‘He was just your typical average guy’

Examining how person descriptions are elicited by frontline police officers

(in press). ***Applied Cognitive Psychology*****.**

Gary Dalton*, Rebecca Milne, Lorraine, Hope, Zarah Vernham & Jordan Nunan

University of Portsmouth

Corresponding author email: Gary.Dalton@port.ac.uk

Author note:

There are no conflicts of interest associated with the research presented in this manuscript.

Data is available upon request from the corresponding author.

Acknowledgements:

We thank the police officers in the participating force, without whom this project could not have been completed.

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. We 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, Lloyd-Jones, & Robinson, 2008). Verbal descriptions of perpetrators are therefore a critical component of the preliminary investigation of a crime (Demarchi & Py, 2009; Meissner, Sporer, & Schooler, 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.

[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.

Perpetrator descriptions provided by witnesses

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, Ask, & Granhag, 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, Martin, & Weber, 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 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-grained detail 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, Koriat, & Weinberg-Eliezer, 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, Brewer and Weber (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, Vagadia, Hope and Gabbert (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, Mycklebust, & Grant, 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, Giebels, Adair, Fennis, & Van Der Zee, 2011; Fisher, 1995), to avoid providing feedback (Semmler, Brewer, & Wells, 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, Memon, & Wright, 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 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, Quas, & Goodman, 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, Nichol, & Loftus, 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, Wilcock, Milne, & Henry, 2009). Indeed, frontline police officers have reported feeling inadequately trained, ill equipped and under pressure when conducting frontline interviews (Dando, Wilcock, & Milne, 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, Fridell, & Lynch, 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, United Kingdom, and Australia; Drover & Ariel, 2015). Early research observed that 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, Farrar, & Sutherland, 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, leading) to gather information (Gabbert, Hope, LaRooy, McGregor, Milne, & Ellis, 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, we 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), we 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). We also expected descriptions provided by witnesses would include more fine grain details than coarse grain details.

Method

Data

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).

[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 we are 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.

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 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 1). 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 vs. 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").

Insert Table 1 about here

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 2, 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 inappropriate. Of the 281 inappropriate questions asked, 44.84% (N_{total} = 126) were leading (see Table 2). An example of a leading question asked by an officer was: "All IC1 [Native White British], all white, all white skinned?".

Insert Table 2 about here

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: Your 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 3 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.

Insert Table 3 about here

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 4). 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 4). 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).

Insert Table 4 about here

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 5). 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 5).

Insert Table 5 about here

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 our 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, Falkner, Trevisan, & McCauley, 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, Wood, & Malpass, 2000). We also observed frontline police officers interrupting 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, we 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, Hope, & Fisher, 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 we 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, Satchell, & Adams-Quackenbush, 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 UK, 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.

Our 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 use of effective questioning techniques for obtaining detailed and reliable information and are aware of existing tools for managing frontline contexts involving multiple witnesses.

References

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

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.

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.

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

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.

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.

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

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

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

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.

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.

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.

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.

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., 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.

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., 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., 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.

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.

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.

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. 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.

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*. Retrieved from https://publicservice.asu.edu/sites/default/files/ppd_spi_feb_20_2015_final.pdf.

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.

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., & 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.

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.

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.

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

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.

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.

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.

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: Oxford University Press.

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). Mahwah, NJ: Lawence Erlbaum.

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.

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.

Wright, A. M., & Alison, L. (2004). Questioning sequences in Canadian police interviews: Constructing and confirming the course of events. *Psychology, Crime and 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.

Table 1. *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?'*

Table 2. *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

Table 3. *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%

Table 4. *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)

Table 5. *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)*