasing the match between encoding and retrieval is necessary to facilitate recall, it is the distinctive quality of the cues that moderates the quantity of information that is retrieved (Goh & Lu, 2012; Nairne, 2002). Further research is needed to explore the use of SGC with single and multiple perpetrators, with future research focusing on the mechanisms behind why the use of SGC may be better for multiple perpetrators than a single perpetrator.

Interestingly, in both experiments, the overall number of coarse-grain details was low and the overall number of fine-grain details was, in comparison, high. This prioritisation of fine-grain details supports previous research using mock-witness accounts (e.g. Brewer et al., 2018; McCallum et al., 2016). One reason why mock-witnesses may be under-reporting coarse-grain details is because they believe the coarse-grain estimates are unhelpful, with the potential to violate social norms of communication. Witnesses might also believe that

frontline police officers would ridicule or scorn them as they believe the police would already know such information (Yaniv & Foster, 1995). Brewer et al. (2018) argued that this potential assumption by witnesses about the informativeness of their knowledge is problematic. For example, frontline police officers attending a scene and carrying out an immediate search for the perpetrator may find coarse-grain details invaluable in narrowing down a search. Promisingly, mock-witnesses were more accurate when they spontaneously reported fine-grain details (e.g. using SGC or a free recall), compared to when using a forced-choice style of reporting (e.g. using PDC). However, the findings for accuracy of coarse-grain details were inconsistent between the single perpetrator event and the multiple perpetrator event, which is likely to be a result of the very low numbers of coarse-grain details reported by mock-witnesses across both experiments. Future research needs to investigate why coarse-grain information is at times under-reported in eyewitness memory accounts and whether mnemonics could be used to enhance the recall of coarse-grain details.

The use of PDC requires a very different task to that of using SGC or a free recall. For example, when using PDC, participants did not need to regulate the precision of their responses as the use of PDC provided pre-set responses. Thus, individuals did not have to monitor the likely accuracy of their memories, nor did they have to control their output to optimise their reports. Additionally, when using the PDC, the items already contained person descriptors that would be classified as fine-grain details (e.g. blue eyes). Participants who simply guessed could therefore have performed better in terms of number of person descriptors, compared to when using SGC or a free recall. Whilst the use of SGC showed an apparent lack of benefit compared to the use of PDC in terms of the number of person descriptors reported for single and multiple perpetrator events, the person descriptors elicited when using SGC were more accurate than those elicited using PDC. Furthermore, whilst the use of SGC showed no benefit to grain size reporting, the use of SGC was again more

accurate for the reporting of fine-grain details than the use of PDC. This accuracy trade-off for overall person descriptors and fine-grain details suggests that the use of SGC would be more appropriate for frontline police officers were accurate person descriptors are vital during the immediate search for a perpetrator. That is, a high number of [inaccurate] person descriptors (e.g. from using PDC) could lead to the arrest of an innocent individual or the actual perpetrator being able to leave the scene.

Conclusion

The use of SGC instead of the 'typical' frontline police style checklist may support the interviewing process by facilitating an open-ended, predominately self-administered style report. This is important during frontline police interviews where scenes are often dynamic and involve multiple witnesses and few police officers. Dalton et al. (2021) has already shown that in the real-world just over 50% of questions asked by frontline police officers to elicit person descriptions were inappropriate or leading. The use of SGC allows witnesses to report details in their own words and is likely to be more accurate for both single and multiple perpetrator events than using person description cues. Notably, for applied contexts, the better accuracy when using SGC gives the police a greater chance of catching the perpetrator(s).

Chapter 5

Thesis General Discussion

The overarching aim of the current thesis was to explore how person descriptions are obtained during frontline interviews, and examine the effectiveness of information elicitation techniques in facilitating the reporting of person description information. Beyond estimator factors (such as poor encoding conditions), there are several reasons why person descriptions may lack diagnostic value. For the purposes of the current research, I focused on strategic monitoring of memory and inadequate interviewing techniques. For this thesis, I began by exploring how person descriptions were elicited in real-world frontline police interviews. Next, I conducted laboratory-based research to examine whether mock-witnesses demonstrated an own-race bias when reporting person descriptions, and I then tested whether mock-witnesses use of self-generated cues facilitated the reporting of person description information, across single and multiple perpetrator events.

This general discussion provides an overview of the main findings, followed by a consideration of theoretical and practical implications for eliciting person descriptions from witnesses during frontline interviews. Finally, the limitations of the current research are presented alongside routes for future research.

Summary of Thesis Findings

In Study 1 (Chapter 2), analyses of real-world body worn video footage showed that when obtaining person description information from witnesses, frontline police officers often asked witnesses a series of short-answer specific questions (e.g. Was it a white male? Southern or northern accent? How old was he? Was he black or white?). This finding supports previous research on police interviewing (Brown et al., 2008; Clifford & George, 1996). These closed-ended questions were often inappropriate and leading. Poor question formulation is known to minimise the amount of perpetrator (and other) information reported by witnesses (e.g. Fisher, 1995; Wright & Alison, 2004). In summary, the findings from Study 1, suggest that the way in which frontline police officers question witnesses when

obtaining a description likely contributes to witnesses providing impoverished, and potentially inaccurate, descriptions of perpetrators. In addition to poor interviewing practices, other errors were observed with frontline police officers interviewing witnesses together, providing witnesses with feedback on their descriptions of perpetrators (e.g. “Yes, you are correct, he did have a London accent”), and interrupting witnesses whilst they were providing their descriptions of the perpetrator. All of these poor practices (e.g. providing feedback, interrupting witnesses’ narrative) have been shown to increase the likelihood of memory contamination, which reduces the accuracy of information reported to frontline police officers by witnesses (see Fisher, 1995). It is therefore not necessarily the case that witnesses are not able to provide detailed person descriptions, but that inadequate interview techniques are being used which limit the detail witnesses provide. This conclusion is consistent with Sporer (1996) who suggested that incomplete perpetrator descriptions should not be considered as evidence that witnesses are unable to provide underlying perpetrator information.

In Study 2 (Chapter 3), I conducted laboratory-based research that explored cross-race differences in the reporting of person descriptions in South Africa, with white and black mock-witnesses. The results of Study 2 indicate that while white mock-witnesses were able to provide more person descriptors than black mock-witnesses, there was no own-race bias in terms of the number of person descriptors reported. Differences were also found in the individual types of person descriptors that were reported by black and white mock-witnesses. For example, white mock-witnesses provided descriptors about the face of the perpetrator more than black mock-witnesses, irrespective of the race of the perpetrator. The current finding is somewhat at odds with earlier research (e.g. Chen & Geiselman 1993; Ellis et al., 1975). However, given the location in which the research was conducted and the increasingly globalised nature of modern life, there are likely a number of explanations for this pattern of results.

In terms of explaining the differences in individual reporting of types of person details, Sporer (2001) argued that participant's familiarity with members of their own-race, means they have learnt to pay greater attention to these specific features as useful cues to distinguish individual members of their own-race. It is therefore possible that participants then use these facial features when evaluating different-race faces, even if a different set of features might be more informative when providing descriptions of different-races. Additionally, irrespective of participant race, an own-race bias was obtained for accuracy of person descriptors. In terms of grain size, white mock-witnesses provided more coarse-grain details than black mock-witnesses and an own-race bias was found for white participants, but not black participants. In terms of accuracy of grain size, an own-race bias was present for both white and black participants in the accuracy of fine-grain details, but not coarse-grain details. The increase in the reporting of fine-grain details for descriptions of perpetrators from a witness's own-race suggests that witnesses describing a perpetrator from their own-race are better at balancing informativeness and accuracy compared with witnesses who describe a perpetrator from a different-race (Koriat & Goldsmith, 1996). Given the limited amount of research to date that has examined cross-race differences in witnesses reporting of person descriptions, Study 2 makes a novel contribution to the literature. However, future research is needed to replicate these findings before any firm conclusions can be made. Future research should aim to extend this area of research to other comparisons and different cross-race groups, whilst also taking into account the role of culture.

Study 3 (Chapter 4), looked to examine whether the use of a mnemonic (self-generated cues) could improve the reporting and accuracy of person description information during a single perpetrator event (Experiment 1) and a multiple perpetrator event (Experiment 2). In Experiment 1, the use of self-generated cues led to less person descriptors being reported than using person description cues, but using self-generated cues and a free recall

led to a similar number of person descriptors being reported by mock-witnesses. For the reporting of grain size details, using person description cues led to more fine-grain details being reported than using both self-generated cues and a free recall; however, the use of person description cues only led to more coarse-grain details than a free recall. In terms of accuracy of person descriptors and accuracy of grain size, the use of self-generated cues and a free recall both outperformed the use of person description cues, with the use of self-generated cues and a free recall achieving similar accuracy rates for both person descriptors and fine-grain details. There were no differences between the accuracy rates of coarse-grain details for the use of self-generated cues, person description cues and a free recall.

In Experiment 2, the use of self-generated cues led to less person descriptors being reported than the use of person description cues, but more person descriptors being reported than the use of a free recall. A potential reason for the increase in person descriptions when using self-generated cues compared to a free recall, could be the initial identification of six-person details that the mock-witnesses were asked to provide, in the self-generated cues condition, prior to recall. Research suggests that an initial list of details could activate stronger memories (Anderson, 1983); thus, facilitating further recall for person descriptions. In terms of accuracy of person descriptors, the use of self-generated cues outperformed the use of person description cues, but there were no differences in accuracy rates when using self-generated cues compared to using a free recall. For the reporting of grain size details, the use of person description cues led to more fine-grain details being reported than the use of self-generated cues or the use of a free recall; however, the use of self-generated cues led to more fine-grain details than the use of a free recall. The use of person description cues and the use of self-generated cues led to more coarse-grain details being reported than the use of a free recall, but there was no difference in the reporting of coarse-grain details when using person description cues compared to using self-generated cues. When considering the

accuracy of grain size reporting, the use of self-generated cues and a free recall outperformed the use of person description cues for accuracy of fine-grain details, with no difference being obtained for the use of self-generated cues versus a free recall. Unexpectedly, the use of a free recall led to better accuracy rates for coarse-grain details than the use of self-generated cues or person description cues. No difference was found between the use of self-generated cues and the use of person description cues in terms of accuracy of coarse-grain details.

For the total number of person descriptors reported, the findings of Study 3 (Experiment 1 and 2; Chapter 4) fail to replicate previous research that has shown a benefit of using self-generated cues (cf. free recall) for person-details in a mock-crime scenario with single perpetrators (Wheeler et al., 2017) and multiple perpetrators (Kontogianni et al., 2018). Additional research is needed to examine whether using mnemonics can improve the reporting of total number of person descriptions by witnesses in both single and multiple perpetrator events, so that frontline police officers can elicit detailed and accurate person descriptions. Specifically, research should aim to examine what the underlying mechanisms are that contribute to a higher accuracy of person descriptors when using self-generated cues (cf. person generated cues), and aim to explain the limitations of their effectiveness in increasing the total number of person descriptors (cf. person generated cues and free recall).

Implications for theory and research on person recall and description

Providing accurate and detailed person descriptions about a previously witnessed perpetrator(s) is a complex process. For a number of psychological reasons (e.g. stress and trauma), even cooperative witnesses do not report all the information they know (Fisher, 2010). Consistent with any memory task, the number of details and accuracy of details reported of a perpetrator(s) is likely to be influenced by a number of factors including those present at encoding. One estimator factor that may influence description performance is cross-race differences between the witness and the perpetrator. Although more than 60

studies have investigated recognition memory for own-race versus different-race faces (for a review see Meissner & Brigham, 2001), very few studies have attempted to explore whether cross-race differences exist in the reporting of person descriptions. Study 2 (Chapter 3) produced conflicting findings with an own-race bias for accuracy but not for the total number of person descriptions reported. Fallshore and Scholar (1995) suggested that the differences between cross-race face recognition and cross-race description may be down to differences in the reliance on configural processing (recognition, i.e. processing not just the shapes of individual features but also processing the relations among the features of the face) versus featural processing (description, i.e. processing individual features of the face; Maurer et al., 2002; Rhodes et al., 1989). However, as there is a dearth of research examining own-race bias for person descriptions, calls have been made for more empirical research to be conducted into the factors that impact the reporting of person description information (e.g. Lindsay, 2012; Meissner et al., 2007). Specifically, there is a need for more research that explores whether the fundamental differences in the process associated with face recognition versus description account for cross-race differences.

The current research also drew on two memory theories to consider (i) how witnesses may regulate the reporting of their memories (Koriat & Goldmsith, 1996) and (ii) how mnemonics may be used to improve the reporting of person descriptions, by capitalising on the associative nature of memory (e.g. Anderson, 1983). If cross-race differences exist when providing descriptions, it might be because witnesses regulate the information they provide about other-races (i.e. witnesses balance the informativeness and accuracy of person descriptions of different-races by monitoring their precision). Witnesses are already unlikely to have extensively encoded the face or person details (regardless of race), so additional complexity as a function of race could increase metacognitive monitoring and result in fewer person description details being provided compared to when there is no cross-race difference

(i.e. the witness and perpetrator are of the same race). When reporting information cooperative witnesses generally aim to be accurate and informative (Horry et al., 2016). However, under conditions of uncertainty, the demands of accuracy and informativeness compete and witnesses may have to reach a compromise between the two (Koriat & Goldsmith, 1996). One method for balancing accuracy and informativeness is to regulate the level of detail (grain size) they report (Goldsmith et al., 2002; Sauer & Hope, 2016). Fine-grain responses are more informative but less likely to be accurate compared with coarse-grain responses (Brewer et al., 2018). When a premium is placed on accurate reporting (e.g. when providing person descriptions to the police), it is likely that witnesses use a higher threshold for accurate reporting and may withhold information that is likely to be incorrect. Thus, witnesses may be strategically regulating their memory to enhance accuracy (Koriat & Goldsmith, 1994, 1996).

The results of Study 3 (Experiments 1 and 2; Chapter 4) suggest that the use of self-generated cues facilitated a higher accuracy with person descriptions than the use of person description cues. However, using self-generated cues led to fewer person descriptors being reported than when using person description cues. Whilst the use of self-generated cues did not differ from using a free recall in terms of frequency of person descriptors within a single perpetrator event, the use of self-generated cues did lead to more person descriptors than when using a free recall within a multiple perpetrator event. The benefit of using self-generated cues for eliciting person descriptors compared to using a free recall in the multiple perpetrator event but not in the single perpetrator event could be explained by the possibility that the person cues in Experiment 2 were more distinctive (e.g. there were more person details in total) than in Experiment 1. Additionally, the task of reporting person descriptors of multiple perpetrators (Experiment 2) may be more challenging and therefore the use of a retrieval technique, such as self-generated cues could be of greater benefit than in an easier

task (e.g. a single perpetrator event; Experiment 1). Research has shown that whilst increasing the match between encoding and retrieval is necessary to facilitate recall, it is the distinctive quality of the cues that moderates how much information is retrieved (Goh & Lu, 2012; Nairne, 2002).

The ability of a witness to regulate the quality and quantity of their reporting is impacted by the type of reporting instructions used (Koriat & Goldsmith, 1996). In Study 3 (Chapter 4), mock-witnesses who used self-generated cues were able to provide more accurate person descriptions and more accurate fine-grain details than mock-witnesses who used person description cues. This advantage might be down to the differences in reporting instructions between the use of self-generated cues and the use of person description cues. Koriat and Goldsmith (1996) examined how participants responded to a general knowledge test when using free-report and forced-report conditions. Participants in the free-report condition were also able to monitor the accuracy of their responses by volunteering answers that exceeded a pre-set criterion of confidence that the answer was correct (*satisficing model;* Ackerman & Goldsmith, 2008). Participants were more effective at monitoring their accuracy when they were able to control whether to volunteer or withhold an answer (*control of report option*). Study 1 showed that frontline police officers in the real-world conducted interviews with witnesses using a forced-choice style of reporting and did not allow witnesses the option of a free-report. This raises important questions around the accuracy of person descriptions elicited in this way. The use of a forced-choice style of reporting (e.g. using a checklist approach) clearly led to errors (e.g. reduced accuracy), which is problematic for frontline policing and outweighs the advantage of using a forced-choice style of reporting to obtain more information from witnesses.

Interestingly, across Studies 2 and 3, the overall number of coarse-grain details reported by mock-witnesses was low and the overall number of fine-grain details was, in

comparison, high. Moreover, the fine-grain details reported by mock-witnesses were often accurate. The Dual-Criterion Model (Ackerman & Goldsmith, 2008) proposes that informativeness mediates reporting, in that even if coarse-grain details are more likely to be accurate (though imprecise; Evans & Fisher, 2011), they may be withheld if assessed by the interviewee as not sufficiently informative. The prioritisation of fine-grain details in Studies 2 and 3 supports previous research using mock-witness accounts (e.g. Brewer et al., 2018; McCallum et al., 2016). One reason why mock-witnesses may under-report coarse-grain details is because they believe the coarse-grain estimates are unhelpful, with the potential to violate social norms of communication. Witnesses might also make assumptions about the information they believe police officers already know and therefore may withhold information (particularly coarse-grain details) to avoid ridicule (Yaniv & Foster, 1995). Brewer et al. (2018) argued that this potential assumption by witnesses about the informativeness of police officers' knowledge is problematic. For example, frontline police officers attending a scene and carrying out an immediate search for the perpetrator may find coarse-grain details invaluable in narrowing down a search. Future research needs to investigate why coarse-grain information is at times under-reported in eyewitness memory accounts and whether mnemonics could be used to enhance the recall of coarse-grain details.

Across all studies within the current thesis, the number of facial features reported by witnesses was low. This is in line with previous literature that shows mock-witnesses are more likely to report general person features (e.g. gender, height, build), than facial features (Demarchi, 2003; Fahsing et al., 2004). The lack of facial features reported throughout all studies in the current thesis suggests that if information is poorly encoded to begin with, the information might not then be available for retrieval despite the use of cues.

Practical implications for obtaining person description information

The findings presented in this thesis make a novel contribution to the wider literature on investigative interviewing with a particular focus on frontline policing, and build on previous research surrounding the reporting of person description information. Person descriptions are important in the early stages of a criminal investigation and can be used to narrow the range of potential suspects in the search for a perpetrator (Farrington & Lambert, 1997). Whilst the quality of information obtained during frontline police interviews is determined by a number of factors, some of which are outside the control of frontline officers (e.g. cross-race differences), some relate directly to the activity of the frontline police officers (e.g. interviewing technique). Psychological factors known to underpin successful interviews with cooperative witnesses include providing retrieval support and promoting accurate and detailed accounts (Hope et al., 2014). However, investigative interviewing is a challenging skill to master (Powell, 2014), and demands a good understanding of memory and communication, as well as other cognitive, social, and environmental factors that may affect the accuracy of witness accounts.

Although frontline police officers carry out the majority of interviews involving person descriptions (albeit on the frontline), these frontline officers have received the least amount of interview training (Dando et al., 2009). Thus, it is possible that frontline officers have not been offered high quality training that would enable them to develop the necessary tools to be able to master the skill of investigative interviewing. Frontline police officers can limit or contaminate witness accounts through the use of misleading or otherwise inadequate questioning techniques (see Chapter 2, Study 1). This inefficient style of interviewing can lead to reduced accuracy in witness reports of perpetrators, as inappropriate questions (e.g. closed questions and leading questions) are often associated with incorrect responses (Fahsing et al., 2004; Granhag et al., 2013). It is important to note that frontline police

officers are often deployed to complex and dynamic scenes where following best-practice interviewing might not be possible. However, there is still considerable room for improvement in the way frontline police officers interview witnesses.

Promisingly, the results of Study 1 demonstrated that when police officers asked open-ended questions, witnesses were able to provide more person description information and these descriptors were more informative and contained more fine-grain details compared to when closed-questions were asked. Fine-grain details tend to be informative and are useful in assisting the police to locate the [correct] perpetrator(s). The findings from Study 1 are consistent with Oxburgh et al. (2010), who found that open questions tended to elicit longer, more detailed, and more accurate responses than those featuring a predominance of closed questions.

Study 2 demonstrated that using a non-leading tool – person description form – was effective in increasing the number of person descriptors (20.6 person descriptors) compared to previous laboratory-based research (average of 10 person descriptors). Using components of the Self-Administered Interview, the person description form aimed to discourage mock-witnesses from guessing and used non-leading questions and cues to facilitate reporting of person description information. Studies using the self-administered interview have reported an increase in the quantity and accuracy of person description details (Gabbert et al., 2009; Krix et al., 2016), as well as overall details (for a review see Horry et al., 2021). This increase in the number of person descriptors is promising and suggests that witnesses can provide underlying perpetrator information as long as frontline police officers (i) discourage witnesses from guessing and (ii) elicit person descriptions using non-leading questions.

Additionally, Study 3 (Experiments 1 and 2; Chapter 4) showed that using self-generated cues improved the accuracy of the person descriptors reported by mock-witnesses for both single and multiple perpetrators when compared to using person description cues.

Whilst the use of person description cues produced more person descriptors across both single perpetrator and multiple perpetrator events, the accuracy of such descriptors was poor, with the use of self-generated cues leading to a higher accuracy rate of person descriptors than the use of person description cues in both single perpetrator events and multiple perpetrator events. One of the primary aims of eliciting a person description on the frontline is the quick apprehension of the perpetrator(s) in the immediate vicinity of the incident. Hence, it is essential that the person descriptions are accurate in order to apprehend the correct perpetrator(s). The reason for the low accuracy rates when using person description cues could be because it forces witnesses to consider particular features (e.g. scars, marks, or peculiarities), even when they have no memory for these features. Feature checklists that are similar to using person description cues have been shown to produce more incorrect person descriptors than a free recall task (see Wogalter, 1991; 1996), which is problematic given that using person description cues is similar to frontline policing in the real-world. Additionally, unlike using person description cues, when witnesses have the option to withhold details they are uncertain about, they can better maintain the accuracy of their memory reports (Evans & Fisher, 2011; Koriat & Goldsmith, 1996). Given that witnessing an event is a subjective experience and that frontline police officers may be unaware of what information is salient to the witness, self-generated cues could be used to support a witness-led largely self-administered report. This would be particularly important when there are multiple witnesses and few frontline police officers to carry out the necessary interviews.

Methodological considerations and future directions

As is often the case with research, there are a number of limitations associated with each of the studies conducted as part of this PhD thesis. Access to Body worn video footage provided a unique insight into the interactions that take place at the frontline of policing, and while the footage showed that frontline police officers ask inappropriate questions to elicit

person descriptions, the absence of ground truth regarding the accuracy of the descriptions provided by witnesses about the perpetrator is problematic. That is, there is no way of knowing whether the descriptions provided by witnesses were correct, or whether the questions impacted the accuracy of the person descriptors reported. Accuracy is essential if frontline police officers are to apprehend the correct perpetrator. Additionally, Study 1 did not measure potential confounds, such as crime type, witness intoxication (see Harvey & Bayless, 2021), individual differences in police interviewer (see Hudson et al., 2018), or police interview training – all of which could impact on the quantity and quality of person descriptions reported by witnesses. These limitations highlight the importance of avoiding drawing casual inferences from field data (Wright, 2006). Police training could impact the types of questions officers asked witnesses and so better training of investigative interviewing and in particular, appropriate question types is essential.

Rapport has been empirically associated with enhanced information elicitation (Abbe & Brandon, 2013; Roberts et al., 2004). However, only one unpublished study has examined rapport and frontline police interviewing, finding that officers often failed to deploy strategies that would help to build rapport with witnesses (Gabbert et al., 2016). Future research should therefore examine the effects of rapport on frontline policing, and specifically the elicitation of person description information. Future research should continue to take advantage of the available BWV footage and use this as a tool to further investigate frontline police interviewing.

Although Study 1 lacked ground truth, the laboratory-based studies (Studies 2 and 3), controlled for this and permitted some analyses of person description accuracy. However, as controlled laboratory-based experiments, these studies could be said to lack ecological validity, due to the experimental design being somewhat artificial and manipulated. Specifically, motivation to provide accurate and informative person descriptions may be

much higher in real-life eyewitness situations compared with laboratory-based settings. Additionally, the conditions in which witnesses observe live events (e.g. intoxication, obstructed view, sudden onset) are likely to be very different than what mock-witnesses observed in the mock-events used in the current thesis. In the current set of laboratory studies (Chapters 3 and 4) mock-witnesses viewed the target perpetrators under optimal encoding conditions, where participants paid full attention and had no distractions. It is therefore possible that the number of reported person descriptors were overestimated. In real-world incidents, witnesses may not have been paying attention, or perhaps were distracted during the encoding of the event. Witnesses may also experience stress or physiological arousal which can divert attention away from the incident (Lane, 2006). Research in the laboratory examining the role of attention in mock-crime scenarios has found that witnesses report weak or partial memory for target incidents under sub-optimal conditions (Hope et al., 2016). Sauer and Hope (2016) examined the strategic regulation of memory reporting following optimal and sub-optimal encoding conditions and found that participants in the divided attention condition volunteered fewer fine-grain responses, and less accurate fine-grain responses, than participants in the full attention condition. Future research should therefore consider the role of attention for successful encoding of perpetrator information, and should look to immersive simulated learning environments (e.g. field-based exercises and table-top simulations) to increase validity.

In addition to the ways in which the studies were designed, there were also potential limitations associated with the coding methodology. The manner in which person descriptions were coded in the current series of studies may have led to more person descriptors being reported than in previous laboratory-based research. For example, in Studies 2 and 3, person information was coded if it included facial features, hair, height, build, clothes, and accessories. However, in previous research, there is no consistency in

terms of description quantity (e.g. the number of physical descriptors contained within a description). Some researchers have focused specially on facial features and hair (e.g. Brown & Lloyd-Jones, 2003), whereas other researchers have included assessments of the whole body (e.g. height, build, weight; van Koppen & Lochun, 1997). The fact that different researchers measure description quantity differently, adds a complexity when trying to compare with other studies. If we are to enhance the comparability and replicability of studies, a recommendation for future researchers would be to fully report coding systems and attempt to standardise how person descriptions are measured. Future researchers should make their coding protocols publically available and pre-register coding schemes and relevant analyses.

Another methodological consideration is the generalisability of the findings of the current thesis and their applicability to all cultural backgrounds. Research conducted around the world has used participants from different races, ethnicities, and cultures; however, the vast majority of the previous research is western-centric, with the majority not including tested hypotheses drawn from cultural theory (Hope et al., 2021). The few studies that have looked at cross-race differences in facial descriptors have been conducted with African and British participants. However, witnesses come from diverse cultural backgrounds. The limited research that has examined culture in more general memory contexts has shown that culture can shape the type of information that is remembered. In one of the few studies to explore culture using an eyewitness paradigm, Anakwah et al. (2020) found that during a free recall task, participants with a collectivist cultural orientation (i.e. the individual is embedded in a complex web of social relationships) reported fewer details about the crime scenes in their memory reports compared to participants with an individualistic cultural orientation (i.e. ties between individuals in a society are relatively loose). The individualism–collectivism cultural dimension may lead to biases in what is considered worthy and informative to report

when witnesses from these cultures are exposed to similar scenes (Boduroglu et al., 2009). In terms of frontline interviewing, it is important that frontline police officers are aware of the role culture may play in memory reports and in particular the reporting of person description information. Future research should look to examine how frontline police officers should conduct their interviews with witnesses from different races and cultures.

The design of Experiments 1 and 2 that comprised Study 3 (Chapter 4) may have been problematic (i.e. when comparing the use of person description cues with using self-generated cues or using a free recall) due to each of the conditions providing different levels of retrieval support. Additional retrieval support is likely to facilitate recall and thus the fact that the use of person description cues had an advantage in terms of the reporting of person descriptors (though not for accuracy) may have simply been that mock-witnesses had more retrieval support in this condition compared to the other two conditions. However, all conditions allowed self-administered reporting, and all mock-witnesses were presented with the same general instructions prior to reporting person descriptions about the perpetrator(s). The rationale for including the use of person description cues, was to see whether current frontline interview techniques were beneficial for the reporting of person description information. The findings suggest that using person description cues is not in line with best-practice interviewing and leads to an increase in the error rate when compared with spontaneous testimony. Future research should explore how best to elicit person descriptions using mnemonics, such as self-generated cues, in order to increase our understanding about the underlying mechanisms of the use of self-generated cues relative to other cues.

Self-reflection: PhD by compilation

The current thesis presented a series of studies in the format of published or publishable articles. Whilst each study stands alone, the PhD thesis, as a whole, tells a story through first identifying a real-world issue in frontline policing (i.e. obtaining person

descriptions from witnesses) and then determining potential factors that inhibit or enhance the reporting of information by witnesses (e.g. cross-race differences). Next, the thesis explored how memory retrieval techniques (e.g. self-generated cues) could be used to increase the number and accuracy of person descriptions reported by witnesses. As an Early Career Researcher, already employed in academia, completing a PhD using a compilation research strategy enabled me to experience the peer review process and prepare myself for submitting to the Research Excellence Framework. PhD by compilation enhanced my career prospects and enabled me to develop my own line of research. For example, I am now part of a team that is supervising a PhD student in the area of body worn video.

Conclusion

Obtaining person descriptions from witnesses is an inescapable necessity for frontline police officers. However, incomplete or inaccurate person descriptions are obviously problematic for frontline policing, as they provide few clues for frontline police officers to find the perpetrator (who may be close by). The findings reported in the current thesis highlight the need for further field research utilising real-world body worn video footage, as well as further empirical investigations into cross-race differences, meta-cognitive monitoring and control strategies, and the use of mnemonics to support the retrieval of memory. Increasing our knowledge of what happens in the field, as well as our theoretical understanding of how person descriptions are elicited and reported, can help to enhance how frontline police interviewing is conducted, whilst also identifying the support that is needed to specifically improve the reporting of person description information.

References

Abbe, A., & Brandon, S. E. (2013). The role of rapport in investigative interviewing: A review. *Journal of Investigative Psychology and Offender Profiling, 10*(3), 237-249. doi: 10.1002/jip.1386.

Ackerman, R., & Goldsmith, M. (2008). Control over grain size in memory reporting--With and without satisficing knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(5), 1224-1245. doi: 10.1037/a0012938.

Ahn, W. K., Brewer, W. F., & Mooney, R. J. (1992). Schema acquisition from a single example. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(2), 391-412. doi: 10.1037/0278-7393.18.2.391

Altman, D. G. (1991). *Practical statistics for medical research.* Chapman and Hall.

Anakwah, N., Horselenberg, R., Hope, L., Amankwah-Poku, M., & van Koppen, P. J. (2020). Cross-cultural differences in eyewitness memory reports. *Applied Cognitive Psychology, 30*(2), 505-515. https://doi.org/10.1002/acp.3637.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22*(3), 261–295. doi: 10.1016/S0022-5371(83)90201-3.

Anderson, S. J., & Conway, M. A. (1993). Investigating the structure of autobiographical memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(5), 1178-1196. doi: 10.1037/0278-7393.19.5.1178.

Ariel, B., Farrar, W. A., & Sutherland, A. (2015). The effect of police body-worn cameras on use of force and citizens' complaints against the police: A randomised controlled trial. *Journal of Quantitative Criminology, 31*(3), 509-535. doi: 10.1007/s10940-014-9236-3.

Association of Chief Police Officers. (2016). True vision association of chief police officers. http://report-it.org.uk/home

Beune, K., Giebels, E., Adair, W. L., Fennis, B. M., & Van Der Zee, K. I. (2011). Strategic sequences in police interviews and the importance of order and cultural fit. *Criminal Justice and Behavior, 38*(9), 934-954. doi: 10.1177/0093854811412170.

Bjorklund, D. F., Dukes, C., & Brown, R. D. (2009). The development of memory strategies. In M. L. Courage, N. Cowan, M. L. Courage, & N. Cowan (Eds.), *The development of memory in infancy and childhood.* (2nd ed., pp. 145–175). Psychology Press.

Boduroglu, A., Priti, S., & Nisbett, R. E. (2009). Cultural differences in allocation of attention in visual information processing. *Journal of Cross-Cultural Psychology, 40,* 349-360. doi: 10.1177/0022022108331005.

Boon, J., & Noon, E. (1994). Changing perspectives in cognitive interviewing. *Psychology, Crime and Law, 1*(1), 59-69. doi: 10.1080/10683169408411936.

Bowles, P. V., & Sharman, S. J. (2014). A review of the impact of different types of leading interview questions on child and adult witnesses with intellectual disabilities. *Psychiatry, Psychology and Law, 21*(2), 205-217. doi: 10.1080/13218719.2013.803276.

Brewer, N., Vagadia, A. N., Hope, L., & Gabbert, F. (2018). Interviewing witnesses: eliciting coarse-grain information. *Law and Human Behavior, 42*(5), 458-471. doi: 10.1037/lhb0000294.

Brewer, W.F., & Nakarmura, G.V. (1984). *The nature and functions of schemas* (report 325). Center for the Study of Reading Technical Report. http://hdl.handle.net/2142/17542.

Brigham, J. C., & Malpass, R. S. (1985). The role of experience and contact in the recognition of faces of own-and other-race persons. *Journal of Social Issues, 41*(3), 139-155. doi: 10.1111/j.1540-4560.1985.tb01133.x.

Brown, C., & Lloyd-Jones, T. J. (2002). Verbal overshadowing in a multiple face presentation paradigm: Effects of description instruction. *Applied Cognitive Psychology, 16*(8), 873-885. doi: 10.1002/acp.919.

Brown, C., & Lloyd-Jones, T. J. (2003). Verbal overshadowing of multiple face and car recognition: Effects of within-versus across-category verbal descriptions. *Applied Cognitive Psychology, 17*(2), 183-201. doi: 10.1002/acp.861.

Brown, C., Lloyd-Jones, T. J., & Robinson, M. (2008). Eliciting person descriptions from eyewitnesses: A survey of police perceptions of eyewitness performance and reported use of interview techniques. *European Journal of Cognitive Psychology, 20(3)*, 529-560. doi: 10.1080/09541440701728474.

Bukach, C. M., Cottle, J., Ubiwa, J., & Miller, J. (2012). Individuation experience predicts other-race effects in holistic processing for both Caucasian and Black participants. *Cognition, 123*(2), 319-324. doi: 10.1016/j.cognition.2012.02.007.

Centrex (2004). *Practical Guide to Investigative Interviewing*. Central Police Training and Development Authority.

Chen, Y. Y., & Geiselman, R. E. (1993). Effects of ethnic stereotyping and ethnically-related cognitive biases on eyewitness recollections of height. *American Journal of Forensic Psychology, 11*(2), 13-19.

Chiroro, P., & Valentine, T. (1995). An investigation of the contact hypothesis of the own-race bias in face recognition. *The Quarterly Journal of Experimental Psychology Section A, 48*(4), 879-894. doi: 10.1080/14640749508401421.

Christianson, B., Elliot, M., & Massey, B. (2015). *The Role of Publications and Other Artefacts in Submissions for the UK PhD (UK Council for Graduate Education)*. http://www.ukcge.ac.uk/article/the-role-of-publications-and-other-artefacts-in-submissions-for-the-uk-phd-201.aspx.

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology, 55*, 591-621. doi: 10.1146/annurev.psych.55.090902.142015.

Clifford, B. R., & George, R. (1996). A field evaluation of training in three methods of witness/victim investigative interviewing. *Psychology, Crime and Law, 2*(3), 231-248. doi: 10.1080/10683169608409780.

College of Policing. (2019). *Obtaining initial accounts from victims and witnesses: Guidelines for first responders.* https://www.college.police.uk/What-we-do/Standards/Guidelines/Initial-accounts/Documents/Initial_Accounts_Guidelines.pdf.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671-684. doi: 10.1016/S0022-5371(72)80001-X.

Dalton, G., Milne, R., Hope, L., Vernham, Z., & Nunan, J. (2021). 'He was just your typical average guy' Examining how person descriptions are elicited by frontline police officers. *Applied Cognitive Psychology, 35*(2), 517-525. doi: 10.1002/acp.3778.

Dando, C., Wilcock, R., & Milne, R. (2008). The cognitive interview: Inexperienced police officers' perceptions of their witness/victim interviewing practices. *Legal & Criminological Psychology, 13*(1), 59-70. doi: 10.1348/135532506X162498.

Dando, C., Wilcock, R., & Milne, R. (2009). The cognitive interview: The efficacy of a modified mental reinstatement of context procedure for frontline police investigators. *Applied Cognitive Psychology, 23*(1), 138-147. doi: 10.1002/acp.1451.

Dando, C., Wilcock, R., Milne, R., & Henry, L. (2009). A modified cognitive interview procedure for frontline police investigators. *Applied Cognitive Psychology, 23,* 698-716. doi: 10.1002/acp.1501.

Davis, M. R., McMahon, M., & Greenwood, K. M. (2005). The efficacy of mnemonic components of the cognitive interview: Towards a shortened variant for time-critical investigations. *Applied Cognitive Psychology, 19*(1), 75-93. doi: 10.1002/acp.1048.

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior, 28*(6), 687-706.

Demarchi, S. (2003). *Conception and evaluation of an interviewing method for person recall based on a social psychology engineering approach.* Unpublished doctoral dissertation, Paris 8 University.

Demarchi, S., & Py, J. (2009). A method to enhance person description: A field study. In R. Bull, T, Valentine & T, Williamson (eds), *Handbook of Psychology of investigative interviewing: current developments and future directions.* (pp. 241-256). John Wiley & Sons, Ltd.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology, 51*(3), 629-636. doi: 10.1037/h0046408.

Douglass, A. B., Brewer, N., Semmler, C., Bustamante, L., & Hiley, A. (2013). The dynamic interaction between eyewitnesses and interviewers: The impact of differences in perspective on memory reports and interviewer behavior. *Law and Human Behavior, 37*(4), 290. doi: 10.1037/lhb0000034.

Drover, P., & Ariel, B. (2015). Leading an Experiment in Police Body-Worn Video Cameras. *International Criminal Justice Review, 25*(1), 80-97. doi: 10.1177/1057567715574373.

Eisen, M., Quas, J. A., & Goodman, G. S. (2002). *Memory and Suggestibility in the Forensic Interview*. Routledge.

Ellis, H. D., Deregowski, J. B., & Shepherd, J. W. (1975). Descriptions of white and black faces by white and black subjects. *International Journal of Psychology, 10*(2), 119-123. doi: 10.1080/00207597508247325.

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1980). The deterioration of verbal descriptions of faces over different delay intervals. *Journal of Police Science and Administration, 8*(1), 101-106.

Evans, J. R., & Fisher, R. P. (2011). Eyewitness memory: Balancing the accuracy, precision and quantity of information through metacognitive monitoring and control. *Applied Cognitive Psychology, 25*(3), 501-508. doi: 10.1002/acp.1722.

Evans, J. R., Meissner, C. A., Brandon, S. E., Russano, M. B., & Kleinman, S. M. (2010). Criminal versus HUMINT interrogations: The importance of psychological science to improving interrogative practice. *The Journal of Psychiatry & Law, 38*(1-2), 215-249. doi: 10.1177/009318531003800110.

Fahsing, I. A., Ask, K., & Granhag, P. A. (2004). The man behind the mask: Accuracy and predictors of eyewitness offender descriptors. *Journal of Applied Psychology, 89,* 722-729. doi: 10.1037/0021-901-.89.4.722.

Fallshore, M., & Schooler, J. W. (1995). Verbal vulnerability of perceptual expertise. Journal of Experimental Psychology: *Learning, Memory, and Cognition, 21*(6), 1608-1623. doi: 10.1037/0278-7393.21.6.1608.

Farrington, D. P., & Lambert, S. (1997). Predicting offender profiles from victim and witness descriptions. In J. L. Jackson & D. A. Bekerian (Eds.), *Offender Profiling: Theory, Research and Practice.* (pp. 133–158). John Wiley & Sons Inc.

Finger, K., & Pezdek, K. (1999). The effect of cognitive interview on face identification accuracy: Release from verbal overshadowing. *Journal of Applied Psychology, 84*(3), 340-348. doi: 10.1037/0021-9010.84.3.340.

Fisher, R. P. (1995). Interviewing victims and witnesses of crime. *Psychology, Public Policy and Law, 1*(4), 732-764. doi: 10.1037/1076-8971.1.4.732.

Fisher, R. P. (2010). Interviewing cooperative witnesses. *Legal and Criminological Psychology, 15*, 25-38. doi: 10.1348/135532509X441891.

Fisher, R. P., Falkner, K. L., Trevisan, M., & McCauley, M. R. (2000). Adapting the Cognitive Interview to enhance long-term (35-years) recall of psychical activities. *Journal of Applied Psychology*, *85*(2), 180-189. doi: 10.1037//0021–9010.85.2.180.

Fisher, R. P., Geiselman, R. E., & Raymond, D. S. (1987). Critical analysis of police interview techniques. *Journal of Police Science and Administration, 15*(3), 177-185.

Fisher, R., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative Interviewing: The cognitive interview*. Charles C. Thomas.

Frenda, S. J., Nichols, R. M., & Loftus, E. F. (2011). Current issues and advances in misinformation research. *Current Directions in Psychological Science*, *20*(1), 20-23. doi: 10.1177/0963721410396620.

Gabbert, F. & Brown, C. (2015). Interviewing for face identification. In T. Valentine, & J. P. Davis, (Eds.), *Forensic Facial Identification: Theory and Practice of Identification from Eyewitnesses, Composites and CCTV.* (pp.17–41). Wiley-Blackwell.

Gabbert, F., & Hope, L. (2013). Suggestibility and memory conformity. *Suggestibility in Legal Contexts: Psychological Research and Forensic Implications*, 63-83.

Gabbert, F., Hope, L., & Fisher, R. P. (2009). Protecting eyewitness evidence: Examining the efficacy of a self-administered interview tool. *Law and Human Behavior, 33,* 298-307. doi: 10.1007/s10979-008-9146-8.

Gabbert, F., Hope, L., La Rooy, D., McGregor, A., Milne, R., & Ellis, T. (2016). *Introducing a PEACE-complaint 'structured interview protocol' to enhance the quality of investigative interviews*. Paper presented at the Ninth International Investigative Interviewing Conference, Porto (June, 2016).

Gabbert, F., MacPherson, I., & Hope, L. (2014). *Adding a new mnemonic to the Cognitive Interview*. In IIIRG Conference, Lausanne.

Gabbert, F., Memon, A., & Allan, K. (2003). Memory conformity: Can eyewitnesses influence each other's memories for an event? *Applied Cognitive Psychology, 17*(5), 533-543. doi: 10.1002/acp.885.

Gabbert, F., Memon, A., & Wright, D. B. (2006). Memory conformity: Disentangling the steps toward influence during a discussion. *Psychonomic Bulletin & Review, 13,* 480-485. doi: 10.3758/ BF03193873.

Garven, S., Wood, J. M., & Malpass, R. S. (2000). Allegations of wrongdoing: The effects of reinforcement on children's mundane and fantastic claims. *Journal of Applied Psychology, 85*(1), 38-49. doi: 10.1037/0021-9010.85.1.38.

Geiselman, R. E., Fisher, R. P., Firstenberg, I., Hutton, L. A., Sullivan, S., Avetissian, I., & Prosk, A. (1984). Enhancement of eyewitness memory: An empirical evaluation of the cognitive interview. *Journal of Police Science and Administration, 12,* 74-80.

Geiselman, R. E., Schroppel, T., Tubridy, A., Konishi, T., & Rodriguez, V. (2000). Objectivity bias in eyewitness performance. *Applied Cognitive Psychology, 14*(4), 323-332. doi: 10.1002/1099-0720(200007/08)14:4<323::AID-ACP652>3.0.CO;2-M.

Goh, W. D., & Lu, S. H. (2012). Testing the myth of the encoding–retrieval match. *Memory & Cognition, 40*(1), 28-39. doi: 10.3758/s13421-011-0133-9.

Goldsmith, M., Koriat, A., & Weinberg-Eliezer, A. (2002). Strategic regulation of grain size memory reporting. *Journal of Experimental Psychology: General, 131*, 73-95. doi: 10.1037/0096-3445.131.1.73.

Granhag, P. A., Ask, K., Rebelius, A., Öhman, L., & Mac Giolla, E. (2013). I saw the man who killed Anna Lindh! An archival study of witnesses' offender descriptions. *Psychology, Crime & Law, 19*(10), 921-931. doi: 10.1080/1068316X.2012.719620.

Harvey, A. J., & Bayless, S. J. (2021). *Alcohol intoxication and the description of faces from memory* [Unpublished manuscript]. Department of Psychology, University of Portsmouth.

Hope, L., & Gabbert, F. (2019). Memory at the sharp end: The costs of remembering with others in forensic contexts. *Topics in Cognitive Science, 11*(4), 609-626. doi:10.1111/tops.12357.

Hope, L., Anakwah, N., Antfolk, J., Brubacher, S. P., Flowe, H., Gabbert, G., Giebals, E., Kanja, W., Korkman, J., Kyo, A., Naka, M., Otgaar, H., Powell, M. B., Selim, H., Skrifvars, J., Sorkpah, I. K., Sowatey, E. A., Steele, L. C., Stevens, L., … Wells, S. (2021). *Urgent issues and Prospects: Examining the role of culture in the*

investigative interviewing of victims and witnesses [Manuscript submitted for publication]. Department of Psychology, University of Portsmouth.

Hope, L., Blocksidge, D., Gabbert, F., Sauer, J. D., Lewinski, W., Mirashi, A., & Atuk, E. (2016). Memory and the operational witness: Police officer recall of firearms encounters as a function of active response role. *Law and Human Behavior*, *40*(1), 23-35. doi: 10.1037/lhb0000159.

Hope, L., Gabbert, F., & Fisher, R. P. (2011). From laboratory to the street: Capturing witness memory using the Self-Administered Interview. *Legal and Criminological Psychology, 16*(2), 211-226. doi: 10.1111/j.2044-8333.2011.02015.x.

Hope, L., Gabbert, F., Fisher, R. P., & Jamieson, K. (2014). Protecting and enhancing eyewitness memory: The impact of an initial recall attempt on performance in an investigative interview. *Applied Cognitive Psychology*, *28*(3), 304-313. doi: 10.1002/acp.2984.

Horry, R., Brewer, N., & Weber, N. (2016). The grain-size lineup: A test of a novel eyewitness identification procedure. *Law and Human Behavior, 40*(2), 147-158. doi: 10.1037/lhb0000166.

Horry, R., Hughes, C., Sharma, A., Gabbert, F., & Hope, L. (2021). A meta-analytic review of the Self-Administered Interview©: Quantity and accuracy of details reported on initial and subsequent retrieval attempts. *Applied Cognitive Psychology*, *35*(2), 428-444. doi: 10.1002/acp.3753.

Howe, M. L., Knott, L. M., & Conway, M. A. (2017). *Memory and miscarriages of justice*. Psychology Press.

Hudson, C. A., Satchell, L. P., & Adams-Quackenbush, N. (2018). It Takes Two: The Round-Robin Methodology for Investigative Interviewing Research. *Frontiers in Psychology, 9*, 2181. doi:10.3389/fpsyg.2018.02181.

Ito, H., Barzykowski, K., Grzesik, M., Gülgöz, S., Gürdere, C., Janssen, S. M. J., Khor, J., Rowthorn, H., Wade, K. A., Luna, K., Albuquerque, P. B., Kumar, D., Singh, A. D., Cecconello, W. W., Cadavid, S., Laird, N. C., Baldassari, M. J., Lindsay, D. S., & Mori, K. (2019). Eyewitness Memory Distortion Following Co-Witness Discussion: A Replication of Garry, French, Kinzett, and Mori (2008) in Ten Countries. *Journal of Applied Research in Memory and Cognition*, *8*(1), 68-77. doi: 10.1016/j.jarmac.2018.09.004.

Jennings, W. G., Fridell, L. A., & Lynch, M. D. (2014). Cops and cameras: Officer perceptions of the use of body-worn cameras in law enforcement. *Journal of Criminal Justice, 42*, 549-556. doi: 10.1016/j.jcrimjus.2014.09.008.

Kask, K., & Bull, R. (2009). The effects of different presentation methods on multi-ethnicity face recognition. *Psychology, Crime & Law*, *15*(1), 73-89. doi: 10.1080/10683160802131131.

Kassin, S. M., Tubb, V. A., Hosch, H. M., & Memon, A. (2001). On the" general acceptance" of eyewitness testimony research: A new survey of the experts. *American Psychologist, 56*(5), 405-416. doi: 10.1037/0003-066X.56.5.405.

Katz, C. M., Choate, D. E., Ready, J. R., Nuňo, L., Kurtenbach, C. M., & Kevin, S. (2014). *Evaluating the impact of officer worn body cameras in the Phoenix police department*. https://publicservice.asu.edu/sites/default/files/ppd_spi_feb_20_2015_final.pdf.

Kebbell, M. R., & Milne, R. (1998). Police officers' perceptions of eyewitness performance in forensic investigations. *The Journal of Social Psychology, 138*(3), 323-330. doi: 10.1080/00224549809600384.

Kellogg, R. T. (2007). Are written and spoken recall of text equivalent? *The American Journal of Psychology, 120*(3), 415-428. doi: 10.2307/20445412.

Kontogianni, F., Hope, L., Taylor, P. J., Vrij, A., & Gabbert, F. (2018). The benefits of a self-generated cue mnemonic for timeline interviewing. *Journal of Applied Research in Memory and Cognition. 7*(3), 454-461. 10.1016/j.jarmac.2018.03.006.

Koriat, A., & Goldsmith, M. (1994). Memory in naturistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General, 123*, 297-315. doi: 10.1037/0096-3445.123.3.297.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103*, 490-517. doi: 10:1037/0033-295x.103.3.490.

Krix, A. C., Sauerland, M., Raymaekers, L. H. C., Memon, A., Quaedflieg, C. W. E. M., & Smeets, T. (2016). Eyewitness evidence obtained with the self-administered interview is unaffected by stress. *Applied Cognitive Psychology, 30*, 103-112. doi: 10.1002/acp.3173.

Kuehn, L. L. (1974). Looking down a gun barrel: Person perception and violent crime. *Perceptual and Motor Skills, 39*, 1159-1164. doi: 10.2466/pms.1974.39.3.1159.

Lane, S. M. (2006). Dividing attention during a witnessed event increases eyewitness suggestibility. *Applied Cognitive Psychology, 20*(2), 199-212. doi: 10.1002/acp.1177.

Lantz, B. (2013). The impact of sample non-normality on ANOVA and alternative methods. *British Journal of Mathematical and Statistical Psychology, 66*(2), 224-244. doi: 10.1111/j.2044-8317.2012.02047.x.

Laughery, K. R., Duval, C., & Wogalter, M. S. (1986). Dynamics of facial recall. *In Aspects of face processing.* (pp. 373-387). Springer.

Lee, T., & Geiselman, R. E. (1994). Recall of perpetrator height as a function of eyewitness and perpetrator ethnicity. *Psychology, Crime and Law, 1*(1), 11-19. doi: 10.1080/10683169408411932.

Lindsay, R. C. L. (2012). *Eyewitness identification: Importance of the description*. Keynote address at the International Investigative Interviewing Research Group (iIIRG), Toronto (May, 2012).

Lindsay, R. C. L., Martin, R., & Webber, L. (1994). Default values in eyewitness descriptions: A problem for the match-to-description lineup foil selection strategy. *Law and Human Behavior, 18,* 527-541. doi: 10.1007/BF01499172.

Loftus, E. F. (1979). The malleability of human memory: Information introduced after we view an incident can transform memory. *American Scientist, 67*(3), 312-320. http://www.jstor.org/stable/27849223

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction – Example of interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13,* 585-589. doi: 10.1016/S0022-5371(74)80011-3.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*(1), 19-31. doi: 10.1037/0278-7393.4.1.19.

Mahr, J. B., & Csibra, G. (2020). Witnessing, remembering, and testifying: Why the past is special for human beings. *Perspectives on Psychological Science, 15*(2), 428-443. doi: 10.1177/1745691619879167.

Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology, 13*(4), 330-334. doi: 10.1037/h0028434.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6*(6), 255-260. doi: 10.1016/S1364-6613(02)01903-4.

McCallum, N., Brewer, N., & Weber, N. (2016). Memorial monitoring and control: How confidence and social and financial consequences affect eyewitnesses' reporting of fine-grain information. *Applied Cognitive Psycholo*gy, *30*, 375-386. doi: 10.1002/acp.3212.

McCallum, N., Brewer, N., & Weber, N. (2019). A measure of perceived informativeness for investigations of eyewitness memory reporting. *Journal of Applied Research in Memory and Cognition*, *8*(2), 214-220.

McGloin, J. M., & Piquero, A. R. (2009). 'I wasn't alone': Collective behaviour and violent delinquency. *Australian & New Zealand Journal of Criminology, 42*, 336-353. doi: 10.1375/acri.42.3.336.

McPhee, I., Paterson, H. M., & Kemp, R. I. (2014). The power of the spoken word: can spoken-recall enhance eyewitness evidence? *Psychiatry, Psychology and Law*, *21*(4), 551-566. doi: 10.1080/13218719.2013.848001.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. Psychology, *Public Policy, and Law*, *7*(1), 3. doi: 10.1037/1076-8971.7.1.3.

Meissner, C. A., Brigham, J. C., & Kelley, C. M. (2001). The influence of retrieval processes in verbal overshadowing. *Memory & Cognition*, *29*(1), 176-186.

Meissner, C. A., Sporer, S. L., & Schooler, J. W. (2007). Person descriptions as eyewitness evidence. In R. Lindsay, D, Ross, J. Read, & M. Toglia. (Eds), *Handbook of Eyewitness Psychology: Memory for People.* (pp. 3-34). Lawrence Erlbaum & Associates.

Meissner, C. A., Sporer, S. L., & Susa, K. J. (2008). A theoretical review and meta-analysis of the description-identification relationship in memory for faces. *European Journal of Cognitive Psychology*, *20*(3), 414-455. doi: 10.1080/09541440701728581.

Memon, A., & Bruce, V. (1985). Context effects in episodic studies of verbal and facial memory: A review. *Current Psychological Research & Reviews*, *4*(4), 349-369.

Memon, A., Meissner, C. A., & Fraser, J. (2010). The Cognitive Interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, and Law*, *16*(4), 340-372.

Milne, R., & Bull, R. (2006). *Investigative interviewing: Psychology and practice*. Wiley.

Ministry of Justice (2011). *Achieving Best Evidence: Guidance on Interviewing Victims and Witnesses and Guidance on Using Special Measures*. Ministry of Justice. http://www.justice.gov.uk/downloads/victims-and-witnesses/vulnerable-witnesses/achieving-best-evidence-criminal-proceedings.pdf.

Myklebust, T., & Alison, L. (2000). The current state of police interviews with children in Norway: How discrepant are they from models based on current issues in memory and communication? *Psychology, Crime, and Law, 6*, 331-351. doi: 10.1080/10683160008409810.

Myklebust, T., & Bjorklund, R. A. (2006). The effect of long-term training on police officers' use of open and closed questions in field investigative interviews of children (FIIC). *International Journal of Investigative Psychology and Offender Profiling, 3*, 165-181. doi: 10.1002/jip.52.

Nairne, J. S. (2002). The myth of the encoding-retrieval match. *Memory, 10*(5-6), 389-395. doi: 10.1080/09658210244000216.

Ng, W. J., & Lindsay, R. C. (1994). Cross-race facial recognition: Failure of the contact hypothesis. *Journal of Cross-Cultural Psychology, 25*(2), 217-232. doi: 10.1177/0022022194252004.

Oxburgh, G. E, Myklebust, T., & Grant, T. D. (2010). The question of question types in police interview: A review of the literature from a psychological and linguistic perspective. *The International Journal of Speech, Language and the Law, 17*(1), 45-66. doi: 10.1558/ijsll.v17i1.45.

Paterson, H. M., Eijkemans, H., & Kemp, R. I. (2015). Investigating the impact of delayed administration on the efficacy of the Self-Administered Interview. *Psychiatry, Psychology and Law, 22*(2), 307-317. doi: 10.1080/13218719.2014.947670.

Pickrell, J. E., McDonald, D. L., Bernstein, D. M., & Loftus, E. F. (2016). Misinformation effect. In R. F. Pohl (Ed.), *Cognitive illusions.* (pp. 416-433). Psychology Press.

Powell, M. B. (2014). Centre for Investigative Interviewing launched. *https://www.deakin.edu.au/research/research-news-and-publications/articles/centre-for-investigative-interviewing-launched.*

Pozzulo, J. D., & Warren, K. L. (2003). Descriptions and identifications of strangers by youth and adult eyewitnesses. *Journal of Applied Psychology, 88*(2), 315-323. doi: 10.1037/0021-9010.88.2.315.

Py, J. & Demarchi, S. (2006). Describe and detect criminal with the cognitive interview. *Revue Québécoise de Psychologie*, *27*, 197-215.

Ready, J. T., & Young, J. T. (2015). The impact of on-officer video cameras on police–citizen contacts: Findings from a controlled experiment in Mesa, AZ. *Journal of Experimental Criminology, 11*(3), 445-458.

Rhodes, G., Brake, S., Taylor, K., & Tan, S. (1989). Expertise and configural coding in face recognition. *British Journal of Psychology, 80*(3), 313-331. doi: 10.1111/j.2044-8295.1989.tb02323.x.

Roberts, K. P., Lamb, M. E., & Sternberg, K. J. (2004). The effects of rapport-building style on children's reports of a staged event. *Applied Cognitive Psychology, 18*(2), 189-202. doi: 10.1002/acp.957.

Satin, G. E., & Fisher, R. P. (2019). Investigative utility of the Cognitive Interview: Describing and finding perpetrators. *Law and Human Behavior, 43*(5), 491. doi: 10.1037/lhb0000326.

Sauer, J., & Hope, L. (2016). The effects of divided attention at study and reporting procedure on regulation and monitoring for episodic recall. *Acta Psychologica, 169*, 143-156. doi: 10.1016/j.actpsy.2016.05.015.

Sauerland, M., & Sporer, S. L. (2011). Written vs. spoken eyewitness accounts: Does modality of testing matter? *Behavioral Sciences & the Law, 29*(6), 846-857. doi: 10.1002/bsl.1013.

Sauerland, M., Krix, A. C., van Kan, N., Glunz, S., & Sak, A. (2014). Speaking is silver, writing is golden? The role of cognitive and social factors in written versus spoken witness accounts. *Memory & Cognition, 42*(6), 978-992.

Scoboria, A., & Fisico, S. (2013). Encouraging and clarifying "don't know" responses enhances interview quality. *Journal of Experimental Psychology: Applied, 19*(1), 72-82. http://dx.doi.org/10.1037/a0032067.

Searcy, J. H., Bartlett, J. C., Memon, A., & Swanson, K. (2001). Aging and lineup performance at long retention intervals: Effects of metamemory and context reinstatement. *Journal of Applied Psychology, 86*(2), 207-214. doi: 10.1037/0021-9010.86.2.207.

Semmler, C., Brewer, N., & Wells, G. L. (2004). Effects of Postidentification Feedback on Eyewitness Identification and Nonidentification Confidence. *Journal of Applied Psychology, 89*(2), 334-346. doi: 10.1037/0021-9010.89.2.334.

Shepherd, E., & Griffiths, A. (2013). *Investigative Interviewing: The Conversation Management Approach*. Oxford University Press.

Shepherd, J. W., & Deregowski, J. B. (1981). Races and faces—a comparison of the responses of Africans and Europeans to faces of the same and different races. *British Journal of Social Psychology, 20*(2), 125-133. doi: 10.1111/j.2044-8309.1981.tb00485.x.

Slone, A. E., Brigham, J. C., & Meissner, C. A. (2000). Social and cognitive factors affecting the own-race bias in Whites. *Basic and Applied Social Psychology, 22*(2), 71-84. doi: 10.1207/S15324834BASP2202_1.

Smith, K., & Milne, B. (2018). Witness interview strategy for critical incidents (WISCI). *Journal of Forensic Practice, 20*(4), 268-278. doi: 10.1108/JFP-03-2018-0007.

Soukara, S., Bull, R., Vrij, A., Turner, M., & Cherryman, J. (2009). What really happens in police interviews of suspects? Tactics and confessions. *Psychology, Crime & Law, 15*(6), 493-506. doi: 10.1080/10683160802201827.

Sporer, S. L. (1992). Post-dicting eyewitness accuracy: Confidence, decision-times and person descriptions of choosers and non-choosers. *European Journal of Social Psychology, 22*(2), 157-180. doi: 10.1002/ejsp.2420220205.

Sporer, S. L. (1996). Psychological aspects of person description. In S.L. Sporer, R.S. Maplass & G. Kohnken (Eds.), *Psychological issues in eyewitness identification*. (pp. 53-87). Lawence Erlbaum.

Sporer, S. L. (2001a). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law, 7*(1), 36-97. doi: 10.1037/1076-8971.7.1.36.

Sporer, S. L. (2001b). The cross-race effect: Beyond recognition of faces in the laboratory. *Psychology, Public Policy, and Law, 7*(1), 170-200. doi: 10.1037/1076-8971.7.1.170.

Sporer, S. L. (2007). Person descriptions as retrieval cues: Do they really help? *Psychology, Crime & Law, 13*(6), 591-609. doi.org/10.1080/10683160701253986.

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

Tham, D. S. Y., Bremner, J. G., & Hay, D. (2017). The other-race effect in children from a multiracial population: A cross-cultural comparison. *Journal of Experimental Child Psychology, 155*, 128-137. doi: 10.1016/j.jecp.2016.11.006.

Thomson, D. M., & Tulving, E. (1970). Associative encoding and retrieval: Weak and strong cues. *Journal of Experimental Psychology, 86*(2), 255-262. doi: 10.1037/h0029997.

Tollestrup, P., Turtle, J., & Yuille, J. (1994). Actual victims and witnesses to robbery and fraud: An archival analysis. In D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments*. (pp. 144–160). Cambridge University Press.

Tullis, J. G., & Benjamin, A. S. (2015). Cue generation: How learners flexibly support future retrieval. *Memory & Cognition, 43*(6), 922-938. doi:10.3758/s13421-015-0517-3

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. (pp. 381–402). Academic Press.

Tulving, E. (1985). How many memory systems are there? *American Psychologist, 40*(4), 385-398. doi: 10.1037/0003-066X.40.4.385.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53*(1), 1-25. doi: 10.1146/annurev.psych.53.100901.135114

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352-373. doi: 10.1037/h0020071.

Tunnicliff, J. L., & Clark, S. E. (2000). Selecting foils for identification lineups: Matching suspects or descriptions? *Law and Human Behavior, 24*(2), 231-258. doi: 10.1023/A:1005463020252.

Valentine, T., Chiroro, P., & Dixon, R. (1995). An account of the own-race bias and the contact hypothesis based on a 'face space' model of face recognition. In T. Valentine (Ed.), *Cognitive and computational aspects of face recognition: Explorations in face space.* (pp. 69-94). Routledge.

van Koppen, P., & Lochun, S. (1997). Portraying perpetrators: the validity of offender descriptions by witnesses. *Law and Human Behaviour, 21,* 661-685. doi: 10.1023/A:1024812831576.

van Mastrigt, S.B., & Farrington, D.P. (2009). Co-offending, age, gender and crime type: Implications for criminal justice policy. *British Journal of Criminology, 49,* 552-573. doi: 10.1093/bjc/azp021.

Vrij, A., Hope, L., & Fisher, R. P. (2014). Eliciting reliable information in investigative interviews. *Policy Insights from the Behavioral and Brain Sciences, 1*(1), 129-136. doi: 10.1177/2372732214548592.

Watkins, O. C., & Watkins, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory, 1*(4), 442-452. doi: 10.1037/0278-7393.1.4.442.

Weber, N., & Brewer, N. (2008). Eyewitness recall: Regulation of grain size and the role of confidence. *Journal of Experimental Psychology: Applied, 14,* 50-60. doi: 10.1037/1076-898X.14.1.50.

Wells, G. L., & Turtle, J. W. (1987). Eyewitness testimony research: Current knowledge and emergent controversies. *Canadian Journal of Behavioural Science, 19*(4), 363-388. doi: 10.1037/h0080000.

Wells, G. L., Lindsay, R. C., & Ferguson, T. J. (1979). Accuracy, confidence, and juror perceptions in eyewitness identification. *Journal of Applied Psychology, 64*(4), 440-448. doi: 10.1037/0021-9010.64.4.440.

Wells, G. L., Rydell, S. M., & Seelau, E. P. (1993). The selection of distractors for eyewitness lineups. *Journal of Applied Psychology, 78*(5), 835-844. doi: 10.1037/0021-9010.78.5.835.

Wheeler, R. L., & Gabbert, F. (2017). Using self-generated cues to facilitate recall: a narrative review. *Frontiers in Psychology, 8,* 18-30. doi: 10.3389/fpsyg.2017.01830.

Wheeler, R. L., Gabbert, F., Hope, L., & Jones, S. (2017, January). *Evaluating a self-generated cue mnemonic to enhance eyewitness retrieval*. Paper presented at SARMAC conference, Sydney, Australia.

Whitley Jr, B. E., & Greenberg, M. S. (1986). The Role of Eyewitness Confidence in Juror Perceptions of Credibility. *Journal of Applied Social Psychology*, *16*(5), 387-409. doi: 10.1111/j.1559-1816.1986.tb01148.x.

Wixted, J. T., Mickes, L., & Fisher, R. P. (2018). Rethinking the reliability of eyewitness memory. *Perspectives on Psychological Science*, *13*(3), 324-335. doi: 10.1177/1745691617734878.

Wogalter, M. S. (1991, September). Effects of post-exposure description and imaging on subsequent face recognition performance. *Proceedings of the Human Factors Society, 35*, 575-579. doi: 10.1518/107118191786754680.

Wogalter, M. S. (1996, October). Describing faces from memory: Accuracy and effects on subsequent recognition performance. *Proceedings of the Human Factors and Ergonomics Society*, *40*, 536-540. doi: 10.1177/154193129604001103.

Wright, A. M., & Alison, L. (2004). Questioning sequences in Canadian police interviews: Constructing and confirming the course of events. *Psychology, Crime & Law, 10*(2), 137-154. doi: 10.1080/1068316031000099120.

Wright, D. B. (2006). Casual and associative hypotheses in psychology: Example from testimony research. *Psychology, Public Policy, & Law, 12*, 190-213. doi: 10.1037/1076-8971.12.2.190.

Yaniv, I., & Foster, D. P. (1995). Graininess of judgment under uncertainty: An accuracy-informativeness trade-off. *Journal of Experimental Psychology: General, 124*(4), 424-432. doi: 10.1037/0096-3445.124.4.424.

Yuille, J. C., & Cutshall, J. L. (1986). A case study of eyewitness memory of a crime. *Journal of Applied Psychology, 71*(2), 291-301. doi: 10.1037/0021-9010.71.2.291.

Yuille, J. C., & Tollestrup, P. A. (1990). Some effects of alcohol on eyewitness memory. *Journal of Applied Psychology, 75*(3), 268-273. doi: 10.1037/0021-9010.75.3.268.

Zaragoza, M. S., & Lane, S. M. (1994). Source misattributions and the suggestibility of eyewitness memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(4), 934-945. doi: 10.1037/0278-7393.20.4.934.

Appendix A: Person Description Instructions (Chapter 3)

Providing a Detailed Person Description

What is the purpose of this form?

This form is designed to enable you provide a detailed person description of the person you saw in the video. The form also aims to capture any uniquely identifying physical features of the individual described.

What should you do now?

If possible, provide the following information about the person you are describing.

Race	Eyes/Ears/Mouth/Nose etc.	Clothing/Shoes
Age	Complexion	Glasses
Gender	Hair colour	Jewellery
Height	Facial hair	Accessories
Weight/build	Scars/Marks/Tattoos	

Please remember that you are the **only person** who has seen the individual in the video. It is therefore important to report **as many details as possible** when giving the description.

In the space below, please write down a detailed physical description of the person you saw in the video Try to describe the person in enough detail so that someone else could use your description to pick them out of a crowd? **Please do not guess** about any detail that you are unsure of.

Useful tip: Research has shown that closing your eyes helps you to remember more details!

In the box below, please try as best as you can to describe the face of the person you saw in the video. Please provide as many details as possible. If you have already provided a description of the face elsewhere in this booklet, take a few moments to consider whether there are any further details you can recall now.

Research has shown that closing your eyes helps you to remember more details!

In the box below, please try as best as you can to describe what the individual in the video was wearing. Please provide as many details as possible. If you have already provided a description of what the individual was wearing elsewhere in this booklet, take a few moments to consider whether there are any further details you can recall now.

Research has shown that closing your eyes helps you to remember more details!

Extending the Description

Even though you have already provided a description of the individual, we want you to carefully consider whether there are any **unique details** that distinguish the suspect from other people.

What should you do now?

Read through the following examples so that you know the type of information that is useful *(note, these are just examples).*

- **Scars / Marks / Tattoos:** (How many? Where? Appearance? Size?)
- **Jewellery:** (Style? Material? Condition? Colour(s)?
- **Teeth:** (Irregular? Chipped? Gold? etc.)
- **Shoes:** (Material? Colour? New or Old? Holes or scuffs? Unusual? etc.)

In the space below, provide a detailed description of any uniquely identifying information that might distinguish this person from other people. As before, it is important to report **as many details as possible** when giving descriptions. Avoid guesses about details you are unsure of.

Research has shown that closing your eyes helps you to remember more details!

Appendix B: Supplemental Materials (Chapter 4)

In this Supplemental Materials section, the instructions given to witnesses are provided (self-generated cues instructions; free recall instructions; person description cues instructions).

Appendix B (i): Self-generated cues instructions

Earlier today you witnessed a video of a crime. Your task is to provide a detailed person description of the person you saw in the video. Please remember that you are the **only person** who has seen the individual. Imagine that the success of this case hinges solely on the description you provide. I have no information about the person in the video other than what you write down. It is therefore important to write down **as many details as possible** when giving the description and take as much time as you require. A description can refer to physical descriptors of an individual but also items of clothing. Try to describe the person in enough detail so that someone else could use your description to pick them out of a crowd. **Please do not guess** about any detail that you are unsure of.

In a moment you will be asked to write down everything you can remember about the appearance of the person in the video. ***Before you do this****, please read and carry out the following instructions that are designed to help you remember as much as you can.*

Without thinking too hard, write down the first six things that you remember seeing or thinking about the person you saw in the video. It doesn't matter what these things are. All that is important is that these are parts of your memory that immediately come to mind when thinking back to the event. Please list them below:

1)

2)

3)

4)

5)

6)

Now, take your time to think about each of the things in your list one at a time. When focusing on each one, think about whether that memory helps you remember other things that also happened in the event. Take as much time as you need to think about each thing separately. Closing your eyes or staring/looking at a blank wall or the floor after reading each one may help you focus and concentrate on this task.

Referring back to the six items you were asked to think about earlier, provide as much detail as possible about the person you saw in the video. **Please do not guess** about any detail that you are unsure of.

[Now, please take your time to answer the next 3 questions]

Referring back to the six items you were asked to think about earlier, describe the face of the person you saw in the video in as much detail as possible. If you have already provided a description of the face you saw, take a few moments to consider whether there are any further details you can recall now.
Referring back to the six items you were asked to think about earlier, describe what the individual in the video was wearing in as much detail as possible. If you have already provided a description of what the individual was wearing, take a few moments to consider whether there are any further details you can recall now.
Even though you have already provided a description of the individual, we want you to carefully consider whether there are any **uniquely identifying details** that might distinguish the suspect from other people (i.e. these are details specific to this person). Referring back to the six items you were asked to think about earlier, provide a detailed description of any **uniquely identifying information** that might distinguish this person from other people. Report **as many details as possible**. Avoid guesses about details you are unsure of.

Appendix B (ii): Free recall instructions

Earlier today you witnessed a video of a crime. Your task is to provide a detailed person description of the person you saw in the video. Please remember that you are the **only person** who has seen the individual. Imagine that the success of this case hinges solely on the description you provide. I have no information about the person in the video other than what you write down. It is therefore important to write down **as many details as possible** when giving the description and take as much time as you require. A description can refer to physical descriptors of an individual but also items of clothing. Try to describe the person in enough detail so that someone else could use your description to pick them out of a crowd. **Please do not guess** about any detail that you are unsure of.

In the space below, please provide as much detail as possible about the person you saw in the video. **Please do not guess** about any detail that you are unsure of.

[Now, please take your time to answer the next 3 questions]

Describe the face of the person you saw in the video, providing as many details as possible. If you have already provided a description of the face, take a few moments to consider whether there are any further details you can recall now.
Describe what the individual in the video was wearing, providing as many details as possible. If you have already provided a description of what the individual was wearing, take a few moments to consider whether there are any further details you can recall now.
Even though you have already provided a description of the individual, we want you to carefully consider whether there are any **unique details** that distinguish the suspect from other people. Provide a detailed description of any uniquely identifying information that might distinguish this person from other people. As before, it is important to report **as many details as possible** when giving descriptions. Avoid guesses about details you are unsure of.

Appendix B (iii): Person description cues instructions

Earlier today you witnessed a video of a crime. Your task is to provide a detailed person description of the person you saw in the video. Please remember that you are the **only person** who has seen the individual. Imagine that the success of this case hinges solely on the description you provide. I have no information about the person in the video other than what you write down. It is therefore important to write down **as many details as possible** when giving the description and take as much time as you require. A description can refer to physical descriptors of an individual but also items of clothing. Try to describe the person in enough detail so that someone else could use your description to pick them out of a crowd. **Please do not guess** about any detail that you are unsure of.

Ethnic appearance

…….

Height

……

Apparent Age

……

Sex

……

Build: Fat / Proportional / Thin / Stocky / Athletic / Heavy / Other

……

Hair Colour: Dark Brown / Light Brown / Fair / Blonde / Grey / White / Black / Ginger / Auburn / Other

……

Eyes: Blue / Brown / Green / Grey / Cast / Staring / Other

……..

Complexion: Fresh / Pal / Ruddy / Tanned / Fair / Freckled / Dark Tone / Mid Tone / Light Tone / Other

……

Facial Hair: Beard / Moustache / Bushy eyebrows / sideburns / other

……

Glasses Worn: Yes / No

Description……...

Marks / Scars / Peculiarities: (Describe below)

……
……
……
……

Suspect's Clothing:

Hat: …… Top Coat: …….

Jacket: …… Jumper: …….

Shirt: ……...Trousers: ……..

Dress: ……..Skirt: ……

Belt: …….Socks: ……

Shoes: ……..Jewellery: ……

Is there anything else you would like to add about the person you saw in the video?

……
……
……
……
……

Appendix C: Ethical Approval Documentation

Each of the four studies presented in this thesis received ethical approval. Study 1 received ethical approval from the University of Portsmouth's Faculty of Humanities and Social Sciences Ethics committee (FHSS), Study 2 received ethical approval from the University of Cape Town's Department of Psychology Research Ethics Committee, and study 3 (Experiment 1 and 2) received ethical approval from the University of Portsmouth's Science Faculty Ethics Committee (SFEC).

Appendix C (i): Ethical approval for Study 1 (Chapter 2)

University of Portsmouth
Faculty of Humanities and Social Sciences
Ethics-fhss@port.ac.uk

27th March 2017

Chairs Action: see document from applicant dated 25th March 2017

Dear Gary Dalton

Study Title:	**Person description information: An examination of frontline communication**
Ethics Committee reference:	**16/17: 31**

Thank you for submitting your documents for ethical review. The Ethics Committee was content to grant a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, revised in the light of any conditions set, subject to the general conditions set out in the attached document.

The Ethics Committee provides a favourable ethical opinion with the following requirements:

1. Ensure that the organisation is fully aware of potential reputational risks. In the view of the Committee more needs to be done to ensure that the organisation is aware of these.

There is no need to submit any further evidence to the Ethics Committee; the favourable opinion has been granted with the assumption of compliance

Appendix C (ii): Ethical approval for Study 2 (Chapter 3)

From: **Johann Louw** <johann.louw@uct.ac.za>
Date: Tue, 5 Apr 2016 at 10:01
Subject: RE: Ethics application
To: Colin Tredoux <colin.tredoux@uct.ac.za>

Go ahead, approved.

Johann Louw

Professor

Department of Psychology

University of Cape Town

Rondebosch 7701

Tel. no. 021 6503414

From: Colin Tredoux
Sent: 24 March 2016 05:23 PM
To: Johann Louw
Subject: Fwd: Ethics application

Hello Johann

I am forwarding this to you since I see that you are referred to in the HOD's recent email as the ethics chair.

Gary will be visiting my research group, and working with us. Hence this application for ethics approval

Please let me know if there is anything else he/we need(s) to declare

Thanks

Appendix C (iii): Ethical approval for Study 3 (Chapter 4)

Gary Dalton
Department of Psychology
University of Portsmouth

Gary.Dalton@port.ac.uk

Science Faculty Ethics Committee
Science Faculty Office
University of Portsmouth
St Michael's Building
White Swan Road
PORTSMOUTH
PO1 2DT

023 9284 3379
ethics-sci@port.ac.uk

20 February 2018

FAVOURABLE ETHICAL OPINION – WITH CONDITIONS

Study Title: "The utility of a self-generated cue mnemonic instruction to facilitate enhanced recall in crimes involving multiple suspects"

Reference Number: SFEC 2018-017

Date Submitted: 05 February 2018

Thank you for submitting your application to the Science Faculty Ethics Committee (SFEC) for ethical review in accordance with current procedures.

I am pleased to inform you that SFEC was content to grant a favourable ethical opinion of the above research on the basis described in the submitted documents listed at Annex A, and subject to standard general conditions (*See Annex B*), with the following specific minor condition and advisory note:

Condition(s)[1]

A. After the data has been collected and digitised, please ensure any hard copies are destroyed.

Advisory Note(s)[2]

i. You may want to conduct a possible power analysis to ensure that the study is adequately powered.

Please resubmit an updated application form incorporating the changes as per the above conditions for the final SFEC records on this application.

[1] The favourable opinion given is dependent upon the study adhering to the conditions stated, which are based on the application document(s) submitted. It is appreciated that Principal Investigators may wish to challenge conditions or propose amendments to these. In that case, please consider the favourable opinion *suspended*, and simply make your case for amending or discarding conditions in writing as you would an application resubmission following ethical review.
[2] The comments are given in good faith and it is hoped they are accepted as such. The PI does not need to adhere to these, or respond to them, unless they wish to.

If you would find it helpful to discuss any of the matters raised above or seek further clarification from a member of the Committee, you are welcome to contact ethics-sci@port.ac.uk who will circulate your queries to SFEC

Please note that the favourable opinion of SFEC does not grant permission or approval to undertake the research. Management permission or approval must be obtained from any host organisation, including the University of Portsmouth or supervisor, prior to the start of the study.

Wishing you every success in your research

John Crossland
Vice Chair Science Faculty Ethics Committee

Annexes

A - Documents reviewed
B - After ethical review - Guidance for researchers

Information:

Professor Becky Milne - Co-Investigator
Professor Lorraine Hope - PhD Supervisor
Rose Barrand - Faculty Administrator

Statement of compliance

SFEC is constituted in accordance with the Governance Arrangements set out by the University of Portsmouth

After Ethical Review

If unfamiliar, please consult the advice After Ethical Review (Annex B), which gives detailed guidance on reporting requirements for studies with a favourable opinion, including, notifying substantial amendments, notification of serious breaches of the protocol, progress reports and notifying SFEC of the end of the study.

Feedback

You are invited to give your view of the service that you have received from the Science Faculty Ethics Committee. If you wish to make your views known please contact the administrator at ethics-sci@port.ac.uk

Appendix D: UPR16 Form

FORM UPR16

Research Ethics Review Checklist

UNIVERSITY OF PORTSMOUTH

Please include this completed form as an appendix to your thesis (see the Research Degrees Operational Handbook for more information

Postgraduate Research Student (PGRS) Information				Student ID:	511053
PGRS Name:	GARY DALTON				
Department:	SCCJ	First Supervisor:	PROF. BECKY MILNE		
Start Date: (or progression date for Prof Doc students)	2012				
Study Mode and Route:	Part-time ☒	Full-time ☐	MPhil ☐	PhD ☐	MD ☐ Professional Doctorate ☐

Title of Thesis:	'He looks like your typical guy': Eliciting person descriptions during frontline interviews
Thesis Word Count: (excluding ancillary data)	33,554

If you are unsure about any of the following, please contact the local representative on your Faculty Ethics Committee for advice. Please note that it is your responsibility to follow the University's Ethics Policy and any relevant University, academic or professional guidelines in the conduct of your study

Although the Ethics Committee may have given your study a favourable opinion, the final responsibility for the ethical conduct of this work lies with the researcher(s).

UKRIO Finished Research Checklist:
(If you would like to know more about the checklist, please see your Faculty or Departmental Ethics Committee rep or see the online version of the full checklist at: http://www.ukrio.org/what-we-do/code-of-practice-for-research/)

a) Have all of your research and findings been reported accurately, honestly and within a reasonable time frame?	YES ☒ NO ☐
b) Have all contributions to knowledge been acknowledged?	YES ☒ NO ☐
c) Have you complied with all agreements relating to intellectual property, publication and authorship?	YES ☒ NO ☐
d) Has your research data been retained in a secure and accessible form and will it remain so for the required duration?	YES ☒ NO ☐
e) Does your research comply with all legal, ethical, and contractual requirements?	YES ☒ NO ☐

Candidate Statement:

I have considered the ethical dimensions of the above named research project, and have successfully obtained the necessary ethical approval(s)

Ethical review number(s) from Faculty Ethics Committee (or from NRES/SCREC):	16/17:31 SFEC 2018-017

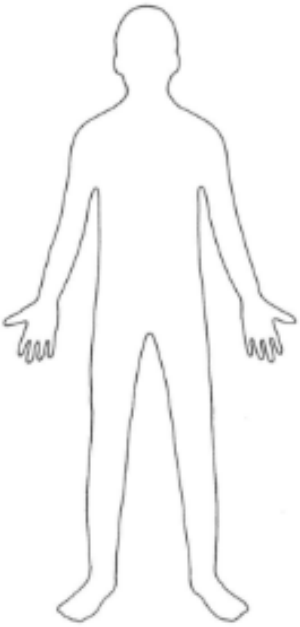
If you have *not* submitted your work for ethical review, and/or you have answered 'No' to one or more of questions a) to e), please explain below why this is so:

Signed *(PGRS)*:	C. Dalton	**Date:** 16.08.2021

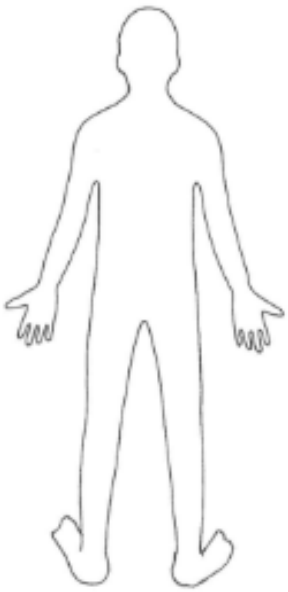
UPR16 – April 2018

Appendix E: Suspect Descriptive Form

SUSPECT/PERSON CONCERNED 1
FRONT



SUSPECT/PERSON CONCERNED 2
BACK



SUSPECT/PERSON CONCERNED 2 34
DESCRIPTIVE FORM

Ethnic appearance as described by witness .....

IC Code ..... Height .....

Apparent Age ..... Sex .....

Build:- * Fat / Proportional / Thin / Stocky / Athletic / Heavy

Other .....

Hair Colour:- * Dark Brown / Light Brown / Fair / Blonde / Grey / White / Black / Ginger / Auburn /

Other .....

Hair Type:- * Bald / Thinning / Receding / Straight / Curly / Wavy / Dyed / Short / Collar length / Shoulder length / Very Long / Wig /

Other .....

Eyes:- * Blue / Brown / Green / Grey / Cast / Staring /

Other .....

Complexion:- * Fresh / Pale / Ruddy / Tanned / Fair / Freckled / Dark Tone / Mid Tone / Light Tone Other .....

Facial Hair:- * Beard / Moustache / Bushy Eyebrows / Sideburns /

Other .....

* Ring where appropriate

Glasses Worn:- * Yes / No Description .....

Accent .....

Words Used .....

* Marks / Scars / Peculiarities:- *(Describe below)*

Suspect's Clothing:-

Hat: ..... Top Coat: .....

Jacket: ..... Jumper: .....

Shirt: ..... Trousers: .....

Dress: ..... Skirt: .....

Belt: ..... Socks: .....

Shoes: ..... Jewellery: .....

Vehicle Used:-

Reg. No.: ..... Make / Model .....

Colour: ..... Type: .....

* Ring where appropriate

Officer's Name/Signature	
Witness's Name/Signature	