The effects of cognitive load for investigative interviewers

Pamela Hanway

December 2020

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

Conducting an investigative interview is cognitively demanding but there is a paucity of research that has examined the effects of cognitive load for interviewers. The overarching aim of the current doctoral research programme was to understand the cognitive processes of interviewing and how cognitive load for interviewers may impact upon their performance. Across five studies, this thesis explored investigative interviewers' experiences of cognitive load and tested the effects of cognitive load on their performance.

The first investigation (**Chapter 2**), a field study, used an Interpretive Phenomenological Approach (IPA) to analyse investigative interviewers' experiences when conducting interviews with children using the Achieving Best Evidence (ABE) guidelines (Ministry of Justice, 2011). Interviewers described factors that contributed to perceived demands, including overarching cognitive processes, such as remembering information and making judgements, and specific factors, such as the emotionality of a case. Interviewers also described the consequences and impact of cognitive load (e.g., forgetting information and being physically and mentally exhausted after interviewing). The first experimental laboratory research (**Chapter 3)** examined the effects of increased cognitive demands on mock-interviewers' perceived cognitive load (PCL), and their recall of information. Under high, moderate, or low cognitive load conditions, participants watched the free narrative of a child witness. Participants in the high and moderate cognitive load conditions reported higher PCL and were less accurate in their recall of the witness's account, than those in the low cognitive load condition. Experimental laboratory study 2 (**Chapter 4)** examined the effects of cognitive load (high or low cognitive load) on the accuracy of interviewers' memory for information

given by five witnesses, and their accuracy for monitoring the source of information, provided by the multiple witnesses. Participants watched the free narratives of the five witnesses. The witnesses described the same crime, but each witness provided unique details about what they had seen. PCL was higher, and the accuracy of participants' memory for the information provided by the witnesses was lower, in the high cognitive load condition than in the low cognitive load condition. Monitoring the source of information was challenging for all participants, regardless of cognitive load condition. The third experimental laboratory study (**Chapter 5)** explored the effects of note taking on participants' perceived cognitive load and their recall of information given by a witness. The moderating effects of Working Memory Capacity (WMC) and access to notes at recall were also examined. Participants took free notes or structured notes and, during the recall task, they either had access, or not, to their notes. There was also a control condition of no note taking. When taking structured notes, with access to notes at recall, participants' free- and cued-recall was more accurate, than it was for those who took free notes and for those who did not take notes, regardless of their WMC level. WMC and access to notes were moderators of PCL at recall. Finally, **Chapter 6** explored, via an online survey, police officers' PCL for different types of interview (i.e., interviews with victims using the ABE guidelines versus interviews with suspects using the PEACE model), for both serious (e.g., rape) and less serious (e.g., theft) crimes. Police officers indicated that they believe cognitive demands are higher when they interview for serious crimes than when they interview for less serious crimes. PCL was rated as higher when police officers conduct an interview with a witness using ABE guidelines than when they conduct an interview with a suspect.

Taken together, this series of results showed that investigative interviewing is cognitively demanding. Factors that contribute to interviewers' cognitive load were

identified and when interviewers experienced cognitive load there was a negative impact on their recall of the information provided by witnesses. Techniques, such as taking structured notes, as well as interviewers' WMC, may moderate the effects of cognitive load on recall accuracy. Implications of this programme of research for investigative interviewers are considered, and options for additional research suggested, in the general discussion of the thesis (**Chapter 7**).

Table of contents

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 - 45, 945

Table of figures and table

Figures

Tables

Acknowledgements

When I began my first career as a police officer at 19 years of age, completing a PhD was not something I thought I could attempt, let alone achieve. However, my police career and instinctive curiosity were key to my embarking on this PhD. 2020 has been a difficult year for many, so I am really proud to have completed my PhD journey at this time. I would like to recognise, and thank, those who have helped me to reach this point.

I really could not have wished for a greater supervisory team. My sincere thanks to Lucy for mentoring me through this PhD. I shall be forever thankful to you for spotting, then developing, the potential in me to undertake and succeed in this venture. Your sustained encouragement on the sunny up-slopes and slippery downslopes has been a source of continual motivation, even when I couldn't get to grips with those pesky apostrophes! I really appreciate your support through these three, or so, years and the laughs along the way. Zarah, you always had time to spare when I just needed a chat or to discuss "a small stats question" and Lorraine, the knowledge and measured insight I gained from you will always be invaluable. You all made time in your very busy schedules for me, you cared about my research, challenged me, and encouraged me to develop as a scholar. The support you have given me, I know, has made me a stronger researcher. Thank you all.

Athena, along with others at the South Coast Doctoral Training Programme, thank you for your support, for seeing my potential to complete this PhD programme of research, for funding the research and for encouraging me to pursue collaborations across disciplines via the doctoral training programme and my internship.

To my PhD buddies: Stefana, I'm so happy we were able to embark on this journey together. In the office, with Marc, Feni and Alex, a little laughter always went a

long way. The perpetual effort, through weeks and months, of reading and re-reading published papers, gathering data, and working on the analyses, was made all the easier by helping each other out when needed. I know we shall remain friends for ever.

For the Psychology Tech Team at Portsmouth, your willingness to help with my 'can you just…' requests helped smooth my path to get to this point. A large round of applause to you all, thank you!

To my husband Rob: Your love and support, and not taking up the offer to shoot me, for taking on this PhD has been wonderful. Honestly, I won't do any more BSc, MSc or PhDs! Thanks for your ever-present calmness, belief and encouragement, and for refuting my continual doubts about my ability. Also, thank you for reminding me, when I was hitting that metaphoric wall, that completing this PhD will be of benefit to others following in my police career footsteps. I love you lots!

Finally, I would like to thank all my family and our friends for their continued love and support: Martyn, Helen, Lucy and the Hanway family, I know you are very proud of me. I would also like to remember those who have sadly passed away, mum, dad, Arnie and recently my brother John, who I know will be looking on with warm smiles and doffing their caps to the first Doctor in the family…

Dissemination

Publications

Hanway, P., Akehurst, L., Vernham, Z., & Hope, L. (2020). The effects of cognitive load during an investigative interviewing task on mock interviewers' recall of information. *Legal and Criminological Psychology.* https://doi.org/10.1111/lcrp.12182

Hanway, P. (2019, March). Cognitive load at interview: The interviewer's perspective. *CREST Security Review, 9.* https://crestresearch.ac.uk/comment/hanway-cognitive-load-at-interview/

Submitted for publication

Hanway, P., Akehurst, L., Vernham, Z., & Hope, L. (2020). "It's the mental gymnastics": A qualitative analysis of cognitive load experienced by investigative interviewers. *Journal of Investigative Psychology and Offender Profiling.*

Hanway, P., Akehurst, L., Vernham, Z., & Hope, L. (2020). "Who said what? The effects of cognitive load on source monitoring and memory for multiple witnesses' accounts. *Applied Cognitive Psychology.*

Presentations

Hanway, P., Akehurst, L., Vernham, Z., & Hope, L. (2020, September). The effects of Working Memory Capacity and note taking on interviewers' recall of information. Paper presented at the *Digital Postgraduate Presentations* (*DiPP) Conference*, Portsmouth, UK.

Hanway, P., Akehurst, L., Vernham, Z., & Hope, L. (2019, December). The effects of cognitive load on investigative interviewer performance. Paper presented at the *ESRC South Coast Doctoral Training Partnership Final Year Conference*. Brighton, UK.

Hanway, P. (2019, August). The effects of cognitive load for investigative interviewers. Paper presented at a *Meeting at Her Majesty's Inspectorate of Constabulary and Fire & Rescue Services*. London, UK.

Hanway, P. (2019, July). The effects of cognitive load for investigative interviewers. Paper presented at a *Seminar at the Crime and Policing Analysis Unit, Home Office*. London, UK.

Hanway, P., Akehurst, L., Vernham, Z., & Hope, L. (2019, June). The effects of cognitive load on investigative interviewer performance. Poster presented at the *13th Biennial Conference of the Society for Applied Research in Memory and Cognition*. Brewster, USA.

Hanway, P., Akehurst, L., Vernham, Z., & Hope, L. (2018, October). The effects of cognitive load on investigative interviewer performance. Poster presented at the *Postgraduate Research Degree Students' Induction Day*. Portsmouth, UK.

Hanway, P., Akehurst, L., Vernham, Z., & Hope, L. (2018, July). Cognitive load for investigative interviewers and it's effect on performance. Paper presented at the *11th Annual Conference of the International Investigative Interviewing Research Group*. Porto, Portugal.

Hanway, P., Akehurst, L., Vernham, Z., & Hope, L. (2018, June). Interviewing vulnerable witnesses: The nature quality and effects of cognitive load for the interviewer. Poster presented at the *28th Annual Conference of the European Association of Psychology and Law*. Turku, Finland.

Hanway, P., Akehurst, L., Vernham, Z., & Hope, L. (2018, April). The effects of cognitive load on investigative interviewer performance. Poster presented at the *1st Science Together Conference*, Portsmouth, UK.

Hanway, P., Akehurst, L., Vernham, Z., & Hope, L. (2017, December). Interviewing child witnesses: Voices from the frontline. Guest Lecture at the *University of West London*, London, UK.

Chapter 1:

General Introduction

General Introduction

Interviewing is central to the success of a criminal investigation (National Policing Improvement Agency, 2009). The outcome of an investigation will often be determined by the accuracy and completeness of information provided during an interview (Milne & Bull, 2006). To assist with the elicitation of accurate and reliable information, techniques and protocols for interviewing have been the focus of research for several decades (e.g., Hope & Gabbert, 2019; Oxburgh et al., 2010). However, research has also identified that, when conducting investigative interviews in practice, interviewers do not always comply with recommended techniques and protocols (Cross & Hershkowitz, 2017; Powell & Barnett, 2015). If an interview is not conducted appropriately there can be serious consequences for the criminal justice process. For example, a witness's testimony may be discredited, or ruled inadmissible in court, or a case may be dismissed by the court (Criminal Justice Joint Inspectorate [CJJI], 2014). Although, psychological research has identified best practices for conducting investigative interviews, there is a paucity of research examining interviewers' experiences while conducting interviews and the resulting effect on their performance.

The overarching aim of the current doctoral research programme was to understand interviewers' cognitive processes and how the cognitive demands of interviewing may affect their performance. Specifically, the aims of the current PhD research, were to i) examine interviewers' experiences of cognitive demands when conducting interviews, ii) test whether the various cognitive demands of interviewing have an impact on interviewers' perceived cognitive load and their recall of information provided by witnesses, iii) explore factors that contribute to, or reduce, cognitive demands, and (iv) inform how best to manage cognitive load in practical interview settings.

Investigative interviewing is a complex social and verbal interaction between an interviewer and witness (Oxburgh & Dando, 2011). Interviewers are required to actively listen to witnesses, remember what is being said, pay attention to witnesses' needs, make decisions about what questions to ask, and identify topics to pursue (Fisher et al., 2014; Hanway & Akehurst, 2018). Interviewing is, therefore, also a complex cognitive task (Fisher et al., 2014; Lafontaine & Cyr, 2016; Powell, 2002). Across five studies, this thesis explored investigative interviewers' experiences of cognitive load (Chapter 2: Field study); tested the effects of increased cognitive demands on mock-interviewers'[1] perceived cognitive load and their recall of information given by a witness (Chapter 3: Laboratory experiment 1); examined the effects of increased cognitive demands, when interviewing multiple witnesses, on perceived cognitive load and the accuracy of memory for information given by the witnesses and the source of information (Chapter 4: Laboratory experiment 2); tested the effects of taking notes during an interview, and the moderating effects of interviewers' working memory capacity, on perceived cognitive load and the accuracy of recall of information given by a witness (Chapter 5: Laboratory experiment 3); and examined differences in perceived cognitive load when investigative interviewers conduct different types of interview with suspects and victims (Chapter 6: Field survey).

In this general introduction, a summary of globally recommended investigative interviewing techniques will first be provided. Second, the cognitive processes experienced by interviewers, along with their theoretical foundations, will be discussed. Third, cognitive load and the limitations of working memory capacity, which may lead

[1] The term 'mock-interviewer' used in this thesis, refers to participants who completed a witness interview observation task. The 'mock-interviewers' were instructed to watch a recorded interview, they did not engage with the 'witness' during the experimental task.

to cognitive load, will be introduced. Fourth, the impact of cognitive load for interviewers will be considered and finally, a brief description of the five studies comprising this thesis will be outlined.

Investigative interviewing techniques, frameworks, and training

Police investigations are often complex enquiries that involve multiple witnesses, victims, and suspects. Each interviewee could provide essential information that goes on to form the basis of an investigation (College of Policing [CoP], 2019). During human intelligence (HUMINT) investigations, information will also be obtained via interviews with informants (Kontogianni et al., 2018). The importance of conducting these interviews in a manner that allows the interviewee the opportunity to provide accurate and reliable information has been the focus of research since the 1980s. In several high-profile cases, techniques (e.g., asking repeated questions) were used to conduct interviews during investigations and the interviews were later found to be flawed, for example, during the McMartin preschool case in the USA (Ceci & Bruck, 1993; Garven et al., 1998) and, in the UK, unreliable information was obtained through interviews with suspects in the Birmingham Six case (Gudjonsson, 1992, 2002). These, and many other cases, highlight the serious impact inappropriate interviewing can have on the criminal justice process.

Understanding of the fallibility of memory, and the impact interviewing can have for the criminal justice process, has informed various interviewing techniques, protocols and frameworks. Techniques and protocols include the Conversation Management (CM) approach (Shepherd & Kite, 1988), the Cognitive Interview (CI; Fisher et al., 1987), and the Enhanced Cognitive Interview (ECI; Fisher & Geiselman, 1992). More recently, researchers have also advocated the use of the timeline technique to enhance the amount of information witnesses recall about events (Hope et al., 2013;

Hope et al., 2019). When reviewing the investigative interviewing research literature, given that a large volume of the research shows efficacy, there are common practices that are generally accepted and recommended (e.g., asking open questions; Milne & Bull, 1999). Building rapport with the interviewee, to enable interviewee cooperation and to elicit their account, is also generally advised (Gabbert et al., 2021).

Some of these recommended techniques (e.g., rapport building and the ECI) are contained within interviewing frameworks, such as, the PEACE model (an acronym for Planning and preparation; Engage and explain; Account; Closure; Evaluation), which is recommended for interviewing suspects and witnesses (Bull & Soukara, 2009; Kassin et al., 2010; Milne & Bull, 1999). The Achieving Best Evidence guidelines (ABE; Ministry of Justice [MoJ], 2011) and the National Institute for Child Health and Human Development protocol (NICHD; Lamb et al., 2018, Orbach et al., 2000), are recommended for use when interviewing children (ABE and NICHD) and vulnerable or intimidated witnesses (ABE). There are some differences in the practical use of these frameworks, for example, ABE guidance includes information for interviewers when preparing the witness for attendance at court and consideration for the provision of support for the witness, whereas the PEACE model and NICHD protocol focus on guidance relating to the interview alone. However, the various frameworks and protocols follow similar interviewing procedures (i.e., preparation, rapport-building, elicitation of a free narrative, questioning and closure) and also recommend the use of various techniques (e.g., using the ECI and asking open questions).

This abundance of techniques, protocols and frameworks have been developed to provide advice and guidance for practitioners on the optimal approach to obtaining a reliable and complete account from witnesses during an investigation (Bull, 2010; Hershkowitz, 2011; Oxburgh et al., 2015). The information obtained during an

interview is often the key element underpinning a prosecution, therefore, the evidential importance of interviews cannot be overstated (CJJI, 2014). In many jurisdictions, including the UK and other areas of Europe, training in the use of recommended interview techniques has been largely standardised (St-Yves et al., 2014). However, despite receiving training, practitioners in the field often fail to comply with best practice guidance. For example, a lack of compliance with instructions to ask open questions was found in the Netherlands (Otgaar et al., 2019) and in the UK poor compliance with ABE guidelines, including insufficient planning for interviews, was identified (CJJI, 2014).

Examination of interviewing techniques in other jurisdictions has also found that best practices were often not adhered to. For example, in Canada, the recommended interviewing techniques were rarely employed, that is, few open questions and too many closed questions were asked by interviewers (Snook & Keating, 2011); in the US, interviewers rarely engaged in positive interviewing techniques and often interrupted witnesses (Schreiber-Compo et al., 2012); and in Australia, despite improvements following training, there remained a significant gap between interviewing recommendations and the application of them in practice (Powell & Barnett, 2015). It is clear, therefore, that although there have been advances in interviewing techniques, and interviewer training, globally recommended techniques and protocols may not always be applied in practice (Lamb, 2016). Recent research, conducted by Hanway and Akehurst (2018) suggested that interviewers' performance and their compliance with best practices may be impacted by the cognitive demands of conducting an investigative interview, while adhering to the recommended guidelines. The cognitive demands of interviewing and the effects on interviewers' performance are the focus of the current research.

Cognitive processing during investigative interviews

Various elements of an interview (e.g., remembering information, thinking of questions to ask, and monitoring the source of information) involve the cognitive processing of information by interviewers. Although, previous research (e.g., Ayers & Reder, 1998 & Johnson et al., 1993) has focused on eyewitness memory errors, cognitive demands may also lead to errors for investigative interviewers. For example, the accuracy of interviewers' memory is vital to ensure that incorrect, or inaccurate, information is not provided to interviewees during questioning. False or incorrect information (i.e., misleading post-event information) can have a negative impact on the accuracy of interviewees' accounts (Ayers & Reder, 1998). Additionally, errors, or disruptions, in monitoring the source of information can have implications for memory, and knowledge, of an event (Johnson et al., 1993). If there are multiple sources of information (i.e., accounts from several witnesses) the interviewer must also accurately attend to the source of the information. These cognitive processes take place when interviewers are managing the content and course of the interview, but cognitive processing of information also takes place pre-interview (Mortimer, 1994).

The planning and preparation phase is one of the most important elements of an investigative interview. This phase relates to 'P' in the PEACE model and is included as a specific phase in the ABE guidelines (MoJ, 2011). It is generally recognised that to obtain accurate and reliable information from a witness, planning for an interview is beneficial (Brandon et al., 2018). When planning an interview, the witness's welfare and additional communication needs should always be considered prior to the interview (Association of Chief Police Officers, 2016). A lack of time available for planning, and the pressure to conduct an interview quickly due to a heavy workload, may impact on interviewers' preparations (Dando et al., 2008). This lack of preparation may increase

cognitive demands for the interviewer, which can result in errors of interviewer's judgement of a witness's account and their analysis of the information that has been provided (Brandon, et al., 2018).

Before and during an interview, rapport building is also viewed as an essential element of investigative interviews (Meissner & Lyles, 2019). Many interviewing frameworks (e.g., ABE and PEACE) recommend that interviewers connect, and build a rapport, with witnesses, particularly with witnesses who are reluctant to engage with an investigation (MoJ, 2011). Rapport building is effective for increasing the quality of information, and reducing the likelihood of errors, provided by witnesses (Vallano & Compo, 2011). However, rapport building is a complex cognitive process, an interviewer uses reasoning and judgement along with their interpersonal skills to understand and interpret witnesses' responses to questioning (Alison et al., 2013). Some frameworks (i.e., ABE and NICHD) recommend that when building rapport with children, and other vulnerable witnesses, attention must be paid to their needs (MoJ, 2011). For example, interviewers should build rapport to put children at ease and to gauge their language and cognitive skills, which will direct the way the interview is conducted (Saywitz et al., 2015). Although, care should be taken when interviewing all interviewees, children and other vulnerable witnesses, can have particular needs (e.g., those related to their knowledge about the proceedings or communication abilities) that an interviewer should also attend to during interviews (Zajac & Brown, 2018). Building rapport may, therefore, add to the cognitive demands of processing information during an interview.

The questioning phase is arguably the key element in many interviewing protocols and frameworks (e.g., PEACE, ABE, and NICHD). The recommendation to ask open questions is included as a specific interviewing technique in most interview

protocols (Fisher et al., 2014). Open questions typically lead to accurate and detailed, free narrative responses from interviewees (Dale et al., 1978; Hershkowitz, 2001). Asking open questions is preferred to asking other types of question (e.g., leading or suggestive questions; Danby et al., 2017). However, interviewers have identified that distinguishing open questions from specific questions (e.g., who, what, or where, type questions) is difficult and asking open questions has been found to be a relatively unfamiliar and novel skill (Wright & Powell, 2006).

In response to asking an open question, interviewers actively listen to witnesses whilst a free narrative account is given (Brandon, et al., 2018). As mentioned previously in this review, interviewers must accurately remember these, often numerous, details and, at the same time, think of questions to ask witnesses, which comply with best practice techniques. These cognitive processes will likely increase cognitive demands for interviewers (Hanway & Akehurst, 2018).

In addition to the core elements of interviewing (e.g., listening, remembering information and questioning), note taking is often a feature of investigative interviewing. In some jurisdictions, audio or video recordings are made when interviewing witnesses and suspects, for example, when using the PEACE model for suspect interviews (Police and Criminal Evidence Act, 1984[2]) and when using the ABE model for interviewing vulnerable witnesses (MoJ, 2011). Even when suspect interviews are being recorded, it is recommended that interviewers take notes (CoP, 2019). When conducting ABE interviews, guidance suggests that interviewers should consider taking notes to assist them during the free narrative phase of the interview (MoJ, 2011). Jansen and colleagues (2017) highlighted that taking notes enabled

[2] Police and Criminal Evidence Act 1984 (PACE) codes of practice - GOV.UK (www.gov.uk)

interviewers to process and respond to the information provided by suspects. Note taking provided a basis for questions that needed to be asked to clarify or challenge interviewees' accounts (Jansen et al., 2017). However, Cauchi and Powell (2009) found that notes taken during interviews with child witnesses were not always accurate; 15% of notes contained one or more errors of commission (i.e., the addition of incorrect information). Similarly, in a study examining 20 forensic interviews with children, Lamb and colleagues (2000) found that 25% of the forensically relevant details in the interviews were not present in the interviewers' notes. When taking notes, several complex cognitive processes involved in language processing are simultaneously activated (i.e., listening, comprehension, and writing), therefore, note taking can be cognitively demanding (Piolat et al., 2005). Completing structured notes (e.g., with sub-headings) may also add complexity to note taking and may be difficult during an investigative interview (MacDonald, 2016). Therefore, taking notes may add to the cognitive demands for interviewers, which may ultimately reduce the accuracy of interviewers' recall of the information elicited from interviewees.

Taken together, the research highlighted here indicates that several elements of an investigative interview involve complex cognitive processes for interviewers. There are several recommendations for interviewing that may impact the cognitive demand for interviewers, for example, if planning and preparations are not completed (Brandon et al., 2018). Rapport building and recommended questioning are complex skills, requiring that interviewers actively listen, remember information and make reasoned judgements (Hanway & Akehurst, 2018). The recommendation that notes should be taken during an interview may also increase cognitive load for interviewers (Piolat et al., 2005).

Working memory and cognitive load

Several of the cognitive processes for interviewers, for example, holding in mind information provided by an interviewee whilst thinking of questions to ask, take place in working memory (WM). WM comprises processing units for visual/spatial and auditory/verbal information, which interact with long term memory, and are necessary for the management of information during complex cognitive tasks (Baddeley, 1992; Baddeley & Hitch, 1974). WM is a functional short-term memory system that enables temporary storage and maintenance of information, which is necessary for many cognitive processes, such as reasoning and language comprehension (Baddeley, 1992). Thus, the role of WM is central to all deliberate cognition, including prose comprehension and learning (Oberauer et al., 2018). Working Memory Capacity (WMC) represents an individual's capacity to i) process information relating to a primary task, ii) maintain the relevant information for the primary task, and iii) access and retrieve information from long-term memory, in the presence of a distraction (Unsworth & Engle, 2007).

The capacity limitations of WM mean that the rehearsal of received sensory information, and the processing of information, may be limited or restricted (van Merrienboer & Sweller, 2010). Controlled processing is needed to complete cognitive tasks that require attention and the management of information (Bargh, 1984). However, this type of processing is slow and effortful and relies on our limited attention capacity (Strayer & Drews, 2007). An attentional bottleneck can occur where attending to one element of information causes other cognitive processes, and the associated information, to be neglected (Strayer & Drews, 2007). High levels of focused attention can be accomplished with effort (Bargh, 1984; Schneider & Shiffrin, 1977), but errors occur if individuals cannot meet the mental demands required to effectively complete

tasks (Paas & van Merrienboer, 1993). During complex tasks, there is an increase in cognitive demand; thus, the amount of mental effort required also increases (Kleider-Offutt et al., 2016). The attentional demands required to perform complex tasks may lead to cognitive load and errors, or a reduction in performance (Engle & Kane, 2004; O'Donnell & Eggemeier, 1986).

Cognitive load, then, refers to the mental workload placed on individuals when they are required to undertake activities (Hart & Staveland, 1988; Van Acker et al., 2018). It signifies WM use and the demands placed on cognitive resources when carrying out multiple and competing tasks (Dias et al., 2018; Engström et al., 2013). Cognitive Load Theory (CLT) identifies three types of load (Sweller, 1988, 1994; Sweller et al., 1998) that are relevant in a variety of applied settings (Galy et al., 2018). The first type, intrinsic load, relates to the load imposed by the fundamental nature of the information being processed and the natural complexity of the task (Schnotz & Kurschner, 2007). The second, extraneous load, is induced by other external factors, such as time pressure (Galy et al., 2012). The third type of load described within the CLT is germane load, which is the load used for learning, the development of skills, and the application of skills in novel situations (Paas et al., 2004). Germane load is required for the construction and automation of schemas for a particular task (Galy et al., 2018).

In a review of the CLT literature, Schnotz and Kürschner (2007) suggest that CLT, as a conceptual framework, is based on knowledge about human cognitive architecture and is useful for empirical research and for research-based practice. However, subjective, physiological, or performance-based methods, to measure cognitive load are generally aimed at measuring the total load experienced and they do not distinguish between intrinsic load, extraneous load and germane load (Schnotz & Kürschner, 2007). The use of modern physiological instruments to gain reliable

measures of cognitive load could be useful for CLT research in more real-world tasks (Johanesson et al., 2020). However, there is no 'silver bullet' in relation to the best physiological measures of cognitive load (Charles & Nixon, 2019). If experimental designs (e.g., Yurko et al., 2010) combine multiple variables that include intrinsic, extraneous, and germane loads, then differentiation of the types of load imposed may be difficult to interpret, but it is important to manage cognitive load regardless of the source (Ayres, 2020).

There is an abundance of literature examining the effects of cognitive load when applied to tasks in several occupational roles. For many studies, the NASA Task Load Index (NASA-TLX) has been used to measure participants' perceived cognitive load (Hart, 2006). The tool was developed as a measure of workload and identifies factors which contribute to cognitive load when completing a variety of tasks (Hart & Staveland, 1988). Six dimensions were selected by Hart and Staveland (1988) for use in the NASA-TLX after extensive analysis of the primary factors that do (and do not) define the subjective experience of workload for a variety of activities ranging from simple laboratory tasks to flying an aircraft (Hart, 2006). The six dimensions, which represent independent clusters of variables, are mental demand, physical demand, temporal demand, frustration, performance, and mental effort. A combination of these dimensions represents the workload experienced by most people performing most tasks (Hart & Staveland, 1988).

Cognitive load, measured using the NASA-TLX, has been associated with poorer performance during laparoscopic surgery (Yurko et al., 2010). During surgeon simulator training, increased errors were made by medical students when they were under increased cognitive load (Dias et al., 2018; Haji et al., 2015). In recruitment settings, when interviewers were placed under a high cognitive load, they had difficulty

gathering information and generated fewer questions, than those in a low cognitive load condition (Nordstrom et al., 1996). Under higher cognitive load conditions, interviewers' decision-making was also impacted (i.e., less time was taken to make decisions, and interviewers used more automatic and less controlled processing; Frieder et al., 2016). These studies show that, in a variety of settings, when cognitive demand exceeds capacity it can lead to cognitive load (Oberauer et al., 2016; Sweller, 2016).

Cognitive load for investigative interviewers

The WM system and WMC, that can result in cognitive load, is relevant when considering the cognitive processing of information by investigative interviewers. Elements of investigative interviewing, as outlined in this review, reveal several intrinsic (e.g., remembering information) and extraneous (e.g., arranging the interview) cognitive processes that must be attended to by interviewers. During training, and with experience, germane load is required for interviewers to apply their knowledge in practice. That is, interviewers are required to, actively listen to witnesses, remember what is being said, pay attention to the witnesses' needs, make judgements about the content of their accounts, make decisions about what questions to ask, identify topics to pursue, and seek clarification from witnesses (Hanway & Akehurst, 2018; Fisher et al., 2014). In addition, when conducting interviews for investigations that include multiple witnesses, to enable differentiation between witnesses' accounts, interviewers must also accurately monitor the source of information (Kontogianni et al., 2018). Monitoring the source of information, that needs to be remembered, is important in many situations as it enables judgements to be made about the information (Johnson et al., 1993).

The current thesis considers how the various elements, which comprise the task of investigative interviewing, may lead to cognitive load for interviewers. The impact that cognitive load may have on interviewers and their performance was also considered. For example, as discussed, planning and preparation are important before an investigative interview, but a lack of planning and preparation due to time and resource constraints may increase the cognitive demands experienced by interviewers (Hanway & Akehurst, 2018). Asking open questions leads to detailed free narrative responses from interviewees (Dale et al., 1978; Hershkowitz, 2001) but, due to an overload of information, interviewers may find it difficult to formulate open questions because their recall of information previously given by a witness may be limited and inaccurate (Hyman-Gregory, 2009). This is an important finding as the impaired recall of interviewers may introduce erroneous information to witnesses, which could have an impact on the subsequent accuracy and reliability of the witnesses' testimonies (Loftus & Pickrell, 1995; Gudjonsson, 2010). In addition, taking notes during an investigative interview is recommended, but the notes taken may not reflect the content of the interview fully or accurately (e.g., Cauchi & Powell, 2009).

Overall, although investigative interviewing is recognised as a cognitively demanding task, cognitive load as perceived and experienced by investigative interviewers, has to date, received little empirical attention. Additionally, several cognitive processes for investigative interviewers likely take place in WM, but the impact that interviewers' WMC has on recall of information provided by a witness has not previously been examined.

Aims and outline of the current thesis

Based on the reviewed literature, this doctoral thesis provides an examination of the effects of cognitive load for investigative interviewers. The thesis consists of five

studies that are described across Chapters 2 to 6. Each study takes a different approach, and applies psychological theories, to examine the effects of cognitive load for investigative interviewers. Each chapter has been written independently of all other chapters so that they can be read and understood individually. Chapter 2 was published as a peer-reviewed journal article and chapters 3 and 4 have been submitted for publication. It is important to note that, as a result, there is some repetition in the discussion of theory throughout the different chapters of this thesis.

The overarching aim of the research was to examine three questions, i) how do interviewers experience cognitive load when conducting interviews; ii) what impact do the various cognitive demands of interviewing (i.e., remembering information, thinking of questions to ask, monitoring the source of information and taking notes during an interview), have on interviewers' perceived cognitive load and their recall of information given by witnesses, and iii) what factors contribute to, or reduce, cognitive load, and which of these factors may help interviewers to manage cognitive load in practical interview settings?

Previous studies have has identified that interviewing is cognitively demanding (e.g., Fisher et al., 2014), but limited research has examined the current research questions, as outlined above. Therefore, initially, a qualitative field study was carried out to seek the perceptions, and to record the interviewing experiences, of police officers in the UK. Three experimental laboratory-based studies were then conducted to examine, in a controlled setting, factors that contribute to the cognitive demands of interviewing. In Chapters 2, 3, 4 and 5, the NASA-TLX measure was used to examine perceived cognitive load when participants completed investigative interviewing tasks. The final study was a survey of police interviewers, which aimed to examine whether

the results identified in the first four studies were experienced across different interviewing situations in the field. A summary of each chapter is provided below.

The first study (Chapter 2) was a field project involving the Interpretive Phenomenological Analysis (IPA) of investigative interviewers' experiences when conducting interviews, using ABE guidelines, with vulnerable witnesses including children. Interviewers described factors, which contributed to their perceived cognitive demands when interviewing. The interviewers also explained the effect these demands had on them and their performance during interviews. Interviewers described how they managed the effects of cognitive load during, and after, interviews. The outcomes of this study informed the four further studies of this doctoral research programme.

The aim of the first laboratory experiment (Chapter 3) was to manipulate the demands of an interview task and to measure whether the manipulation had an effect on the perceived cognitive load of the participants and whether there were differences in the participants' recall of information that had been provided by a witness. Based on previous research regarding cognitive load in other occupational settings, it was hypothesised that increasing the cognitive demands during an interview task would result in higher perceived cognitive load and a reduction in the accuracy of participants' recall of information provided by a witness.

The second laboratory experiment (Chapter 4) built on the findings in Chapter 2 that police officers reported that interviewing multiple witnesses in the same case increased the cognitive demands for interviewers, which may have an impact on their performance (i.e., they may introduce erroneous information to witnesses). As such, this study examined the accuracy of participants' recognition of information provided by multiple witnesses and the accuracy of the participants' monitoring of the source of the

information. The cognitive demands of the interview task were manipulated and participants' perceived cognitive load during their tasks was also measured.

The third laboratory experiment (Chapter 5) explored the effects of note taking on participants' perceived cognitive load and their recall of information given by a witness. The moderating effects of WMC limitations were also explored. As outlined above, Piolat and colleagues (2005) suggested that note taking can be cognitively demanding, however, it can also aid recall (Jansen et al., 2017). In addition, the findings from the first study of this doctoral work indicated that some interviewers reported they always take notes during an interview as they cannot recall the information that is provided, whereas others reported that they do not take notes as they find it distracting and they can remember what has been said by witnesses without taking notes. Thus, the objective of this study was to examine the individual differences of participants in terms of WMC, and the impact of note taking, on their perceived cognitive load and their recall of information given by a witness.

The fifth and final study for this doctoral programme of research (Chapter 6) was a survey study. The three laboratory experiments outlined above (Chapters 3, 4, 5) focused on elements of an investigative interview that are common across most interview situations. The field study (Chapter 2) was conducted with interviewers who were trained in the use of the ABE guidelines and focused on the cognitive demands of interviewing vulnerable witnesses including children. The survey for the final study explored differences between the reported experiences of police officers when they interview witnesses, using the ABE guidelines, and when they interview suspects, using the PEACE model, for serious (e.g., rape) or less serious (e.g., theft) crimes. Factors, which were identified in the field and experimental studies, that may contribute to cognitive load when conducting different types of interview were examined.

The final chapter (Chapter 7) of this thesis provides a general discussion and overview of the key findings of this programme of research. Theoretical limitations of the research, along with practical implications for interviewing in applied settings, and suggestions for future research of cognitive load for investigative interviewers, are also provided.

Chapter 2:

"It's the mental gymnastics":

A qualitative analysis of cognitive load experienced by investigative interviewers

This chapter is under review as:

Hanway, P., Akehurst, L., Vernham, Z., & Hope, L. (2020). "It's the mental gymnastics": A qualitative analysis of cognitive load experienced by investigative interviewers. *Journal of Investigative Psychology and Offender Behaviour.*

Abstract

The current study aimed to examine factors, as described by investigative interviewers, that contribute to cognitive load during interviews with vulnerable witnesses. Semi-structured interviews were conducted with eight trained investigators from one UK police force. They were asked about their thoughts concerning, and experiences of, investigative interviewing with vulnerable witnesses. An Interpretative Phenomenological Approach was used to analyse the interview transcripts. Emergent themes were identified including overarching cognitive processes, such as remembering information and making judgements, and specific factors associated with interviewing that influence cognitive demands such as the emotionality of a case. The consequences and impact of cognitive load (e.g., forgetting information and being physically and mentally exhausted after interviewing) also emerged as themes. Despite some investigative interviews being complex and demanding, it is suggested that cognitive load could be managed by reducing extraneous load (e.g., time pressures) and increasing interviewers' skills, so aspects of interviewing become more automatic.

Introduction

A number of cognitive processes must be attended to when investigators are conducting interviews with witnesses. For example, interviewers are required to actively listen to witnesses, remember what is being said, pay attention to the witnesses' needs, make judgements about the content of their accounts, and make decisions about what questions to ask (Hanway & Akehurst, 2018). Best practice guidelines recommend that open-ended questions should be asked when building rapport with vulnerable witnesses and that open questions should be favoured, and (mis)leading questions avoided, throughout interviews (e.g., Ministry of Justice [MoJ], 2011). These different elements of interviewing are essential but can be cognitively demanding (i.e., mentally taxing) for interviewers (Frieder et al., 2016). There is a paucity of research examining the demands on interviewers that may lead to cognitive load while conducting interviews. The goal of the current study was to explore, via the reported perceptions and experiences of trained interviewers, factors that may contribute to an increased cognitive load. Issues, as described by the participants, were analysed to better understand the impact cognitive load has on interviewers in the field.

Cognitive load: Definition and effects

Cognitive load, in this context, signifies the demands placed on mental resources when carrying out multiple and competing tasks (Dias et al., 2018; Engström et al., 2013). It relates to variations in mental workload for individuals (e.g., as a result of their experience or the task difficulty) when they undertake an activity (Van Acker et al., 2018). Working memory capacity is relevant to understanding cognitive load because working memory and attentional processes are related to the maintenance and/or suppression of information (Engle, 2002). Working memory comprises processing units for visual/spatial and auditory/verbal information, which interact with long term

memory, and are necessary for the management of information during complex cognitive tasks (Baddeley, 1992). However, our working memory is limited (Engle et al., 1999). This limitation is widely acknowledged and can lead to cognitive load (Oberauer et al., 2016, Sweller, 2016).

Cognitive load, therefore, refers to the working memory capacity that is allocated to meet the demands of a task (Paas et al., 2003). It comprises three categories; intrinsic load, which is the load induced by the inherent nature of the task and the information being processed; extraneous load, which is induced by factors external to the task, including the situation, work organisation, or time pressures; and germane load, which is defined as the load placed on working memory during schema formation and automation (Sweller, 1994). Germane load also refers to the application of skills and knowledge in unique situations. These three types of load are additive and when combined, to maintain performance, they should not exceed available working memory capacity (Galy et al., 2012).

There is an abundance of literature examining the effects of cognitive load when applied to tasks in a number of occupational roles. For example, cognitive load has been associated with poorer performance during laparoscopic surgery (Yurko et al., 2010). During surgeons' simulator training, increased errors were made by medical students under increased cognitive load (Dias et al., 2018; Haji et al., 2015). In recruitment settings, when interviewers were placed under a higher cognitive load they had difficulty gathering information and generated fewer questions than those in a lower cognitive load condition (Nordstrom et al., 1996). Under higher cognitive load conditions, interviewers' decision-making was also impacted (i.e., less time was taken to make decisions, and they used more automatic and less controlled processing; Frieder et al., 2016). These examples show that the capacity to perform complex tasks relies on

an ability to access and retain task-relevant information over time and to selectively process that information during the task (Fougnie & Marois, 2006).

Interviewing vulnerable witnesses

The elicitation of information from vulnerable witnesses is a unique and complex task that requires specialist skills in investigative interviewing (Powell et al., 2005). Guidance and protocols have been developed, which document approaches for investigative interviewers to use when they conduct interviews with vulnerable witnesses. For example, the Achieving Best Evidence guidelines (ABE; MoJ, 2011) and the National Institute for Child Health and Human Development protocol (NICHD; Lamb et al., 2018, Orbach et al., 2000) are recommended for use when interviewing children (ABE and NICHD) and vulnerable or intimidated witnesses (ABE). The guidance identifies additional factors that need to be considered before and during an interview, including the age, communication and cognitive abilities, and vocabulary of the interviewee (Zajac & Brown, 2018). The witness's emotional state, relationship with the alleged perpetrator, their overall sexual knowledge and experiences (if the case is related to sexual offences), and any significant recent stressors and life events (e.g., bereavement or domestic violence) must also be considered by interviewers (MoJ, 2011). Using the guidance appropriately (e.g., by asking open questions) is associated with an increase in the amount and accuracy of information elicited from witnesses (Brown et al., 2015; Danby et al., 2017). However, interviewers do not always comply with these best practice recommendations. For example, planning the interview is not always completed as required and leading questions may be asked (Criminal Justice Joint Inspectorate [CJJI], 2014; Lafontaine & Cy, 2016; Powell et al., 2010).

Non-compliance with guidance can have serious consequences for the criminal justice process, for example, a witness's testimony may be discredited or ruled

inadmissible in court or a case may even be dismissed by the court (CJJI, 2014). The cognitive demands on interviewers have been identified as a possible barrier to compliance with best practices (Hanway & Akehurst, 2018). Previous research examining the effects of increased cognitive load on interviewers' judgements of emotionality and the veracity of a victim's statement found greater reliance on stereotypic expectations and more use of heuristic strategies for making judgements under high cognitive load, in comparison to no load, conditions (Ask & Landstrom, 2010). These findings suggest that more automatic and less controlled processing was used by interviewers when under higher cognitive load, which may have a negative impact on the quality of interviews (Frieder et al., 2016). However, understanding how these demands may impact on interviewers' cognitive load in practice requires further exploration.

The current study

To further understand the challenges that are experienced by interviewers when conducting interviews with vulnerable witnesses, the factors, described by the participants, which likely contribute to cognitive load were examined and the impact of these factors for interviewers were explored. An Interpretive Phenomenological Analysis (IPA) approach was used to investigate field-based interviewers' perceptions and their experiences of cognitive load when interviewing vulnerable witnesses. This inductive approach was selected as no previous research has explored how interviewers experience and deal with the complexity of cognitive demands when conducting these specialist interviews. The current research questions were (i) what factors of an interview contribute to the cognitive load for interviewers, and (ii) what are the effects of cognitive load on interviewers and their performance during interviews?

Method

Research design

Qualitative research aims to discover the interpretations and perspectives of participants regarding a particular issue (Dixon-Woods et al., 2005). For the current study, an IPA approach was chosen to investigate cognitive load. This type of analysis requires the researcher to pay critical attention to how a phenomenon is being experienced and presented by participants; enabling the researcher to enhance and develop current knowledge (Findlay, 2014). Using the IPA approach, the subjective everyday experiences of people in the situation of interest was specifically sought (Smith et al., 2009). This exploration was coupled with a subjective and reflective process of interpretation (Reid et al., 2005). A process of interpretation was applied in the current study, i.e., the participants' interpretation of their experience (hermeneutics) and the researcher's interpretation of the participants' description of their experiences (double hermeneutics; Smith et al., 2009). In sum, during their interviews, the participants made sense of their experiences and then the Principal Investigator (PI) made sense of the participants' descriptions of them.

The PI was formerly a police detective, with experience of interviewing vulnerable witnesses. Her prior experience enabled communication with the participants in a more meaningful way. For example, when police policies and practices were referred to by participants, they were easily understood without additional explanation being necessary. The PI took care to ask open questions and to examine the full transcripts systematically in order to remain as objective and open-minded as possible during both data collection and later analysis (Smith et al., 2009).

Participants

The current research questions focused on participants' experiences of interviewing vulnerable witnesses. Therefore, a purposive homogenous sample of eight police officers who had been trained to conduct interviews with vulnerable witnesses using the ABE guidelines (MoJ, 2011) was recruited. A sample of eight participants allowed for examination of similarities, and differences, between individuals' experiences and perceptions, whilst at the same time ensuring the PI was not overwhelmed with data. A sample of six to eight participants is appropriate for an IPA study of this nature (Pietkiewicz & Smith, 2014).

Participants were recruited from one UK police force. In line with the force research participation requirements, support for the research project and permission to approach suitably trained participants were obtained in advance. A force interviewing advisor contacted qualified ABE interviewers and provided them with information regarding the research study. Potential participants were told that the research aimed to investigate police officers' experiences when conducting interviews with vulnerable witnesses using techniques outlined in the ABE guidance (MoJ, 2011). They were informed that the goal of the research was to gain an understanding of conducting these interviews from the interviewers' perspectives to inform further research and improve training for police officers. Potential participants were advised that their participation was entirely voluntary and that if they agreed to take part, they would be interviewed individually for approximately 60 minutes and that their interview would be recorded. The onus was on potential participants to notify the PI that they were willing to participate. Arrangements were then made for their interview to take place at a mutually agreed time and place.

The sample of eight participants comprised five females and three males aged between 37 to 62 years (M = 47.00 years, SD = 7.93 years). The number of years since participants had been trained to interview vulnerable witnesses using ABE guidelines ranged from 4 to 24 years (M = 14.00 years, SD = 7.09 years). The length of police service of the participants ranged from 12 to 42 years (M = 23.63 years, SD = 9.15 years).

Interviews

Participants were interviewed by the PI in a private office at each participant's workplace. Anonymity, data security and confidentiality were discussed with the participants. Any questions were answered and written informed consent was obtained prior to the commencement of the interviews.

To obtain detailed first-person accounts about the participants' experiences, one-to-one interviews were conducted (Turner, 2010). During the interviews, a semi-structured interview protocol was followed by the PI (see Appendix A.1). Initially, the PI informed participants that she would like them to give as much information as they could about what ABE interviewing was like for them. To encourage participants to fully express their viewpoints and experiences, open-ended questions were asked (Turner, 2010). For example, "Tell me what the initial part of an interview is like from your perspective?" or "What is it like for you when you ask open-ended questions?". Open prompts were also used (e.g., "Tell me more about that?" or "Give me an example of that?"). A semi-structured interview was chosen as it allowed for a dialogue between the PI and participants. The discussion enabled rich, full accounts of the participants' experiences to be elicited (Pietkiewicz & Smith, 2014).

All interviews were audio-recorded using a digital recorder. Following completion of each interview, participants were asked if they wished to clarify any

points and were given the opportunity to ask questions. At the conclusion of the interviews, participants were thanked and debriefed. The interviews ranged in duration from 33.41 minutes to 52.49 minutes (M = 41.24 minutes, SD = 6.02 minutes).

Data analysis

The PI transcribed verbatim each of the audio-recorded interviews. Utterances by the interviewer and participants were included to allow content analysis of the participants' accounts. As the IPA approach focuses primarily on the verbal content of participants' accounts, details of the length of pauses and speech disturbances were not documented in the transcripts. The entire transcript for each participant was then analysed according to IPA principles (i.e., to explore, investigate, and interpret, the experiences of the research participants when they are conducting investigative interviews with vulnerable witnesses).

The eight transcripts were read and re-read, one at a time, and content relevant to the aims of the study (i.e., factors that likely contribute to cognitive load and the impact of these factors for the interviewer) was identified. Initial notes and comments were made regarding the transcript wording, including descriptive features (e.g., key words or phrases used by participants), linguistic features (e.g., any metaphors used), and conceptual features (e.g., participants' understanding of the issues they were describing). This enabled the PI to interpret, through experience and professional knowledge, the meanings that participants made of their experiences (as recommended by Smith et al., 2009). The next step was the identification of emergent topics within individual transcripts. Recurrent topics were identified across the transcripts and the descriptions of these were clustered into discrete themes (Starks & Brown-Trinidad, 2007). These were then condensed, and common elements were fitted together. This integration resulted in super-ordinate and sub-ordinate themes, which represented the

participants' subjective experiences of interviewing vulnerable witnesses (see Table 2.1). An illustration of a transcript and theme development is provided in Appendix A.2.

Results

Interpretative Phenomenological Analysis (IPA) is an inductive process, as such, there was no previously established hypothesis to be tested. The findings presented here are supported by evidence in the form of quotes and excerpts from the participants' accounts. The PI used her knowledge of the types of questions recommended for investigative interviewing and reflected on the information provided by participants to draw inferences, through interpretation, from the comments made. For example, one participant said, "I'm particularly conscious of the type of questions I'm asking and how they are being perceived by the person I'm speaking to [] conscious about no leading questions, not introducing stuff that the child has not introduced themselves" (P3). This was interpreted as the interviewer being focused on the cognitive process of asking questions that comply with their training, i.e., not asking leading questions, and being aware that an error may impact their interview, which adds to the cognitive demands of the interview. The quotes and results summary were returned to participants for their feedback. All eight participants responded that they agreed with the accuracy of the summary of their interview. Each quote used in this chapter is followed by a participant number (P1, P2, etc.). The symbol [] indicates that redundant words have been removed.

Using the IPA approach, and the process of reflection and drawing inferences, three super-ordinate and seven sub-ordinate themes were identified. These themes distinguish factors that were described by participants and interpreted by the PI, as contributing to cognitive load for the participants when they interview vulnerable

witnesses. The themes also include an interpretation of the effects of cognitive load on the participants. The themes are summarised in Table 2.1.

Table 2. 1

Number of participants who referred to the identified themes during their interviews.

	Participants who mentioned a theme								
	P1	P2	P3	P4	P5	P6	P7	P8	Total
1 Overarching cognitive demands of interviewing									
1A The interviewer's cognitive processing	√	√	√	√	√	√	√	√	8
1B The investigative context	√	√	√		√	√	√	√	7
2 Specific factors that influence cognitive demands									
2A The interviewer's perceptions of their own ability and individual pressures	√	√	√	√	√	√	√	√	8
2B The interviewee's needs and their contribution	√	√	√	√	√	√	√	√	8
3 Consequences and impact of cognitive demands									
3A Performance in interviews	√	√	√	√	√	√	√	√	8
3B Effect of cognitive demands on the interviewer	√	√	√	√	√	√	√	√	8
3C Strategies to prepare for and conduct interviews	√	√		√	√	√	√	√	7

Super-ordinate theme 1: Overarching cognitive demands of interviewing

Participants described common overarching factors that influence all investigative interviews with vulnerable witnesses. These included the multiple cognitive processes that interviewers experience and specific features of complex cases they are investigating.

Sub-ordinate theme 1A: The interviewer's cognitive processing during an interview.

When participants explained their experience of interviewing vulnerable witnesses, they described a number of cognitive processes that they attend to during the interviews. These were paying attention and actively listening to the witness, monitoring their own behaviour, making decisions regarding what questions to ask, judging the ability, reliability and needs of the interviewee, and making reasoned judgements about the relevance and reliability of the accounts. For example, participants said;

> *"I think that the most important thing that you can do is actively listen"* (P4)
>
> *"you have to remember what they've told you [] you have to think ahead in relation to [] points to prove your offences you are trying to formulate the questions you want to ask them before they've finished [] you pick up on one thing and think right I need to remember to ask them that because [] that might prove something [] whilst you're listening to them at the same time"* (P2)
>
> *"you're trying to understand what they're telling you [] trying to remember the words that they've used to tell you, and the phrases, you are trying to put it into a structure and break it up into how you are going to cover it, you are thinking is there anything there that doesn't fit with what we've expecting, is there [] something that [] isn't quite right, have they brought up areas that I need to talk to them about that I can [] cross off my [] list, so all of this is going on and then what questions am I going to ask and which bit am I going to cover first"* (P1)
>
> *"you're listening to them and taking everything in but then in the back of your mind you're trying to remember everything and what I need to come back to"* (P8)
>
> *"I think about the question and I think about how to put that question to them without influencing or steering them down a particular path"* (P6)

Sub-ordinate theme 1B: The investigative context of the interview.

Participants described the additional demands they experienced when interviewing vulnerable witnesses in a variety of investigative contexts in comparison to other types of interview. For example, the timeliness of the interview and the evidential importance of the interview were described. Participants also identified that information may be required for operational requirements, or that specific details may be required from the interviewee. For example, participants said,

> *"you're trying to remember [] making sure you're sticking to best practice, structure, content and question type, topic selection, using the witnesses words [] not paraphrase [] balance what the SIO [Senior Investigating Officer] wants from the enquiry, the aim for your interview [] remember how to do that [] and keep to your structure"* (P1)
>
> *"the pressure to get the information straight away, the pressure to identify this person [] the level of offence you are dealing with [] there is a lot of scrutiny [] a lot of people are going to listen to it afterwards [] it is a very pressured role"* (P7)
>
> *"just to get in the place because it is a secure mental health unit, just to manage all the people [] a social worker [] and solicitor"* (P1)
>
> *"you sometimes (ask) [] can this not wait until tomorrow, it'll be better for the victim or witness, it'd be better for me, I'd be more mentally ready for it, but the pressure of having to get it in there, I don't think they get as good a product"* (P7)

Super-ordinate theme 2: Specific factors that influence cognitive demands.

Participants explained that their experience of investigating a diverse range of cases, which all have different characteristics (e.g., the complexity or emotional content of the case), varies with each interview they conduct. Their comments reflected that the individual differences of interviewees, as well as their own abilities, affected each

participant's experience of investigative interviews with vulnerable witnesses. All interviewees and interviewers are different, which makes each interaction unique. For example, participants said,

> *"you've got to go with that person and every one is different every single interview you do is different you know" (P4)*
>
> *"I think some interviews, it's easier depending on the circumstances and the nature of the offence and the relationship between the victim and the offender and the scene" (P7)*

Sub-ordinate theme 2A: The interviewer's perceptions of their own ability and individual pressures.

Participants described their experience of interviewing in terms of how they perceive their own cognitive ability, for example, whether they thought they had a good or poor memory. Participants also described their different interviewing styles and emotional control, which were identified as having an impact on their experience of interviewing vulnerable witnesses. The personal pressure that interviewers experienced from insufficient planning and preparation time, being under scrutiny, and their personal circumstances, were also identified as factors that had an impact on how they conduct interviews.

> *"I've got a terrible shocking memory [] so I do worry about remembering the exact phrases, I'm a bit pedantic about it [] I do have to take notes" (P1)*
>
> *"I had the ability to remember what they were saying, so I found that quite easy" (P2)*
>
> *"I felt a great deal of apprehension [] and pressure [] but over the years [] and with more experience [] I find I don't have that level of apprehension that I used to have" (P6)*

"I do have a note pad but when I've come out I've usually written the time and name, I don't really take a lot of notes because I found I'm not listening properly if I do" (P4)

"that interview had a lot of elements in it [] identifying with your own child and also the time constraints" (P5)

Sub-ordinate theme 2B: The interviewee's needs and their contribution.

The specific needs of interviewees, their openness, willingness, and ability to engage, were discussed by participants as having an impact on interviews. Building rapport with interviewees was also identified as an important factor.

"sometimes you'll ask somebody, tell me everything and they'll give you a really lovely a to z that is perfect [] that is a completely different task to the person who then says [] a million and one things and doesn't actually cover the offence" (P1)

"I did have to interrupt her, she just kept going on and on and on [] I was very conscious of what I needed to get out of the interview [] she probably wanted to off load again about what happened [] it was pressurised" (P5)

"I think the easier ones are where they [] just talk to you [] that makes it easier [] you've got a difficulty when you are trying to draw teeth, that's difficult, when they are just giving you one word and you [] are trying desperately" (P2)

Super-ordinate theme 3: Consequences and impact of cognitive demands

Participants explained that conducting interviews with vulnerable witnesses can have consequences for themselves and the way they conduct their interviews. They described the effects they experience and the outcomes on their performance when conducting their interviews. Participants also described some procedures and strategies they use to manage the cognitive demands during the interviews they conduct.

***Sub-ordinate theme 3A: Performance in interviews*.**

As well as describing the multiple mental demands of interviews, participants also explained how these factors impact on their performance. When conducting interviews with vulnerable witnesses, participants reported straying from best practice guidelines, making errors in their interviews such as, forgetting information and misunderstanding the interviewee.

"I'm still a little bit old school and I may drift off from the new model [] because I've been doing it the other way for too long really [] I'm concentrating more on [the] new model rather than what I'm there to achieve" (P4)

"Very often [the interviewer] who has been put under pressure, lack of time, lack of planning and preparation time, they're rushing, they've not thought it through, and they don't know the job well enough to ask the right questions and they produce a poor interview as a result" (P1)

"if there's several witnesses and victims and [] several accounts off various other people [] you are under pressure and [] thinking has she just told me that or is that what I already know" (P7)

"it can be harder when there are several victims who don't know that person you've got several descriptions and that might leak in" (P7)

"I said it once in an interview, the nickname of [the offender] and I thought she's never told me he was called (name redacted), I know that from (another interviewee, name redacted), so I think it's hard" (P7)

"I just learnt to sort of listen to key bits better [] it doesn't always work; someone can tell me something and because I've not written it down, I'll forget" (P8)

"I find that because I'm thinking about the actual phrases and the actual structure, what am I going to do with what they are telling me, how am I going to break this

up and divide it up what areas do I need to think about I actually miss understanding what they're telling me" (P1)

Sub-ordinate theme 3B: Effect of cognitive demands on the interviewer.

Participants described the impact on themselves of conducting interviews with vulnerable witnesses who often describe harrowing events. Participants explained they were often mentally, physically and emotionally drained after conducting interviews.

"I think that type of interviewing is exhausting – you are both mentally and physically exhausted [] so the impact of it is huge [] the mental strain it puts on you to stay alert [] the stress it has on you, the strain is huge" (P6)

"I just think it's the mental gymnastics that you [] have to do, 'cause you are keeping a track of where you are, you're looking at the action detail, you're thinking have I covered every piece of action in that, hard work of the brain like that is very tiring" (P3)

"I would say they take an absolute chunk out of you [] it's a funny kind of tiredness [] your brain is tired, your emotions are quite often tired [] there's just certain things that you just never forget [] it's really emotional, so you [] are knackered [] that's the impact of dealing with some of these (cases)" (P1)

"I think you are mentally exhausted, sometimes it will play on my mind you know things we've spoken about, I've even had dreams of the offenders" (P7)

Sub-ordinate theme 3C: Strategies to prepare for and conduct interviews.

Participants highlighted the importance of using the recommended protocol during their interviews. They identified different strategies they use when conducting interviews with vulnerable witnesses to manage the effects of the cognitive demands. These included the importance of planning and preparation, note taking (for some, but not others), managing errors and mentally structuring interviews.

"there's no such thing as a perfect interview, you always slip up with the odd clumsy question at times, but it is about recovery and it's about not constantly asking poor questions" and "I'm a stickler for planning and prep" (P3)

"I follow a set routine through interviews [] my only way really of dealing with a lot of (information) is note taking [] I tend to [] write down their free recall [] as much verbatim as I can [] their phrases in topic boxes [] that's my notes to help me through the rest of the interview" (P1)

"it helps to set parameters, it is good because you can keep on track and keep focused, focus your mind as where you're going so you are not jumping (back and forth) [] you should have some kind of structure [] so it's not all over the place" (P4)

"(I) relied very much on the loggist [] it suited me that I didn't write any notes and I still don't like writing notes in the interview" (P2)

Discussion

An IPA approach was used to examine factors, described by the participants who were trained investigative interviewers, which likely contribute to an increased cognitive load when they conduct investigative interviews with vulnerable witnesses and victims. This approach allowed identification of a number of overarching cognitive processes during the course of investigative interviews, including actively listening, paying attention, remembering information and making judgements. Participants described the cognitive demands of interviewing in the context of an investigation (e.g., the evidential importance and timeliness of interviewing). Specific factors associated with interviewing vulnerable witnesses that may increase cognitive load were also identified. These included the diverse range of characteristics of individual investigations, such as the complexity or emotional content of the case and the management of uncooperative interviewees. The participants' perceptions of their own

ability and the interviewees' contribution during interviews were also identified as factors that influence the cognitive demands of an interview. Participants described the consequences and effects of the cognitive demands they experienced, which may impact upon their performance. For example, they perceived that forgetting information, or misunderstanding the interviewee, influenced their performance. In addition, participants described the impact that conducting interviews with vulnerable witnesses can have on them, for example, they experienced being physically and mentally exhausted after interviews. When describing their experiences, participants also identified certain strategies they use to manage the cognitive demands of investigative interviewing.

Factors identified as contributing to cognitive load

Participants experienced features of cognitive load when they conducted interviews with vulnerable witnesses. They described making conscious and controlled decisions during the interview, i.e., being conscious of how they are asking questions, judging the quality of information, and deciding which elements to follow up with further questions. These complex intrinsic (inherent) features of interviewing require controlled processing. Completing the task, therefore, requires additional cognitive resources, is effortful, and is more cognitively demanding than for more automatic processing (Kahneman, 2012; Kleider-Offutt et al., 2016).

Participants also explained that other intrinsic factors, such as differences in opportunity for rapport building with the interviewee and the degree of case sensitivity, varied across interviews. During an interaction, making judgements of emotions and decoding social cues from faces and verbal stimuli requires substantial working memory resources and can be cognitively demanding (Phillips et al., 2008). When the number of elements needed in working memory to complete a task exceeds capacity,

cognitive load can result (Paas et al., 2003). Participants in the current study explained that they were required to make judgements about interviewees, decode social cues, and judge the content of their accounts. Although the perceptual processes involved in the discrimination of emotions are relatively automatic and make little demand on working memory, the decision-making processes do place demands on working memory. Detecting and responding to emotions are critical factors that influence cognitive load during a social task (Phillips et al., 2008).

Organising and processing novel information (e.g., during an investigative interview) becomes increasingly complex as the amount of unique information increases (Sweller, 1994). As supported by the current participants, each investigative interview is unique and interactive and can be lengthy, but the onus is always on the interviewer to maximise the quality and quantity of information provided by the interviewee (Powell et al., 2005). Therefore, in lengthy or complex interviews, even experienced interviewers will be placed under increasing cognitive load. Participants in the current study also described extraneous (additional) features of interviewing, such as being under scrutiny, following best-practice guidance and time pressure, which they experienced as contributing to the cognitive demands of the interview. If cognitive capacity is expended on these extraneous features, less capacity is available to complete the intrinsic elements (Leppink & van den Heuvel, 2015). Therefore, as intrinsic and extraneous factors are additive it is likely that the current participants' experience of extraneous factors contribute further to their cognitive load.

Consequences and effects of cognitive load

Participants in the current study explained that conducting interviews with vulnerable witnesses can be mentally, physically and emotionally exhausting. An outcome of cognitive load is cognitive fatigue, which is described as a decline in

cognitive resources due to sustained cognitive demands (Borragan et al., 2017). The exhausting nature of interviewing described by participants is indicative of signs of cognitive fatigue resulting from cognitive load. This is problematic for interviewers as the feelings of exhaustion may hamper cognitive performance (Borragán et al., 2017). High cognitive load can have an effect on impression formation and judgement of interviewees (Nordstrom et al., 1996). Therefore, when under cognitive load, and cognitive fatigue, interviewers may misinterpret their perceptions through reliance on more automatic processes, which can introduce bias into judgements (Kahneman, 2012).

Cognitive load can impact on the amount of information recalled by the interviewer, which may also impact on their performance (Galy et al., 2012). The complex cognitive tasks required to complete an interview can result in less information being obtained from an interviewee and difficulties for the interviewer when attempting to integrate information (Nordstrom et al., 1996). Information provided by a witness often needs to be confirmed or clarified by asking closed or specific questions (MoJ, 2011). Participants in the current study understood that care should be taken when asking these types of questions, so as not to introduce information not previously mentioned by an interviewee. However, they described that the use of inappropriate questions does happen during their interviews. When under high cognitive load, the interviewers' ability to form information-gathering questions may be limited and they may rely on pre-interview information which may result in leading questions (Nordstrom et al., 1996).

Participants described that, with experience, they were able to manage some challenges of investigative interviewing. Experience is important for learning and developing task-relevant knowledge and skills (Anderson, 1995). Expertise develops as

learners combine simple ideas with more complex ones (van Merrienboer & Sweller, 2010). The extent to which interviewers have developed schemas (frameworks to organise information, i.e., blocks of knowledge) for processing task-related information will also determine how cognitively demanding they find interviewing (Frieder et al., 2016). However, if an interview involves increased intrinsic and extraneous cognitive load, then cognitive load may also result in reduced germane load capacity. That is, there may be a reduction in the ability to apply general skills and knowledge to novel situations (Galy et al., 2012). The interviewer may have insufficient working memory capacity to enable the formation of schemas for the specific task.

It may be that interviewers' working memory capacity and their ability to manage intrinsic and extraneous cognitive load during complex interviews has an impact on their performance. Organising information, or knowledge, that needs to be processed in working memory only develops for those aspects of performance that are consistent across task situations, such as routines for operating medical equipment (van Merrienboer & Sweller, 2010). Automatic processing is then developed as a function of practice. However, participants in the current study described that the interviews they conduct are all different. Therefore, it may be that they are able to develop some skills that help reduce or manage load over time and with interview experiences, but they are unable to develop sufficient schemas for diverse interviews. With further practice, and to reduce intrinsic cognitive load, interviewers may build schemas that incorporate elements of the task, for example, how to ask open questions or memorising many different open question stems. Processing the schema then becomes less cognitively demanding (van Merrinboer & Sweller, 2010). However, it should be noted that controlled processing is also important during an interview. Over-reliance on heuristics

or automatic processing of information (i.e., a 'one size fits all' strategy) could be very detrimental to interview quality.

In terms of strategies to handle the information provided by witnesses during their interviews, several participants mentioned note taking. Note taking is a complex activity that involves interweaving both comprehension and production processes and can be considered in terms of the major cognitive effort required in specific contexts or situations (Piolat et al., 2005). Although note taking is cognitively demanding, it can improve the encoding of information and improve recall of information (Meise & Leue, 2019). However, note taking may also interfere with the different tasks required of the interviewer (i.e., listening, memorising and maintaining the information that is given by the interviewee; Horowitz & FosterLee, 2001). Multiple cognitive processes must be coordinated in rapid succession for note taking to be successful. Note takers must simultaneously comprehend, evaluate, sort, and write down information (Piolat et al., 2005). Thus, the additional cognitive processes involved in note taking could impose significant demands on the limited resources of working memory, which may account for the descriptions of extra load experienced by some participants when they take notes.

Limitations

The IPA approach is a suitable method when seeking perceptions and understanding of situations that are complex, poorly understood or previously unexplored (McCormack & Joseph, 2018). However, conducting IPA research has limitations in terms of sample size and generalisability of findings. For the current research, the sample was a suitable size for IPA research and, as recommended, the sample was homogenous (i.e., participants were from one UK police force). Using this sampling method enabled the experiences of the participants to be recorded along with

aspects of their experience that were shared with other participants within the force. Nevertheless, it is acknowledged that other interviewers and those from different jurisdictions may have alternative perceptions and experiences. Therefore, despite being consistent with prior qualitative research and the IPA approach, the generalisability of these results is likely limited.

Future research

Using an IPA approach in this study has opened up new lines of research that are informed by practitioners' experiences. For example, as working memory capacity is linked with cognitive load, it would be interesting to explore whether working memory capacity has an impact on perceived cognitive load and the performance of interviewers. Future studies might also usefully examine, under experimental conditions, the effect of cognitive load on interviewers' recall of information and whether note taking would help interviewers manage their cognitive load.

Conclusion

To date, the cognitive load experienced by investigative interviewers has received little empirical attention. Based on the experiences described by investigative interviewers, the current research identified factors that contribute to the experience of cognitive load. The intrinsic factors of interviewing, such as making conscious controlled decisions and processing novel information, can be cognitively demanding. Extraneous factors, such as time pressure, may further increase cognitive load. Despite interviewing being often complex and demanding, cognitive load could be managed, for example, through the reduction of extraneous load and interviewers undergoing further training to increase skills, so aspects of interviewing become more automatic.

Chapter 3:

The effects of cognitive load during an investigative interviewing task on mock interviewers' recall of information.

This chapter has been published as:

Hanway, P., Akehurst, L., Vernham, Z., & Hope, L. (2020). The effects of cognitive load during an investigative interviewing task on mock interviewers' recall of information. *Legal and Criminological Psychology*.

Abstract

Although investigative interviewers receive training in interviewing techniques, they often fail to comply with recommended practices. Interviewers are required to actively listen, accurately remember information, think of questions to ask, make judgements, and seek clarification, while conducting interviews with witnesses, victims or suspects. The current study examined the impact of increased cognitive load on mock interviewers' recall of a witness's account. Participants took the role of an investigative interviewer in one of three conditions, high cognitive load (HCL), moderate cognitive load (MCL), or no cognitive load (NCL). Participants watched a video recorded free narrative of a child witness during which they followed condition-relevant task instructions. Each participant rated their perceived cognitive load during their task and then recalled (free- and cued-recall) the content of the witness's account.

Participants in the HCL and MCL conditions perceived higher cognitive load and demonstrated poorer performance on the free-recall task than those in the NCL condition. Participants in the HCL condition demonstrated poorer performance on the cued-recall task compared to participants in the NCL condition. The cognitive demands required to complete an investigative interview task led to an increased perceived cognitive load and had a negative impact on recall performance for mock interviewers. Accurately recalling what has been reported by a witness is vital during an investigation. Inaccurate recall can impact on interviewers' questioning and their compliance with recommended interviewing practices. Developing and practising interview techniques may help interviewers to better cope with the high cognitive demands of investigative interviewing.

Introduction

Despite having knowledge of, and receiving training in, recommended interviewing techniques, interviewers in real-world settings do not always follow best practice guidelines (Criminal Justice Joint Inspectorate [CJJI], 2014; Powell & Barnett, 2015; Schreiber-Compo et al., 2012). Such guidelines generally recommend a range of interviewing techniques that have been developed from decades of international research for use in criminal investigations, child protection enquiries and intelligence-gathering settings. For example, the PEACE model (an acronym for Planning and preparation; Engage and explain; Account; Closure; Evaluation) is recommended for interviewing suspects and witnesses (Bull & Soukara, 2009; Kassin et al., 2010; Milne & Bull, 1999). Similarly, the Achieving Best Evidence guidelines (ABE; Ministry of Justice [MoJ], 2011) and the National Institute for Child Health and Human Development protocol (NICHD; Lamb et al., 2018, Orbach et al., 2000) have been developed, and are recommended for interviewing vulnerable witnesses[3].

There are, therefore, an abundance of guidelines, which provide advice to practitioners for the optimal approach to obtaining precise and complete statements from interviewees (Bull, 2010; Hershkowitz, 2011; Oxburgh et al., 2015). However, adhering to these guidelines remains a challenge for investigative interviewers (Lamb, 2016; Schreiber-Compo et al., 2012). This may be because interviewing is a complex cognitive task for the interviewer (Lafontaine & Cyr, 2016; Powell, 2002). In an exploratory study, the cognitive load experienced by interviewers was identified as a possible barrier to compliance with recommended techniques (Hanway & Akehurst,

[3] Vulnerable witnesses for the purpose of this chapter include witnesses or victims who may be vulnerable due to their age, intellectual or communication difficulties, or intimidated witnesses (MoJ, 2011).

2018). Contrary to recommendations, interviewers' cognitive burden may result in them interrupting the witness or asking questions that have already been answered (Schreiber-Compo et al., 2012). However, as noted by Kleider-Offutt et al. (2016), the impact of multiple cognitive demands for investigative interviewers has not been empirically examined. The current study explored the cognitive demands of a witness interview observation task and tested the effects of cognitive load on the recall of a witness's account.

Cognitive Load and Task Performance

Cognitive load is the mental workload placed on individuals when they are required to undertake activities (Hart & Staveland, 1988; Van Acker et al., 2018). It signifies working memory use and the demands placed on cognitive resources when carrying out multiple and competing tasks (Dias et al., 2018; Engström et al., 2013). The capacity limitations of working memory mean that without the rehearsal of received sensory information, the processing of information is restricted (van Merrienboer & Sweller, 2010). This can lead to an attentional bottleneck where attending to one element of information causes other cognitive processes, and the associated information, to be neglected (Strayer & Drews, 2007).

Controlled processing is needed to complete cognitive tasks that require attention and the management of information (Bargh, 1984). However, this type of processing is slow and effortful and relies on our limited attention capacity (Strayer & Drews, 2007). High levels of focused attention can be accomplished with effort (Bargh, 1984; Schneider & Shiffrin, 1977), but errors occur if an individual cannot meet the mental demands required to effectively complete the tasks (Paas & van Merrienboer, 1993). Additionally, during complex tasks, there is an increase in cognitive demand; thus, the amount of mental effort required also increases (Kleider- Offutt et al., 2016).

The attentional demands required to perform complex tasks may lead to cognitive load and errors, or a reduction in performance (Engle & Kane, 2004; O'Donnell & Eggemeier, 1986).

Cognitive Load Theory (CLT) identifies three types of load (Sweller, 1988, 1994; Sweller et al., 1998) that are relevant in a variety of applied settings (Galy et al., 2018). The first type, intrinsic load, relates to the load imposed by the fundamental nature of the information being processed and the natural complexity of the task (Schnotz & Kurschner, 2007). The second, extraneous load, is induced by other external factors, such as time pressure (Galy et al., 2012). The third type of load described within CLT is germane load, which is the load used for learning, the development of skills, and the application of skills in a novel situation (Paas, et al., 2004). Notably, germane load is required for the construction and automation of schemas for a particular task (Galy, et al., 2018).

Cognitive Load in Investigative Interviews

For investigative interviewers, there are several inherent (i.e., intrinsic) features of interviewing that may contribute to a cognitive load, including the generation of questions, identifying topics to pursue, and seeking clarification from interviewees. Interviewers are required to actively listen to, and accurately remember, what interviewees are saying (Fisher et al., 2014). They may also be required to take notes and formulate hypotheses to account for the events described. As such, interviewers must attend to multiple cognitive processes (Kleider-Offutt et al., 2016). At the same time, they are required to adhere to best practice guidance, such as, building rapport and forming appropriate questions (Hanway & Akehurst, 2018).

Open questions typically lead to detailed, free narrative responses from interviewees (Dale et al., 1978; Hershkowitz, 2001). Hence, asking open questions is an

important feature of an investigative interview (Danby et al., 2017). Interviewers must then accurately remember the often-numerous details provided by interviewees, but interviewers' recall of information may be limited and inaccurate (Hyman-Gregory, 2009). The interviewer may introduce this erroneous information to the witness, which may have an impact on the subsequent accuracy and reliability of the witness's testimony (Loftus & Pickrell, 1995; Gudjonsson, 2010). In doing this, interviewers can affect the amount and quality of evidence provided by witnesses (Brown & Lamb, 2015; Gudjonsson, 2010).

In sum, obtaining accurate and detailed accounts from witnesses during investigative interviews can be difficult (Hope & Gabbert, 2019; La Rooy & Dando, 2010). Interviewers hold information provided by witnesses in their memory, whilst at the same time assessing that information, thinking of questions to ask, and identifying the correct order in which to ask those questions (i.e., which topic to ask questions about first; Hanway & Akehurst, 2018). The complex cognitive functions required to complete these tasks are likely to have an impact on interviewers' performance and their judgements (Ask & Landstrom, 2010; Nordstrom et al., 1996).

The Current Research

The current research examined the effect of increased cognitive demands on participants' perceived cognitive load during a witness interview observation task. The tasks for each condition were designed to replicate the cognitive demands present during an investigative interview (i.e., to listen to the witness, remember information, judge information and think of questions to ask; Fisher et al., 2014; Hanway & Akehurst, 2018). The effect of increased cognitive demands on the amount and accuracy of information recalled from a witness's statement by participants who took on the role of interviewers was explored.

Based on previous cognitive load research (e.g., Dias et al., 2018; Nordstrom et al., 1996), it was hypothesised that during the interview and recall tasks participants in a high cognitive load (HCL) condition would report higher perceived cognitive load (PCL) compared to those in a moderate cognitive load (MCL) condition, who would report higher PCL than those in a no cognitive load (NCL) condition. Second, it was hypothesised that participants in the HCL condition would recall fewer details and would have a lower accuracy rate for their free recall of a witness's statement, than those in the MCL condition, who would recall fewer details and have a lower accuracy rate than those in the NCL condition. Third, it was predicted that participants in the HCL condition would have a lower percentage accuracy score when answering questions about a witness's statement than those in the MCL condition, who would have lower percentage accuracy score when answering questions about a witness's statement than those in the NCL condition.

Method

Design

For this independent-groups study, there was one between-subjects factor, cognitive load, with three levels: high cognitive load (HCL); moderate cognitive load (MCL); and no cognitive load (NCL; control). The dependent variables were perceived cognitive load (PCL), the amount and accuracy of statement details provided by participants during free recall, and the accuracy of their cued recall.

Participants

A priori G*power analysis (Faul et al., 2009) for an omnibus one-way ANOVA with three groups, indicated that a sample size of 102 participants was required. This was based on power = 0.95, a large effect size of $f = 0.40$, and the traditional alpha = .05. A large effect on recall accuracy was predicted on the basis of research showing large

effects of working memory capacity on memory accuracy (e.g., Jarrold et al., 2011) and large effects of cognitive load on recall accuracy for the spoken word (e.g., Hunter & Pisoni, 2018).

102 participants, staff and students, were recruited via a university participant pool and workplace advertisements at the university. Participants were invited to take part in a study that examined what it is like to be an investigative interviewer. No monetary incentives were offered to participants, but first year undergraduate psychology students were offered course credit for their participation. Participants attended for one test session, which lasted approximately 45 minutes. Only adults with English as a first or primary language were recruited. The aim of the study was to assess participants' recall of information provided by a witness, when under varying degrees of cognitive load. Therefore, as experience can have an impact on task performance when under cognitive load (Paas, et al., 2004), prior investigative interviewing experience was an exclusion criterion.

The sample comprised 68 females and 34 males. Participants were aged 18 to 71 years (M_{age} = 25.95 years, SD = 10.02, the median age was 22 years). To ensure equal numbers of participants (N = 34) in each condition, they were pseudo-randomly allocated to one of the three conditions (HCL, MCL, NCL). Data from one participant was removed from the analysis as their responses suggested a poor understanding of the task and a z-score for accuracy rate of the witness's account was an outlier at -3.41 (Field, 2013). Data from two further participants were removed due to recording equipment failure. The final sample, therefore, comprised 99 participants who were aged 18 to 71 years (M_{age} = 26.03 years, SD = 10.09, median age = 22 years) [4]. There

[4] Two participants in the study were aged 71 years old. All other participants were aged 18 to 54 years. There was no significant difference in age between the three conditions, F (2, 96) = 1.25, p = .293, η^2_p = .03. Analyses were conducted with and without these two participants' data, which revealed no

were 67 females and 32 males. For the final analyses there were 34 participants in the High Cognitive Load (HCL) condition, 33 in the Moderate Cognitive Load (MCL) condition and 32 in the No Cognitive Load (NCL) condition.

Materials

Stimulus event

To enable an accurate reflection of a real-world interview, the interview room setting, interview procedure, and recording of the interview, were designed to correspond with published guidance for interviewing child witnesses (MoJ, 2011). An eight-year-old child witness was interviewed about an event she had experienced (a recent birthday party). The witness was given an open prompt by the interviewer (i.e., "*Please tell me everything you can remember about the party you went to*"). This question and the witness's subsequent free recall were digitally recorded. The recording of the interview captured a head and shoulders view of the witness. The child's recorded free recall account lasted for 6 minutes and 30 seconds.

Perceived Cognitive Load measure

To measure participants' perceived cognitive load, the National Aeronautics and Space Administration, Task Load Index (NASA-TLX) was used. This questionnaire combines information about the magnitude and source of six related factors to derive a sensitive and reliable estimate of workload (Hart & Staveland, 1988). The NASA-TLX uses a multi-dimensional rating scale questionnaire to evaluate participants' subjective ratings of mental workload; the scale items are mental demand, physical demand, temporal demand, performance, effort and frustration. These items were selected following analysis of the primary factors that do (and do not) define a subjective

differences in the results. For completeness, the data of all participants, including those aged 71 years old, were included in the analyses and reported results.

experience of workload (Hart, 2006). Each item is measured on a 20-point scale from low to high (except for performance which is measured on a scale from good to poor). A weighted score is obtained by completing 15 pairwise comparisons of the six scale items. For each pair, one item is selected that is more relevant for the participant when completing the task (Hart & Staveland, 1988; see Appendix B.1 for NASA-TLX App examples). For this study and following the scoring procedure devised by Hart and Staveland (1988), a PCL score out of 100 was calculated by multiplying each scale item score (rating score) by the number of times that item was selected in the pairwise comparisons (adjusted score); the six weighted item scores were then totalled and divided by 15 to obtain an overall PCL score. The NASA-TLX was designed to be used during, or immediately after, a task and has been widely used in a variety of settings to measure the cognitive load perceived by participants when they complete a task (e.g., Hart, 2006; Rizzo et al., 2016).

Procedure

After reading the information sheet and providing written informed consent, participants were allocated to one of the three conditions: HCL, MCL, or NCL. The lead author conducted the research and followed written instructions for all conditions. The experimenter was aware of each participant's condition. To reduce experimenter effects, instructions for each condition were read out verbatim from a written script and all questions were asked verbatim from a prepared script. All participants were instructed to take the role of a police interviewer and were informed that a child had witnessed an event, which the participant needed to investigate. Participants were asked to watch and listen to the witness's recorded interview and were informed that they would be asked some questions after they had watched the interview. In the HCL condition participants were given the following additional instructions, *"Whilst*

watching the interview, I would like you to consider carefully what the witness is telling you so that you clearly understand the witness's experience of the event she is describing. Your other task is to identify follow-up questions to ask the witness once she has given her statement. So, whilst you are listening to the child, please think about the wording of your questions and in what order the questions should be asked". In the MCL condition, participants were given the following additional instructions; *"Whilst watching the interview, I would like you to consider carefully what the witness is telling you, so that you clearly understand the witness's experience of the event she is describing".* In the NCL (control) condition, no further instructions were given to participants.

After receiving their specific instructions, all participants watched the recorded interview on a computer screen wearing headphones to reduce distractions. Immediately after watching the interview with the child witness, all participants completed the first PCL measure (i.e., they recorded their perceived cognitive load during the interview task, using the NASA-TLX scale presented via an android tablet application). Participants then carried out a 15-minute distraction task, which required them to work through some unrelated number puzzles.

Following the distraction task, participants were asked to recall as much information, in as much detail as they could, from the witness's recorded statement. After participants finished their free recall, they were asked if there was anything further that they could recall about the interview. Once participants had completed the free recall task, they were asked 40 cued recall questions about the content of the witness's interview (e.g., "*What did the witness say was 'quite tricky'?*"; "*Who drove the witness home?*"; see Appendix B.2 for 20 cued recall questions). The order of these questions was randomised across participants. All participants were audio recorded whilst they gave their free narrative and answered the cued recall questions.

Participants then completed a second self-report of their PCL for the recall task (i.e., their perceived cognitive load when they were recalling the child's statement and answering the 40 questions). This was again completed using the NASA-TLX scales.

For completeness, as participants in the HCL condition had been asked to think about questions to ask the witness, they were then asked to write down 10 follow-up questions they would ask the witness if they were the investigator in the case. To ensure all participants completed the same tasks, those in the MCL and NCL conditions were also asked to write down 10 questions they would like to ask the witness[5].

Finally, participants were asked to rate, using 7-point scales their confidence in their memory accuracy, from [1] not at all confident to [7] extremely confident; the extent to which they felt motivated to remember the content of the child's interview, from [1] not at all motivated to [7] extremely motivated; the extent to which they found remembering the child's statement easy or difficult, from [1] very easy to [7] very difficult; and the extent to which they found coming up with questions easy or difficult, from [1] very easy to [7] very difficult. Participants in the HCL condition were also asked to rate how motivated they were to think about questions whilst they were listening to the child's statement, from [1] not at all motivated to [7] extremely motivated.

As a manipulation check, participants were then asked to write down the instructions they were given by the researcher before they watched the child's account. Demographic details including age and gender were also recorded. A verbal debrief was provided for all participants and they were thanked for their time and effort.

[5] Mean time (in seconds) for writing down 10 follow-up questions; HCL, $M = 256.03$ ($SD = 72.64$); MCL, $M = 260.81$ ($SD = 72.55$); NCL, $M = 285.18$ ($SD = 65.39$). The differences were not significant $F(2, 96) = 1.61$, $p = .205$, $\eta^2_p = .03$).

Coding

Free recall coding

Verbatim transcripts of the participants' audio recorded free recall of the witness's statement were coded for quantity and accuracy of details reported. Details were coded as person, action, object, setting or temporal details. For example, participant accounts were coded as follows "Amelia (1-person) trotted (1-action) on her horse (1-object) in the stables (1-setting)". If the participant mentioned a detail relating to time (e.g., "at the end of the day") it was coded as a temporal detail. Consistent with previous research and to facilitate assessment of overall accuracy, details were coded as correct, incorrect or confabulations (Wright & Holliday, 2007). A detail was deemed (i) correct, if it was present in the witness's account and was correctly reported by the participant (e.g., "she was called Amelia"); (ii) incorrect, if a reported detail was discrepant from the witness's account (e.g., participant recalls "pull the reins back to go" but the witness actually said "pull the reins back to stop"); and (iii) confabulated, if a reported detail was mentioned in the participant's account which was not mentioned at all by the witness (e.g., the participant reported "they got into a car" but the witness did not mention a car at all during her account). Accuracy rate for the free recall accounts was calculated by dividing the total number of correct details reported by the total number of details reported (i.e., correct plus incorrect plus confabulations). Additionally, to assess indicators of uncertainty in participants' recall of the witness's account, ambiguities were coded (e.g., "*I'm not sure, it was something like…*").

Inter-coder reliability for the free recall accounts was assessed by selecting 20 interview transcripts (20%), which were coded by an independent scorer. Intra-class correlation coefficients (ICC) using absolute agreement were computed for the following measures: total details [r (19) .97, p <.001]; correct details [r (19) .95, p <

.001]; incorrect details [r (19) .83, p < .001]; confabulations [r (19) .90, p < .001]; ambiguities [r (19) .84, p < .001]. This analysis indicated that the inter-coder reliability was 'good' for the coding of incorrect details and ambiguities, and 'excellent' for the coding of total details, confabulations, and correct details (Koo & Li, 2016).

Cued recall coding

Answers to 40 cued recall questions were scored as; fully correct (e.g., in relation to the location of the event, "*Pink Mead Farm*": 2 points), partially correct (e.g., "*Mead stables*": 1 point), don't know response (0 points), and incorrect (e.g., "*Crofton stables*": -1 point). Total accuracy could therefore range from -40 (all questions answered incorrectly) to 80 (all answers fully correct). The scores were added, and a percentage accuracy score for each participant was calculated.

Results

Manipulation check

All 99 participants passed the manipulation check and accurately reported their instructions. As per their instructions, participants in the NCL condition confirmed they were required to watch the interview carefully and participants in the MCL condition confirmed they were to watch the interview and consider what the witness was saying. Participants in the HCL condition confirmed that they were asked to think of questions to ask the witness, as if they were the interviewer in the case, and to watch the interview carefully.

Hypothesis testing

To examine the current hypotheses, a series of between-groups ANOVAs were conducted.

Perceived Cognitive Load

For the 'encoding of interview' task that the participants were first asked to undertake, Levene's test indicated that the assumption of homogeneity of variance for PCL scores had been violated, F (2, 96) = 3.94, p = .023. Therefore, the more robust Welch equality of means test was examined. As predicted, there was a significant difference in PCL scores between the three conditions; F (2, 62.10) = 7.70, p = .001, with a large effect size, η^2_p = .20 (see Table 3.1). Tukey HSD post-hoc comparisons showed there was no significant difference between PCL scores for participants in the HCL and MCL conditions (p = .209). However, participants in the HCL and MCL conditions scored higher for PCL than those in the NCL condition (HCL, p < .001; MCL, p = .033). For the 'recall' task, there was no significant difference between the three conditions in terms of PCL scores, F (2, 96) = 1.21, p = .304, η^2_p = .02 (see Table 3.1).

Table 3. 1

Mean PCL scores during the 'encoding the interview' and 'recall' tasks for each condition.

	PCL for 'encoding the interview' task		PCL for 'recall' task	
Condition	*M* (*SD*)	95% CI	*M* (*SD*)	95% CI
HCL	53.08 (14.89)	[47.88, 58.27]	68.26 (12.71)	[63.35, 68.35]
MCL	46.21 (13.73)	[41.34, 51.08]	65.64 (12.11)	[61.34, 69.93]
NCL	35.77 (20.28)	[28.46, 43.08]	63.50 (12.66)	[58.94, 68.07]

Free recall

With respect to the total number of free recall details reported about the witness's statement, there were no significant differences between the three experimental conditions, F (2, 96) = 2.20, p = .117, η^2_p = .04 (see Table 3.2). In terms of accuracy rate of the details recalled, there was a difference between the three conditions with a large effect size, F (2, 96) = 8.54, p < .001, η^2_p = .15. Post-hoc comparisons of percentage accuracy indicated that there was no significant difference in percentage accuracy for participants in the HCL condition compared with those in the MCL condition (p = .476). However, percentage accuracy for participants in the HCL condition was lower than for those in the NCL condition, (p < .001). Accuracy was also lower for those in the MCL condition compared with those in the NCL condition, (p = .015), as shown in Table 3.2. For details of free-recall accuracy mean scores for correct details, incorrect details, confabulations, and ambiguity, see Appendix B.3)

Table 3. 2

Total number of details recalled and accuracy rate for each condition during the free recall task.

	Total details recalled		Accuracy rate	
Condition	*M* (*SD*)	95% CI	*M* (*SD*)	95% CI
HCL	116.5 (44.19)	[101.4, 131.6]	0.91 (0.39)	[0.90, 0.93]
MCL	137.3 (32.77)	[125.7, 148.9]	0.92 (0.42)	[0.91, 0.94]
NCL	134.0 (52.90)	[115.0, 153.1]	0.95 (0.37)	[0.94, 0.97]

Cued recall questions

For the accuracy of cued recall question responses, there was a difference between the three conditions for percentage accuracy score, with a large effect size, F

(2, 96) = 7.87, p = .001, η^2_p = .14. Tukey HSD post-hoc comparisons indicated that percentage accuracy score for participants in the HCL condition was not significantly different from those in the MCL condition (p = .114). The percentage accuracy score for participants in the MCL condition was also not significantly different from those in the NCL condition (p = .130). However, percentage accuracy score for participants in the HCL condition was significantly lower than for those in the NCL condition (p <.001; see Table 3.3). For details of cued-recall mean scores for correct, partially correct, incorrect, and don't know responses, see Appendix B. 3.

Table 3. 3

Percentage accuracy scores for the questions task across the three conditions.

Condition	*M*(*SD*)	95% CI
HCL	48.90(14.59)	[43.81, 53.99]
MCL	55.91(13.62)	[51.08, 60.74]
NCL	62.81(14.52)	[57.58, 68.05]

Motivation, confidence and task difficulty

A series of Pearson's correlations were calculated to determine whether the dependent variables of motivation, confidence and task difficulty were correlated with each other. There were significant, but moderate, correlations between the majority of variables (see Appendix B.3). Therefore, the assumption of an absence of multicollinearity was met, and to reduce Type 1 error, a one-way between-groups MANOVA was conducted to investigate differences between the conditions for participants' motivation, confidence, and how difficult they found the tasks. The MANOVA indicated that there was no significant multivariate effect: Wilks' λ = .95, F (8, 186) = .62, p = .764, η^2_p = .03 (for details of scores across each of the dependent

variables for each condition, see Appendix B.3). There were no significant differences at the univariate level.

Exploratory Analysis

As the confirmatory analysis showed that increased cognitive demand for participants in the HCL and MCL conditions was associated with increased perceived cognitive load during the 'encoding the interview' task and also a lower recall accuracy for the free recall and questions tasks, further exploratory analyses were conducted. A Pearson's correlation showed that there was a relationship between PCL and accuracy of free recall, r = -.279, p = .003. When the sample was split by condition, a linear regression analysis indicated that in the HCL condition, PCL was a predictor of participants' free recall accuracy rate (β = - .40, p = .018) accounting for 16% of the variance. However, PCL was not a predictor of free recall accuracy for participants in the MCL (β = -.08, p = .653), or NCL conditions (β <.001, p = .1.00), see Figure 3.1. PCL was also not a predicator of cued recall percentage accuracy scores across any of the conditions (HCL, β = -.042, p = .815; MCL, β =-.121, p = .502; NCL, β = -.047, p = .797).

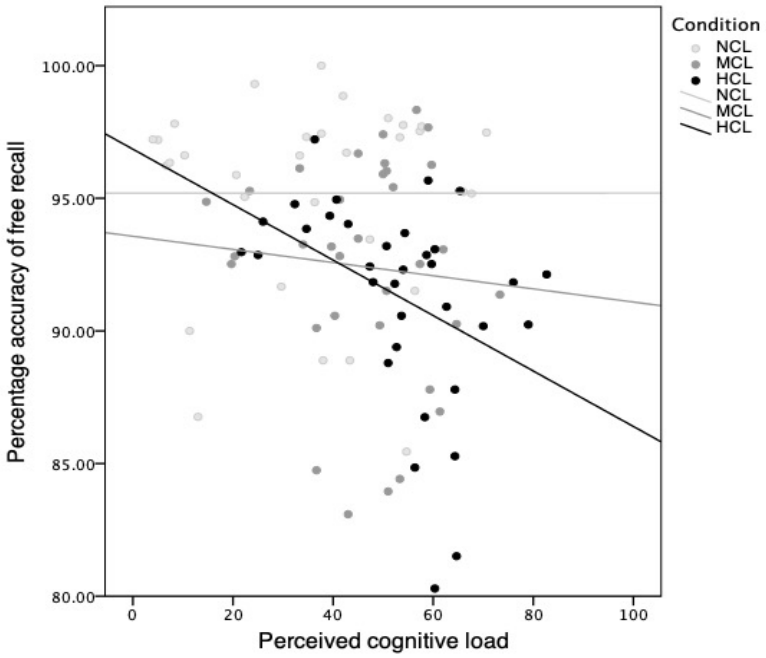


Figure 3. 1 Linear regression analysis with PCL as a predictor of free recall accuracy rates for each condition (HCL; MCL; NCL). Data points for the three groups are indicated with their associated line of best-fit plot.

Discussion

The effects of increased cognitive demands on perceived cognitive load and subsequent recall of an interviewee's account in a witness interview observation task were examined. As predicted, participants who were required to complete tasks that are intrinsic to investigative interviewing (i.e., listening, remembering, judging the information provided and generating follow-up questions to ask) perceived a higher cognitive load than did participants who were required to complete tasks with fewer cognitive demands (i.e., merely watching and listening to a witness's statement). Participants who were asked to complete more cognitively demanding tasks were less accurate, when freely recalling information provided by the witness, than those who were asked to perform less cognitively demanding tasks. Additionally, when asked cued questions about the witness's account, interviewees who completed more demanding cognitive tasks than those asked to perform fewer cognitively demanding tasks while

watching the interview, provided less accurate responses. Taken together, these results suggest that the demands placed on the participants' cognitive resources when carrying out the multiple tasks of an investigative interview resulted in a reduction in performance on the tasks.

In exploratory analyses, a relationship between PCL and recall accuracy rate was found. When participants' scores for the three conditions were examined separately, the relationship was moderated by the tasks undertaken by participants (i.e., for the HCL condition, higher levels of perceived cognitive load predicted performance in terms of free recall accuracy). When more controlled and focused attention was required for the task of generating questions to ask, there was an increase in perceived cognitive load and a reduction in performance. The reduction in recall performance may have been due to a limited capacity to carry out multiple cognitive tasks in working memory (Kahneman, 1973; Reisberg, 2007). However, more automatic processes (i.e., listening and watching the witness) were less affected by cognitive load (Schneider & Shiffrin, 1977). This research provides the first empirical evidence that increased cognitive demands inherent in an investigative interviewing task result in higher perceived cognitive load as well as reduced recall performance for participants.

For the current experimental task, which was designed to reflect real-world interviewing procedures, participants were asked to focus on certain intrinsic features of interviewing, including listening, remembering information and thinking of questions to ask. Whilst the current experimental design included a manipulation of cognitive load based on realistic processes for interviewers, it is recognised that investigative interviewing in the field is a complex task and likely requires more cognitive processing than was required for the participants. In practice, interviewers are required to build rapport, interact with the witness, and consider other aspects of the case (Schreiber-

Compo et al., 2012). Interviews, therefore, occur in a social context, whereby interviewers also perceive witnesses' actions and make judgements about their credibility, reliability and wellbeing (Ask & Landstrom, 2010; Hanway & Akehurst, 2018). These extraneous factors, and that of time pressure (i.e., temporal demand), were not present during the current study. However, cognitive load is additive (Leppink et al., 2015). Therefore, the additional factors identified as present when conducting investigative interviews, will likely contribute to a higher cognitive load for interviewers in practice (Hanway & Akehurst, 2018; Nordstrom et al., 1996).

Cognitive Load Theory suggests that automatic processing relies on schemas to reduce effort (Paas et al., 2004). With training, and skill development, more schemas are potentially built. However, if a task is cognitively demanding, and the intrinsic and extraneous load exceeds capacity, then there is little opportunity to form these schemas (Schnotz & Kurchner, 2007). Cognitive load, therefore, may also have an impact on interviewers' skill development. It may be that, despite their training and knowledge of best practice guidance, the intrinsic and extraneous cognitive demands imposed on investigative interviewers each time they conduct a unique interview leaves little capacity for building schemas. Consequently, interviewers are not afforded the opportunity to rely on more automatic processing and they experience significant cognitive load. Thus, interviewers do not always comply with their training (CJJI, 2014; Cross & Hershkowitz, 2017; Powell & Barnett, 2015).

For this study, the aim was to examine the effect of holding information in mind while judging that information and thinking of questions to ask a child witness. The aim was also to reduce extraneous load not directly related to the task. Note-taking can be cognitively demanding in itself and may divide attention between listening to the witness, formulating questions and recording information (Piolat et al., 2005; Schreiber-

Compo, et al., 2012). Therefore, in the HCL condition, participants were not permitted to note down the questions they were thinking about whilst they were listening to the child. An inevitable limitation of this design was that we could not be sure what participants were thinking during their task. To mitigate this limitation, and to ensure participants had understood their instructions, a manipulation check after the recall phase to check participants' understanding of what they had been asked to do was included. Future research might examine the effects of note taking for the interviewer.

While the design of this study replicated some of the cognitive demands experienced by interviewers during real-world interviews, a limitation is that the participants were novice interviewers, who had not received any training in investigative interviewing. As such, the current findings may have limited generalisability to trained or experienced interviewers. However, interviewers in the real world are also required to think about, and comply with, their training when undertaking interviews, which may increase their cognitive load (Hanway & Akehurst, 2018; Schreiber-Compo, 2012). Considering this, and the additional intrinsic and extraneous factors, it is possible that interviewers in the field will experience more cognitive load than the novice participants in the current study. In turn, interviewers' performance in the field may be impacted to a greater extent than was the case for participants in the current experiment. Further research should focus on aspects of investigative interviewing in context. It would be interesting to explore the impact that training and experience has on interviewers' cognitive load as well as the effects of cognitive load on other aspects of interviewer performance, such as, the types of questions asked. As some of the variation seen in the current study may be accounted for by individual differences in cognitive ability, this may also be an interesting area for further research, for example, individual differences in working memory capacity

(Engle, 2002). Finally, the sample size estimation may also be a limitation for this study. The sample size was based on a predicted large effect size, which has practical relevance in an applied setting. The approach was considered to be appropriate and in line with similar research in the investigative interviewing literature (e.g., Hoogesteyn et al., 2020; Kontogianni et al., 2018).

The current findings suggest that the cognitive demands required to complete an investigative interview can lead to an increased cognitive load and a reduction in recall accuracy of what was said by an interviewee, which may have an impact on interviewers' questioning and compliance with recommended interviewing practices. Providing interviewers with the opportunity to develop and practise their techniques, so that skills relating to interviewing become more automatic, along with better management of factors which may contribute to additional cognitive load, such as time pressure, may help interviewers to better cope with the high cognitive demands of investigative interviewing.

Chapter 4:

Who said what? The effects of cognitive load on source monitoring and memory for multiple witnesses' accounts.

This chapter is under review as:

Hanway, P., Akehurst, L., Vernham, Z., & Hope, L. (under review). "Who said what? The effects of cognitive load on source monitoring and memory for multiple witnesses' accounts. *Applied Cognitive Psychology.*

Abstract

This research examined the effects of cognitive load on memory and source monitoring accuracy for information provided by multiple witnesses. Using a witness interview observation paradigm, participants watched five witnesses' accounts of the same crime, under conditions of high cognitive load (HCL) where load was induced via interviewer-relevant tasks (e.g., formulating questions), or no cognitive load (NCL). Each witness provided unique details about what they saw. When asked about account details, and which witness had provided each detail, mock interviewers' memory accuracy was lower in the HCL condition than in the NCL condition. Source monitoring accuracy was poor regardless of cognitive load condition.

Introduction

Crimes are often complex events involving multiple witnesses, victims and suspects. Information is frequently elicited, by the same interviewer, from multiple people. The interviewer is then required to accurately remember the information that has been provided by the different people in order to ask appropriate questions, make decisions, and pursue further enquiries (College of Policing [CoP], 2019). Even when accounts are recorded, interviewers may not have timely recourse to the information they need when they need it. Thus, it is crucial that interviewers accurately remember who has provided specific details relating to a crime. Clearly, this is a cognitively demanding task, particularly when considered in conjunction with the other demands of interviewing, such as building a rapport with the witness, paying attention to the witness's needs, and seeking clarification of the information provided (Fisher et al., 2014; Hanway & Akehurst, 2018). In the current research, the effects of increasing cognitive demands, during a witness interview observation task, on participants' perceived cognitive load, the accuracy of their memory for detailed information provided by witnesses, and their accuracy for monitoring the source of information provided by multiple witnesses were examined.

Information provided by witnesses during their interviews forms the basis of an investigation and the accuracy and completeness of that information can determine the outcome of an investigation. When conducting investigative interviews there are several ways in which an interviewer's inaccurate recall of information provided by a witness could negatively affect an investigation. For example, an interviewer's inaccurate representation of a witness's account, when recapping evidence either during an interview or in subsequent legal discussions, can be damaging to both the investigation and the forward criminal justice process (Criminal Justice Joint

Inspectorate [CJJI], 2014). In particular, inaccurately recording, or omitting, information from an account can lead to a loss of potentially important investigative leads (Gregory et al., 2011).

To evaluate the accuracy and completeness of interviewers' memory of events described by witnesses, research has examined the nature and accuracy of details recorded when interviewers take notes during an interview. For example, Cauchi and Powell (2009) found that notes made during interviews with child witnesses were not always accurate; 15% of notes contained at least one or more errors of commission (i.e., the addition of incorrect information). Similarly, in a study examining 20 forensic interviews with children, Lamb and colleagues (2000) found that 25% of the forensically relevant details in the interviews were not present in the interviewers' notes. An evaluation of interviewers' reports after conducting cognitive or structured interviews with adult witnesses uncovered 81% and 78% accuracy respectively (Kohnken et al., 1994). Warren and Woodall (1999) found that interviewers' reports were 98% accurate but only accounted for, on average, 68% of the information provided by a witness. These studies indicate that information provided by witnesses is not always accurately or fully recorded in interviewers' notes. However, in such cases, it is not clear whether information was originally encoded by interviewers, but not remembered, or whether interviewers did not encode the information in the first place.

In addition to accurately recalling the information provided by a single witness, when conducting investigations with multiple witnesses, interviewers must also accurately monitor the source of information. This is important to enable differentiation between witnesses' accounts and, in doing so assess, among other things, what information is novel and what has been corroborated by others or by physical evidence. Monitoring the source of remembered information is important in many everyday

situations and enables judgements to be made about that information (Johnson et al., 1993). For example, when a person witnesses an event with at least two people involved they rely on source memory to recollect who performed what action (Kleider et al., 2008). Errors, or disruptions, in monitoring the source of information can have implications for memory, and knowledge, of an event (Johnson et al., 1993).

The Source Monitoring Framework (SMF) conceptualises how people distinguish memories from different sources (Lindsay, 2007) and is used to explore the mechanisms by which memories are attributed to particular events or origins (Johnson et al, 1993). Research on the SMF shows that accurately identifying the source of information depends on the quality and characteristics of the activated memory (Mitchell & Johnson, 2000). For example, in eyewitness identification studies it has been argued that, when making an identification, a witness must differentiate between their thoughts and feelings when encoding an event/crime, from other features, such as familiarity with one of the line-up members (Lindsay, 2007). The mental reinstatement of these representations of memory can range from general feelings of familiarity to memory for specific features (Evans & Wilding, 2012). The subjective awareness of the memory for an event is, therefore, central to understanding the SMF (Mitchell & Johnson, 2000). A subjective recollection of, or familiarity with, an event can distinguish between memories for specific events and more general memories of the event. That is, it is possible to know about something without remembering the event that provided the knowledge (Gardiner et al., 2002). 'Remembering', then, refers to a 'conscious recollection' of an event, whereas, 'knowing' occurs with a 'sense of familiarity' but without conscious recollection of the event (Tulving, 1985).

Research examining the SMF in applied contexts has generally focused on the source monitoring (SM) accuracy of witnesses' recall (e.g., Lindsay, 2014); and

memories of repeated or multiple events as recalled by children (Poole et al., 2014) and adults (Willen et al., 2015). SM in these areas is important as recognition of the source of information can lead to successful recognition of other details about an event (Reisburg, 2007). SM is also important when interviewing multiple witnesses, for example, if a witness provides information during an interview, and the interviewer recalls that a previous witness has provided the same or similar information, the interviewer then also needs to consider the accuracy and source of the previous information. Memory errors may occur when interviewers consider information from multiple sources. For example, the source of specific information may be misidentified (Hanway & Akehurst, 2018). To date, however, research has not focused on the SM accuracy of interviewers when interviewing multiple witnesses.

Accurately monitoring the source of information when conducting multiple interviews is challenging. The cognitive process of monitoring the source is added to the inherent cognitive demands of interviewing, which may lead to a cognitive load for interviewers. Cognitive load is an indicator of working memory use and the demands placed on cognitive resources when carrying out multiple and competing tasks (Dias et al., 2018; Engström et al., 2013; Van Acker et al., 2018). In the course of an investigative interview, interviewers are required to hold information provided by witnesses in memory, at the same time as assessing that information, thinking of questions to ask, and identifying the correct order in which to ask those questions (i.e., which topic to ask questions about first; Hanway & Akehurst, 2018). These features of interviewing contribute to the cognitive demands of investigative interviewing and require interviewers to attend to multiple cognitive processes (Kleider-Offutt et al., 2016). When completing interviewing tasks, an increase in cognitive demands for interviewers

can negatively impact their recall of information and increase their perceived cognitive load (see Chapter 3).

Cognitive load due to the capacity limitations of working memory means that received sensory information may not be rehearsed and the processing of the information can be restricted (Van Merriënboer & Sweller, 2010). For example, interviewers may neglect to listen carefully to an answer being given by a witness because they are thinking about the next question they should ask. This divided attention may also increase SM errors if useful source specific information is not linked with the memory of a particular piece of information (Mitchell & Johnson, 2000). Cognitive load that results from attending to multiple cognitive tasks during an interview may, therefore, increase errors in memory recall and SM for interviewers.

The current study

The current research examined the effects of increasing cognitive demands on participants' perceived cognitive load (PCL), the accuracy of their memory for information provided in the accounts of multiple witnesses, their accuracy in identifying the source of information and their subjective experience of their memory and their monitoring of the source of information.

We are not aware of any other research that has explored the effects of cognitive load on interviewers tasked with interviewing multiple witnesses, nor are we aware of research examining SM accuracy for information provided by different witnesses across interviews. However, based on previous research regarding cognitive load, recall performance, and SM, it was predicted that in a high cognitive load (HCL) condition participants would report increased levels of PCL, have lower accuracy scores for memory of witnesses' accounts, and lower SM accuracy, when compared with those in a no additional cognitive load (NCL) condition (Hypothesis 1). For interviewers'

subjective experience of remembering details, and of SM, it was predicted that in the HCL condition, participants would report a lower proportion of '*remember*' and '*know*' responses, and a higher proportion of '*guess*' responses, when compared with those in the NCL condition (Hypothesis 2).

Method

Design

For this preregistered (see Appendix C1) independent-groups study, there was one between-subjects factor, cognitive load, with two levels: high cognitive load (HCL) and no additional cognitive load (NCL; control). The dependent variables were perceived cognitive load (PCL), memory accuracy, and SM accuracy. Participants' subjective experiences of recognising information and the source of the information were also measured. The tasks for each condition were designed to replicate cognitive demands that have been identified as inherent in investigative interviewing (i.e. to listen to the witness, remember information, judge information and think of questions to ask; Fisher et al., 2014; Hanway & Akehurst, 2018).

Participants

A priori G*power analysis for a one-way *t*-test, with two independent groups based on an alpha level of 0.05 and a medium effect size of .50, indicated that a sample size of 102 participants was required for this study to have sufficient power at 0.80 (Faul et al., 2009). The sample of 102 participants comprised 65 females and 37 males, who were aged 18 to 45 years (M_{age} = 21.02 years, *SD* = 4.38 years). 91% of the sample reported their nationality as British. Of the 102 participants, 99 reported English as their first language and the other three reported English as their primary language. The participants, who were university staff and students, were recruited via a local participant pool and advertisements placed in university buildings. First year

undergraduate psychology students were offered course credit for their participation. Other participants were offered a monetary incentive of £5.00 for taking part. All participants were informed that they would be required to attend for one 45-minute session during which they would take the role of a police interviewer investigating a crime that had been witnessed by several people.

Materials

Stimulus interviews

Five student actors served as 'witnesses' to a fictitious crime event. The witnesses were all females. (Photographs of the five witnesses are provided in Appendix C.2). They were each given a script to learn relating to the 'crime', which involved a man attacking a woman in a nightclub. Each of the witnesses acted as if they were present in the nightclub where the attack took place. They were asked to describe the event as naturally as possible, including a number of unique and generic details provided in their script. To prevent order effects, each witness described the crime event in a different order (e.g., one witness described the man and then the assault and another described details of the assault and then of the man). The witnesses were each interviewed separately about the crime event and their interviews were audio and video recorded.

During their interviews, the witnesses were asked the same open question and they each provided a free recall narrative containing four unique details and 16 generic details of the crime. That is, 20 details in total were mentioned by each of the five witnesses but each witness provided four unique details. For example, one witness provided four unique details about the incident (i.e., they went to *'Clouds'* nightclub, she was with '*four*' friends, she was on the dance floor with '*Chloe'*, and the man punched the lady on the '*left side*' of her head). The four other witnesses provided generic

information about these unique details (i.e., they went to a nightclub, she was with her friends, she was on the dancefloor with a friend, and the man punched the lady on her head). A matrix table of the unique details is provided in Appendix C.3.

Each of the five interviews lasted for approximately two minutes (*M* = 2.08 minutes, *SD* = 15.38 seconds). The five interviews, presented in the same order, were combined into one media file. Each interview was separated by a screen with the witness number, one through to five, introducing each witness. The break between interviews lasted for 5 seconds. The media file, which lasted for 11 minutes 12 seconds, was played in totality as stimulus material for each participant.

Perceived Cognitive load (PCL) measure

To measure participants' perceived cognitive load, the National Aeronautics and Space Administration, Task Load Index (NASA-TLX) was used. This questionnaire combines information about the magnitude and source of six related factors to derive a sensitive and reliable estimate of workload (Hart & Staveland, 1988). The NASA-TLX was designed to be used during, or immediately after, a task. It has been widely used in a variety of settings (e.g., air traffic control and medical training) to measure the cognitive load perceived by participants when they have completed a task (Hart, 2006; Rizzo et al., 2016).

The NASA-TLX multi-dimensional rating scale questionnaire evaluates participants' subjective workload ratings. The scale items comprise mental demand, physical demand, temporal demand, performance, effort and frustration (Hart, 2006). Each item is measured on a 20-point scale from low to high (except for performance which is measured on a scale from good to poor). A weighted score is obtained by completing 15 pairwise comparisons of the six scale items. For each pair of scale items, one item is selected that the participant feels is more relevant for them when

completing the task (Hart & Staveland, 1988). For this study and following the scoring procedure devised by Hart and Staveland (1988), an overall PCL score out of 100 was calculated. Each scale item score (rating score) was multiplied by the number of times that item was selected in the pairwise comparisons (giving an adjusted score); the six adjusted item scores were then totalled and divided by 15 to obtain the overall PCL score which could range from 0 to 100.

Memory and Source Monitoring Task

Participants were presented with 20 questions, in a random order, about the 20 unique details provided by the five witnesses. They were given multiple-choice answers (four choices) with just one definitive answer for each question. For example, one question was "*what colour jacket was the suspect wearing*?" The answer options were *Red/Blue/Black/White,* and the correct answer was '*black*' (see Appendix C.4 for questions).

For the source monitoring task, participants were presented with their answer to each of the 20 multiple-choice questions (e.g., previously you were asked the question '*How many friends did the witness say she went out with*?' The answer you gave to this question was '*four'*) they were then asked to indicate which witness provided the information (i.e., witness 1, 2, 3, 4, or 5). As each witness had given a unique set of details during their interview, there was only one source, (i.e., correct answer) for each unique detail.

Subjective experiences of remembering and source monitoring

Participants' subjective experiences during the memory task and source monitoring task were measured as the proportions of remember, know, or guess (R/K/G) responses they gave relating to their answers to the 20 unique detail questions and the 20 SM questions. The two forms of recognition memory, recollection

(remember) and familiarity (know) were included as response options to distinguish between memories for specific details and more general memories of the details. 'Guess' was added as a response criterion as, when measuring participants' subjective awareness, 'guess' can prevent 'know' being selected when participants actually guessed the answer (Gardiner et al., 2002). Including 'guess' also helped to confine R–K judgements to confidently recognized items, which can encourage more accurate 'know' responses (Eldridge et al, 2002). Instructions for completing the R/G/K task were drawn from the definitions outlined by Williams and Lindsay (2019). The definitions were as follows:

REMEMBER. You have an experience of recollection for the exact answer. This could include being consciously aware of some aspect or aspects of what was experienced at the time the answer was given in the interview phase (e.g., aspects of the physical appearance of the witness, or of something that happened, or of what you were thinking or doing at the time). In other words, you should choose "Remember" if you have a sense of yourself in the past and/or the question brings back to mind a particular association, image, or thought, from the time of interviews. *For example, if you see someone on the street, you may think, "Who is that? Oh yes, it's the person I saw in the queue in the book shop. I remember thinking what a funny hat they had on…"*

KNOW. You feel that you just know that the answer was a detail you heard in the interviews phase, but you cannot consciously recollect anything about its actual occurrence or what was experienced at the time of its occurrence. In other words, you should choose "Know" if you know the answer was a detail you were given, but you cannot recollect any details associated with seeing or hearing it. *For example, if you see someone on the street, you may think "Who is that? I know I've seen that person before, but I don't recall where that would have been…"*

GUESS. You do not have any memories or feelings associated with the question, and you are simply guessing that the answer was one of the details given in the interviews phase.

Procedure

Ethical approval was obtained from the Science Faculty Ethics Committee of a UK University. On arrival at the lab, all participants were required to read an information sheet and gave written informed consent. To ensure equal numbers (N = 51) in each of two conditions, participants were pseudo-randomly allocated to one of the two conditions, high cognitive load (HCL) or no cognitive load (NCL). All participants were instructed to take the role of a police interviewer and told that their task was to investigate a reported crime of assault on a 23-year-old woman. Participants were informed that several witnesses to the crime had been interviewed regarding the incident, the interviews had been recorded and they would view the recorded witness interviews.

Participants were asked to watch the interviews carefully as if they were the interviewer and listen to everything the witnesses said. They were also informed that they would be asked some questions after they watched the interviews. Participants in the NCL condition were given no further instructions. For the HCL condition, participants were also instructed that whilst watching the interviews, they should consider carefully what the witnesses said so that they could clearly understand the witnesses' knowledge of the event. They were also informed that an additional task was to identify follow-up questions to ask each witness once they had given their accounts. Hence, whilst listening to each of the witnesses, participants in the HCL condition were instructed to think about what further information they would like to obtain from the witnesses to help the investigation and how they should word their questions to obtain

that information. To ensure compliance with this aspect of their task, participants in the HCL condition were provided with a pen and paper and were asked to write their questions down while each interview was in progress.

After receiving instructions specific to their experimental condition, all participants were asked to wear headphones to reduce distractions and they watched the five recorded interviews on a computer screen. Immediately after watching all five of the interviews, all participants completed the PCL measure using the NASA-TLX scale, which was presented to participants via a tablet application. They then carried out a 5-minute distraction task that required them to work through some unrelated number puzzles.

Following the distraction task, all participants were asked 20 multiple-choice questions about the 20 details that had been provided in the accounts given by the witnesses. The questions were presented on a computer using Qualtrics survey software and were presented in random, differing orders for each participant. For each question, participants were provided with four answer choices and they were able to select just one answer. Once participants had selected their answer, they were asked to provide their subjective experience of their memory of the information (i.e., did they R/K/G their answer). For clarity, the definitions of R/K/G were presented to participants on a laminated card and they were able to refer to the definitions throughout the time they were answering the R/K/G questions.

Participants were then presented with the answers they gave to the 20 multiple-choice questions in turn and were asked SM questions for each of the 20 answers they had provided. They were asked which of the five witnesses had provided the information (i.e., witness 1 to 5) relating to each multiple-choice question. When completing this section, a picture of each witness (1 to 5) was given to participants on a

laminated sheet which they could refer to when selecting their responses. After each SM question was answered, participants were asked to provide their subjective experience of their SM memory (i.e., did they R/K/G their answer).

Finally, participants were asked to rate for each witness 1 to 5, using 7-point scales, *'I am confident in the accuracy of my memory for information given by witness (1 to 5)'* from [1] strongly disagree to [7] strongly agree; '*I was motivated to remember the content of the account given by witness (1 to 5)'*, from [1] strongly disagree to [7] strongly agree; *'I found remembering the content of the witness' account (1 to 5)'*, from [1] extremely easy to [7] extremely difficult. Participants in the HCL condition were also asked *'I was motivated to think about questions whilst I was listening to the account of witness (1 to 5)*', from [1] strongly disagree to [7] strongly agree.

As a manipulation check, participants were asked to write down the instructions they were given by the researcher before they watched the witnesses' accounts. They were also asked if they personally knew any on the five witnesses. Demographic details including age, gender, nationality, ethnicity and first language, were also recorded. On completion, a verbal debrief was provided for all participants and they were thanked for their time and effort. Participation in the study lasted approximately 40 minutes.

Results

Manipulation check

All 102 participants passed the manipulation check and accurately reported their instructions. As per their instructions, participants in the NCL condition confirmed they were required to watch the interviews carefully and participants in the HCL condition confirmed that they were asked to watch the interviews carefully and to think of questions to ask the witnesses. One participant in the NCL condition indicated that she knew one of the witnesses. Analyses were conducted with and without this participant's

data, which revealed no differences in the results. For completeness, the data of all 102 participants were included in the analyses and are reported in the results.

Hypothesis testing

Perceived cognitive load (PCL)

A *t*-test was conducted to examine differences between conditions in PCL scores. As predicted, there was a difference in PCL scores between the two conditions. Participants in the HCL condition scored higher (M = 57.63, SD = 11.41) than those in the NCL condition (M = 50.59, SD = 12.15), t (100) = 3.02, p = .003, 95% CI [2.41, 11.67], d = 0.60. In sum, participants in the HCL condition who were required to think of questions whilst watching the witness interviews reported higher perceived cognitive load than did those in the NCL condition who merely watched the interviews.

Memory accuracy for unique details

The percentage accuracy scores for recognition of unique details of the witnesses' accounts were calculated as the percentage of correct responses to the 20 multiple-choice questions. A *t*-test was conducted to examine differences between conditions for percentage accuracy of recognised information. As predicted, there was a difference in percentage accuracy between the two conditions. Percentage accuracy for unique details was lower in the HCL condition (M = 77.16, SD = 11.80), than in the NCL condition (M = 83.24, SD = 9.63), t (100) 2.85, p = .005, 95% CI [1.85, 10.31], d = .57.

Accuracy of source monitoring

First, the percentage accuracy scores for monitoring the source of the unique details were calculated as the percentage of correct responses to the 20 SM questions, irrespective of whether the answer provided for the unique detail had been correct (i.e., monitoring the source of information without considering the accuracy of unique detail questions). A *t*-test was conducted to examine differences between conditions for

percentage accuracy of SM. The predication that there would be a difference in SM scores was not supported. There was no significant difference in accuracy of SM between participants in the HCL condition (M = 39.71, SD = 12.90) and those in the NCL condition (M = 43.04, SD = 12.33), t (100) = -1.33, p = .185, 95% CI [-8.29, 1.63], d = .27.

Second, percentage accuracy scores were then calculated for correct answers to the SM questions following a correct answer to the unique detail questions (i.e., recognition of both the unique detail and the source of the detail was correct). A *t*-test was conducted to examine differences between conditions for percentage accuracy of SM and correct details. There was no significant difference in accuracy of SM between participants in the HCL condition (M = 34.41, SD = 12.56) and those in the NCL condition (M = 38.14, SD = 12.33), t (100) = 1.44, p = .154, 95% CI [-1.42, 8.87], d = .30.

Subjective experiences of remembering

Participants' subjective experience of their memory for unique details and SM were measured as the proportions of remember, know, or guess (R/K/G) responses they gave relating to their answers to the 20 memory questions and the 20 SM questions.

Memory of unique details. Pearson's correlations indicated that there were significant, but moderate, associations between the proportion of R/K/G responses related to participants' subjective experience of their answers to the unique detail questions (remember and know, r (102) = -.78, p < .001; remember and guess r (102) = -.65, p <.001). There was no correlation between know and guess responses, r (102) = .03, p = .760. As a result, the assumption of an absence of multicollinearity was met and to reduce Type 1 error, a one-way between-groups MANOVA with Condition (HCL vs. NCL) as the only factor was conducted. The proportion of remember, know and guess responses about the unique details were the three dependent variables. Separate

MANOVAs were conducted for all answers, correct answers, and incorrect answers, that participants provided to the unique detail questions.

For all answers to subjective experience of the unique detail answers, the MANOVA revealed a significant multivariate main effect for condition, Wilks $\Lambda = .86$, F (3, 98) = 7.93, $p = .001$, $\eta^2_p = .14$. As predicted, the univariate main effects revealed a difference in the proportion of remember and guess responses between HCL and NCL conditions. Those in the HCL condition reported 'remembering' fewer of their answers to the unique details questions and 'guessing' more of their answers than did those in the NCL condition. Contrary to the prediction, there was no significant difference between conditions for the reported subjective experience 'know' (see Table 4.1).

To examine any differences between conditions for R/K/G responses, that may be dependent on the accuracy of recognition for details, analyses were conducted for correct and incorrect answers to recognition of the unique details. For correct answers to unique details, the MANOVA revealed no significant multivariate main effect for condition, Wilks $\Lambda = .95$, F (3, 98) = 1.84, $p = .146$, $\eta^2_p = .05$. For incorrect answers to the unique detail questions, the MANOVA also revealed no significant multivariate main effect for condition, Wilks $\Lambda = .99$, F (3, 98) = .40, $p = .754$, $\eta^2_p = .01$ (see Table 4.1).

Table 4. 1

Mean and standard deviation scores for R/K/G responses including results for all answers, correct answers and incorrect answers to unique detail questions for each condition.

Answers	**R/K/G**	**HCL**		**NCL**				
		Mean	**SD**	**Mean**	**SD**	***t(1,100)***	***p***	***d***
All answers	Remember	.50	.16	.60	.17	9.66	.002	.61
	Know	.25	.11	.23	.15	.81	.370	.15
	Guess	.25	.11	.17	.10	15.19	<.001	.77
Correct answers	Remember	.62	.17	.71	.19	2.38	.019	.50
	Know	.25	.14	.22	.17	-1.07	.289	.19
	Guess	.13	.09	.08	.07	-3.22	.002	.63
Incorrect answers	Remember	.07	.13	.03	.12	-1.57	.120	.32
	Know	.24	.23	.27	.32	.65	.517	.11
	Guess	.67	.29	.66	.34	-.27	.791	.03

Memory for sources of information. For recognition of the source of information, irrespective of the accuracy of answers to the unique details questions, Pearson's correlations indicated that there were significant, but moderate, associations between the proportions of remember, know, and guess responses out of 20 for answers to recognising the sources of information questions: for remember and know, r (102) = -.49, p < .001; for remember and guess r (102) = -.59, p <.001; and for know and guess, r (102) = .41, p <.001. As a result, the assumption of an absence of multicollinearity was met and to reduce Type 1 error, a one-way between-groups MANOVA with Condition (HCL vs. NCL) as the only factor was conducted. Remember, know and guess sources of information were the three dependent variables.

For the MANOVA for recognition of the source of information, there was no significant multivariate main effect for Condition, Wilks Λ = .96, F (2, 99) = 2.35, p = .101, η^2_p = .05. To examine any differences between condition for R/K/G responses that may be dependent on the accuracy of SM, additional exploratory analyses were conducted for correct and incorrect answers to recognition of the source of information. For correct answers, the MANOVA revealed a significant multivariate main effect for condition, Wilks Λ = .91, F (3, 98) = 3.23, p = .026, η^2_p = .09. The univariate main effects revealed a difference in the proportion of guess responses for the source of information between the HCL and NCL conditions. Those in the HCL reported 'guessing' more answers to the SM questions than did those in the NCL condition, there was no significant difference between conditions for the reported subjective experiences of 'remember' and 'know' (see table 4.1). For incorrect answers to recognition of the unique details, the MANOVA revealed there was no significant multivariate main effect for condition, Wilks Λ = .97, F (3, 98) = 1.13, p = .342, η^2_p = .03.

Table 4. 2

Mean and standard deviation scores for R/K/G responses including results for all answers, correct answers and incorrect answers to source monitoring questions for each condition.

Answers	**R/K/G**	**HCL**		**NCL**				
		Mean	**SD**	**Mean**	**SD**	***t(1,100)***	***p***	***d***
All answers	Remember	.25	.17	.29	.18	1.30	.257	.23
	Know	.33	.15	.37	.16	1.06	.305	.26
	Guess	.41	.16	.34	.17	4.72	.032	.43
Correct answers	Remember	.37	.23	.40	.22	.70	.485	.13
	Know	.33	.15	.37	.16	1.06	.305	.26
	Guess	.29	.21	.21	.15	-2.17	.033	.44
Incorrect answers	Remember	.18	.18	.20	.19	.50	.615	.11
	Know	.34	.17	.36	.21	.66	.514	.11
	Guess	.48	.20	.44	.22	-1.06	.291	.19

Correct unique detail and correct source monitoring answers. For recognition of the source of information, when correct responses were given for unique details and the source of information, Pearson's correlations indicated that the correlations between proportions of remember, know, and guess responses were small: for remember and know, r (102) = -.19, p =.60; for remember and guess r (102) = -.23, p =.021; and for know and guess, r (102) = -.002, p =.982. The assumption of an absence of multicollinearity was not met, therefore, *t*-tests were conducted to explore differences between the HCL and NCL conditions. The *t*-tests revealed there were no significant differences in the proportion of remember, know, or guess responses for the source of information between the HCL and NCL conditions when a correct answer to unique

detail and correct source of information was given (see Table 4.3).

Table 4. 3

Mean and standard deviation scores along with t-test results across each of the dependent variables for each condition.

	HCL		**NCL**		***F*(1,100)**	***p***	***d***
	Mean	**SD**	**Mean**	**SD**			
Remember details	.15	.12	.16	.10	.13	.715	.09
Know details	.13	.09	.15	.09	1.31	.256	.22
Guess details	.06	.08	.07	.07	.47	.497	.13

Confidence, ease of remembering, and motivation

In the post recall questionnaire, the dependent variables were confidence, ease of remembering the witnesses' accounts, motivation to remember the accounts, and for the HCL condition, motivation to think of questions to ask the witnesses. As participants' scores for each witness (1-5) contributed the same weight to each of the dependent variables, a composite score for each variable was calculated as the mean of participants' scores for each witness (1-5) for each variable. A series of Pearson's correlations were conducted to determine whether the dependent variable composite scores were correlated with each other (see Appendix C.5). Only two of the variables were moderately correlated, therefore, the assumption of an absence of multicollinearity was not met. A series of *t*-tests were conducted for each dependent variable with condition (HCL vs. NCL) as the only factor. To reduce the risk of Type 1 errors a Bonferroni adjustment was made (i.e., the alpha level of .05 was divided by the number of tests to be performed [4] to give an alpha of .013). Differences between participants' motivation, confidence, and their ease of remembering the account for

each witness were investigated.

For participants' confidence in the accuracy of their memories for witnesses' accounts, there was no difference between the HCL (M = 4.47 SD = .91) and NCL (M = 4.69 SD = 1.11) conditions, t (100) = -1.09, p = .278, 95% CI [-0.62, 0.18], d = 0.22. For participants' ratings of ease of remembering witnesses' accounts, participants in the HCL condition rated remembering the accounts as more difficult than did those in the NCL condition, HCL M = 3.93 (SD = 1.17), NCL M = 3.38 (SD = 1.00), t (100) = 2.59, p = .011, 95% CI [1.22, 0.98], d = 0.51, where ratings of 1 = extremely easy and ratings of 7 = extremely difficult. For participants' motivation to remember the witnesses' accounts, participants in the NCL condition rated their motivation to remember the accounts as higher than did those in the HCL condition, HCL (M = 5.05 SD = .95), NCL (M = 5.51 SD = .87), t (100) = -2.53, p = .013, 95% CI [-0.82, -0.10], d = 0.51, where 1 = strongly disagree and 7 = strongly agree. Participants in the HCL condition were also asked to rate their motivation to think of questions to ask while listening to the witnesses' accounts, M = 5.56 SD = 1.00, where 1 = strongly disagree and 7 = strongly agree.

Discussion

The effects of increasing cognitive demands on participants' perceived cognitive load (PCL) and their memory for details in accounts that were provided by multiple witnesses were examined. The findings demonstrated that, as predicted, participants' perceived a higher cognitive load when they were required to complete additional cognitive tasks (i.e., considering the information provided and thinking of questions to ask whilst watching and remembering the content of the witnesses' accounts), compared with simply watching and remembering the content of the witnesses' accounts. Participants who were asked to think about follow-up questions were less accurate for their memory of unique details that had been given by each witness than

those who were not asked to think about questions. Contrary to the prediction, there were no differences between conditions for SM accuracy with respect to which of the witnesses provided the specific details.

For participants' subjective ratings of their memory for unique details, in the HCL condition, participants who were required to think of questions during their task, reported a lower proportion of 'remember' responses and a higher proportion of 'guess' responses, when compared with those in the NCL condition who merely watched the interviews. This pattern was repeated when correct answers to unique detail questions were given with lower proportions of 'remember' responses, and higher proportions of 'guess' responses in the HCL condition compared to the NCL condition. However, there were no differences in R/K/G responses for incorrect answers. When considering participants' subjective experience of SM, for all responses and when participants gave a correct response to the SM questions, those in the HCL condition reported a higher proportion of 'guess' responses for their subjective experience of SM, when compared with those in the NCL condition. However, there were no differences when an incorrect response was given. There were no differences in 'remember' and 'know' responses across conditions when all responses were considered nor when participants had given correct or incorrect responses to the SM questions.

To test participants' memory of the witnesses' accounts, the current research used a recognition, rather than recall, task. It was demonstrated that impaired *recognition* memory accuracy for unique details in witness accounts when under high cognitive load, broadly replicates previous research showing impaired performance by mock-interviewers on *recall* tasks when under high cognitive load (see Chapter 3). Recognition memory is generally more accurate, and less effortful, than recall (Yonelinas, 2002). Therefore, it was interesting that participants in the current study

performed poorly under a high cognitive load compared with those in the relatively low load condition. If interviewers cannot recognise information because it has not been encoded, or it is not available for retrieval, then the amount and accuracy of information recalled by interviewers will inevitably be reduced (e.g., Cauchi & Powell, 2009; Lamb et al., 2000; Warrren & Woodall, 1999). The current results indicated this might occur due to the demands of the concurrent cognitive tasks that interviewers undertake.

In relation to the cognitive tasks required during interviews with multiple witnesses, an interviewer must attend to the information provided by a witness and consider if that, or similar, information has been previously given. The interviewer will also then decide who provided the information and if the information matched the witness's information. Some of these SM decisions are rapid and automatic requiring less conscious thought, however, other decisions are more effortful and require conscious decision making (Johnson et al, 1993). Yonelinas (2002) suggested that the conscious recollection of information (remembering) is more likely to be negatively affected by divided attention than is perceived familiarity of information (knowing). In the current study, participants who were required to complete additional cognitive tasks (i.e., thinking of questions to ask) *remembered* fewer unique details than did those who merely watched the witnesses' interviews, however, there were no differences in *know* responses across cognitive load conditions. As familiarity (knowing) is less sensitive to the effects of divided attention, this may explain the current findings.

Considering the process of source monitoring in practical settings, interviewers have described that when dealing with complex crimes, they are often required to interview multiple witnesses about the same crime or the same offender (Hanway & Akehurst, 2018). When conducting interviews with multiple witnesses, interviewers will rely on their source memory to identify who provided specific information (Kleider

et al., 2008). In the current study, no difference in accuracy for monitoring the source of information between participants in the HCL and NCL conditions was found. However, it is worthy of note that all participants performed poorly on the SM task (i.e., when correct answers to recognition task questions were given, SM accuracy was only 34% for the HCL condition and 38% for the NCL condition). The low SM accuracy scores indicate that, during divided attention tasks, such as interviewing, it is likely that useful source specific information may not be linked to the details of the information and increased SM errors may result (Mitchell & Johnson, 2000). In addition, as rates were low there may have been an impact on results for the R/K/G analysis for correct identification of the source of information (i.e., floor effects with the rates too low to detect any differences).

The low SM accuracy rates mirror those of previous studies, which suggest people are more likely to confuse memories from similar sources (Lindsay et al., 1991). For the current study, the sources of information (i.e., the five witnesses) were similar, in that, they were of the same ethnicity, age and gender. The five witnesses also provided similar details that differed in specificity. That is, one witness provided a unique detail (e.g., "we went to 'Clouds' nightclub"), whereas the four other witnesses provided generic information (e.g., "we went to a nightclub"). Thus, low SM rates may also be due to the nature of the information (i.e., generic or specific) that was provided by the five witnesses.

For the current study, although participants in the HCL condition wrote down their follow-up questions, no participants made any notes of what was said by the witnesses while they were watching the interviews. Actively attending to aspects of source information during an event (e.g., by noting down what is being witnessed) has been shown to enhance source memory (Lindsay, 2007). However, as note taking can be

cognitively demanding and may divide attention further (e.g., Piolat et al., 2005), the aim was to focus participants' attention on the cognitive tasks of thinking about questions and remembering information. It is acknowledged that, in operational settings, interviewers may take notes during interviews. These notes, along with other written witness statements, may be available during subsequent interviews to aid interviewers recall. However, this is not always the case, for example in intelligence gathering contexts, when interviewing at the scene of an incident, or when interviewing child witnesses. Trained interviewers have also identified that note taking can be distracting during an interview (see Chapter 2). Future research should examine whether note taking would have an impact on interviewers' PCL or their accuracy for memory for unique details and the source of information.

The nature of the tasks that were completed for the current study (i.e., passively watching the account of an interviewee and a recognition memory task) are not akin to those experienced by investigators in the field. Interviewers in practice would be engaged with interviewees during their interviews and would not have additional information available (e.g., response options) to prompt their recognition memory. To increase the generalizability of the current findings it is recommended that future research examine whether the current results replicate with trained investigative interviewers.

In sum, the current research has highlighted that keeping track of information provided by multiple witnesses when they give their account of a crime is challenging for interviewers. This challenge is reflected in impaired memory accuracy and impaired ability to monitor the source of specific information. To reduce such errors, interviewers should be trained to not only carefully attend to information provided by witnesses, but also to attend to the source of information. This may assist interviewers'

later recall of key information when interviewing additional witnesses to the same event.

Chapter 5:

The effects of mock-interviewers' note taking and working memory capacity on perceived cognitive load and recall of a witness's account of a crime.

Abstract

Cognitive load has an impact upon interviewers' recall of information provided by witnesses (see Chapters 3 and 4). Taking notes during investigative interviews can be cognitively demanding but it can also improve memory for information. Further, processing information during an interview takes place in working memory, but a limited working memory capacity (WMC) may impact performance. The current study examined the effects of note taking and working memory capacity on interviewers' perceived cognitive load and their recall of information provided by a witness. Participants took free notes or structured notes and they either had access, or not, to their notes when they recalled a witness's account. A third group of participants (the control group) did not take notes whilst listening to the witness. The different note taking conditions had no effect on interviewers' PCL when observing the witness's interview, but WMC and access to notes were moderators of PCL at recall. Regardless of WMC level, participants' free- and cued-recall was more accurate when they took structured notes and they had access to their notes at recall than it was for any other groups. To reduce errors in their recall of information gleaned from witnesses, it is recommended that interviewers should take structured notes as this did not impact on participants' PCL and may be beneficial for recall for all interviewers.

Introduction

Investigative interviewing is a complex cognitive undertaking. Interviewers are required to complete multiple competing tasks, including remembering the information provided by a witness, thinking of questions to ask and taking notes (Hanway & Akehurst, 2018). However, carrying out multiple cognitive tasks may lead to increased cognitive load, which can impact on interviewers' recall of information (see Chapters 3 and 4). In addition, as individuals have a limited working memory capacity, the rehearsal of received information may be restricted (Van Merriënboer & Sweller, 2010). Completing multiple cognitive tasks during an interview may be, therefore, limited by individual capacity to process information in working memory. Using a witness interview observation methodology, the current research examined the effects of note taking on recall of a witness's account of a crime and explored whether working memory capacity, and access to notes during retrieval, were associated with the amount and accuracy of details recalled.

The cognitive demands of investigative interviewing

An interviewer's role is to refrain from interrupting the witness and to engage in active listening (Brandon et al., 2018). However, analysis of interviewers' performance in the field found that interviewers' often interrupt witnesses, ask suggestive and leading questions (Schreiber Compo et al., 2012; Snook & Keating, 2011) and fail to comply with evidence-based interviewing practices, such as the use of open questions (Lamb, 2016). For example, a review of eight studies examining investigative interviewing practices found that, even after interview training, investigators asked more questions that were deleterious to memory recall (e.g., suggestive or leading questions), than non-deleterious questions (e.g., open questions; Launay & Py, 2015). It

seems, therefore, that applying skills learnt during interviewer training in practical interview settings can be difficult (Duron & Cheung, 2016; Mount & Mazerolle, 2020).

Recent research by Hanway and Akehurst (2018) noted that, when conducting an investigative interview in practice, interviewers are required to hold information provided by witnesses in memory, at the same time as assessing that information, thinking of questions to ask, and identifying the correct order in which to ask those questions. These features of interviewing require interviewers to attend to multiple cognitive processes (Kleider-Offutt et al., 2016). When increased cognitive demands were required to complete an interview task (i.e., thinking of questions and remembering information, as opposed to, being required to just remember information), participants perceived a higher cognitive load and the accuracy of their recall of information provided by a witness was reduced (see Chapters 3 and 4). A reduced recall of information and increased cognitive load may also impact on interviewers' compliance with best practices (Hanway & Akehurst 2018).

Working memory and cognitive load

Working memory (WM) is a functional short-term memory system that enables temporary information storage and maintenance, which is necessary for many cognitive processes such as reasoning and language comprehension (Baddeley, 1992). The role of WM is central to all deliberate cognition, such as prose comprehension and learning (Baddeley & Hitch, 1974; Oberauer et al., 2018). WM consists of four components; (i) the central executive component serves as an attentional control system and determines what information will be entered into, and maintained, in the (ii) visuospatial sketchpad and the (iii) phonological loop components (Baddeley & Logie, 1999). The fourth component, the episodic buffer, forms links between these systems and long-term memory storage (Baddeley et al., 2019). Working memory capacity

(WMC) represents an individual's capacity to i) process information relating to a primary task, ii) maintain the relevant information for the primary task, and iii) access and retrieve information from long-term memory, in the presence of a distraction (Unsworth & Engle, 2007). Interviewers' WMC may, therefore, have an impact when conducting an investigative interview and taking notes.

It is generally accepted that the components of WM allow task-relevant information to be accessed during the execution of complex cognitive tasks (Conway et al., 2005). In WM, mental representations are temporarily available for use during thought and action (Cowan, 2017). However, individual differences in capacity and the use of WM, for reasoning and language comprehension, have also been identified (Unsworth & Engle, 2007). These individual differences and the limited capacity of WM are related to an individual's ability to maintain a limited amount of information in a state that can be accessed during a task (Engle, 2002; Rosen & Engle, 1998). In experimental word list studies, working memory capacity has been shown to be a predictor of both free and cued recall (Unsworth, 2009).

When dual tasks are carried out, cognitive resources can be overloaded and there may be a reduction in performance on both task components (Hart & Staveland, 1988). For example, divided attention at encoding can reduce accuracy of recall (Marsh et al., 2017). In terms of investigative interviewing, interviewers are required to hold information in mind, assess the information, consider it with other information that is known, ask relevant follow-up questions, and take notes (Hanway & Akehurst, 2018). When carrying out these multiple and competing tasks, during an interview, the demands placed on individuals' WM may also lead to a cognitive load (Van Acker et al., 2018). Cognitive load that results from attending to multiple cognitive tasks during an interview may, therefore, increase recall errors made by interviewers.

Note taking during interviews

Notes taken during investigative interviews are often used by interviewers to help them to remember witnesses' accounts and/or write up reports or statements. Interviewers may also use notes during interviews to decide on topics to pursue and questions to ask (Shepherd, 2007). As such, note taking is an important component of an investigative interview (Gregory et al., 2011). To ask appropriate questions without misleading the witness, an interviewer's notes need to be an accurate and true reflection of what has been said; they may also be used as evidence in the courtroom (Ministry of Justice [MoJ], 2011). MacDonald (2016) found, during three studies, that when interviewers took notes, they recalled more information about the content of the interview, than interviewers who did not take notes. There was a further improvement of interviewers' recall of information when they were given access to their notes at recall (MacDonald, 2016). However, although generally accurate, interviewers' notes are often incomplete. Lamb and colleagues (2000) found that when interviewers took notes during investigative interviews with children, 25% of forensically relevant details, and 18% of details central to the investigation, were not noted. This level of omission suggests that interviewers may filter information provided by witnesses, which may be due to a lack of encoding of the information or, although encoded, an interviewer may have deemed some information to be peripheral and unnecessary (Fisher & Schreiber, 2017). A lack of detail, especially peripheral information, may reduce the topics and questions covered during an interview (Gregory et al., 2011).

The amount and accuracy of details recalled may be influenced by the style of note taking. Flanigan and Titsworth (2020) found that the lack of efficiency during long hand note taking meant that fewer details were recorded when compared with more structured methods. That said, MacDonald (2016) found that more organised methods

of note taking (e.g., using a 'spidergram') did not improve interviewers' recall performance when compared with conventional (i.e., free) note taking. Topic boxes can be used for note taking during an interview. This method groups relevant details systematically into topics (Shepherd, 2007). Macdonald (2016) suggests that completing more structured types of notes may add complexity to note taking and may be more difficult to complete during an investigative interview. However, it is also argued that non-linear note taking has a more cognitively optimal knowledge management system that allows for better integration of new information (Makany et al., 2009). It may be that structured note taking is more cognitively demanding but can improve the processing of information, and recall, if an individual has the cognitive capacity to conduct an interview and take notes.

The current study

The current research examined the effects of no note taking, free note taking and structured note taking, on participants' perceived cognitive load (PCL) during a witness interview observation task. Participants' recall of information provided by the witness was also examined. We are not aware of previous studies that have examined these factors but based on previous research regarding cognitive load, note taking, and working memory capacity, the following predictions were made. Firstly, it was hypothesised that participants in a Structured Note Taking (SNT) condition would report increased levels of PCL for an interview observation task compared to those in a Free Note Taking (FNT) condition, who in turn would report increased levels of PCL for the interview observation task than those in a No Note Taking (NNT) condition (Hypothesis 1). It was also predicted that participants in a NNT condition would report fewer details and have lower accuracy scores for their free and cued recall about a witness's account than those in a FNT condition, who in turn would report fewer details

and have lower percentage accuracy for their recall of the witness's account than those in a SNT condition (Hypothesis 2). For Hypothesis 3, it was predicted that participants in a 'no access to notes' condition would report fewer details and have lower percentage accuracy for their recall of the account given by the witness than those in an 'access to notes' condition. For Hypothesis 4, it was further predicted that participants in a NNT condition would report increased levels of PCL for the recall tasks compared to those in a FNT condition, who in turn would report increased levels of PCL for the recall tasks when compared to those in a SNT condition. Finally, it was predicted that participants in a 'no access to notes' condition would report increased levels of PCL for the recall of the interview task compared to those in an 'access to notes' condition (Hypothesis 5).

Method

Design

The method and data analysis for this study were preregistered (see Appendix D.1) and followed a similar design to those used in Chapters 3 and 4. For the independent groups study there were two between-subject factors. The first between-subject factor was Note Taking with three levels, No Note Taking (NNT; control), Free Note Taking (FNT) and Structured Note Taking (SNT). The second between-subjects factor was Access to Notes, with two levels, access to notes and no access to notes. The Access to Notes factor only applied to the FNT and SNT (i.e., not the NNT) levels of the note taking condition, thus, this was not a fully crossed 3 x 2 design. Therefore, to examine differences between the conditions, and following previous research with similar design structure (e.g., Thorley et al., 2016), the current data were analysed by comparing the five experimental conditions to address Hypotheses 1 to 5. Moderation analyses were conducted to examine interactions between the note taking conditions and access to notes with WMC as a moderator. The dependent variables were perceived

cognitive load (PCL) when performing an interview observation task and PCL for two subsequent recall tasks. In addition, the amount and accuracy of free recall details and the accuracy of participants' cued recall were also measured.

Participants

Participants, who were university staff and students, were recruited via a local participant pool and advertisements placed in university buildings. Undergraduate psychology students were offered course credit for their participation. Other participants were offered a monetary incentive of £10.00 for taking part. All participants were informed that they would be required to attend for two 45-minute sessions, and during the second session they would take the role of a police interviewer investigating a crime.

Perceived cognitive load was a key measure for this study and, as this measure may be affected by age-related cognitive processing capacity (Salthouse, 2009), participants under 18 years and over 60 years of age were excluded from participating. A further key dependent variable involved understanding verbal outputs recorded in English. Therefore, to ensure language comprehension, and to ensure that language skills did not interfere with the study outcomes, English was required as a first or primary language.

A G*power analysis for a one-way ANOVA, with three groups based on an alpha of .05, power of .80 and a medium effect size of .35, gave a desired sample size of 130. For the moderation analysis (i.e., a linear multiple regression), to achieve .80 power, with a small effect size of .15, and an alpha of .05, the sample size would need to be 77 (Faul et al., 2009). The final sample of 130 participants comprised 97 females and 33 males, who were aged 18 to 57 years (M_{age} = 22.81 years, SD = 7.00 years). Of the 130

participants, 119 (91.5%) reported English as their first language with the remaining 11 participants reporting English as their primary language.

Materials

Stimulus interview

A male drama student was recruited to act as a witness to a fictitious crime. This actor was given a script to learn, which related to the fictitious crime event. The script described a man attacking, and injuring, a busker outside a train station, the man then threatened the witness with a knife before fleeing from the scene. To enable an accurate reflection of a real-world interview, the interview room setting, interview procedure, and recording of the interview, were designed to correspond with published guidance for interviewing witnesses (i.e., ABE guidance; MoJ, 2011). The interview was digitally recorded. The view of the witness was captured from behind the interviewer, with the main frame of the recording showing the head and body of the witness. During the interview, the student actor was asked an open question (i.e., *'please tell me everything you can remember about the incident'*) and he provided a free recall account containing details of the crime. The interview lasted for six minutes.

Working Memory Capacity (WMC) tests

Participants' working memory capacity (WMC) scores were measured using shortened complex span tasks (i.e., OSPAN, SSPAN, RSPAN; Foster et al., 2015). During each test participants were given a sequence of items to remember (e.g., letters for the OSPAN test) and between presentation of each 'to be remembered' item participants completed a distraction test (e.g., a simple mathematical puzzle for the OSPAN test; see Appendix D.2). Each sequence comprised 2 to 7 items and the span of the number of items was randomized (e.g., 3 items, then 5 items, then 2 items, etc.). After each sequence of items participants completed a recall task to measure their WMC. One block

(3 sequences of 2-7 items) for each of the span tests (i.e., OSPAN, SSPAN, RSPAN) was conducted. The tests were presented on a computer using E-Prime software. One block for each test was presented as this shortened version allows for accuracy in assessing WMC, while minimising the administration time of the tests (Foster at al., 2015). These WMC tests take approximately 30 to 40 minutes to complete, therefore, to avoid participant fatigue, tests were completed at least one day prior to completion of the interviewing tasks.

Perceived Cognitive Load (PCL)

Participants' perceived cognitive load was measured using a subjective rating scale questionnaire, the National Aeronautics and Space Administration Task Load Index (NASA-TLX). This questionnaire provides a sensitive and reliable estimate of workload (Hart & Staveland, 1988). The NASA-TLX was designed to be used during, or immediately after, a task. It has been widely used, in a variety of settings (e.g., in air traffic control and medical settings), to measure cognitive load as perceived by participants during a range of tasks (Hart, 2006; Rizzo et al., 2016).

For the current study, the multiple item NASA-TLX scale measured participants' ratings of mental demand, physical demand, temporal demand, performance, effort and frustration, during the task. For the rating score, each item was measured on a 20-point scale from low to high (except for performance which was measured on a scale from good to poor). A weighted score was obtained by completing 15 pairwise comparisons of the six scale items. For each pair of scale items, one item was selected that the participant felt was more relevant for them when completing the task than was the other item in the pair (Hart & Staveland, 1988). Each scale item score (i.e., the rating score) was multiplied by the number of times that item was selected in the pairwise comparisons giving an adjusted rating score. The six adjusted rating scores were then

totalled and divided by 15 to obtain an overall PCL score out of 100, with higher scores reflecting perceptions of higher cognitive load. The NASA-TLX scale was presented on a tablet device, after participants had completed each task (i.e., after the interviewing task and again after the two recall tasks).

Procedure

In Session 1, participants completed the WMC complex span tasks. The tests were presented on a computer using E-Prime and, on completion, arrangements were made for participants to attend for their second session at least one day later.

When participants attended for Session 2, they were randomly allocated to one of the five experimental conditions (FNT [access to notes] vs. SNT [access to notes] vs FNT [no access to notes] vs SNT [no access to notes] vs NNT). 26 participants were allocated to each of the five conditions. All participants were informed that a crime had occurred, which they were to investigate, and they were instructed to take the role of a police investigator during an interviewing task. Participants were asked to watch a recorded interview of a witness giving his account of a crime. They were instructed to watch and listen to the recording carefully and to think of questions they would like to ask the witness if they were the interviewer in the case. They were also instructed to think about how, and in what order, they would ask the questions. These instructions were designed to replicate cognitive tasks that are required during investigative interviews (i.e., listening to a witness, remembering information, judging information, and thinking of questions to ask; Fisher & Geiselman, 2010; Hanway & Akehurst, 2018). Participants were informed that they would be asked some questions about the witness's account after they had watched the interview.

For the SNT condition, the researcher provided participants with a sheet of pre-printed note paper containing four separate boxes with titles; person/description

details; location/setting details; action details; and object details (see Appendix D.3). For this condition, participants were given the following instructions: *"Here is a structured note pad and a pen, please feel free to make notes regarding the witness's account while you are watching the recording. These are your notes so you may record any information you wish".* Participants were given time to read and understand the format of the note pad and any questions they had were answered. In the FNT condition, the researcher provided participants with a blank sheet of note paper and gave the following instructions: *"Here is some note paper and a pen, please feel free to make notes regarding the witness's account while you are watching the recording. These are your notes so you may record any information you wish".* In the NNT (control) condition no instructions concerning note taking were given to the participants and they were not given a pen or paper.

Once participants were satisfied with their instructions, they watched the recorded interview of the witness on a computer screen, wearing headphones to reduce any distractions. When the interview recording was finished, the researcher took the notes from those in the SNT and FNT conditions. All participants were then asked to complete the NASA-TLX questionnaire in relation to the interview observation task they had just completed in order to record their PCL. They then carried out a 5-minute distraction task that required them to work through some number puzzles.

Following the distraction task, participants were asked, with one open question, to verbally recall as much information, in as much detail as they could, about the witness's account. Half of the participants who had taken notes (those in the SNT and FNT conditions) were given their notes to help them with their recall tasks. The other half of participants in these conditions were not given their notes. After participants finished their free recall, they were asked if there was anything further that they could

recall about the interview. Once participants had completed the free recall task, they were asked 40 cued recall questions about the content of the witness's account (e.g., "*What time did the witness say he finished work?*" *"What was the name of the bar the witness went to for a drink?"*; *see* Appendix D.4). The order of these questions was randomised across participants. All participants were audio recorded whilst they gave their free narrative and answered the cued recall questions.

Participants then completed the second NASA –TLX questionnaire in relation to the recall tasks in order to record their PCL for this second task. For completeness, as participants had been asked to think of questions to ask the witness, they were all asked to write down questions they would like to ask the witness if they were the investigator in this case.

Finally, participants completed a post-experiment questionnaire. They were asked to rate, using 7-point scales, their confidence in their memory accuracy from [1] not at all confident to [7] extremely confident, their motivation to remember the witness's account from [1] not at all motivated to [7] extremely motivated, and the extent to which they found remembering the content of the witness's account easy or difficult, from [1] very easy to [7] very difficult.

As a manipulation check, participants were asked to write down the instructions they were given before watching the witness's account. Demographic details were obtained including participants' age, gender, nationality, ethnicity and first language. A verbal debrief was provided to all participants and they were thanked for their time and effort. Participation in the study took approximately 45 minutes for each of the two sessions.

Coding

Free recall

The audio recordings of participants' free recall of the witness's account were transcribed verbatim and coded for details reported. Consistent with previous research (e.g., Wright & Holliday, 2007) and to facilitate assessment of overall accuracy, details were coded as correct, incorrect or confabulations. A detail was deemed correct if it was present in the witness's account and was correctly reported by the participant (e.g., *"he finished work at five o'clock"*); incorrect if a reported detail was discrepant from the witness's account (e.g., participant recalls *"the man had black hair"* but the witness actually said *"short dirty brown hair"*); and confabulated if a reported detail was mentioned in a participant's account which was not mentioned at all by the witness (e.g., the participant reported *"the man's behaviour was quite strange"* but the witness did not mention the man behaving strangely). A proportion accuracy rate was calculated by dividing the total number of correct details reported by a participant by the total number of details reported (i.e., correct plus incorrect plus confabulations).

One coder, Rater 1, blind to the condition-allocation of each participant, rated all transcripts. Inter-coder reliability for the free recall accounts was assessed by selecting 26 interview transcripts (20%), which were coded by a second independent scorer, Rater 2. Intra-class correlation coefficients (ICC) using absolute agreement were computed for the following measures: total details [r (25) = .97, p <.001]; correct details [r (25) = .95, p < .001]; incorrect details [r (25) = .83, p < .001]; confabulations [r (25) = .92, p < .001]; accuracy [r (25) = .97, p < .001]. These analyses indicated that the inter-coder reliability was 'good' for the coding of incorrect details, and 'excellent' for the coding of total details, correct details, confabulations, and accuracy (Koo & Li, 2016). As

the analyses indicated that the inter-coder reliability was 'excellent' or 'good' for further analyses the coding of Rater 1 was used.

Cued recall

Participants' answers to 40 cued recall questions were scored by one rater, blind to the condition-allocation of participants, as; fully correct with one point (e.g., *'What time did the witness say he finished work?'*, response *"Five"*) or fully correct with two points (e.g., *'What did the witness say the lady outside the train station was wearing?'*, response *"a dress and a coat"*). Two-point answers could be scored as partially correct with one point (e.g., participant responds *"a dress"*). Don't know responses were scored zero, and incorrect answers scored minus one point (e.g., *"What colour hair did the lady outside the train station have?",* response *"brown"*, when the correct answer was *"blonde"*). There were 24 two-point questions and 16 one-point questions, the total accuracy could therefore range from -40 (all questions answered incorrectly) to 64 (all questions answered correctly and in full). The scores were added, and a percentage accuracy score for each participant was calculated (i.e., score/64 x 100).

Results

Manipulation check and data screening

All participants passed the manipulation check by correctly reporting, as appropriate for their condition, the instructions they had received at the outset of Session 2.

The data were checked for missing values and outliers. Any score that was an extreme outlier (i.e., beyond 3 SDs from the mean) was removed from the analysis. One participant's score for PCL (interview observation task) was 3.22 SDs from the mean score. For the 'number of details recalled' measure, one participant's score was 3.46 SDs from the mean and for the 'free recall accuracy' measure, two participants' data were

3.26 and 3.39 SDs from the mean. These four data points were excluded from the current analyses. (Results for the descriptive statistics pre-removal of data are reported in Appendix D.5).

Working Memory Capacity (WMC) Scores

Each participant's scores on their OSPAN, RSPAN and SSPAN tests were added to give their WMC score (Conway et al., 2005). A one-way between groups ANOVA was conducted to examine whether there were any differences between the conditions for WMC. There was no significant difference in WMC scores between the five conditions, $F(4, 125) = 1.24$, $p = .297$, $\eta^2_p = .04$ (see Table 5.1).

Table 5. 1

Mean and SD scores for WMC for each condition.

Note taking condition	*N*	*Mean*	*SD*
FNT (Access to notes)	26	37.46	7.83
SNT (Access to notes)	26	34.77	9.09
FNT (No access to notes)	26	35.62	6.99
SNT (No access to notes)	26	32.77	7.77
NNT	26	34.88	7.15

Hypothesis testing

Perceived Cognitive Load

A one-way between groups ANOVA was conducted to examine differences between note taking conditions (FNT [access to notes] vs. SNT [access to notes] vs FNT [no access to notes] vs SNT [no access to notes] vs NNT) for perceived cognitive load (PCL) during the interview observation and recall tasks. Contrary to Hypothesis 1, there

were no differences in PCL between the five conditions for the observation task, F (4, 124) = .68, p = .606, η^2_p = .02 (see Table 5.2). Further, contrary to Hypothesis 4, there were no differences in PCL between the five conditions for the recall task, F (4, 125) = .50, p = .738, η^2_p = .02 (see Table 5.2).

Table 5. 2

Mean, SD and CI scores for PCL for the interview observation task and recall tasks for each condition.

	PCL for 'interview observation' task		PCL for 'recall' task	
Condition	*M* (*SD*)	95% CI	*M* (*SD*)	95% CI
FNT (access)	58.09 (12.10)	[53.20, 62.98]	64.36 (9.86)	[60.38, 68.34]
SNT (access)	60.94 (13.87)	[55.33, 66.54]	60.58 (11.69)	[55.85, 65.30]
FNT (no access)	55.45 (14.62)	[49.54, 61.35]	63.71 (13.35)	[58.31, 69.10]
SNT (no access)	59.95 (11.40)	[55.35, 64.55]	63.09 (12.58)	[58.01, 68.17]
NNT	58.52 (12.29)	[53.45, 63.59]	64.56 (10.32)	[60.40, 68.73]

Free Recall

A one-way between groups ANOVA was conducted to examine differences between the note taking conditions (FNT [access to notes] vs. SNT [access to notes] vs FNT [no access to notes] vs SNT [no access to notes] vs NNT) for the total amount of details that were recalled by participants during their free recall. Contrary to the prediction, there were no differences between the five conditions, F (4, 121) = 1.43, p = .228, η^2_p = .05 (see Table 5.3). A one-way between groups ANOVA was conducted to examine differences between note taking conditions (FNT [access] vs. SNT [access] vs FNT [no access] vs SNT [no access] vs NNT) for the proportion of accurate details

recalled by participants during their free recall. There was a difference in free recall accuracy between conditions, F (4, 121) = 2.48, p = .047, η^2_p = .08 (see Table 5.3). Tukey post-hoc comparisons indicated that, as predicted, those in the SNT (access) condition were significantly more accurate than those in the NNT condition (p = .022, d = .86). No other post-hoc comparisons reached significance (for analyses, see Appendix D.5).

Table 5. 3

Mean, SD and CI scores for amount of free recall details and accuracy of free recall details for each condition.

	Amount of details		Accuracy of details	
Condition	*M* (*SD*)	95% CI	*M* (*SD*)	95% CI
FNT (access)	110.81 (27.23)	[97.96, 123.65]	.73 (.09)	[.70, .77]
SNT (access)	121.96 (26.88)	[110.97, 133.05]	.76 (.07)	[.73, .79]
FNT (no access)	111.88 (23.86)	[101.80, 121.95]	.72 (.07)	[.69, .75]
SNT (no access)	112.04 (28.52)	[100.27, 123.81]	.73 (.07)	[.70, .75]
NNT	103.42 (27.23)	[92.43, 114.42]	.70 (.08)	[.66, .73]

Cued recall

A one-way between groups ANOVA was conducted to examine differences between note taking conditions for accuracy of cued recall when answering 40 questions. There was a difference in cued recall accuracy between the five conditions, F (4, 124) = 4.26, p = .003, η^2_p = .12 (see Table 5.4). Tukey post-hoc comparisons indicated that, as predicted, those in the SNT (access) condition were significantly more accurate than those in the NNT condition (p = .001, d = 1.29) and those in the FNT (access) condition were more accurate than those in the NNT condition (p = .051, d = .61). No other pot-hoc analyses reached significance (see Appendix D.6).

Table 5. 4

Mean, SD and CI percentage accuracy scores for cued recall across the five conditions.

Condition	*M*(*SD*)	95% CI
FNT (access)	64.73 (15.13)	[58.61, 70.84]
SNT (access)	71.25 (10.54)	[66.90, 75.60]
FNT (no access)	66.50 (13.03)	[61.24, 71.76]
SNT (no access)	66.06 (14.60)	[60.15, 71.94]
NNT	56.25 (13.07)	[50.97, 61.53]

Access to Notes

A series of independent groups *t*-tests were conducted to examine differences between the access to notes conditions (i.e., access to notes and no access to notes) for each of the dependent variables (i.e., PCL 'interview task', PCL 'recall task', total amount of free recall details, accuracy of free recall detail*s* and accuracy of cued recall). Contrary to the predictions, there were no differences between conditions (access vs no access to notes) for any of the dependent variables (for t-tests results see Appendix D.5).

Moderation analysis

Moderated regression analyses were conducted to identify when, or for whom, effects between the predictor (i.e., the note taking conditions SNT vs FNT vs NNT) and outcome variables (i.e., PCL and recall performance) existed. Whether WMC moderated the effects of note taking on PCL during the interview observation task was examined. The moderating effects of WMC and access to notes on the effect of note taking on PCL during the recall task were also examined. Finally, the moderating effects of WMC and access to notes on the effect of note taking on recall performance were investigated.

These moderation analyses were exploratory and therefore no predictions were made regarding the direction of any effects.

Moderation analyses were performed using PROCESS, which, based on a regression path-analytic framework, is used to conduct moderation analyses in one conditional process model (Hayes, 2018). A series of linear regression models were estimated to explore any moderator effects. Separate coefficients for each regression equation were estimated and tested to examine the effect of the predictor variable on the outcome variables, when moderators were introduced. Note taking condition (SNT vs FNT vs NNT) was the predictor variable, which was dummy coded using the multi-categorical indicator selection in PROCESS. The outcome variables were PCL, the amount and accuracy of free recall details and the accuracy of cued recall responses. One of the moderator variables was participants' WMC score measured using three complex span tasks (OSPAN, RSPAN, SSPAN). To analyse regression interactions for continuous variables, the 'pick-a-point' approach is the most popular method to probe interactions (Bauer & Curran, 2005). To probe values of the WMC moderator that are within the bounds of the measurements, percentile points one SD above and one SD below the mean 50th percentile were selected. This 'pick-a-point' approach enables examination of the distribution of the moderator score (Hayes, 2018). For the current analysis, a high WMC score of 43 (at the 84th percentile point), a medium WMC score of 36 (at the 50th percentile point), and a low WMC score of 26 (at the 16th percentile point) were used. At each percentile point, the score is representative of, rather than participants' actual, scores at that point (Hayes, 2018). Access to notes was a second moderator (i.e., access and no access). An association between two variables, the predictor and outcome, are moderated when the size or sign of the effect depends on a third variable or set of variables, this can be difficult to interpret but visualisation of the

results can be useful (Hayes, 2018). Figures for each of the moderation analyses have, therefore, been included within the following sections.

Perceived cognitive load with WMC as moderator

WMC was a moderator of the effects of note taking (SNT, FNT and NNT) on PCL during the interview observation task was investigated (see Figure 5.1). For participants in the SNT condition there appeared to be higher PCL scores at the low WMC point when compared with PCL scores for the NNT and FNT conditions and no difference in PCL scores between conditions at the medium and high WMC points. However, moderation analysis indicated that this interaction was not significant, b = 5.05, t (129) = 1.06, p = .292, 95% CI [-4.40, 14.50], d = .19.

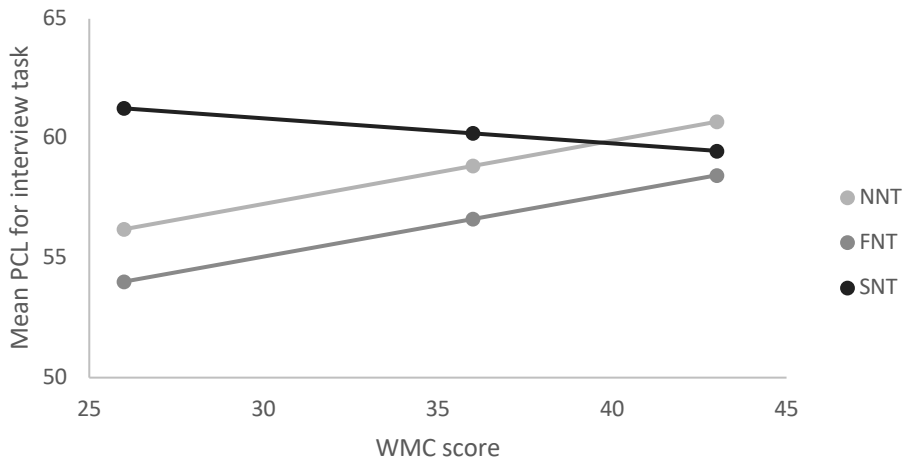


Figure 5. 1 Moderation analysis with note taking condition (SNT, FNT and NNT) as a predictor of PCL, for 'interview observation', with WMC as a moderator. WMC scores are at the 16th, 50th and 84th percentile points.

In the following sections of moderation analyses, firstly results are presented for the single WMC moderator. The second, moderation analysis presented in each section shows results when two moderators, WMC and access to notes, are included in the analysis.

Perceived cognitive load at recall with WMC as moderator

WMC as a moderator of the effects of note taking (SNT, FNT and NNT) on PCL during the recall tasks was investigated (see Figure 5.2). For participants in the SNT condition, the PCL scores appeared lower at the high WMC point, when compared with PCL scores for the NNT and FNT conditions at the high WMC point. There were no differences between conditions for PCL at the low or medium WMC points. Moderation analysis showed this interaction was not significant, b = -6.65, t (130) = -1.58, p = .117, 95% CI [-14.98, 1.68], d = .28.

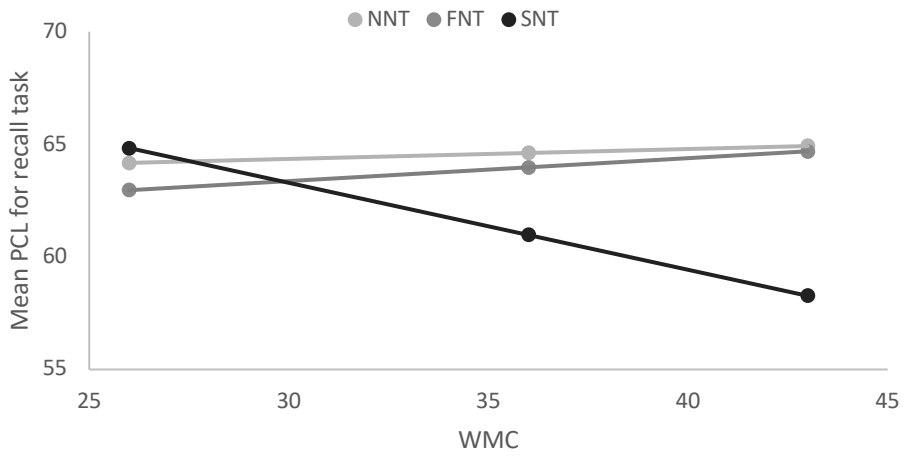


Figure 5. 2 Moderation analysis with note taking condition (SNT, FNT and NNT) as a predictor of PCL, for 'recall task', with WMC as a moderator. WMC scores are at the 16th, 50th and 84th percentile points.

Perceived cognitive load with WMC and access to notes as moderators

WMC as a moderator of the effects of note taking (SNT, FNT and NNT) on PCL during the recall task was investigated separately for when participants had access to notes and for when they did not (see Figures 5.3a and 5.3b). For participants who did not have access to their notes during recall, there appeared to be a difference in PCL for participants in the SNT condition compared with those in the FNT condition. Regression

analysis showed that the note taking condition predicted PCL, b = 32.78, *t* (104) = 2.02, *p* = .046, 95% CI [.59, 64.96], *d* = .40 and that WMC predicted PCL, *b* = 1.48, *t* (104) = 1.99, *p* = .050, 95% CI [.01, 2.96], *d* = .39. There was a significant moderation effect between note taking and WMC with no access to notes, *F* (1, 96) = 4.44, *p* = .038, η^2_p = .04, but there was no moderation effect when participants had access to notes *F* (1, 96) = .04, *p* = .839, η^2_p <.01.

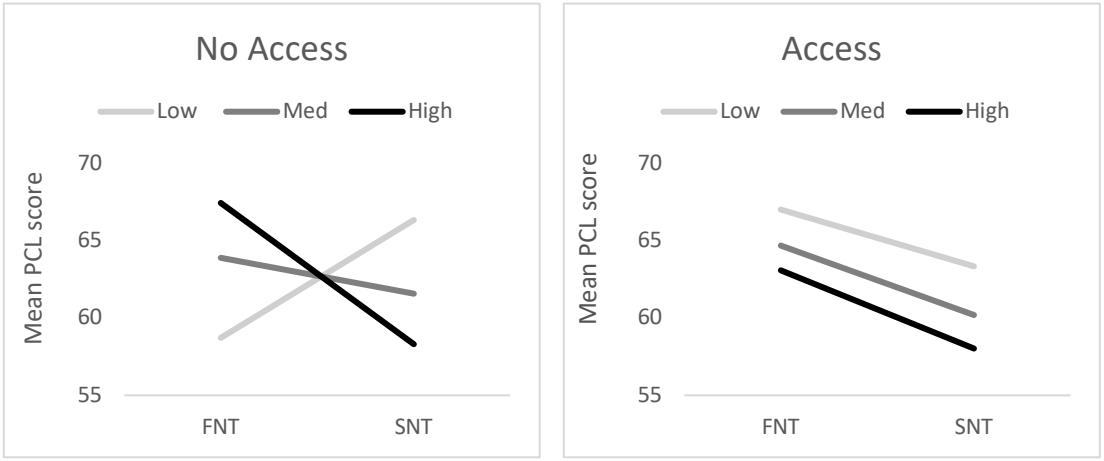


Figure 5. 3 5a (No access to notes) and 5.3b (Access to notes). Moderation analysis with note taking condition (SNT, FNT and NNT) as a predictor of PCL, for 'recall task', with WMC and access to notes as moderators. WMC scores are at the 16th (low), 50th (medium) and 84th (high) percentile points.

Amount of details recalled with WMC as moderator

WMC as a moderator of the effects of note taking (SNT, FNT and NNT) on the amount of details recalled during the free recall task (see Figure 5.4) was investigated. WMC was not a significant predictor of amount of details recalled, *b* = .33, *t* (126) = .42 *p* = .674, 95% CI [-1.20, 1.86], *d* = .07. Examination of the conditional effects of note taking at different WMC points showed a significant effect between SNT and NNT conditions at the medium WMC point, *t* (126) = 2.15, *p* = .034, 95% CI [1.13, 28.13], *d* = .40. No other conditional effects or interactions were significant (*p*-values ranged from .096 to .816).

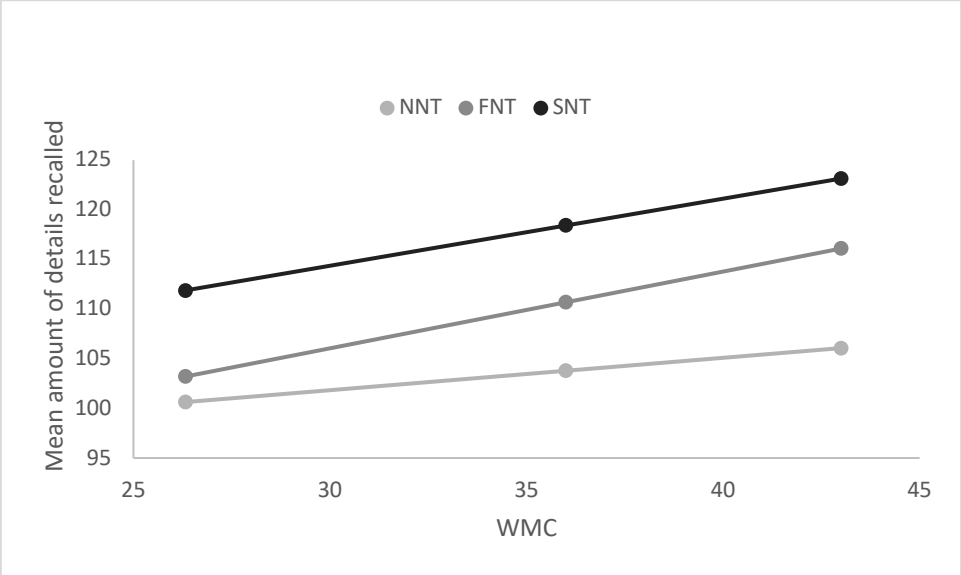


Figure 5. 4 Moderation analysis with note taking condition (SNT, FNT and NNT) as a predictor of the amount of free recall details, with WMC as a moderator. WMC scores are at the 16th, 50th and 84th percentile points.

Amount of details recalled with WMC and access to notes as moderators

WMC was also examined as a moderator of the effects of note taking on the amount of details recalled separately for those with access to notes and for those with no access to notes (see Figures 5.5a and 5.5b). There appeared to be a difference in the amount of details recalled in the SNT condition when compared with the FNT condition with access to notes. Analysis showed the moderation effect was not significant, F (1, 92) = .96, p = .329, η^2_p = .01. Examination of the conditional effects of note taking at different WMC points were not significant, at the low WMC point with access to notes, b =22.48, t (100) = 1.78, p = .079, 95% CI [-2.66, 47.62], d = .36 and at the medium WMC point with access to notes, b = 13.21, t (100) = 1.68, p = .097, 95% CI [-2.44, 28.87], d = .34. For no access to notes, the moderation effect was also not significant, F (1, 92) = .59, p = .446, η^2_p = .01.

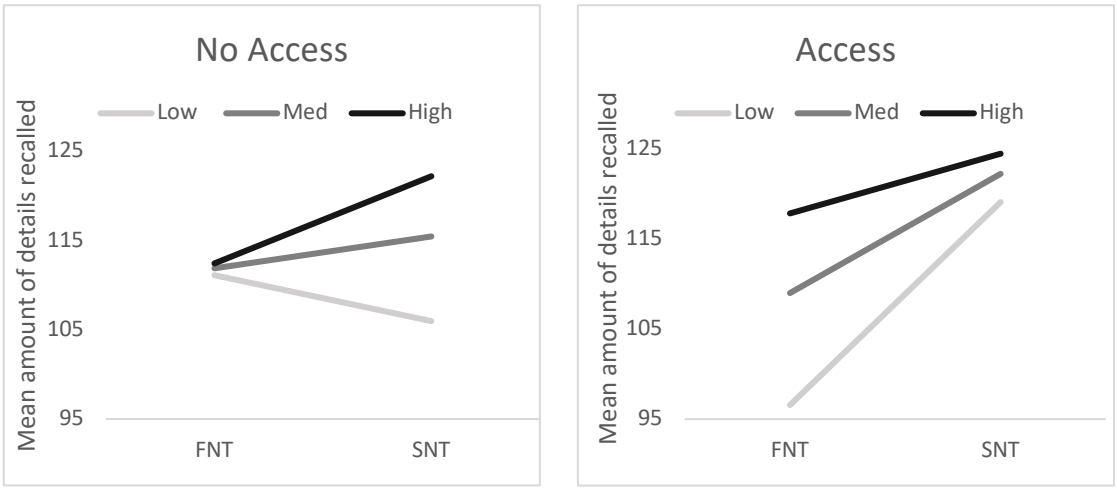


Figure 5. 5a (No access to notes) and 5.5b (Access to notes). Moderation analysis with note taking condition (SNT, FNT and NNT) as a predictor of the amount of free recall details, with WMC and access to notes as moderators. WMC scores are at the 16th (low), 50th (medium) and 84th (high) percentile points.

Accuracy of details recalled with WMC as moderator

A moderation analysis was conducted for accuracy of free recall of information as an effect of the note taking condition, with WMC as the moderator (see Figure 5.6). There appeared to be a difference in the accuracy of details recalled in the SNT condition when compared with the NNT condition. Analysis showed that WMC was not a significant predictor of the accuracy of details recalled, b = .01, t (126) = .53 p = .598, 95% CI [-.003, .005], d = .09. Examination of the conditional effects of note taking at different WMC points showed a significant effect between SNT and NNT conditions at the medium WMC point, t (126) = 2.52, p = .013, 95% CI [.01, .08], d = .45. No other conditional effects or interactions were significant (p-values ranged from .072 to .406).

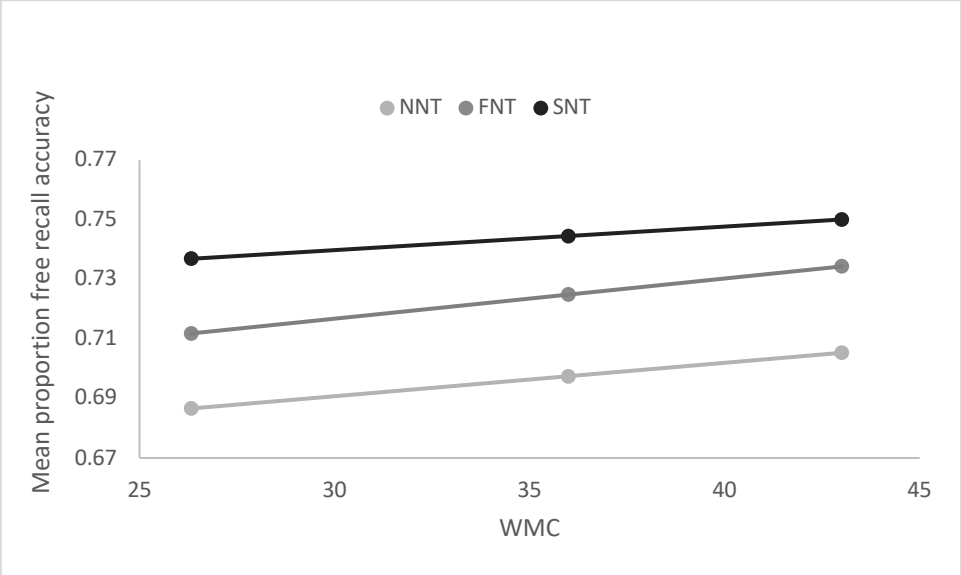


Figure 5. 6. Moderation analysis with note taking condition (SNT, FNT and NNT) as a predictor of the accuracy of free recall details, with WMC as a moderator. WMC scores are at the 16th, 50th and 84th percentile points.

Accuracy of details recalled with WMC and access to notes as moderators

WMC was examined as a moderator of the effects of note taking on accuracy of free recall details separately for those with access to notes and those who did not have access to notes (see Figures 5.7a and 5.7b). For no access to notes, there appeared to be no difference in accuracy of details recalled in the SNT condition, but there appeared to be a difference in the FNT condition, however, analysis showed the moderation effect was not significant, F (1, 92) = .41, p = .523, η^2_p <.01. There were no further significant moderation effects (p-values ranged from .183 to .856).

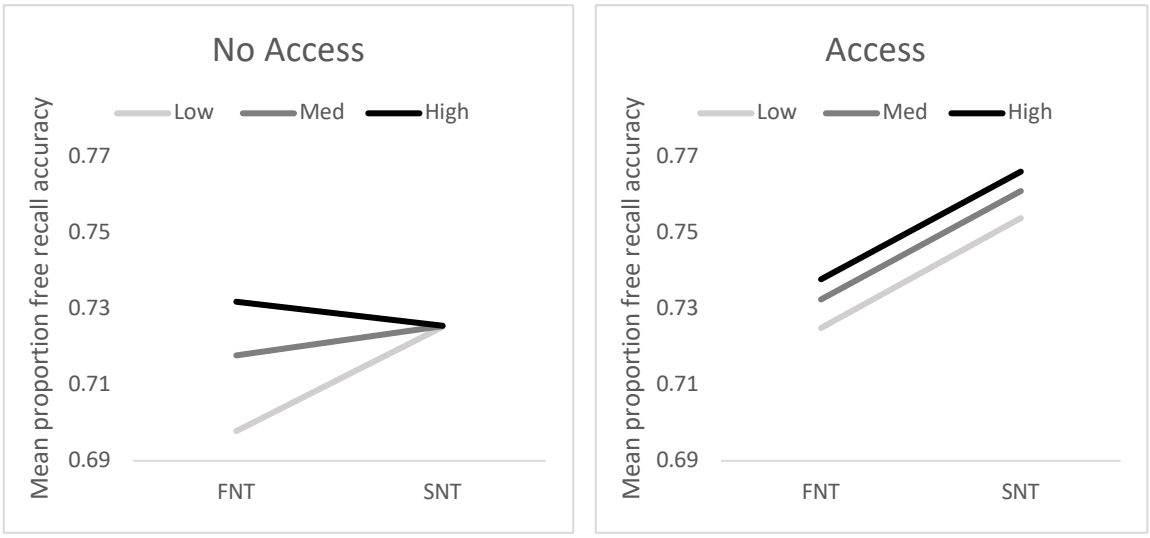


Figure 5. 7. 7a (No access to notes) and 7b (Access to notes). Moderation analysis with note taking condition (SNT, FNT and NNT) as a predictor of the accuracy of free recall details, with WMC and access to notes as moderators. WMC scores are at the 16th (low), 50th (medium) and 84th (high) percentile points.

Accuracy of cued recall with WMC as moderator

A moderation analysis was also conducted for accuracy of cued recall responses (see Figure 5.8). There appeared to be a difference in accuracy of cued recall responses in the FNT condition when compared with the NNT conditions at the high WMC point, and no difference between FNT and NNT at the low WMC point (Figure 5.8). Analysis showed this model was significant, F (5,123) = 5.37, p <.001, η^2_p = .18. The moderation effect between FNT and WMC was not significant, b = .79, t (129) = 1.81 p = .073, 95% CI [-.08, 1.66], d = .32.

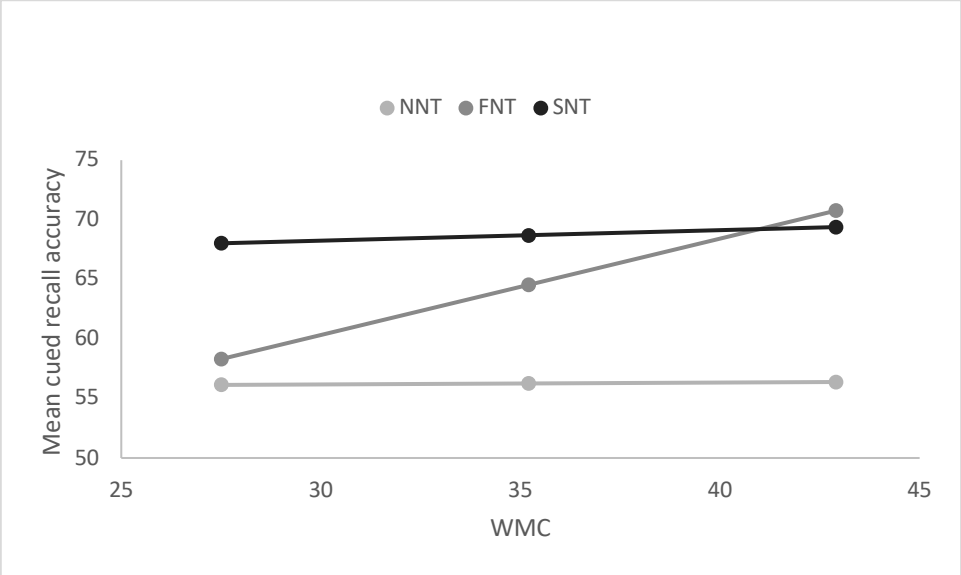


Figure 5. 8 Moderation analysis with note taking condition (SNT, FNT and NNT) as a predictor of the accuracy of cued recall responses, with WMC as a moderator. WMC scores are at the 16th, 50th and 84th percentile points.

Examination of the conditional effects of note taking at different WMC points showed significant effects between SNT and NNT conditions at the low, medium and high WMC points and between FNT and NNT at the medium and high WMC points (see Table 5.5).

Table 5. 5

Results of t-tests for conditional effects of moderation analyses.

Conditions	WMC point	*t* (129)	*p*	CI [LL, UL]	*d*
SNT vs NNT	Low	2.43	.017	2.19, 21.37	.43
	Medium	3.92	<.001	6.19, 18.81	.69
	High	2.74	.007	3.61, 22.42	.48
FNT vs NNT	Low	.19	.851	-9.28, 11.24	.03
	Medium	2.83	.006	2.68, 15.14	.50
	High	3.16	.002	4.40, 23.52	.56

Accuracy of cued recall with WMC and access to notes as moderators

Access to notes and WMC were examined together as moderators of the effects of note taking on accuracy of cued recall responses (see Figures 5.9a and 5.9b). There appeared to be no difference in accuracy of cued recall responses in the SNT condition when WMC and access to notes were included as moderators, but there was a difference in the FNT condition. Moderation analysis showed that the model was significant, F (7, 95) = 2.33, p = .031, η^2_p =.15. Note taking condition predicted cued recall accuracy, b = 36.50, t (103) = 2.06, p = .043, 95% CI [1.26, 71.75], d = .41, as did WMC, b = 1.89, t (103) = 2.31 p = .023, 95% CI [.26, 3.51], d = .46, and there was a moderation effect between note taking and WMC for cued recall accuracy scores, b = -1.06, t (103) = -2.08 p = .040, 95% CI [-2.06, -.05], d = .41.

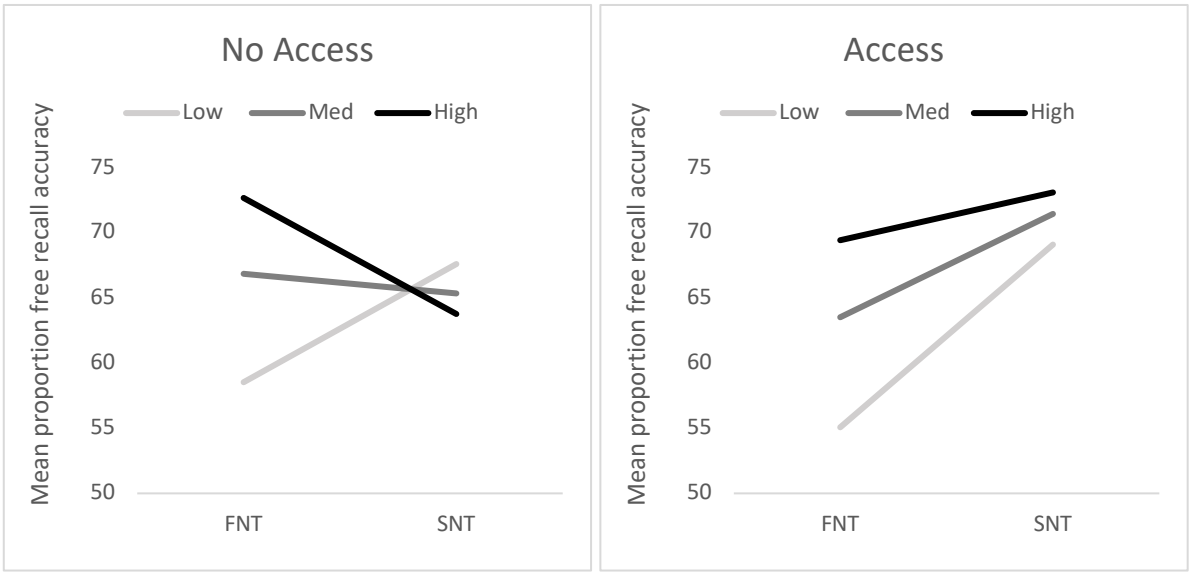


Figure 5. 9. 10a (No access to notes) and 5.9b (Access to notes). Moderation analysis with note taking condition (SNT, FNT and NNT) as a predictor of the accuracy of cued recall responses, with WMC and access to notes as moderators. WMC scores are at the 16th (low), 50th (medium) and 84th (high) percentile points.

The interaction between note taking and WMC with no access to notes was significant, F (1, 95) = 4.34, p = .040, η^2_p =.04, but for access to notes the interaction was not significant, F (1, 95) = 1.87, p = .175, η^2_p = .02. Examination of the conditional effects of note taking at different WMC points with access to notes showed significant differences between SNT and NNT conditions at the low WMC point, b = 14.00, t (103) = 2.36, p = .020, 95% CI [2.23, 25.77], d = .47. and the medium WMC point, b = 7.92, t (103) = 2.17, p = .03, 95% CI [.66, 15.18], d = .43. No other conditional effects reached significance (p-values ranged from .107 to .690).

Post experiment questionnaire

A series of Pearson's correlations were calculated to determine whether the dependent variables of motivation, confidence and task difficulty were correlated with each other. There were significant, but moderate, correlations between some of the variables (for Pearson's correlation results, see Appendix D.5). Therefore, the assumption of an absence of multicollinearity was met, and to reduce Type 1 error, a one-way between-groups MANOVA was conducted to investigate differences between the conditions (FNT [access to notes] vs. SNT [access to notes] vs FNT [no access to notes] vs SNT [no access to notes] vs NNT) for participants' self-reported motivation, confidence, and ratings of how difficult they found the tasks. The MANOVA indicated that there was a significant multivariate effect: Wilks' λ = .74, F (20, 402.26) = 1.88, p = .012, η^2_p = .07 (for MANOVA results across each of the dependent variables see Appendix D.5). The only significant difference at the univariate level was for confidence in remembering the witness's account, F (4, 125) = 5.95, p < .001, η^2_p = .16. Tukey post-hoc comparisons indicated that participants in the SNT (access) condition provided higher confidence ratings (M = 5.31, SD = .68) compared with those in the NNT (M = 4.19, SD = 1.02; p <.001, d = 1.31). Also, participants in the SNT (no access) condition

provided higher confidence ratings (M = 4.92, SD = .93) compared with those in the NNT (M = 4.19, SD = 1.02; p = .021, d = .76). Finally, participants in the FNT (no access) condition provided higher confidence ratings (M = 4.96, SD = .72) compared with those in the NNT (M = 4.19, SD = 1.02; p <.001, d = .89), where scores of 1 = not at all confident and 7 = extremely confident.

Discussion

The current research examined the effects of note taking on recall of a witness's account of a crime. Working Memory Capacity (WMC), and access to notes during recall, were examined as moderators, potentially influencing the amount and accuracy of details recalled. The current findings suggested that there were no significant effects of note taking condition on participants' perceived cognitive load (PCL) when observing a witness's interview nor when they were recalling the content of the interview. That is, when participants were required to listen to the witness, consider what was being said and think of questions to ask, making structured notes, free notes or not taking notes made no difference to the cognitive demands perceived by the participants. When participants' working memory capacity (WMC) and access to notes were considered as moderators, however, the note taking condition was a predictor of PCL scores. In the structured note taking (SNT) condition, participants who had no access to notes and low WMC, rated PCL higher than those in the free note taking (FNT) condition. Participants, with no access to notes and high WMC, rated lower PCL in the SNT condition than those in the FNT condition.

The effects of note taking on participants' recall of information provided by a witness, in terms of the amount and accuracy of free recall details and the accuracy of cued recall responses were also examined. For the total amount of details provided during free recall, contrary to the predictions, there were no significant differences

across the five note taking conditions. Further, when participants had access to their notes at recall, for participants in the FNT and SNT conditions, there were no significant differences in the amount of details provided in free recall. When WMC and access to notes were examined as moderators, participants in the SNT condition, with low and medium WMC, recalled a greater number of details during free narrative, than those in the FNT condition but there was no difference in the number of recalled details between conditions for high WMC. In terms of the accuracy of free recall, participants were more accurate in the SNT condition than in the NNT condition. WMC was a significant predictor of the difference in free recall accuracy between the SNT and FNT conditions but only for medium WMC.

For the accuracy of cued recall responses, participants in the SNT and FNT conditions were more accurate than those in the NNT condition, but there was no difference in cued accuracy between those in the SNT and FNT conditions. When WMC was considered as a moderator of the effects of note taking condition on cued recall accuracy, participants in the FNT and SNT conditions, with medium and high WMC, were more accurate than those in the NNT condition. For participants in the SNT condition with low WMC, scores were more accurate than for those in the NNT condition. There was no difference in accuracy of cued recall responses as a function of whether or not participants were given access to their notes. However, the moderation analysis showed that the note taking condition and WMC scores predicted cued recall accuracy and there was a moderation effect. WMC was a significant moderator when participants did not have access to their notes, but the effect was not significant when participants had access to their notes.

In terms of the PCL findings, the lack of differences between the note taking and no note taking conditions was interesting. Previous studies suggest that note taking can

be cognitively demanding (Piolat et al., 2005). In addition, applying a new method of note taking (such as in the structured note taking condition for the current study) has previously been found to challenge participants' cognitive resources (MacDonald, 2016). However, an increased perceived cognitive load for participants was not observed. The current findings did reveal an influence of WMC as a moderator of PCL. So, although overall there was no increase in perceived cognitive load, there may be individual differences which impact PCL. This finding may contribute to understanding why some interviewers express that they find note taking during an investigative interview to be demanding while others note that, for them, note taking is not demanding at all (e.g., see Chapter 2). One interesting possibility is that interviewers with a higher WMC are better able to manage the cognitive demands of interviewing and note taking. Future research should explore whether interviewers in practical settings perceive a high cognitive load when they take notes. Examining any impact of note taking on their interview performance would also be interesting.

For cued recall, accuracy was higher in both note taking conditions than it was in the no note taking condition. This is consistent with previous research which showed that using notes to generate police reports assisted officers' recall of additional details (Gregory et al., 2011). Structured note taking and access to notes has also been shown to improve juror members' recall of information (Thorley et at, 2016). When taking structured notes, and with access to the notes, the current results showed there was increased recall accuracy for those with higher WMC. With structured note taking and no access to notes free recall accuracy also increased for those with lower WMC. It may be that, when taking structured notes, participants record more specific details as they are not making long hand notes (Flanigan & Titsworth, 2020). The use of headings for structured note taking, rather than freestyle note taking, may also increase the

processing of information and allow internal connections between concepts, which may assist recall (Hope et al., 2014; Mayer, 1984). Performance may also be increased through a deeper understanding and integration of the information during note taking (Makany et al., 2009). MacDonald (2016) advises that interviewers should ensure they have access to their notes when their memory for a witness's account is being relied upon, but the current results suggest this may not always be necessary.

Participants with higher WMC achieved higher scores for free recall and cued recall responses, than those with medium or lower WMC. Individual differences in WMC, as measured by the complex span tasks, are due to variation in people's ability to maintain and process information in WM (Unsworth et al., 2013). Note taking can impose cognitive demands on the limited resources of working memory (Piolat et al., 2005). The current study, showed that with lower WMC, recall accuracy was lower than with higher WMC (i.e., for those with lower WMC information could not be actively maintained in WM). When participants' attention was divided between recording interview information in notes, formulating questions for the witness, and listening to the witness's responses, limited resources impacted upon the processing of information (Kolk et al., 2002). It may be that those individuals with lower WMCs were poorer at maintaining items in WM and were poorer at using cues to guide the search process of their memory (Unsworth et al., 2013). During cued recall, externally presented and internally generated cues have to be combined to focus the search for relevant items; it could be that individuals with low WMC were poorer at retrieving relevant information because of a poorer discrimination process at retrieval (Unsworth & Engle, 2007). Thus, for some people, forensically relevant details may be missed when taking notes (Lamb et al, 2000). This has important practical implications as 'missed' details may reduce the

topics and questions covered by an interviewer later in an interview or during a subsequent interview (Gregory et al., 2011).

The current sample comprised lay people who were not trained investigative interviewers. As such, the findings may not be generalizable to trained interviewers who use note taking during the course of their interviews. Interviewers in real-world settings often have more cognitive demands to deal with than did the current participants. Additional factors that may have an influence on cognitive load for interviewers, before and during an interview, include the age, communication skills, cognitive abilities, and vocabulary of an interviewee (Zajac & Brown, 2018). A witness's emotional state, relationship with the alleged perpetrator, their overall sexual knowledge and experiences (if the case is related to sexual offences), and any significant recent stressors and life events (e.g., bereavement or domestic violence) may also impact the interviewer (MoJ, 2011). Nonetheless, the current laboratory results contribute to understanding the effects of note taking, and WMC, on individuals' recall of information provided during an interview task. The findings may be beneficial for investigative interviewers, as the potential variation in cognitive demands of interviewing, and taking notes, may have an impact on interviewers' performance. Training and practice for interviewers in both interviewing and note taking techniques is vital to reduce their cognitive burden during an interview.

Only immediate recall was examined in the current study. The effect of taking notes may be stronger if there was a delay between an investigative interview (and note taking) and recalling the information provided by an interviewee. A delay before interviewers' recall may occur if, for example, they are preparing for a second interview later in the day or the following day with the same witness or another witness involved

in an investigation. Future research should be conducted with practitioners and should include recall of information over a longer period of time.

Investigative interviewing is a cognitively demanding task that can lead to a cognitive load (see Chapter 3). In the field, investigators are often required to deal with complex crimes and note taking is an essential element of interviewing witnesses. During an interview the effects of taking notes on PCL may be modest and moderated by individual WMC limits but note taking can impact on recall of information. To reduce errors in their recall of information gleaned from witnesses, interviewers should carefully attend to the information provided when taking notes. The general finding that improved recall using structured notes was not moderated by WMC indicated that structured note taking may be beneficial for all interviewers, regardless of their WMC.

Chapter 6:

Interviewers' perceptions of cognitive demands for investigative interviews with suspects and witnesses

Abstract

The current research aimed to explore police officers' perceived cognitive load for different types of interview (i.e., interviews with victims versus suspects and for serious and less serious crimes). In England and Wales, interviewers are trained to conduct interviews using the PEACE model and the ABE guidelines (College of Policing, 2019). During previous research, a small sample of interviewers described their belief that ABE interviews with vulnerable witnesses were more cognitively demanding than conducting other types of interviews, such as PEACE interviews with suspects (Hanway & Akehurst, 2018). Interviewers also explained that the complexity and seriousness of a case (i.e., the level of the offence being investigated) influenced cognitive demands during interviewing (see Chapter 2).

An online survey utilising a within-subjects design was conducted. Differences between police officers' reported experiences when they conduct four types of interview (i.e., interviews with victims of serious or less serious crimes using the ABE guidelines and interviews with persons suspected of serious or less serious crimes using the PEACE model) were examined. Interviewers reported increased perceived mental effort when conducting ABE interviews, than when conducting PEACE interviews. Interviewers also reported increased perceived mental demand when interviewing for serious offences, than when interviewing for less serious offences. It is suggested that these factors may contribute to cognitive load. The increased demands may be due to, for example, the sensitive nature of the interview (e.g., when interviewing for crime of rape, rather than theft).

Introduction

Investigation is a core duty of policing. Interviews conducted with victims, witnesses and suspects are central to the success of the investigation process (National Policing Improvement Agency [NPIA], 2009). However, investigative interviews are a cognitively demanding task (Fisher et al., 2010). The cognitive demands of interviewing have been found to increase perceived cognitive load when completing an interview observation task (i.e., actively listening to an interviewee, remembering the information provided and thinking of questions to ask) and are associated with a reduction in the accuracy of recall of an interviewee's account (Chapter 3). However, little research has examined cognitive load as perceived by investigative interviewers themselves. The aim of the current research was to explore, via an online survey, police officers' perceived cognitive load for different types of interview (i.e., interviews with victims versus suspects) and for serious and less serious crimes.

Cognitive load when interviewing

Cognitive load refers to the mental workload placed on individuals when they are required to undertake activities (Hart & Staveland, 1988; Van Acker et al., 2018). It signifies working memory use and the demands placed on cognitive resources when carrying out multiple and competing tasks (Dias et al., 2018; Engström et al., 2013). Cognitive load, and the cognitive demands of various tasks, has been examined in a variety of occupational settings (e.g., medicine; Dias et al. [2018] and air traffic control; Bittner et al. [1989]). The NASA Task Load Index (NASA-TLX) was developed as a measure of workload and identifies factors which contribute to cognitive load when completing a variety of tasks (Hart & Staveland, 1988). Six dimensions for the NASA-TLX were selected after extensive analysis of the primary factors that do (and do not) define the subjective experience of workload for a variety of activities ranging from

simple laboratory tasks to flying an aircraft (Hart, 2006). The six dimensions, which represent independent clusters of variables, are mental demand, physical demand, temporal demand, frustration, performance, and mental effort. A combination of these dimensions represents the workload experienced by most people performing most tasks (Hart & Staveland, 1988). The NASA-TLX measure has been used to examine perceived cognitive load when completing investigative interviewing tasks (e.g., Chapter 2).

During an interview, interviewers are required to actively listen to interviewees, accurately remember what they are saying, take notes, and formulate hypotheses to account for the events described (Fisher et al., 2014). Completing these cognitive tasks can be cognitively demanding and relies on interviewer's available cognitive capacity, but we all have a limited working memory capacity to perform cognitive tasks (Baddeley & Logie, 1999; Kahneman, 1973). As such, if the cognitive demands required to complete an interview exceed an individual's cognitive capacity then their performance during an interview may be impaired (see Chapter 3).

Cognitive load when completing various tasks, including investigative interviewing, can be explored through consideration of three categories of load, (i.e., intrinsic, extraneous and germane load; Galy et al., 2018). Intrinsic load relates to the load imposed by the nature of the information being processed and the natural complexity of the task (Schnotz & Kurschner, 2007). Extraneous load is induced by external factors, such as time pressure (Galy et al., 2012) and germane load is required for learning, the development of skills, and the application of skills in a novel situation (Paas, et al., 2004). Cognitive load is cumulative, and these three categories of load contribute to overall load.

Features of intrinsic load during investigative interviewing may include the type, seriousness, and complexity of the crime being investigated, whether notes are taken or not, and the age or communication abilities of a witness (see Chapter 2). A lack of time to plan and prepare for an interview, and time pressure experienced during an interview, may impact extraneous load (Hanway & Akehurst, 2018). Finally, germane load relates to the application of interviewers' training and knowledge in novel interviewing situations, which may impact on the cognitive demands of interviewing (see Chapter 2). The current research examined the effects of cognitive load when conducting investigative interviews with victims and witnesses.

Conducting investigative interviews

To obtain the most accurate and reliable evidence when conducting an investigative interview, various interviewing techniques and frameworks have been developed. For example, the Cognitive Interview (CI; Fisher et al., 1987) and the PEACE model for interviewing witnesses and suspects (Bull & Soukara, 2009; Kassin et al., 2010; Milne & Bull, 1999) have been advocated by researchers and practitioners. For interviewing children and other vulnerable witnesses and victims, prominent approaches include the Achieving Best Evidence guidelines (ABE; Ministry of Justice [MoJ], 2011), and the National Institute for Child Health and Human Development protocol (NICHD; Orbach et al., 2000). In many countries, training in the use of these techniques and other interview frameworks has been provided (for reviews of the training, see Fisher & Schreiber, 2007; Lamb et al., 2018, Clarke et al., 2011).

Despite interviewers receiving training, and having knowledge of best practice techniques, adherence to guidelines remains a challenge for investigative interviewers (Criminal Justice Joint Inspectorate, 2014; Powell & Barnett, 2015; Schreiber-Compo et al., 2012). Lamb (2016) suggests that the knowledge obtained through research-based

best practices is not being transferred to interviewing in practice, that is, interviewers are not converting their training to good quality interviews. Extensive training and case follow-up, including access to field interviews for evaluation, have been recommended to improve interviewer performance (Powell & Barnett, 2015). However, it may be that the cognitive demands of conducting an investigative interview, while adhering to guidelines, impact on interviewers' performance (Hanway & Akehurst, 2018). For example, a lack of preparation can affect interviewers' compliance with best practices (Brandon, 2018), but a lack of time for planning and preparation has been identified by police officers as a regular problem (see Chapter 2).

In England and Wales, interviewers undertake comprehensive training to conduct interviews using the PEACE model and the ABE guidelines (College of Policing [CoP], 2019). The PEACE model is an acronym for Planning and preparation; Engage and explain; Account; Closure; Evaluation (Milne & Bull, 1999). ABE guidelines follow a similar format to the PEACE model, and both include the development of rapport, asking open-ended questions and probing through questioning as best practice. In England and Wales, recordings are made of all PEACE interviews with suspects, but taking notes through a structured note taking process is also recommended for interviewers (CoP, 2019). ABE interviews are also recorded, and the guidance suggests that interviewers should consider taking brief notes to assist them during the free narrative phase of interviews (MoJ, 2011).

There are, then, similarities in recommended practice for conducting interviews using the ABE guidelines and PEACE model. For example, they use a similar format, they include the same recommended questioning techniques, and similar note taking recommendations are included in guidelines for their use. That said, there are also differences across the two interview types. It is likely that children, more so than

adults, are interviewed using the ABE guidance. Further, it is likely that adults, more so than children, are interviewed as suspects using the PEACE model. Furthermore, common practice dictates that PEACE interviews are often conducted by two police officers and ABE interviews with only one interviewing officer. In previous studies, interviewers described a perception that ABE interviews are more cognitively demanding than conducting other types of interviews, such as PEACE interviews (Hanway & Akehurst, 2018). Interviewers have also explained that the complexity and seriousness of a case (i.e., the level of the offence) increased cognitive demands when interviewing regardless of the status of the interviewee (see Chapter 2). Yet, no previous research has explored the similarities, or differences, in perceived cognitive demand when conducting these different types of interview nor the similarities and differences inherent in each type of interview that might explain these differences.

The current study

The current study, an online survey, examined investigative interviewers' perceived cognitive load when they conduct interviews in police operational settings with suspects (using the PEACE model) and victims (using the ABE guidelines). Specifically, the aim was to investigate how interviewers perceive the cognitive demands of interviewing depending on the type of interview conducted (i.e., interviews with suspects using the PEACE model versus ABE interviews with victims), and the gravity of the offence being investigated (i.e., serious crimes versus less serious crimes). Based on the previous findings outlined in this thesis, it was predicted that respondents would rate more highly the factors that contribute to cognitive load (i.e., mental demand, temporal demand, frustration, performance and mental effort; Hart & Staveland, 1988), when conducting an ABE interview than when conducting a PEACE interview (Hypothesis 1), and they would perceive higher cognitive load for conducting

interviews for serious crimes (e.g., rape or wounding) when compared with less serious crimes (e.g., theft or robbery; Hypothesis 2). Whilst it is possible to speculate regarding the reasons for potential differences in perceived cognitive load across interview types, another aim of this study was to collect data regarding similarities and differences between interview types in terms of the age of interviewees, how many police officers are involved, the prevalence of note taking, and how much planning is generally achieved, to assess whether these factors may influence perceived cognitive load.

Method

Design

The current study utilised an online survey and within-subjects design. All potential respondents were asked if, in the six months prior to the study, they had conducted an ABE interview with a victim of a serious crime and a victim of a less serious crime. They were also asked if, in the same time period, they had conducted a PEACE interview with the suspect of a serious crime and the suspect of a less serious crime. Examples of serious crimes (child sexual abuse, rape, wounding) and less serious crimes (robbery, theft, burglary) were provided to respondents for clarification. Participants responded, 'yes' or 'no', to these four scenarios questions. If potential respondents answered 'no' to all four scenario questions (i.e., they had not completed any of the types of interview in the previous six months), they were thanked for their time and were asked no further questions. If respondents answered 'yes' to any of the scenario questions, they were asked further questions (see below) about each interview type they had conducted.

Hence, there were two independent variables, each with two levels; Interview Type ('ABE' interview vs. 'PEACE' interview); and Offence Gravity (serious crime vs. less serious crime). The interview types were selected as, in England and Wales, these are

audio (PEACE) and video (ABE) recorded, and the recordings are used as evidence in court proceedings. Therefore, the importance of these types of interview, evidentially, is comparable. The gravity of the crime (i.e., serious and less serous), were included as there may be differences in how these crimes impact the perceived cognitive demands for the interviewer.

The dependent variables were measures of cognitive load (i.e., mental demand, temporal demand, frustration, performance and mental effort) for interviewers' 'experience of cognitive load' when interviewing. To understand the context of the interviews, questions about 'features of the interview' (i.e., the age of the interviewee, if another interviewer was present during the interview, if they took notes, and if they had sufficient time to plan for the interview) were also asked.

Participants

Power analysis for factorial designs is often a challenge for designs with several within-subject factors and current software packages may not accurately calculate power analyses (Lakens & Caldwell, 2019). However, power to detect an effect is much higher when using a within-subjects design than it is for between-subjects designs, given the same sample size (Thompson & Campbell, 2004). A priori G*power analysis for a repeated measures ANOVA with one group and four measurements indicated that based on a power of alpha = 0.80, a large effect size of $f = 0.23$ (calculated from a η^2_p of .05), the traditional alpha = .05, and correlation among the repeated measures of 0.5, a sample size of 28 would be required.

75 respondents completed the survey. The sample comprised 40 females, 26 males and 9 respondents who did not answer questions about gender, age etc. Of the 66 responses to demographic questions, 11 (16.7%) respondents were aged from 25 to 34 years, 32 (48.5%) were aged from 35 to 44 years, 18 (27.3%) were aged 45 to 54 years,

and 5 (7.6.%) were aged 55 to 64 years. The length of police service of the respondents ranged from 2 to 42 years (*M* = 16.17 years, *SD* = 7.66 years). The number of interview courses respondents had attended ranged from 1 to 10 courses (*M* = 3.48 courses, *SD* = 1.78 courses). Respondents were asked how many interviews they had conducted in the six months prior to completing the survey. For ABE interviews, 53 (70.7%) reported that they had conducted 1-10 interviews and 13 (17.3%) had conducted 11-20 interviews. For the number of PEACE interviews, 37 (49.3%) reported that they had conducted 1-10 in the six months prior to the study, 16 (21.3%) had conducted 11-20 interviews and 12 (17.3%) had conducted 21 or more interviews.

Materials

Experience of cognitive load measures

The six dimensions of workload that contribute to cognitive load are mental demand, physical demand, temporal demand, frustration, performance, and mental effort. A combination of these dimensions represents the workload experienced by most people performing most tasks (Hart & Staveland, 1988). For the current study, one question was asked for each of the dimensions 'mental demand', 'temporal demand', 'frustration', and 'mental effort'[6]. In addition, four questions were included for the dimension 'performance'. The NASA-TLX questionnaire includes, in its description of performance, 'success' and 'satisfaction'. We split these two aspects of performance as respondents in the current study may perceive that they were 'successful' in achieving the aims of an interview but, at the same time, they may not be 'satisfied' with their performance. In addition, interviewers' performance in terms of 'compliance with guidance' (i.e., they followed the ABE guidelines or PEACE model) and their

[6] Physical demand was the only dimension of the NASA-TLX not included as it is not relevant for the investigative interview setting.

performance in achieving the 'aims of the interview' (i.e., to obtain accurate and reliable information from an interviewee) were viewed as distinct. Therefore, we also split these two aspects of performance across different questions. In sum, the four questions relating to performance enabled an examination of differences in respondents' perceptions of their 'success' and 'satisfaction' in 'complying with the guidance' and regarding 'achieving the aims of the interview'. In total, eight questions were asked, which represented dimensions of the NASA-TLX measures of cognitive load (see Table 6.1). Respondents were asked to move a slider between 'Low' and 'High' or 'Good' and 'Poor' (on a range from 0 to 100) to the point that represented their experience of each dimension for the interview they had in mind (see appendix E.1).

Table 6. 1

Measures of respondents' experiences of cognitive load.

Question		Sliding scale	
1	Mental demand. How much mental and perceptual activity was required (e.g. thinking, deciding, remembering, etc.), i.e. was the interview easy, simple, straight-forward (low) or demanding, complex, exacting (high)?	Low 0	High 100
2	Temporal demand. How much time pressure did you feel due to the rate of pace at which the interview occurred, i.e. was the pace slow and leisurely (low) or rapid and frantic (high)?	Low 0	High 100
3	Frustration. How frustrated were you during the interview, i.e. were you content, satisfied and relaxed (low) or irritated, stressed and annoyed (high) during the interview?	Low 0	High 100
4	Performance (Success-Compliance). How successful do you think you were in complying with ABE guidance/PEACE model for this interview?	Poor 0	Good 100
5	Performance (Success-Aims). How successful do you think you were in accomplishing the aims for the interview, e.g. obtaining a full and accurate account from the interviewee?	Poor 0	Good 100
6	Performance (Satisfied-Compliance). How satisfied were you with your performance in complying with ABE guidance/PEACE model for this interview?	Poor 0	Good 100
7	Performance (Satisfied-Aims). How satisfied were you with your performance in accomplishing the aims of this interview, e.g. obtaining a full and accurate account from the interviewee?	Poor 0	Good 100
8	Mental effort. How hard did you have to work mentally to achieve your level of performance for this interview?	Low 0	High 100

Additional questions about the interviews

Respondents answered the following questions about features of each interview they selected to report on, i) *"what was the age of the interviewee?"* (adult or child); ii) *"was there another interviewer present?"* (yes or no) and iii) *"did you take notes during the interview?"* (yes or no). Interviewers were also asked to rate the extent to which they agreed with the following statement "*I felt I had sufficient time to plan and prepare"*, using a 5-point scale (i.e., Strongly agree [1]; Somewhat agree [2]; Neither agree nor disagree [3]; Somewhat disagree [4]; Strongly disagree [5]). For the full survey see Appendix E.1.

Procedure

Agreement for the study was obtained from six police forces in England, who were contacted by the first author. Once authority to contact participants had been obtained, participants were recruited through third parties (i.e., force gatekeepers). Each gatekeeper was asked to circulate a link to the online survey, by email, to interviewers who had received 'ABE' and 'PEACE' interview training. In the email to potential respondents, it was made clear that the gatekeeper was taking no role in the research beyond circulation of the weblink to the survey. The survey was administered through the online software Qualtrics. Written information about the study was provided and participants' consent obtained prior to their progression to the survey questions. To ensure that participants had recent experience of interviewing, a criterion that participants should have conducted each of their 'ABE' and/or 'PEACE' interviews within the past six months was included.

Respondents were asked if, in the six months prior to the study, they had conducted interviews, using the ABE guidance, with the victim of a serious and/or a less

serious crime. They were also asked if, in the same time period, using the PEACE model, they had conducted interviews with the suspect of a serious and/or a less serious crime. It was anticipated that participants would have experience of at least two of the four types of interview. To maximise the total number of responses, the number of interview types required to have been conducted was not specified. If potential respondents specified that that they had not completed any of the four types of interview in the previous six months, they were thanked for their time and were asked no further questions. For each type of interview respondents specified they had conducted in the six months prior to the study, respondents were asked to think carefully about the last interview they had conducted and to indicate the offence that was being investigated. They were asked to keep that interview in mind whilst answering the related 'experience of cognitive load' and 'features of the interview' questions. To minimise practice effects, the order of presentation for each type of interview, was counterbalanced across respondents.

Finally, all respondents were asked to indicate their age, gender, length of police service, and the region in the UK of their police force. They were also asked to indicate the number, and type, of interview training courses they had attended and the number of ABE and PEACE interviews they had conducted in the six months prior to completing the survey. The survey took around 25 minutes to complete and at the conclusion all respondents were debriefed and thanked for their time.

Results

Of the 75 respondents, 65 had completed, in the six months prior to the survey, an ABE interview with a victim of a serious crime (e.g., rape or wounding); 38 had conducted an ABE interview with a victim of a less serious crime (e.g., theft or robbery); 66 a PEACE interview with a suspect of a serious crime (e.g., rape or wounding), and 53

a PEACE interview with a suspect of a less serious crime (e.g., theft or robbery). 32 respondents had completed all four types of interview, 17 had completed three types of interview, 17 had competed two types of interview and 9 had only completed one of the types of interview under investigation.

It was predicted that respondents would rate more highly the factors that contribute to cognitive load when conducting an ABE interview than when conducting a PEACE interview (Hypothesis 1), and they would perceive higher cognitive load for conducting interviews for serious crimes (e.g., rape or wounding) when compared with less serious crimes (e.g., theft or robbery; Hypothesis 2). To test these hypotheses data from the 32 participants who had completed all four interview types were analysed.

Experience of cognitive load

Two-way repeated measures ANOVAs were conducted to compare the impact of each Interview Type ('ABE' interview versus 'PEACE' interview), and Offence Gravity (serious crime vs. less serious crime), on each of the cognitive load measures. For these analyses, participant responses were included for those who had responded to questions about all four interview types (n = 32). The analyses showed there were no Interview Type X Offence Gravity interaction effects for any of the experience of cognitive load questions (p-values ranged from p =.058 to p = .924). Results for the main effects are shown in Table 6.2.

Table 6. 2

Repeated measures ANOVA results for the main effects of Interview Type, and Offence Gravity, on experiences of cognitive load

Question (DV)	Factor	F (1,31)	p	$\eta^2{}_p$
Mental demand	Interview Type	.00	.970	< .001
	Offence Gravity	7.60	.010*	.20
Temporal demand	Interview Type	2.41	.131	.07
	Offence Gravity	.06	.811	.002
Frustration	Interview Type	1.07	.310	.03
	Offence Gravity	.15	.706	.01
Performance – success in complying with ABE/PEACE	Interview Type	2.85	.101	.08
	Offence Gravity	.56	.462	.02
Performance – success in accomplishing the aims of the interview	Interview Type	15.48	< .001**	.33
	Offence Gravity	3.03	.092	.09
Performance – satisfied in complying with ABE/PEACE	Interview Type	1.59	.217	.05
	Offence Gravity	.29	.597	.01
Performance – satisfied in accomplishing the aims of interview	Interview Type	10.94	.002*	.26
	Offence Gravity	2.67	.112	.08
Mental effort	Interview Type	6.00	.020*	.16
	Offence Gravity	7.41	.011	.19

*Note. ** significant at the p = 001 level; * significant at the p = .05 level.*

There was a significant main effect of Interview Type on respondents' ratings of performance for success in accomplishing the aims of the interview, respondents rated their success higher for conducting ABE interviews than for conducting PEACE interviews. There was a significant main effect of Interview Type on performance for

being satisfied with accomplishing the aims of the interview, respondents rated their satisfaction higher when conducting ABE interviews than PEACE interviews. There was also a main effect of Interview Type on mental effort, respondents rated mental effort higher when conducting ABE interviews than PEACE interviews. There was a significant effect of Offence Gravity on respondents' ratings of mental demand, respondents rated mental demand higher when interviewing for serious crimes than for less serious crimes. There was also a significant effect of Offence Gravity on mental effort, respondents rated mental effort higher when interviewing for serious crimes than for less serious crimes (see Table 6.3 for mean and SD scores).

Table 6. 3

Mean and SD for effects of Interview Type and Offence gravity on DVs (n = 32)

	ABE		**PEACE**	
	Serious	Less serious	Serious	Less serious
Question (DV)	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
Mental demand	75.03 (19.53)	65.56 (21.87)	75.41 (19.25)	65.41 (20.35)
Temporal demand	39.00 (24.49)	40.75 (25.85)	46.97 (25.74)	43.69 (22.49)
Frustration	31.53 (25.22)	31.81 (22.67)	33.81 (23.27)	36.38 (20.52)
Performance – success in complying with ABE/PEACE	84.91 (9.16)	84.31 (14.86)	80.75 (16.70)	79.03 (16.22)
Performance – success in accomplishing the aims of the interview	81.56 (18.99)	83.19 (13.96)	74.81 (22.91)	60.66 (29.02)
Performance – satisfied in complying with ABE/PEACE	82.69 (12.50)	82.87 (15.52)	80.16 (18.24)	77.84 (17.33)
Performance – satisfied in accomplishing the aims of interview	82.34 (15.70)	81.38 (14.36)	74.91 (22.22)	65.87 (25.75)
Mental effort	79.41 (15.88)	70.44 (20.30)	72.19 (18.50)	64.81 (21.61)

Additional questions about the interviews

A further aim of this study was to collect data regarding similarities and differences between interview types in terms of the age of interviewees, how many police officers are involved, the prevalence of note taking, and how much planning is generally achieved. For the following analyses data for each interview type were

included, i.e., ABE (serious) *n* = 65; ABE (less serious) *n* = 38; PEACE (serious) *n* = 66; and PEACE (less serious) *n* = 53.

For each interview type, the number and percentage of interviews that were conducted, i) with an adult or child, ii) when another interviewer was present or not, and iii) where the interviewer took notes during the interview or not, were calculated (see Table 6.4).

Table 6. 4

Responses to the age of interviewee, other interviewer present and note taking questions.

		ABE		PEACE	
		Serious (*n*= 65)	Less serious (*n*= 38)	Serious (*n*= 66)	Less serious (*n*= 53)
Question	Response	*N* (%)	*N* (%)	*N* (%)	*N* (%)
What was the age of interviewee?	Adult	37 (56.9)	21 (55.3)	63 (95.5)	44 (83.0)
	Child	28 (43.1)	17 (44.7)	3 (4.5)	9 (17.0)
Was there another interviewer present?	Yes	12 (18.5)	6 (15.8)	57 (86.4)	36 (67.9)
	No	53 (81.5)	32 (84.2)	9 (13.6)	17 (32.1)
Did you take notes during the interview?	Yes	40 (61.5)	25 (65.8)	40 (60.6)	36 (67.9)
	No	25 (38.5)	13 (34.2)	26 (39.4)	17 (32.1)

Cochran's Q tests were conducted to examine differences across the four interview types. For the age of the interviewee, Cochran's Q test showed that the proportions of adult or child interviewees were significantly different across the four conditions, $\chi 2$ (32) = 17.71, p <.001. When considering responses to whether another interviewer was present or not, Cochran's Q test showed that the proportions of yes or no responses was significantly different across the four interview conditions, $\chi 2$ (32) =

45.00, p <.001. When considering whether the interviewer took notes during the interview or not, the test showed that the proportions of yes or no responses were not different across the four conditions, $\chi 2$ (32) = 0.47, p =.927.

Post-hoc analysis to examine the direction of differences was conducted using McNemar tests. As multiple tests were completed, the Bonferroni-corrected alpha value to achieve statistical significance was calculated by dividing the alpha value of .05 by the three dependent variable measures (adjusted alpha = .017). There were more ABE interviews with a child (43.1%) than PEACE interviews with a child (4.5%) when examining the age of the witness for serious crimes, p <.001. For less serious crimes, there were more ABE interviews with a child (44.7%) than PEACE interviews with a child (17.0%), p <.001. There were fewer ABE interviews conducted with another interviewer present (18.5%) then PEACE interviews with another interviewer present (86.4%), for serious crimes, p <.001. Fewer ABE interviews were conducted with another person present (15.8%) than PEACE interviews with another person present (65.8%), for less serious crimes, p <.001.

A 5-point scale (i.e., Strongly agree [1]; Somewhat agree [2]; Neither agree nor disagree [3]; Somewhat disagree [4]; Strongly disagree [5]) was used for respondents to indicate to what extent they had been able to prepare for interviews. A two-way repeated measures ANOVA was conducted to compare opportunity to plan for the interview across Interview Type and Offence Gravity. The repeated measures ANOVA showed there was no significant difference across Interview Type, F (1,31) = 1.37, p = .250, η^2_p = .42 and no significant difference for Offence Gravity, F (1,31) = .67, p = .418, η^2_p = .02 in terms of ratings of opportunity to plan for interviews.

Discussion

The current research examined the effects of conducting different types of investigative interview on interviewers' perceived cognitive load. As predicted, the *mental demand* (e.g., thinking and remembering) of interviewing for serious crimes was rated higher than the mental demand of interviewing for less serious crimes. For *success, and satisfaction, in achieving the aims of an interview*, the interviewers' ratings were higher when they were thinking about an interview they had conducted when they used ABE guidelines than it was when they were thinking about interviewing using the PEACE model. For *mental effort* (i.e., the mental effort required to achieve their level of performance), interviewers' ratings when considering an ABE interview were higher than when considering a PEACE interview. Respondents also rated mental effort higher when they conducted interviews for serious crimes than for less serious crimes.

These finding are consistent with interviewers' experience that ABE interviewing is mentally demanding and can often lead to perceptions of cognitive load (Hanway & Akehurst, 2018). Investigative interviewing is a complex social and verbal interaction for the witness and interviewer (Oxburgh & Dando, 2011). Interviewers are required to actively listen to witnesses, remember what is being said, pay attention to witnesses' needs, make decisions about what questions to ask, and identify topics to pursue (Fisher et al., 2014; Hanway & Akehurst, 2018). It may be that aspects of an ABE interview, for example attending to a witness's needs (i.e., if they are a child), are more demanding than when interviewing a suspect (i.e., an adult) using the PEACE model.

The current results showed that conducting investigative interviews relating to serious crimes was perceived to be a more demanding cognitive process than conducting interviews about less serious crimes. Interviewing for a serious crime (e.g., rape) can be difficult as the interviewer may need to consider additional elements, for

example, the nature of the relationship between the victim and suspect or the degree to which the interviewee is able to describe the intimate nature of the events. These elements are less likely to be a consideration when interviewing for a less serious crime, such as theft. Thus, the gravity of the offence, could contribute to the cognitive demands of interviewing.

For the current study, interviewers were instructed to consider their performance in achieving the aims of an interview (i.e., obtaining accurate and reliable information) and their compliance with the ABE guidance or the PEACE model. Respondents rated their success, and satisfaction, in achieving the aims of the interview higher for their ABE interviews, than for their PEACE interviews. They also rated the mental effort required to complete ABE interviews higher than for PEACE interviews, for both serious and less serious crimes. If the cognitive demands of an interview exceed an interviewer's capacity to complete the multiple tasks required, then performance on the task may be reduced (Frieder et al., 2016; Hanway & Akehurst, 2018; Nordstrom, 1996). The current results suggest that, contrary to there being a reduction in perceived performance, respondents rated that they obtained accurate and reliable information during ABE interviews, and interviews for serious crimes, with greater success than for PEACE interviews, and less serious crimes. However, to achieve this level of performance additional mental effort was required. It is the combination of dimensions (i.e., mental demand, temporal demand, frustration, performance and mental effort) that represent the overall workload experienced by an individual (Hart & Staveland, 1988). Thus, poorer perceived performance may not result, if ratings were low for one or more of the other dimensions. For example, it may be that interviewers did not experience increased temporal demand or frustration during their interviews, so their performance was less impacted by cognitive load.

Interviewing is a multifaceted task where interviewers complete multiple cognitive processes (Lafontaine & Cyr, 2016). Four features (i.e., note taking, solo interviewing, age of the interviewee and planning time) of ABE and PEACE interviewing for serious and less serious crimes were also considered. Despite recommendations to take notes during ABE and PEACE interviews, interviewers reported taking notes in only around two thirds (i.e., from 61% to 68%) of all the interviews considered during this study. Previous research has revealed that some interviewers take notes during interviews to assist their recall of witnesses' accounts, but others do not as they find note taking distracting (see Chapter 2). The current research supports this finding that not all interviewers take notes. Thus, note taking (or not) may be more dependent on the preferences of individual interviewers rather than the context of interviews.

When a solo interviewer conducts interviews, attending to the cognitive demands of interviewing, without the support from a second interviewer, may increase cognitive load for ABE interviewing compared with that required for PEACE interviewing, when it is more likely that two interviewers are present. There are additional complexities and unique concerns arising when interviewing children due to their limited cognitive and social development (Brubacher et al., 2014). The purpose of the ABE guidelines is to provide guidance, additional to that for PEACE interviews, for interviewing of children and other vulnerable witnesses (CoP, 2019). The additional demands of interviewing children, without another officer being present, may increase cognitive demands when conducting ABE interviews.

In terms of planning and preparation for interviews, for the current research, interviewers' ratings showed they generally agreed that they were given sufficient time to plan and prepare for all types of interview. That is, the mean score for "*I felt I had sufficient time to plan and prepare"* was less than 2 (where 1 = strongly agree and 5 =

strongly disagree) for all conditions. Although, a lack of available time for planning can impact on an interviewer's preparations, and this could increase errors in judgement regarding an interviewee's account (Brandon, et al., 2018; Dando et al., 2008), the current findings suggest this was not present for the interviewers who took part in this research. Overall, the current findings relating to the features of different types of interview, do not indicate whether these features increased the perceived cognitive demands for interviewers, but it is suggested that the interview context should be considered when examining the effects of cognitive load for interviewers.

As the current study involved repeated measures, practice and fatigue effects were a concern. To reduce these effects, the order of presentation of the four different interview conditions was counterbalanced across respondents. There were benefits to this repeated measure design, for example, the effects of individual differences across participants, such as their level of experience and /or training, were minimized and thus the findings were likely due to the different interview types than individual differences. Across all conditions, the current sample size allowed for meaningful statistical analyses. There were a limited number of interviewers who completed questions for all four interview types. That said, the effects were relatively large, and the sample size was sufficient to detect the large effects.

In terms of the measures of performance, previous research found that, when interviewing children, interviewers' perception of their performance was measured though the amount of detail provided by the child (Wright & Powell, 2007). Respondents in the current study may have considered that witnesses in their ABE interviews provided more information, which they believed to be accurate and reliable, than suspects in their PEACE interviews, who they may have perceived as providing less information, which was less accurate and reliable. There were no differences between

the four interview types for interviewers' perceived compliance with ABE or PEACE guidelines. This may reflect the lack of a relationship between interviewers' perceived and their actual ability to comply with best practices (Warren et al., 1999; Wright et al, 2008). For the current study, the interviewers' performance in adherence to protocols was unknown. To understand the contribution perceived performance makes to interviewers' overall cognitive load, further research, using objective measures of interviewers' performance, such as measuring the number of details provided, should be considered.

In terms of generalisability, the current study focused on respondents who regularly undertake two types of interview that are recommended in England and Wales. Similar frameworks (e.g., NICHD) are used globally and have corresponding training, therefore, the current research is likely to be applicable in many countries with established protocols or frameworks for the effective interviewing of victims and suspects. However, further research in different jurisdictions, would be of benefit to improve knowledge and understanding of interviewers' perceptions of cognitive load across diverse interview settings that employ different frameworks.

In sum, there is a paucity of research examining cognitive load for investigative interviewers. Current findings suggest that ABE interviews and interviewing for serious offences likely increases cognitive load for interviewers, but further research is required to examine the extent to which cognitive load may impact interviewers' performance. Alternative measures of performance, such as the types of questions that are asked during an interview or interviewers' compliance with best practices, should be explored. Closer examination of the effects on cognitive load and performance of interviewing in pairs, or solo, would be an interesting line of research as would the impact of the type of training and number of training courses attended, on perceptions

of cognitive load. Current findings suggest that it may be beneficial to conduct interviews with child witnesses, for serious crimes, in pairs when planning and preparation indicate that this would be a suitable arrangement.

Chapter 7:

General Discussion

General Discussion

The overarching aim of the current thesis was to understand the cognitive processes for investigative interviewers and how cognitive load may impact interviewers' performance. Specifically, the aims of the current research were to; i) examine interviewers' experiences of cognitive demands when conducting interviews, ii) test whether the various cognitive demands of interviewing have an impact on interviewers' perceived cognitive load and their recall of information provided by witnesses, iii) explore factors that contribute to, or reduce, cognitive demands, and (iv) inform how best to manage cognitive load in practical interview settings. This discussion provides an overview of the main research findings. The associated theoretical and practical implications are discussed, and limitations of the current research are considered together with suggestions for future research. The current programme of research makes a unique contribution to existing investigative interviewing literature by, for the first time, providing a qualitative and quantitative analysis of the effects of cognitive load for investigative interviewers.

Summary of findings

The aim of the field study (Chapter 2) was to explore interviewers' experiences of cognitive load when they conduct interviews with children and vulnerable witnesses, which require specialist interviewing skills (Powell et al., 2005). The study explored factors, as described by serving police officers, that they perceived contribute to cognitive load and examined the impact of these factors for interviewers in practice. The analysis showed that interviewers described common overarching factors, which increased cognitive load and influenced their investigative interviews with vulnerable witnesses. Multiple cognitive processes (e.g., remembering information, thinking of questions to ask, and how to ask those questions without leading the witness), were

identified as intrinsic features of their interviews. Interviewers also explained that operational demands, such as time pressure and taking notes during interviews, further contributed to the cognitive demands of interviewing. Carrying out tasks, such as those identified in this field study, that require controlled processing, are effortful and more cognitively demanding than more automatic processing (Kahneman, 2012; Kleider-Offutt et al., 2016). Thus, additional cognitive tasks that constitute extraneous load, along with tasks contributing to intrinsic load, can result in experiences of cognitive load for interviewers. The current research also observed that interviewers described conducting interviews with children and vulnerable witnesses as mentally, physically and emotionally exhausting, which may impact their performance during an interview. Perhaps this is not surprising as an outcome of cognitive load is cognitive fatigue, which is described as a decline in cognitive resources due to sustained cognitive demands (Borragan et al., 2017). This qualitative research informed three experimental laboratory studies, which were conducted to measure cognitive load, in a controlled setting, and the impact of various cognitive demands on performance during interview tasks.

In the first laboratory study (Experiment 1; Chapter 3), the NASA-TLX was used for the first time to measure *perceived cognitive load* during an investigative interviewing task. The cognitive load for mock-interviewers was manipulated through the task-related instructions given to participants. When participants were instructed to think of questions to ask while watching an interview (high cognitive load condition), the cognitive demands led to higher perceived cognitive load, than was the case for participants who merely watched the interview (low cognitive load condition). Furthermore, participants in the high cognitive load condition exhibited poorer recall accuracy with regard to an interviewee's account than did those in the low cognitive

load condition. This poorer recall accuracy may be due to divided attention, where attending to one element of information (e.g., thinking of questions to ask) caused other cognitive processes (e.g., remembering information) to be neglected (Strayer & Drews, 2007). The results also showed that when the high, moderate and low cognitive load conditions were manipulated, as predicted, participants reported corresponding levels of high, moderate and low perceived cognitive load, as measured by the NASA-TLX. The NASA-TLX was, therefore, considered to be a suitable measure of perceived cognitive load in this experiment and was used as a measure of cognitive load for each of the subsequent experiments in this programme of research. In the two further laboratory studies, additional features of interviewing (i.e., interviewing multiple witnesses and note-taking during interviews) were examined.

Given that interviewers in the qualitative field study (Chapter 2) explained they are often required to interview several witnesses about the same event, the second laboratory study (Experiment 2; Chapter 4) examined the effects of increased cognitive demands during an interview observation task with multiple witnesses. The effects of increased cognitive demands on participants' perceived cognitive load, the accuracy of their memory for information provided by multiple witnesses, and their accuracy for monitoring the source of information, were tested. As with Experiment 1 (Chapter 3), results showed that, when thinking of questions to ask while watching a witness's interview, participants' perceived cognitive load was higher and their recall accuracy was lower, than it was for participants who were simply asked to watch the interviews. Although there were no differences in participants' accuracy for monitoring the source of information across cognitive load conditions, participants generally performed poorly in terms of source monitoring regardless of the cognitive load manipulation. These findings highlighted that keeping track of who provided what information, when

multiple witnesses give an account of the same crime, is challenging for interviewers. The results of the first two experimental studies supported the experiences described by the police officers who participated in the field study (Chapter 2).

For the first two laboratory experiments (Chapters 3 and 4), participants were not given the option to make notes whilst watching witnesses give their accounts. However, note taking is a common feature of investigative interviewing and is, in fact, recommended (College of Policing [CoP], 2019; Ministry of Justice [MoJ], 2011). Note taking can, in itself, be cognitively demanding and may further increase the demands on interviewers (MacDonald, 2016; Piolat et al., 2005). Therefore, the cognitive demand of taking notes was identified as the focus for the third laboratory study (Experiment 3; Chapter 5). In addition, as cognitive load results from increased demand on the working memory system, Working Memory Capacity (WMC) was also measured for this experiment. Specifically, whether individuals' WMC influenced the effects of taking notes on perceived cognitive load and recall accuracy was examined. WMC and access to notes were considered as moderators of the effect of note taking on perceived cognitive load and the accuracy of participants' recall of information from a witness. There was no difference in perceived cognitive load between those who took notes, and those who did not, during an interview task when all participants were asked to think of questions and remember a witness's account. Note taking had a positive impact on participants' recall of information (i.e., the accuracy of recall increased), and the benefits for recall increased when taking structured notes rather than free notes. When WMC was examined as a moderator, a higher WMC was associated with lower perceived cognitive load when recalling information from the witness's account. Participants recall was more accurate when they had access to structured notes (rather than free notes) regardless of their WMC. In sum, the types of notes taken and individual differences in

WMC may be regulating factors for perceived cognitive load and recall performance for investigative interviewers.

For the final study (Chapter 6), a survey was designed for police officers, who undertake different types of interview during investigations for different types of crime, to capture their experiences and perceptions of the cognitive demands identified in previous studies (Chapters, 2, 3, 4 and 5). Interviewers responded that, for both serious and less serious crimes, conducting an interview using the ABE guidelines required more mental effort than was required for conducting interviews with suspects using the PEACE model. Interviewing using the ABE guidelines was more cognitively demanding and was more likely to include child interviewees and to be conducted by a solo interviewer, than interviewing suspects using the PEACE model. The additional cognitive demands of interviewing children, or without another officer being present, may increase cognitive load when conducting ABE interviews. These findings were consistent with the experiences described by interviewers in the field study (Chapter 2), in that, interviewers described interviewing children and vulnerable witnesses as particularly demanding.

The current research was the first to measure perceived cognitive load and the effects for investigative interviewers. Taken together, the findings indicate that increased cognitive load can impact on interviewers' performance in terms of their recall of information provided by witnesses. The reduction in performance, due to cognitive load, is consistent with previous work indicating that cognitive load impacts performance in different settings (e.g., for surgeons; Dias et al, 2018). This thesis contributes to the existing investigative interviewing literature by providing insight and an understanding of the effects of cognitive demands for investigative interviewers. The

theoretical and practical implications of the research presented in this thesis are discussed below.

Theoretical implications

The current findings indicated that the temporary storage and manipulation of information was undertaken during interviewing tasks (e.g., thinking of questions of ask, whilst considering how to ask the questions in a way that complies with best practice; Chapter 2), and that there was increased perceived cognitive load when these tasks were completed. For the laboratory experiments (Chapters 3, 4, 5), participants were required to hold information in mind whist also thinking of questions to ask and, for the third experiment, taking notes. Working Memory (WM), that is, the cognitive mechanism supporting the temporary storage and manipulation of information is, therefore, an important factor to consider when examining the cognitive processing of information (Baddeley, 1992). An individual's capacity (i.e., their Working Memory Capacity; WMC) to access and maintain information in memory should also be considered when examining the impact of cognitive load on performance (Kleider-Offutt et al., 2016; Unsworth and Engle, 2004).

WM comprises processing units for visual/spatial and auditory/verbal information, which interact with long term memory, and are necessary for the management of information during complex cognitive tasks (Baddeley, 1992). WMC represents an individual's ability to i) process information relating to a primary task, ii) maintain the relevant information for the primary task, and iii) access and retrieve information from long-term memory, in the presence of a distraction (Unsworth & Engle, 2007). This theoretical background supports the finding that higher WMC led to greater accuracy for remembering a witness's account (Chapter 5). Additional cognitive processes (i.e., taking notes and thinking of questions) likely distracted participants

from remembering the information being provided by a witness, but with higher WMC the effects were reduced, and more information was remembered. Unsworth and Engle (2007) suggest that those with lower or higher WMC do not differ on all cognitive tasks, but that poorer performance on some memory tasks may be due to a neglect of the task goal (i.e., not focusing on remembering the information). Individuals with lower WMC may remember less information because they are using their WMC to hold information that is irrelevant to the memory task (Kane et al., 2001). For investigative interviewers, attending to the information provided by an interviewee with less focus on irrelevant information (e.g., the structure of the interview, or how to ask their questions; Chapter 2), will likely increase interviewers' memory for the information provided.

WMC is, therefore, relevant when considering cognitive load as experienced by investigative interviewers. Cognitive load can result when the number of elements required to complete a task exceeds WMC (Paas et al., 2003). Cognitive Load Theory (CLT; Sweller, 1988) has also informed this programme of research by providing a framework for understanding the types of load (i.e., intrinsic, extraneous and germane load) that may impact cognitive load experienced by investigative interviewers. Previous research (e.g., Hanway & Akehurst, 2018; Fisher et al., 2014) identified some cognitive processes, which are required for the task of investigative interviewing (e.g., remembering the information provided by an interviewee). Extending this knowledge, the current research identified that some features of interviewing are intrinsic, or inherent to the task (e.g., judging the quality of information and deciding which elements to follow up with further questions), which can increase cognitive demands. The cognitive processing of information was reflected in interviewers' descriptions of their conscious and controlled decisions, for example, being conscious of how they are asking questions (Chapter 2). This type of conscious processing requires control and is

more cognitively demanding than more automatic processing (Kahneman, 2012). Interviewers also described extraneous (additional) features of interviewing, such as being under scrutiny, following best-practice guidance and time pressure, which they suggested contributed to the cognitive demands of interviewing. If cognitive capacity is expended on these extraneous features, then less capacity is available to complete the intrinsic elements (Leppink & van den Heuvel, 2015). Additionally, increased intrinsic and extraneous cognitive load may result in reduced germane load capacity. That is, for interviewers, there may be a reduction in their ability to apply general skills and knowledge to novel situations (Galy et al., 2012).

For the current research, Experiments 1 and 2 (Chapters 3 and 4) tested the effects of increased intrinsic cognitive demands (i.e., remembering information and thinking of questions) and found there was a reduction in the accuracy of interviewers' memory for information when completing additional cognitive tasks. That is, when cognitive demands were increased, in addition to there being reduced recall accuracy (Chapter 3), the accuracy of recognition memory was also reduced (Chapter 4). This reduction in memory performance is likely to have been due to a limited capacity to carry out multiple cognitive tasks in working memory (Kahneman, 1973; Reisberg, 2007). It may be that controlled processing (e.g., thinking of questions) was difficult, but more automatic processes (e.g., merely listening to and watching a witness) were less affected by cognitive load (Schneider & Shiffrin, 1977). Results showed that with increased cognitive demands during interviews, there was a negative impact of multiple witnesses' accounts on the accuracy of participants' source monitoring (SM), which was concerning. Some SM decisions that are rapid and automatic require less conscious thought (e.g., when a person 'remembers' the source of information), however, other decisions are more effortful and require conscious decision making (e.g., when a person

has a familiarity and 'knows' or 'guesses' the source; Johnson et al, 1993). The conscious recollection of information, rather than a familiarity with the information, is more likely to be negatively affected by divided attention (Yonelinas, 2002). Thus, if controlled processing is required to accurately remember information, dividing attention to complete other aspects of an interview (e.g., thinking of questions) is likely to increase cognitive demands, which can have an impact on memory accuracy and source monitoring.

The findings of the current research mirror previous research (e.g., Dias at al., 2018), which found cognitive load has a negative effect on performance, when applied to tasks in a number of occupational roles (Hart, 2006). For example, cognitive load was associated with poorer performance during laparoscopic surgery and increased errors were made by medical students undertaking surgical training (Dias et al., 2018; Haji et al., 2015; Yurko et al., 2010). Additionally, in recruitment settings, when under higher cognitive load, interviewers' decision-making was impacted (i.e., less time was taken to make decisions, and they used more automatic and less controlled processing; Frieder et al., 2016).

Practical implications for investigative interviewers

Information provided by witnesses, during their interviews, forms the basis of an investigation and the completeness of that information can determine the outcome of the investigation (Milne & Bull, 2006). The current programme of research examined interviewers' perceived cognitive load and the accuracy of their recall of information, when completing various interviewing tasks. To obtain accurate and reliable information from interviewees, interviewers must ask questions based on information provided by the interviewees. Therefore, interviewers must be able to accurately recall what interviewees have told them. If interviewers cannot remember information

because it has not been encoded, or it is not available for retrieval, then the amount and accuracy of information recalled by interviewers will inevitably be reduced (e.g., Cauchi & Powell, 2009; Lamb et al., 2000; Warrren & Woodall, 1999). Reduced recall of information may result in lines of enquiry being missed, or interviewees not being challenged about elements of their accounts (CoP, 2019). Failure to accurately remember information also adds weight to the view that interviewers can affect the amount and quality of evidence provided by witnesses (Brown & Lamb, 2015; Gudjonsson, 2010). Of course, interviewers can elicit more information from witnesses by asking follow-up questions. However, accurately attending to information when it is initially provided by witnesses is vital as it will likely improve an interviewer's questioning of an interviewee (i.e., they will be more likely to ask facilitative questions that elicit 'extra' reliable details, and they will be less likely to ask repeated questions). Understanding the cognitive processes for interviewers, including how and why cognitive load and errors in recall may occur, is important to enhance investigative interviewers' performance.

Previously, mixed results were found regarding the benefits of taking notes for enhancing memory (Kobayashi, 2006). The current research found that when taking structured notes and with access to notes at recall, there was increased accuracy for remembering a witness's account compared with not taking notes. This may be due to an external encoding effect when making notes (Kobayashi, 2006). Interviewers may record specific details when making structured notes, which aid their recall. Applying a new method of note taking has previously been found to challenge participants' cognitive resources (MacDonald, 2016). In the current research however, there was no increase in perceived cognitive load due to the novel structured note taking method (when it was compared with a free note taking approach). That is, there was no

difference in perceived cognitive load for those interviewers who took structured notes and those who took free notes, or those who did not take any notes. WMC was shown to account for some individual differences in the accuracy of interviewers' recall of information. However, it was found that taking structured notes benefited recall, without increasing perceived cognitive load, irrespective of WMC. Structured types of note taking are likely, therefore, to be advantageous for increasing the accuracy of investigative interviewers' memory of interviewees' accounts.

For interviewers with a low WMC, focusing on elements of an interview that can reduce the burden on working memory, and thus reduce cognitive load, is particularly important, but reducing cognitive load can also benefit those with higher WMC. For example, interviewers may find that learning to take structured notes in a way that is 'automatic' (i.e., they are not having to think about how to make the notes) will reduce the cognitive demands of note taking. Interviewers will then have more capacity to complete other cognitive tasks, such as, making decisions about what questions to ask or identifying topics to pursue (Fisher et al., 2014; Hanway & Akehurst, 2018). Reducing the cognitive demands of intrinsic features of an interview can then improve interviewers' performance.

Interviewers are required to build rapport, interact with witnesses, and consider other aspects of a case (Schreiber-Compo et al., 2012). Interviews, therefore, occur in a social context, whereby interviewers perceive witnesses' actions and make judgements about their credibility, reliability and wellbeing (Ask & Landström, 2010; Hanway & Akehurst, 2018). These factors, and the extraneous factor of time pressure (i.e., temporal demand), were not manipulated during the current programme of research. However, as Leppink and colleagues (2015) suggest, cognitive load is additive. Therefore, extraneous factors, which were identified in the field study (Chapter 2) as

being present when conducting investigative interviews, will likely contribute to a higher cognitive load for interviewers in practice (Hanway & Akehurst, 2018; Nordstrom et al., 1996). Some cognitive demands are necessarily present for all types of investigative interviews and they cannot be avoided. For example, thinking of questions to ask and remembering information are central elements of interviewing. That said, cognitive load induced by these intrinsic factors may be reduced through training and the development of schemas.

Schemas are unconscious mental structures and processes that contain generic knowledge, and they underlie key aspects of cognitive processing and skill development (Brewer & Nakamura, 1984). Cognitive Load Theory suggests that the automatic processing of information relies on schemas to reduce mental effort (Paas et al., 2004). For investigative interviewers, a reduction of cognitive load could be achieved by learning and building schemas for elements of an interview. For example, the phrasing and typology of asking open questions requires the development of skills (Oxburgh et al., 2010). Building schemas for multiple open question stems (e.g., "explain more about...", "tell me what happened when…", or "describe the…") may then increase skill development. If these question stems come to mind automatically during questioning (i.e., without thinking about how to ask a question in an open way) then less cognitive effort will be required. It should be noted that, if a task is cognitively demanding (i.e., the intrinsic and extraneous load exceed capacity), then there is little opportunity to form schemas (Schnotz & Kurchner, 2007). Therefore, allowing time during training, and refresher training, for schema development of this type of interviewing skill is essential. In doing so, asking open questions will become more automatic and will likely improve interviewers' ability to ask these types of questions, which will then be an indicator of an expert interviewer (Oxburgh et al., 2010). Additionally, factors that

contribute to extraneous load (e.g., time pressure) may be managed through the application of investigation management techniques, such as allowing interviewers sufficient time for planning and preparation. With training, and skill development (e.g., being skilled at taking structured notes or having in-depth knowledge of the ground rules of ABE interviews), more schemas for investigative interviewing are potentially built, which will reduce cognitive load for interviewers.

Methodological considerations and limitations

For the current doctoral programme of research, investigative interviewers' experience of cognitive load was explored. The Interpretative Phenomenological Approach (IPA) employed in the field study (Chapter 2) is a suitable method when seeking perceptions and understanding of situations that are complex, poorly understood or previously unexplored (McCormack & Joseph, 2018). However, it is acknowledged that conducting IPA research has limitations in terms of sample size and generalisability of findings. For this topic of study, the sample was a suitable size for IPA research but other interviewers, and those from different jurisdictions, may have alternative perceptions and experiences of interviewing children and other witnesses. Therefore, despite being consistent with prior qualitative research and the IPA approach, the generalisability of the results for this study, are likely limited. That said, the results of the study provided a detailed information-base with which to inform further research.

A limitation of the design for the first laboratory study (Chapter 3) was that, as participants were not permitted to take notes during their task, it was not clear what they were thinking during the task. To mitigate this limitation, and to ensure participants had understood their instructions, a manipulation check was included after the recall phase to check participants' understanding of what they had been asked to do.

This limitation was also addressed in the second laboratory study (Chapter 4) as participants were asked to make a note of their questions, when instructed to do so. For the third laboratory study (Chapter 5), the issue of note taking during an interview and the cognitive demands of taking notes was specifically addressed.

A further limitation of the laboratory studies was that ecological validity for aspects of the designs was low and elements of the designs (e.g., watching interviewees rather than actually questioning interviewees) were artificial. However, this was a necessary process to standardise and control the cognitive demands that were placed on the participants. The use of pre-recorded free narrative accounts for the interview observation tasks allowed details provided by the witnesses to be controlled, so the 'ground truth' of the witnesses' accounts was known. The amount of information provided by the witnesses was also consistent, thereby standardising participants' cognitive load when watching the accounts. When designing the experimental studies, the cognitive demands for interviewers when they hold information in mind and process information, both experienced by interviewers during real-world interviews, was replicated. That is, the participants were required to listen to the interviewees, think of questions to ask and, for the third experimental study, take notes. It is accepted that interviewers in practice are required to complete more tasks (e.g., engaging with the interviewee), therefore, the experimental studies, by necessity, did not replicate interviewing in operational police settings. However, it is important to note that perceived cognitive load increased when participants in the experiments were required to complete just a few additional cognitive tasks as instructed by the researcher. It likely follows that interviewers in practice, who have far more cognitive demands to manage, will also perceive increased cognitive load.

It could also be argued that, for the laboratory studies, participants were novice interviewers who had not received any training in investigative interviewing, which may not reflect the real-life experiences of investigative interviewers. The experimental studies were purposefully conducted with non-trained interviewers so that they would all have an equal level of experience (i.e., no experience). Skills improve with training and, had trained interviewers been recruited, some interviewers may have performed better on the interview tasks than others, but it would be difficult to know what schemas for interviewing had been built by individual participants. Training participants and controlling for their subsequent skill level was not practical for the current experimental studies. As such, the current findings may have limited generalisability to trained or experienced interviewers.

For the current research programme, perceived cognitive load was measured using subjective rating scales (the NASA-TLX scales). The distinction of intrinsic, extraneous, and germane load was considered when designing the studies, but total perceived cognitive load was measured. This measurement technique does not differentiate between the three cognitive load components (Paas et al., 2003). In addition, subjective measures of cognitive load can become confounded with perceptions of task demand or task performance (Pickup et al., 2005). However, more objective measurement strategies are not a 'silver bullet' for measuring cognitive load (Charles & Nixon, 2019). The NASA-TLX was employed in the current programme of research as it was designed to be used during, or immediately after, a task and has been widely used in a variety of settings to measure the cognitive load perceived by participants when they complete a variety of tasks (e.g., Hart, 2006; Rizzo et al., 2016).

Future directions

It is recognised that, for the three laboratory studies (Chapters 3, 4 and 5), the current research examined, in controlled settings, a limited aspect of investigative interviewing (i.e., interviewers' memory for information provided by interviewees). Future research should examine the effects of cognitive load on other aspects of interviewer performance, and their compliance with best practices. It would be useful to examine whether cognitive load also has an impact on, for example, the types of questions asked by interviewers or the decisions made by them regarding the truthfulness of information provided by interviewees. Rapport building and the cooperativeness of witnesses were identified as contributing to cognitive demands for interviewers (Chapter 2). Further research should focus on the effects of rapport building during interactions with interviewees on cognitive load, types of questions asked, or interviewers' recall of information. In addition, further experimental research using objective measures of cognitive load, such as physiological measures (i.e., EEG and pupillometry) and dual-task paradigms, could be considered for future studies, but it is likely that these measures would be very difficult to employ during an interviewing task. It is worthy of note that individual physiological parameters may not provide a single true measure of the cognitive load experienced in response to a task (Charles & Nixon, 2019). Development of a scale to measure cognitive load, specifically for investigative interviewers, would be useful. A combination of objective and subjective measures may then advance this line of research.

Cognitive load experienced by interviewers in field settings should also be examined in future research. For example, skills can improve with training, therefore some interviewers' performance may improve, even when cognitive demands are high. It would be interesting to explore the impact that training, experience, and the

development of skills, has on experienced interviewers' perceived cognitive load and their performance. Does training increase cognitive load for interviewers as they then have more to think about in terms of adhering to what they have learnt? Or is cognitive load for interviewers reduced over time as their skills become more automatic?

Conclusion

Across this programme of doctoral research, the concept of cognitive load for investigative interviewers was explored and tested. In sum, investigative interviewers described factors of an interview that, in their view, contribute to cognitive load and it was found that cognitive load can have an impact on investigative interviewers' performance. The findings reported in this thesis contribute to the understanding of interviewers' cognitive processing of information and the impact that cognitive load can have for investigative interviewers. Understanding the cognitive demands, helps us to appreciate why, despite training, interviewers often find it difficult to comply with best practice interview techniques and guidelines. That is, the cognitive demands of completing an interview may exceed interviewers' capacity when they undertake an interview which may in turn reduce their performance. More research should be conducted that examines the impact of cognitive load for interviewers on additional performance measures (e.g., the types of questions asked). However, this thesis proposes that the negative effects of cognitive load can be managed through the use of techniques, such as structured note taking, that may benefit interviewers irrespective of their working memory capacity. Providing interviewers with additional training on i) the impact of cognitive load when interviewing, ii) errors in interviewers recall that may occur, for example, when monitoring the source of information, and iii) the development of schemas that increase the automaticity of cognitive processing for investigative interviewers, is also recommended for consideration.

References

Anderson, J. R. (1995). *Learning and memory: An integrated approach*. John Wiley & Sons

Ask, K., & Landström, S. (2010). Why emotions matter: Expectancy violation and affective response mediate the emotional victim effect. *Law and human behavior*, *34*(5), 392-401. https://doi.org/10.1007/s10979-009-9208-6

Alison, L. J., Alison, E., Noone, G., Elntib, S., & Christiansen, P. (2013). Why tough tactics fail and rapport gets results: Observing Rapport-Based Interpersonal Techniques (ORBIT) to generate useful information from terrorists. *Psychology, Public Policy, and Law*, *19*(4), 411. Doi:10.1037/a0034564

Association of Chief Police Officers (ACPO), National Investigative Interviewing Strategic Steering Group (2013). *Advice on the structure of visually recorded witness interviews (2nd Edition)*. Retrieved from http://library.college.police.uk/docs/APPREF/ACPO-Witness-Interview-Structure-2013.pdf

Ayres, P. (2020). Something old, something new from cognitive load theory. *Computers in Human Behavior*, *113.* Doi.org/10.1016/j.chb.2020.106503.

Ayers, M. S., & Reder, L. M. (1998). A theoretical review of the misinformation effect: Predictions from an activation-based memory model. *Psychonomic Bulletin & Review*, *5*(1), 1-21. https://doi.org/10.3758/BF03209454

Baddeley, A. (1992). Working memory: The interface between memory and cognition. *Journal of cognitive neuroscience*, *4*(3), 281-288. Doi.org/10.1162/jocn.1992.4.3.281

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic press. https://doi.org/10.1016/S0079-7421(08)60452-1

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2019). From short-term store to multicomponent working memory: The role of the modal model. *Memory & cognition*, *47*(4), 575-588. https://doi.org/10.3758/s13421-018-0878-5

Baddeley, A. D., & Logie, R. H. (1999). *Working memory: The multiple-component model.* In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (p. 28–61). Cambridge University Press. https://doi.org/10.1017/CBO9781139174909.005

Bargh, J. A. (1984). Automatic and conscious processing of social information. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition,* Vol. 3, pp. 1-43). Lawrence Erlbaum Associates Publishers.

Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate behavioral research*, *40*(3), 373-400. Doi.org/10.1207/s15327906mbr4003_5.

Bittner, A. V., Byers, J. C., Hill, S. G., Zaklad, A. L., & Christ, R. E. (1989) Generic workload ratings of a mobile air defense system (LOS-FH). *Proceedings of the 33rd Annual Meeting of the Human Factors and Ergonomics Society*, 1476-1480. HFES.

Borragán, G., Slama, H., Bartolomei, M., & Peigneux, P. (2017). Cognitive fatigue: A Time-based Resource-sharing account. *Cortex*, *89*(February), 71–84. https://doi.org/10.1016/j.cortex.2017.01.023

Brandon, S. E., Wells, S., & Seale, C. (2018). Science-based interviewing: Information elicitation. *Journal of Investigative Psychology and Offender Profiling*, *15*(2), 133-148. Doi.org/10.1002/jip.1496

Brewer, W. F., & Nakamura, G. V. (1984). The nature and functions of schemas. *Center for the Study of Reading Technical Report; no. 325*.

Brown, D. A., & Lamb, M. E. (2015). Can children be useful witnesses? It depends how they are questioned. *Child Development Perspectives*, *9*(4), 250-255. https://doi.org/10.1111/cdep.12142

Brubacher, S. P., Powell, M. B., & Roberts, K. P. (2014). Recommendations for interviewing children about repeated experiences. *Psychology, Public Policy, and Law*, *20*(3), 325. Doi.org/10.1037/law0000011

Bull, R. (2010). The investigative interviewing of children and other vulnerable witnesses: Psychological research and working/professional practice. *Legal and Criminological Psychology*, *15*, 5–23. doi.org/10.1348/014466509X440160

Bull, R. & Soukara, S. (2009). A set of studies of what really happens in police interviews with suspects in Lassiter, G. D. Meissner C. A. *Interrogations and confessions: Research, practice, and policy*. American Psychological Association. https://doi.org/10.1037/12085-005

Cauchi, R., & Powell, M. B. (2009). An examination of police officers' notes of interviews with alleged child abuse victims. *International Journal of Police Science & Management*, *11*(4), 505-515. https://doi.org/10.1350/ijps.2009.11.4.147

Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological bulletin*, *113*(3), 403. https://doi.org/10.1037/0033-2909.113.3.403

Charles, R. L., & Nixon, J. (2019). Measuring mental workload using physiological measures: a systematic review. *Applied ergonomics, 74*, 221-232. oi.org/10.1016/j.apergo.2018.08.028

Clarke, C., Milne, R., & Bull, R. (2011). Interviewing suspects of crime: The impact of PEACE training, supervision and the presence of a legal advisor. *Journal of investigative psychology and offender profiling*, *8*(2), 149-162. https://doi.org/10.1002/jip.144

College of Policing (April, 2019). Investigation process. https://www.app.college.police.uk/app-content/investigations/investigation-process/

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic bulletin & review*, *12*(5), 769-786. https://doi.org/10.3758/BF03196772

Cowan, N. (2017). The many faces of working memory and short-term storage. Psychonomic Bulletin & Review, 24, 1158–1170. http://dx.doi.org/10.3758/s13423-016-1191-6

Criminal Justice Joint Inspectorate (CJJI; 2014). Achieving best evidence in child sexual abuse cases; A joint inspection. https://www.justiceinspectorates.gov.uk/cjji/wp-content/uploads/sites/2/2014/12/CJJI_ABE_Dec14_rpt.pdf

abuse, *25*(4), 347-362. doi.org/10.1080/10538712.2016.1161687

Cross, T. P., & Hershkowitz, I. (2017). Psychology and child protection: Promoting widespread improvement in practice. *Psychology, Public Policy, and Law*, *23*(4), 503. http://dx.doi.org/10.1037/law0000141

Dale, P. S., Loftus, E. F., & Rathbun, L. (1978). The influence of the form of the question on the eyewitness testimony of preschool children. *Journal of Psycholinguistic Research*, *7*(4), 269-277. doi.org/10.1007/BF01068110

Danby, M. C., Sharman, S. J., Brubacher, S. P., Powell, M. B., & Roberts, K. P. (2017). Differential effects of general versus cued invitations on children's reports of a repeated event episode. *Psychology, Crime & Law*, *23*(8), 794-811. https://doi.org/10.1080/1068316X.2017.1324028

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

Dias, R. D., Ngo-Howard, M. C., Boskovski, M. T., Zenati, M. A., & Yule, S. J. (2018). Systematic review of measurement tools to assess surgeons' intraoperative cognitive workload. *British Journal of Surgery*. https://doi.org/10.1002/bjs.10795

Dixon-Woods, M., Agarwal, S., Jones, D., Young, B., & Sutton, A. (2005). Synthesising qualitative and quantitative evidence: a review of possible methods. *Journal of health services research & policy*, *10*(1), 45-53. doi: 10.1177/135581960501000110.

Duron, J. F., & Cheung, M. (2016). Impact of repeated questioning on interviewers: learning from a forensic interview training project. *Journal of child sexual abuse*, *25*(4), 347-362. doi.org/10.1080/10538712.2016.1161687

Eldridge, L. L., Sarfatti, S., & Knowlton, B. J. (2002). The effect of testing procedure on remember-know judgments. *Psychonomic Bulletin & Review*, *9*(1), 139-145.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current directions in psychological science*, *11*(1), 19-23. https://doi.org/10.1111/1467-8721.00160

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *Psychology of learning and motivation*, *44*, 145-200. http://englelab.gatech.edu/publications/2004/exectutive-attention2c-working-memory-capactiy-2c-and-a-two-factor-theory-of-cognitive-control.pdf

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of experimental psychology: General*, *128*(3), 309. https://doi.org/10.1037/0096-3445.128.3.309

Engstrom, M., Landtblom, A. M., & Karlsson, T. (2013). Brain and effort: brain activation and effort-related working memory in healthy participants and patients with working memory deficits. *Frontiers in human neuroscience*, *7*, 140. https://doi.org/10.3389/fnhum.2013.00140

Evans, L. H., & Wilding, E. L. (2012). Recollection and familiarity make independent contributions to memory judgments. Journal of Neuroscience, 32(21), 7253-7257. doi.org/10.1523/JNEUROSCI.6396-11.2012

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, *41*(4), 1149-1160. doi:10.3758/BRM.41.4.1149

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage.

Findlay, L. (2014). Engaging phenomenological analysis. *Qualitiative Research in Psychology*, 121–141. https://doi.org/10.1080/14780887.2013.807899

Fisher, R. P., Compo, N. S., Rivard, J., & Hirn, D. (2014). Interviewing witnesses. In T. J. Perfect (Ed.), *The Sage Handbook of Applied Memory* (pp. 559–578). Sage Publications Ltd. https://www.researchgate.net/profile/Nadja_Compo/publication/288259307_Interviewing_witnesses/links/56b353f608ae3d06a2664555.pdf

Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing: The cognitive interview. - PsycNET*. Springfield: Charles C. Thomas. Retrieved from https://psycnet.apa.org/record/1992-98595-000

Fisher, R. P., & Geiselman, R. E. (2010). The cognitive interview method of conducting police interviews: Eliciting extensive information and promoting therapeutic jurisprudence. *International journal of law and psychiatry*, *33*(5-6), 321-328. oi.org/10.1016/j.ijlp.2010.09.004

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. P., Ross, S. J., & Cahill, B. S. (2010). Interviewing witnesses and victims. *Forensic psychology in context: Nordic and international approaches*, 56-74. http://faculty.washington.edu/sjross2/documents/FisherRoss&10-interviewing%20witnesses%20and%20victims.pdf

Fisher, R. P., & Schreiber, N. (2017). Interview protocols to improve eyewitness memory. *The handbook of eyewitness psychology*, *1*.

Flanigan, A. E., & Titsworth, S. (2020). The impact of digital distraction on lecture note taking and student learning. *Instructional Science*, 1-30. doi.org/10.1007/s11251-020-09517-2

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & cognition*, *43*(2), 226-236. doi.org/10.3758/s13421-014-0461-7

Frieder, R. E., Van Iddekinge, C. H., & Raymark, P. H. (2016). How quickly do interviewers reach decisions? An examination of interviewers' decision-making time across applicants. *Journal of Occupational & Organizational Psychology*, *89*(2), 223–248. https://doi.org/10.1111/joop.12118

Gabbert, F., Hope, L., Luther, K., Wright, G., Ng, M., & Oxburgh, G. (2021). Exploring the use of rapport in professional information-gathering contexts by systematically mapping the evidence base. *Applied Cognitive Psychology*, *35*(2), 329-341. https://doi.org/10.1002/acp.3762

Galy, E., Cariou, M., & Mélan, C. (2012). What is the relationship between mental workload factors and cognitive load types?. *International Journal of Psychophysiology*, *83*(3), 269-275. doi.org/10.1016/j.ijpsycho.2011.09.023

Galy, E., Paxion, J., & Berthelon, C. (2018). Measuring mental workload with the NASA-TLX needs to examine each dimension rather than relying on the global score: an example with driving. *Ergonomics*, *61*(4), 517-527. doi.org/10.1080/00140139.2017.1369583

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (2002). Recognition memory and decision processes: A meta-analysis of remember, know, and guess responses. *Memory*, *10*(2), 83–98. https://doi.org/10.1080/09658210143000281

Garven, S., Wood, J. M., Malpass, R. S., & Shaw III, J. S. (1998). More than suggestion: The effect of interviewing techniques from the McMartin Preschool case. *Journal of applied psychology*, *83*(3), 347.

Gregory, A. H., Compo, N. S., Vertefeuille, L., & Zambruski, G. (2011). A comparison of US police interviewers' notes with their subsequent reports. *Journal of Investigative Psychology and Offender Profiling*, *8*(2), 203–215. https://doi.org/10.1002/jip.139

Gudjonsson, G. H. (1992). *The psychology of interrogations, confessions and testimony*. John Wiley & Sons.

Gudjonsson, G. H., & MacKeith, J. A. C. (2002). The 'Guildford Four'and the 'Birmingham Six'. *The Psychology of Interrogations and Confessions: A Handbook*, 445-457.

Gudjonsson, G. H. (2010). Psychological vulnerabilities during police interviews. Why are they important?. *Legal and criminological Psychology*, *15*(2), 161-175. doi.org/10.1348/135532510X500064

Haji, F. A., Rojas, D., Childs, R., de Ribaupierre, S., & Dubrowski, A. (2015). Measuring cognitive load: Performance, mental effort and simulation task complexity. *Medical Education*, *49*(8). https://doi.org/10.1111/medu.12773

Hanway, P., & Akehurst, L. (2018). Voices from the front line: police officers' perceptions of real-world interviewing with vulnerable witnesses. *Investigative Interviewing: Research and Practice*, *9*(1), 14-33.

Hart, S. G. (2006, October). NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 50, No. 9, pp. 904-908). Sage publications. doi.org/10.1177/154193120605000909

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology*, *52*I, 139–183. doi.org/10.1016/S0166-4115(08)62386-9

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications.

Hershkowitz, I. (2001). Children's responses to open-ended utterances in investigative interviews. *Legal and Criminological Psychology*, *6*(1), 49-63. doi.org/10.1348/135532501168190

Hershkowitz, I. (2011). Rapport building in investigative interviews of children. In Lamb, M. E., La Rooy, D., Malloy, C. & Katz, C. (Eds.), *Children's Testimony: A Handbook of Psychological Research and Forensic Practice* (2nd ed., pp. 109–128). John Wiley & Sons Ltd.

Hoogesteyn, K., Meijer, E., & Vrij, A. (2020). Examining witness interviewing environments. *Journal of Investigative Psychology and Offender Profiling*. 1-12. DOI: 10.1002/jip.1549

Hope, L., Eales, N., & Mirashi, A. (2014). Assisting jurors: Promoting recall of trial information through the use of a trial-ordered notebook. *Legal and criminological psychology*, *19*(2), 316-331. https://doi.org/10.1111/lcrp.12003

Hope, L., & Gabbert, F. (2019). Developments and Innovations in Evidence-Based Investigative Interviewing. *Evidence-based Investigative Interviewing: Applying Cognitive Principles*, 42. https://doi.org/10.1111/lcrp.12003

Hope, L., Gabbert, F., Kinninger, M., Kontogianni, F., Bracey, A., & Hanger, A. (2019). Who said what and when? A timeline approach to eliciting information and intelligence about conversations, plots, and plans. *Law and human behavior*, *43*(3), 263. https://doi.org/10.1037/lhb0000329

Hope, L., Mullis, R., & Gabbert, F. (2013). Who? What? When? Using a timeline technique to facilitate recall of a complex event. *Journal of Applied Research in Memory and Cognition*, *2*(1), 20-24. https://doi.org/10.1016/j.jarmac.2013.01.002

Horowitz, I. A., & ForsterLee, L. (2001). The effects of note-taking and trial transcript access on mock jury decisions in a complex civil trial. *Law and Human Behavior*, *25*(4), 373-391. https://doi.org/10.1023/A:1010655602400

Hunter, C. R., & Pisoni, D. B. (2018). Extrinsic Cognitive Load Impairs Spoken Word Recognition in High- and Low-Predictability Sentences. Ear and hearing, 39(2), 378–389. https://doi.org/10.1097/AUD.0000000000000493

Hyman Gregory, A. (2009). Investigative Interviewing and Memory: How Accurate Are Interviewers' Recollections of Investigative Interviews? (Doctoral dissertation). Florida International University.

Jarrold, C., Tam, H., Baddeley, A. D., & Harvey, C. E. (2011). How does processing affect storage in working memory tasks? Evidence for both domain-general and domain-specific effects. Journal of experimental psychology. Learning, memory, and cognition, 37(3), 688–705. https://doi.org/10.1037/a0022527

Jansen, R. S., Lakens, D., & IJsselsteijn, W. A. (2017). An integrative review of the cognitive costs and benefits of note-taking. *Educational Research Review*, *22*, 223-233. doi.org/10.1016/j.edurev.2017.10.001

Johannessen, E., Szulewski, A., Radulovic, N., White, M>, Braund, H., Howes, D., Rodenburg, D., & Davies, C. (2020). Psychophysiologic measures of cognitive load in physician team leaders during trauma resuscitation. *Computers in Human Behavior*, *111*. doi.org/10.1016/j.chb.2020.106393.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. Psychological bulletin, 114(1), 3. https://doi.org/10.1037/0033-2909.114.1.3

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Prentice-Hall.

Kahneman, D. (2012). *Thinking, fast and slow*. Penguin.

Kane, M.J., Bleckley, M.K., Conway, A.R.A., Engle, R.W. (2001). A controlled-attention view of working-memory capacity. Journal of Experimental Psychology: General, 130, 169–183. DOI: 10.1037//0096-3445.130.2.169

Kassin, S. M., Appleby, S. C., & Perillo, J. T. (2010). Interviewing suspects: Practice, science, and future directions. *Legal and Criminological Psychology*, *15*(1), 39-55. doi.org/10.1348/135532509X449361

Kleider-Offutt, H. M., Clevinger, A. M., & Bond, A. D. (2016). Working memory and cognitive load in the legal system: Influences on police shooting decisions, interrogation and jury decisions. Journal of Applied Research in Memory and Cognition, 5(4), 426-433. doi.org/10.1016/j.jarmac.2016.04.008

Kleider, H. M., Pezdek, K., Goldinger, S., & Kirk, A. (2008). Repetition and Dual Coding in Procedural Multimedia Presentations. *Applied Cognitive Psychology*, *22*, 1–20. https://doi.org/10.1002/acp

Kobayashi, K. (2006). Combined Effects of Note-Taking/-Reviewing on Learning and the Enhancement through Interventions: A meta-analytic review. *Educational Psychology*, *26*(3), 459-477. https://doi.org/10.1080/01443410500342070

Kohnken, G., Thurer, C., & Zoberbier, D. (1994). The cognitive interview: Are the interviewers' memories enhanced, too? *Applied Cognitive Psychology, 8*, 13-24. doi.org/10.1002/acp.2350080103

Kolk, N. J., Born, M. P., Van Der Flier, H., & Olman, J. M. (2002). Assessment center procedures: Cognitive load during the observation phase. *International Journal of Selection and Assessment*, *10*(4), 271-278. doi.org/10.1111/1468-2389.00217

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. doi.org/10.1016/j.jarmac.2018.03.006

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine*, *15*(2), 155-163. https://doi.org/10.1016/j.jcm.2016.02.012

Lafontaine, J., & Cyr, M. (2016). A Study of the Relationship between Investigators' Personal Characteristics and Adherence to Interview Best Practices in Training. *Psychiatry, Psychology and Law*, *23*(5). doi.org/10.1080/13218719.2016.1152925

Lakens, D., & Caldwell, A. R. (2019). Simulation-Based Power-Analysis for Factorial ANOVA Designs. https://doi.org/10.31234/osf.io/baxsf

Lamb, M. E. (2016). Difficulties translating research on forensic interview practices to practitioners: Finding water, leading horses, but can we get them to drink?. *American Psychologist*, *71*(8), 710. https://doi.org/10.1037/amp0000039

Lamb, M. E., Brown, D. A., Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2018). *Tell me what happened: Questioning children about abuse* (2nd edition). Wiley-Blackwell.

Lamb, M. E., Orbach, Y., Sternberg, K. J., Hershkowitz, I., & Horowitz, D. (2000). Accuracy of Investigators' Verbatim Notes of Their Forensic Interviews with Alleged Child Abuse Victims. *Law and Human Behavior*, *24*(6). https://doi.org/10.1023/A:1005556404636

La Rooy, D., & Dando, C. (2010). Witness interviewing. In Towl, G. J. & Crighton, D. A. (Eds.), *Forensic psychology* (pp. 195–209). BPS and Blackwell Publishing Ltd.

Launay, C., & Py, J. (2015). Methods and aims of investigative interviewing of adult witnesses: An analysis of professional practices. *Pratiques psychologiques*, *21*(1), 55-70. doi.org/10.1016/j.prps.2014.11.001

Leppink, J., & van den Heuvel, A. (2015). The evolution of cognitive load theory and its application to medical education. *Perspectives on medical education, 4*(3), 119-127. https://doi.org/10.1007/s40037-015-0192-x

Lindsay, D. S. (2007). Autobiographical Memory, Eyewitness Reports, and Public Policy. *Canadian Psychology*, *48*(2), 57–66. https://doi.org/10.1037/cp2007007

Lindsay, D. S. (2014). Memory source monitoring applied. *The SAGE handbook of applied memory*, 59-75.

Lindsay, D. S., Johnson, M. K., & Kwon, P. (1991). Developmental changes in memory source monitoring. Journal of Experimental Child Psychology, 52(3), 297-318. doi.org/10.1016/0022-0965(91)90065-Z

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric annals*, *25*(12), 720-725. doi.org/10.3928/0048-5713-19951201-07

MacDonald, S. (2016). *The effect of note taking on memory for details in investigative interviews* (Doctoral dissertation, Memorial University of Newfoundland).

Makany, T., Kemp, J., & Dror, I. E. (2009). Optimising the use of note-taking as an external cognitive aid for increasing learning. *British Journal of Educational Technology*, *40*(4), 619-635. doi.org/10.1111/j.1467-8535.2008.00906.x

Marsh, J. E., Patel, K., Labonté, K., Threadgold, E., Skelton, F. C., Fodarella, C., ... & Vachon, F. (2017). Chatting in the face of the eyewitness: The impact of extraneous cell-phone conversation on memory for a perpetrator. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *71*(3), 183. doi.org/10.1037/cep0000101

Mayer, R. E. (1984). Aids to text comprehension. *Educational psychologist*, *19*(1), 30-42. doi.org/10.1080/00461528409529279

McCormack, L., & Joseph, S. (2018). *PHENOMENA: A 9-Step Guide to Avoiding Pitfalls When Doing Interpretative Phenomenological Analysis (IPA)—IPA and the "Lived" Experience of Complex Trauma*. SAGE Publications Ltd. http://dx.doi.org/10.4135/9781526429681

Meise, J., & Leue, A. (2019). Quality of written record following mock eyewitness testimony: Note taking should be a minimum standard! *Journal of Investigative Psychology and Offender Profiling*. https://doi.org/10.1002/jip.1522

Meissner, C. A., & Lyles, A. M. (2019). IX investigations: The importance of training investigators in evidence-based approaches to interviewing. *Journal of Applied Research in Memory and Cognition*, *8*(4), 387-397. https://doi.org/10.1016/j.jarmac.2019.07.001

Milne, B., & Bull, R. (1999). *Investigative interviewing: Psychology and practice*. Wiley.

[redacted]

Ministry of Justice. (2011). Achieving Best Evidence in Criminal Guidance on interviewing victims and witnesses, and guidance on using special measures, (March), 68–99. http://www.justice.gov.uk/downloads/victims-and-witnesses/vulnerable-witnesses/achieving-best-evidence-criminal-proceedings.pdf

Mitchell, K. J., & Johnson, M. K. (2000). Source monitoring: Attributing mental experiences. In E. Tulving & F.I.M. Craik (Eds.). *Source*, 179–195.

Mortimer, A. (1994). *Cognitive processes underlying police investigative interviewing behaviour.* Unpublished PhD thesis, University of Portsmouth, https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.386953

Mount, D., & Mazerolle, L. (2020). Investigative interviewing skills in policing: examining the transfer of training into workplace practices. *Policing: An International Journal*. doi.org/10.1108/PIJPSM-12-2019-0182

National Policing Improvement Agency (2009). National investigative interviewing strategy. Bedfordshire: NPIA and Association of Chief Police Officers. http://library.college.police.uk/docs/npia/BP-Nat-Investigative-Interviewing-Strategy-2009.pdf

Nordstrom, C. R., Williams, K. B., & LeBreton, J. M. (1996). The effect of cognitive load on the processing of employment selection information. *Basic and Applied Social Psychology*, *18*(3), 305–318. doi.org/10.1207/s15324834basp1803_4

Oberauer, K., Farrell, S., Jarrold, C., & Lewandowsky, S. (2016). What limits working memory capacity? *Psychological bulletin*, *142*(7), 758. https://doi.org/10.1037/bul0000046

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D., Conway, A., Cowan, N., ... & Ma, W. J. (2018). Benchmarks for models of short-term and working memory. Psychological Bulletin, 144(9), 885. doi.org/10.1037/bul0000153

O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology. In Boff, K. R., Kaufman, L. & Thomas, J. P. (Eds.), Handbook of perception and human performance, Vol. 2. Cognitive processes and performance (pp. 1-49). John Wiley & Sons.

Orbach, Y., Hershkowitz, I., Lamb, M. E., Sternberg, K. J., Esplin, P. W., & Horowitz, D. (2000). Assessing the value of structured protocols for forensic interviews of alleged child abuse victims. *Child abuse & neglect*, *24*(6), 733-752. doi.org/10.1016/S0145-2134(00)00137-X

Otgaar, H., La Rooy, D., Horselenberg, R., Hershkowitz, I., de Ruiter, C., Blezer, L., ... & Kollau, R. (2019). Assessing the quality of child investigative interviewing in the Netherlands. *Applied Cognitive Psychology*, *33*(5), 889-897. https://doi.org/10.1002/acp.3521

Oxburgh, G., & Dando, C. J. (2011). Psychology and interviewing: what direction now for our quest for reliable information? *The British Journal of Forensic Practice*, *13*(2), 135–144. https://doi.org/10.1108/14636641111134378

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

Oxburgh, G., Myklebust, T., Grant, T., & Milne, R. (Eds.). (2015). *Communication in investigative and legal contexts: Integrated approaches from forensic psychology, linguistics and law enforcement*. John Wiley & Sons.

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional science*, *32*(1), 1-8. doi.org/10.1023/B:TRUC.0000021806.17516.d0

Paas, F. G., & Van Merriënboer, J. J. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human factors*, *35*(4), 737-743. doi.org/10.1177/001872089303500412

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational psychologist, 38*(1), 63-71. https://doi.org/10.1207/S15326985EP3801_8

Phillips, L. H., Channon, S., Tunstall, M., Hedenstrom, A., & Lyons, K. (2008). The role of working memory in decoding emotions. *Emotion*, *8*(2), 184. doi: 10.1037/1528-3542.8.2.184.

Pickup, L., Wilson, J. R., Sharpies, S., Norris, B., Clarke, T., & Young, M. S. (2005). Fundamental examination of mental workload in the rail industry. *Theoretical issues in ergonomics science*, *6*(6), 463-482. https://doi.org/10.1080/14639220500078021

Pietkiewicz, I., & Smith, J. A. (2014). A practical guide to using Interpretative Phenomenological Analysis in qualitative research psychology. *Psychological Journal*, *20*(1), 7–14. DOI: 10.14691/CPPJ.20.1.7

Piolat, A., Olive, T., & Kellogg, R. T. (2005). Cognitive effort during note taking. *Applied cognitive psychology*, *19*(3), 291-312. https://doi.org/10.1002/acp.1086

Poole, D. A., Dickinson, J. J., & Brubacher, S. P. (2014). Sources of unreliable testimony from children. *Roger Williams University Law Review*, *2*, 382–410. https://docs.rwu.edu/rwu_LR/vol19/iss2/4

Powell, M. B. (2002). Specialist training in investigative and evidential interviewing: Is it having any effect on the behaviour of professioanls in the field? *Psychiatry, Psychology and Law*, *9*(1), 44–55. doi.org/10.1375/pplt.2002.9.1.44

Powell, M. B., & Barnett, M. (2015). Elements underpinning successful implementation of a National Best Practice Child Investigative Interviewing Framework. *Psychiatry, Psychology and Law*, *22*(3), 368–377. https://doi.org/10.1080/13218719.2014.951112

Powell, M. B., Fisher, R. P., & Wright, R. (2005). Investigative Interviewing. In N. Brewer & K. D. Williams (Eds.), *Psychology and Law* (pp. 11–42). The Guildford Press.

Powell, M., Wright, R., & Clark, S. (2010). Improving the competency of police officers in conducting investigative interviews with children. *Police Practice & Research*, *11*(3), 211–226. https://doi.org/10.1080/15614260902830070

Reid, K., Flowers, P., & Larkin, M. (2005). Exploring lived experience. *The Psychologist*, *18*(1), 20–23.

Reisberg, D. (2007). *Cognition : exploring the science of the mind*. W. W. Norton.

Rickards, J. P., & Friedman, F. (1978). The encoding versus the external storage hypothesis in note taking. *Contemporary Educational Psychology*, *3*(2), 136-143. doi.org/10.1016/0361-476X(78)90020-6

Rizzo, L., Dondio, P., Delany, S. J., & Longo, L. (2016, September). Modeling mental workload via rule-based expert system: a comparison with NASA-TLX and workload profile. In *IFIP International Conference on Artificial Intelligence Applications and Innovations* (pp. 215-229). doi.org/10.1007/978-3-319-44944-9_19

Rosen, V. M., & Engle, R. W. (1998). Working memory capacity and suppression. *Journal of memory and language*, *39*(3), 418-436. doi.org/10.1006/jmla.1998.2590

Salthouse, T. A. (2009). When does age-related cognitive decline begin?. Neurobiology of aging, 30(4), 507-514.doi.org/10.1016/j.neurobiolaging.2008.09.023

Saywitz, K. J., Larson, R. P., Hobbs, S. D., & Wells, C. R. (2015). Developing rapport with children in forensic interviews: Systematic review of experimental research. *Behavioral Sciences & the Law*, *33*(4), 372-389. DOI: 10.1002/bsl.2186

Shepherd, E. (2007). *Investigative interviewing: The conversation management approach*. Oxford University Press.

Shepherd, E., & Kite, F. (1988). Training to interview. Policing, 4(4), 264-280.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. Psychological review, 84(1), 1. doi.org/10.1037/0033-295X.84.1.1

Schnotz, W., & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational psychology review*, *19*(4), 469-508. doi.org/10.1007/s10648-007-9053-4

Schreiber Compo, N., Hyman Gregory, A., & Fisher, R. (2012). Interviewing behaviors in police investigators: A field study of a current US sample. *Psychology, Crime & Law*, *18*(4), 359-375. https://doi.org/10.1080/1068316X.2010.494604

Smith, J. A., Flowers, P., & Larkin, M. (2009). *Interpretative phenomenological analysis: theory, method and research*. Sage.

Snook, B., & Keating, K. (2011). A field study of adult witness interviewing practices in a Canadian police organization. *Legal and Criminological Psychology*, *16*(1), 160-172. doi.org/10.1348/135532510X497258

Starks, H., & Brown Trinidad, S. (2007). Choose your method: A comparison of phenomenology, discourse analysis, and grounded theory. *Qualitative health research*, *17*(10), 1372-1380. https://doi.org/10.1177/1049732307307031

Strayer, D. L., & Drews, F. A. (2007). Cell-phone–induced driver distraction. *Current Directions in Psychological Science*, *16*(3), 128-131. https://doi.org/10.1111/j.1467-8721.2007.00489.x

St-Yves, M., & Meissner, C. A. (2014). Interviewing suspects. *Investigative interviewing: The essentials*. Thomson Reuters Ltd.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science*, *12*(2), 257-285. doi.org/10.1016/0364-0213(88)90023-7

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction*, *4*(4), 295-312. doi.org/10.1016/0959-4752(94)90003-5

Sweller, J. (2016). Working memory, long-term memory, and instructional design. *Journal of Applied Research in Memory and Cognition*, *5*(4), 360-367. https://doi.org/10.1016/j.jarmac.2015.12.002

Sweller, J., Van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, *10*(3), 251–296. doi.org/10.1023/a:1022193728205

Thompson, V. A., & Campbell, J. I. (2004). A power struggle: Between-vs. within-subjects designs in deductive reasoning research. *Psychologia*, *47*(4), 277-296. doi.org/10.2117/psysoc.2004.277

Thorley, C., Baxter, R. E., & Lorek, J. (2016). The impact of note taking style and note availability at retrieval on mock jurors' recall and recognition of trial information. *Memory*, *24*(4), 560-574. doi.org/10.1080/09658211.2015.1031250

Tulving, E. (1985). Memory and consciousness. Canadian Psychology/Psychologie canadienne, 26(1), 1-12. http://dx.doi.org/10.1037/h0080017

Turner III, D. W. (2010). Qualitative interview design: A practical guide for novice investigators. *The qualitative report*, *15*(3), 754. DOI. 10.46743/2160-3715/2010.1178

Unsworth, N. (2009). Variation in working memory capacity, fluid intelligence, and episodic recall: A latent variable examination of differences in the dynamics of free recall. *Memory & Cognition*, *37*(6), 837-849. doi.org/10.3758/MC.37.6.837

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2013). Working memory capacity and retrieval from long-term memory: The role of controlled search. *Memory & Cognition*, *41*(2), 242-254. doi.org/10.3758/s13421-012-0261-x

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: an examination of simple and complex span and their relation to higher order abilities. *Psychological bulletin*, *133*(6), 1038. doi.org/10.1037/0033-2909.133.6.1038

Vallano, J. P., & Compo, N. S. (2011). A comfortable witness is a good witness: Rapport-building and susceptibility to misinformation in an investigative mock-crime interview. *Applied cognitive psychology*, *25*(6), 960-970. https://doi.org/10.1002/acp.1789

Van Acker, B. B., Parmentier, D. D., Vlerick, P., & Saldien, J. (2018). Understanding mental workload: from a clarifying concept analysis toward an implementable framework. *Cognition, Technology and Work*, *20*(3), 351–365. doi.org/10.1007/s10111-018-0481-3

Van Merriënboer, J. J., & Sweller, J. (2010). Cognitive load theory in health professional education: design principles and strategies. *Medical education*, *44*(1), 85-93. https://doi.org/10.1111/j.1365-2923.2009.03498.x

Warren, A. R., & Woodall, C. E. (1999). The reliability of hearsay testimony: How well do interviewers recall their interviews with children? *Psychology, Public Policy, and Law*, *5*(2), 355. https://doi.org/10.1037/1076-8971.5.2.355

Willén, R. M., Granhag, P. A., Strömwall, L. A., & Fisher, R. P. (2015). Facilitating particularization of repeated similar events with context-specific cues. *Scandinavian journal of psychology*, *56*(1), 28-37. https://doi.org/10.1111/sjop.12180

Williams, H. L., & Lindsay, D. S. (2019). Different definitions of the nonrecollection-based response option(s) change how people use the "remember" response in the remember/know paradigm. *Memory and Cognition*, *47*(7), 1359–1374. https://doi.org/10.3758/s13421-019-00938-0

Wright, A. M., & Holliday, R. E. (2007). Enhancing the recall of young, young–old and old–old adults with cognitive interviews. *Applied Cognitive Psychology*, *21*(1), 19-43. doi.org/10.1002/acp.1260

Wright, R., & Powell, M. B. (2006). Investigative interviewers' perceptions of their difficulty in adhering to open-ended questions with child witnesses. *International Journal of Police Science & Management*, *8*(4), 316-325. https://doi.org/10.1350/ijps.2006.8.4.316

Yonelinas A.P. (2002). The nature of recollection and familiarity: A review of 30 years of research. Journal of Memory & Language. 46. pp 441–517. doi:10.1006/jmla.2002.2864

Yurko, Y., Scrbo, M. W., Prabhu, A. S., Acker, C. E., & Stefanidis, D. (2010). Higher mental workload is associated with poorer laparoscopic performance as measured by the NASA-TLX tool. *Simulation in Heathcare. 5(5),* 267-271. doi: 10.1097/SIH.0b013e3181e3f329.

Zajac, R., & Brown, D. A. (2018). Conducting successful memory interviews with children. *Child and Adolescent Social Work Journal*, *35*(3), 297-308. doi.org/10.1007/s10560-017-0527-z

Appendices

List of appendices

Appendix A: Supplementary materials (Chapter 2: Field study)

Appendix B: Supplementary materials (Chapter 3: Laboratory study 1)

Appendix C: Supplementary materials (Chapter 4: Laboratory study 2)

Appendix D: Supplementary materials (Chapter 5: Laboratory study 3)

Appendix E: Supplementary materials (Chapter 6: Online field survey)

Appendix F: Ethical approval (Chapters 2, 3, 4, 5, 6 and UPR16 form)

Appendix A: Supplementary materials (Chapter 2: Field study)

A.1 Interview protocol

A.2 Sample interview transcript

A.1 Interview protocol.

Welcome and thank you

Thank you for agreeing to take part in this research project. As outlined in the Participant Information Sheet, my research is about police officers' experiences when conducting interviews with children and other vulnerable witnesses using the guidelines outlined in the Achieving Best evidence protocols (2011). I am interested in this because gaining an understanding from the practitioner's perspective may inform further research and improve training for police officers in this field.

Do you have any questions?

Reiterate anonymity, data security, right to withdraw and limits of confidentiality

Provide consent form to sign - Provide demographics sheet for completion

Check interviewee happy to start - Switch on recorder

This interview is about your views when you are conducting ABE interviews. I am really interested in your thoughts and your experiences when you are interviewing. So like you inform witnesses during the ABE interview, you hold the information I am interested in and you are the expert. There are no right or wrong answers, but I'd like you to give me as much information as you can about what ABE interviewing is like for you.

To start, please can tell me a little about how you came to be an ABE trained interviewer.

What is it like for you, when you conduct interviews with vulnerable witnesses?

Tell me what the initial part of an interview is like from your perspective.

What is it like for you when you ask open questions?

What is it like when you have planned and prepared to conduct the interviews? And tell me what it is like for you if you don't plan and prepare.

Tell me more about the information you have before the interview and how you use that during the interview.

If you could think back to an interview that stands out for you, can you tell me about that?

Prompts

How does that impact you and the interview?

Can you tell me more about that?

Could you give me an example of…?

What did you mean by…?

How did that make you feel?

What strategies do you have to mitigate...?

The following areas will be considered for follow-up questions to elicit further information;

Cognition – Task of interviewing, thinking, processing, decision making

Mental Effort - Level of effort when conducting interviews

Temporal demand – Time pressure during your interviews

Physical effects – Tiredness, frustration, concentration

Planning and preparation – How much time, how often not enough

Ending the interview:

And how long you have been conducting ABE interviews in the course of your

investigations?

How many interviews do you generally conduct, weekly or monthly?

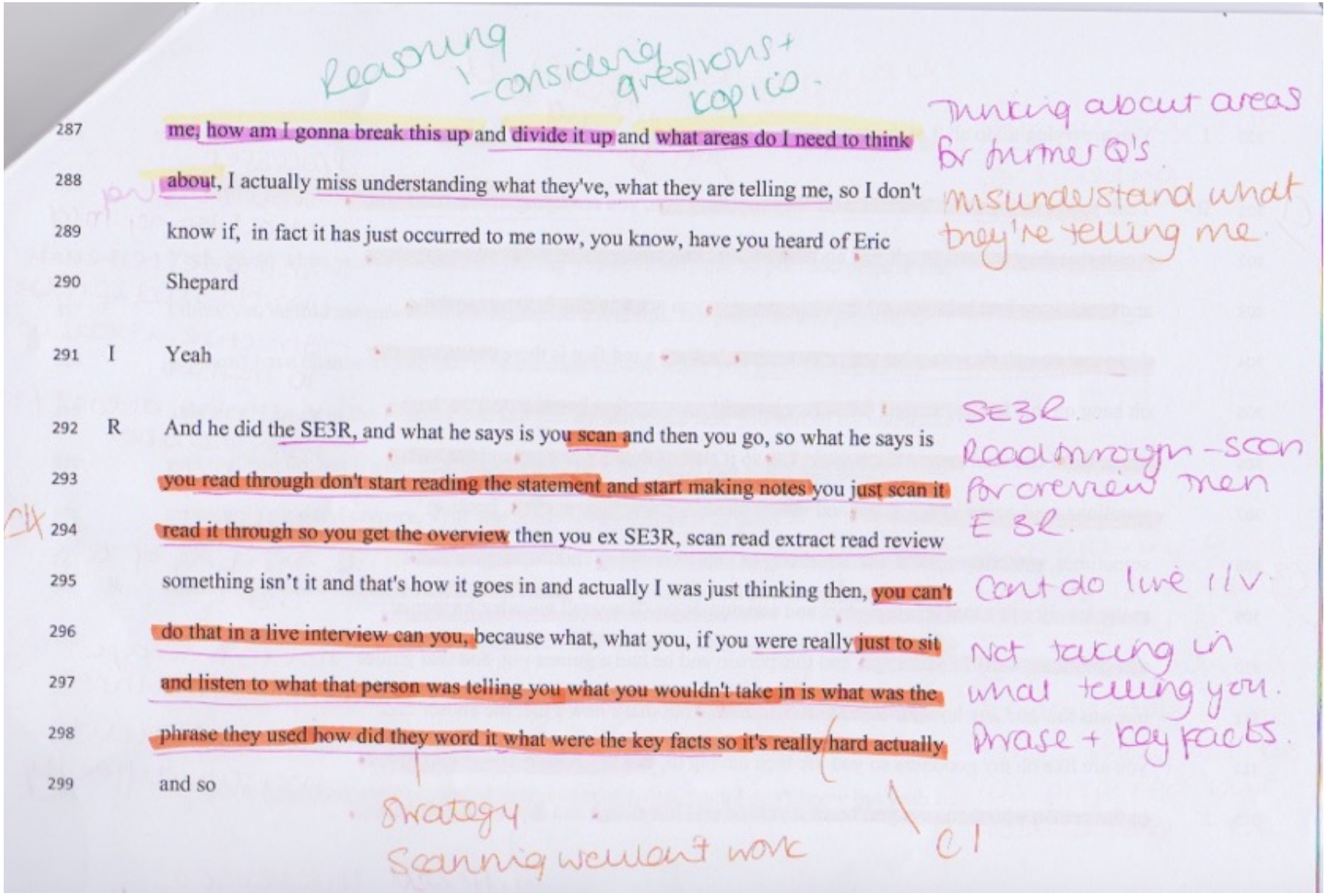
Is there anything you would like to add?

Do you have any questions?

Are you OK to finish?

Switch off recorder

Verbal debrief



Reasoning - considering questions + topics.

Thinking about areas for further Q's misunderstand what they're telling me.

SE3R. Read through - scan for overview then E 3R.

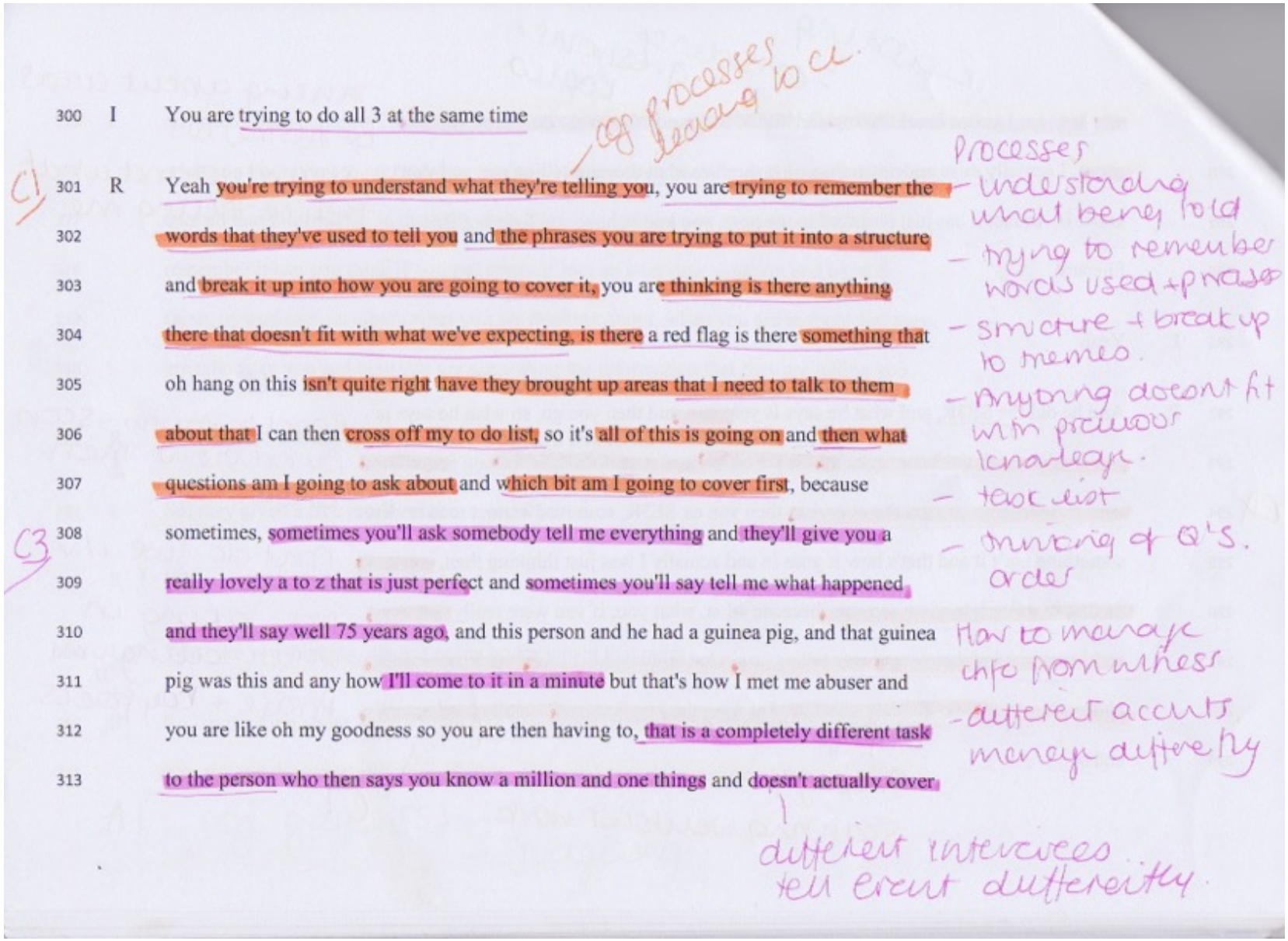
Can't do live i/v.

Not taking in what telling you. Phrase + key facts.

Strategy Scanning wouldn't work

CI

287 me, how am I gonna break this up and divide it up and what areas do I need to think
288 about, I actually miss understanding what they've, what they are telling me, so I don't
289 know if, in fact it has just occurred to me now, you know, have you heard of Eric
290 Shepard
291 I Yeah
292 R And he did the SE3R, and what he says is you scan and then you go, so what he says is
293 you read through don't start reading the statement and start making notes you just scan it
294 read it through so you get the overview then you ex SE3R, scan read extract read review
295 something isn't it and that's how it goes in and actually I was just thinking then, you can't
296 do that in a live interview can you, because what, what you, if you were really just to sit
297 and listen to what that person was telling you what you wouldn't take in is what was the
298 phrase they used how did they word it what were the key facts so it's really hard actually
299 and so



300 I You are trying to do all 3 at the same time

301 R Yeah you're trying to understand what they're telling you, you are trying to remember the
302 words that they've used to tell you and the phrases you are trying to put it into a structure
303 and break it up into how you are going to cover it, you are thinking is there anything
304 there that doesn't fit with what we've expecting, is there a red flag is there something that
305 oh hang on this isn't quite right have they brought up areas that I need to talk to them
306 about that I can then cross off my to do list, so it's all of this is going on and then what
307 questions am I going to ask about and which bit am I going to cover first, because
308 sometimes, sometimes you'll ask somebody tell me everything and they'll give you a
309 really lovely a to z that is just perfect and sometimes you'll say tell me what happened
310 and they'll say well 75 years ago, and this person and he had a guinea pig, and that guinea
311 pig was this and any how I'll come to it in a minute but that's how I met me abuser and
312 you are like oh my goodness so you are then having to, that is a completely different task
313 to the person who then says you know a million and one things and doesn't actually cover

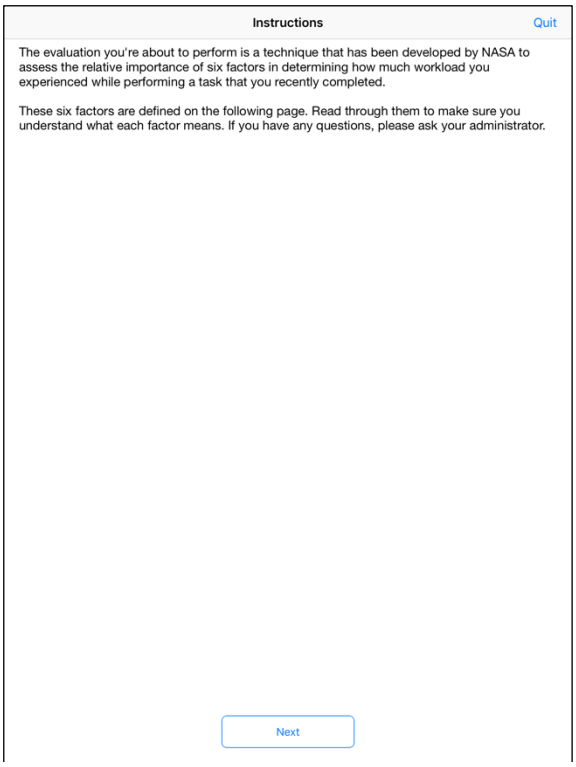
Appendix B: Supplementary materials (Chapter 3: Laboratory study 1)

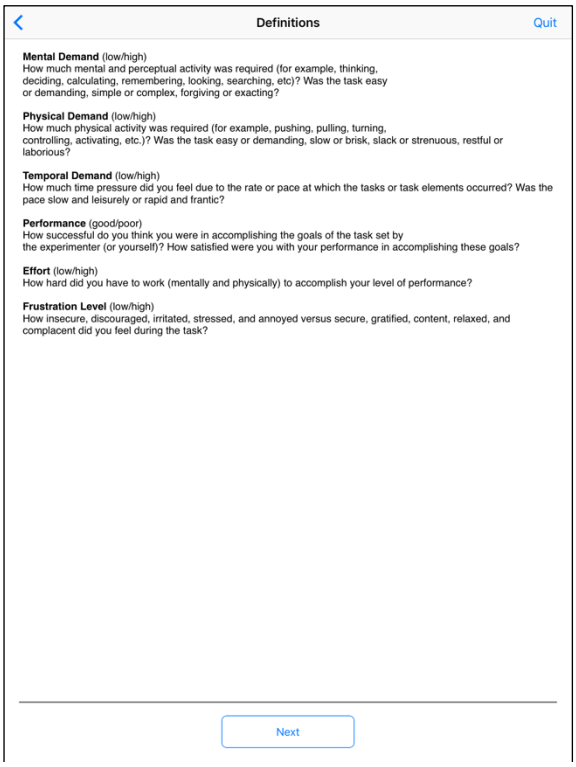
B.1 NASA – TLX app examples

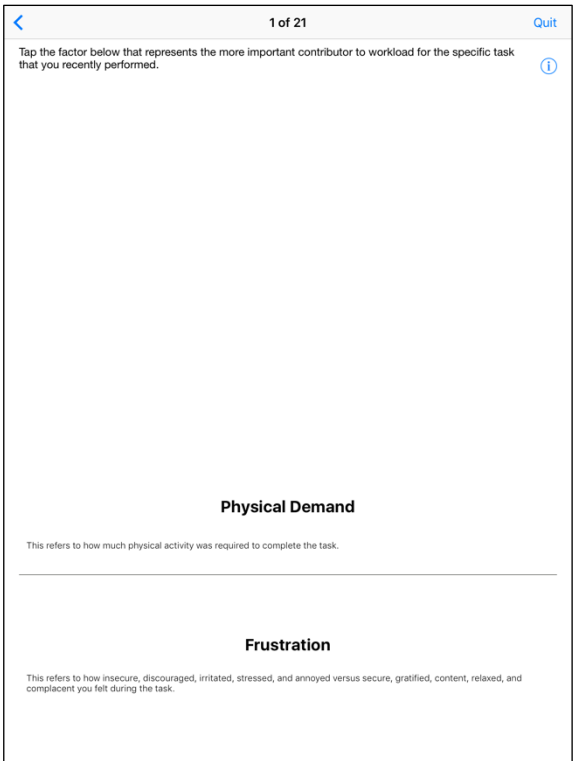
B.2 Cued recall questions

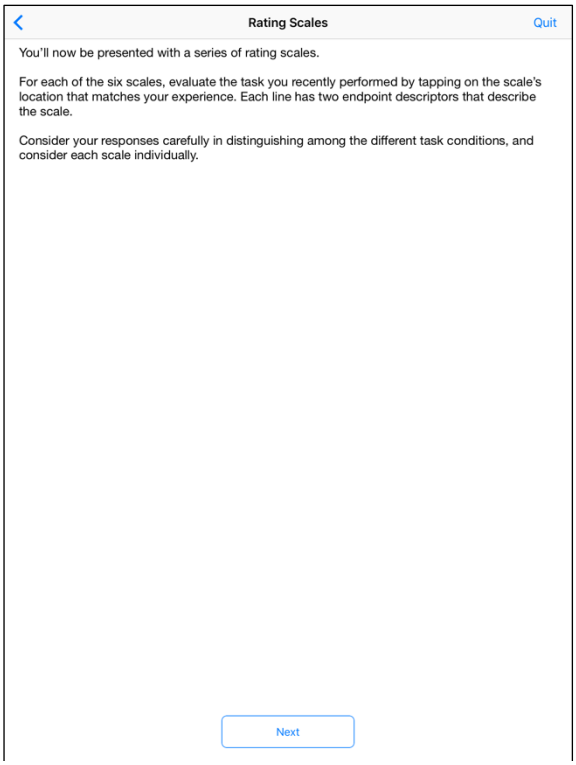
B.3 Analyses: Free and cued recall accuracy scores; confidence, motivation, and task difficulty correlations; questionnaire scores.

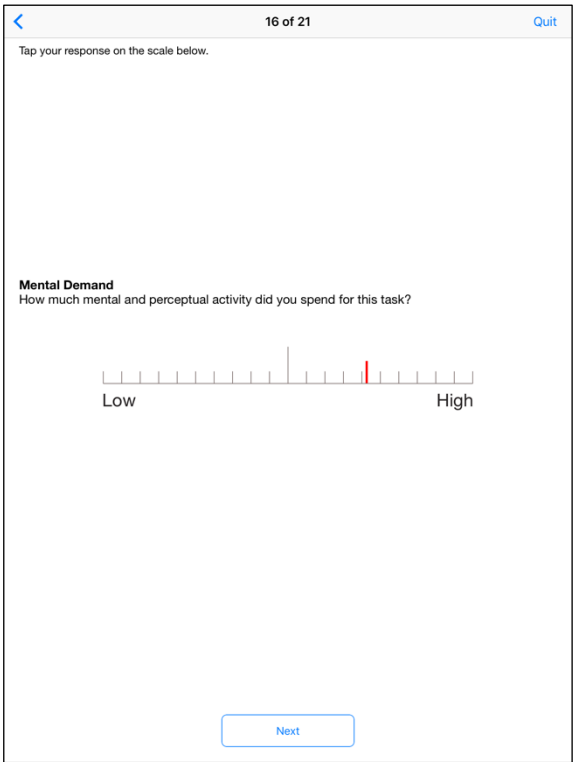
B.1 NASA-TLX app examples

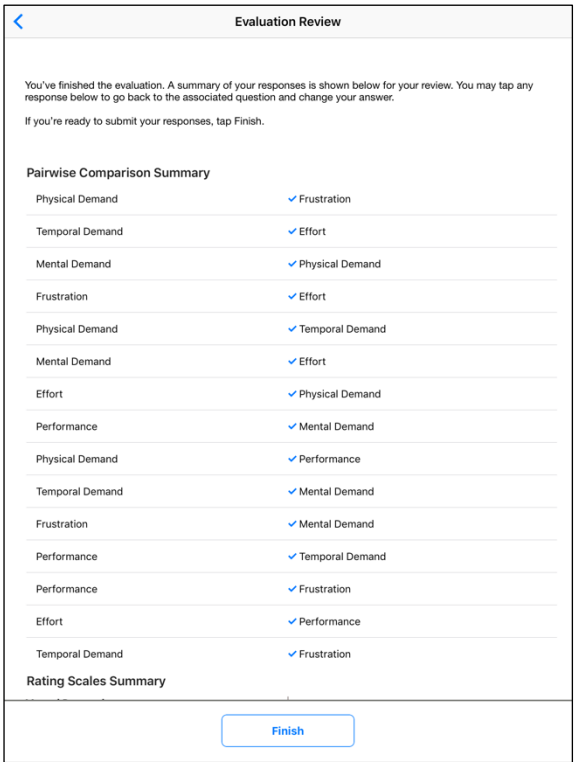









16 of 21
Quit
Tap your response on the scale below.
Mental Demand
How much mental and perceptual activity did you spend for this task?
Low
High
Next


Evaluation Review
You've finished the evaluation. A summary of your responses is shown below for your review. You may tap any response below to go back to the associated question and change your answer.
If you're ready to submit your responses, tap Finish.
Pairwise Comparison Summary
Physical Demand ✓ Frustration
Temporal Demand ✓ Effort
Mental Demand ✓ Physical Demand
Frustration ✓ Effort
Physical Demand ✓ Temporal Demand
Mental Demand ✓ Effort
Effort ✓ Physical Demand
Performance ✓ Mental Demand
Physical Demand ✓ Performance
Temporal Demand ✓ Mental Demand
Frustration ✓ Mental Demand
Performance ✓ Temporal Demand
Performance ✓ Frustration
Effort ✓ Performance
Temporal Demand ✓ Frustration
Rating Scales Summary
Finish

B.2 Cued recall questions

Questions and answers from stimulus video

1 What colour was the witness's hair?
Blonde/light brown

2 How did the witness describe the horses when she arrived?
A bit smelly

3 What kind of party was it?
Horse riding

4 What was the name of the place the witness said she went to?
Pink Mead Farm

5 Where did the witness go when she first arrived?
Reception

6 What was the witness told she was not allowed to wear?
Flat heeled boots

7 What was in the stables when they first went in?
Some horses but still more to come

8 What did they all do when a big horse came in?
All gasped when a big one came in

9 What was the name of the birthday girl?
Tahlia

10 What was the name of the pony of the person whose birthday it was?
Tofty

11 What was Amelia's horse called?
Storm

12 How did the witness describe the horse called Storm?
A small pony

13 What was the name of the witness' pony?
Whisper

14 What did the witness say they did after they first got on their ponies?
Had a little slow walk around the stables just to get used to being on a horse

15 What was the name the witness gave to the first game they played?
Like musical statues but you were on a horse

16 What did the lady shout to make them stop the horses?
Red

17 What did they do with the reins to make the horse stop?
Pull back

18 What did the signal amber mean?
To do a slow walk

19 What did it mean when the lady shouted green?
Do a trot

20 How many times did the witness say she kicked the pony to make it trot?
3 times

21 Where did the witness say she had to kick the pony to make it trot?
On its side

22 What did the witness say was quite tricky?
Stand up and sit down

23 Why did the witness say it was OK if you fell off?
They had a helmet

24 Why did the witness say some of the others were scared or worried?
They were going to fall off
25 Why did the witness say you wouldn't fall off?
Because you're holding the reins
26 What was the colour of the instruction to stop the horses?
Red
27 What did they have to do if the horse didn't stop on the red command?
Horse and you go into the middle
28 What number did the witness say she was out in the first game?
Fourth
29 Who won the first game?
Tahlia
30 What did they do after the first game?
We had a few races
31 What were they rewarded with if they cheered on the horses?
Extra points
32 What was the witness wearing?
Stripey top, legging/jeans and brown boots
33 What did the birthday girl do when the others were in the middle?
Trotted around the stable while they sang happy birthday to her
34 What did the others do when the birthday girl was trotting round them?
Went into the middle of the stable
35 What did they do after singing happy birthday the second time?
The lady handed out prizes
36 Who had the purple rosette?
Tahlia
37 What colour rosettes did everybody except the birthday girl have?
Pink
38 What did the witness say she did when she went to the reception?
We got our helmets on and I put my boots back
39 What type of birthday cake did they have?
Unicorn
40 Who drove the witness home?
Amelia and her mum

B.3 Analyses: Free and cued recall accuracy scores; confidence, motivation, and task difficulty correlations; questionnaire scores.

The total number of free recall details, and cued recall accuracy scores, for each condition are reported in the main analyses (Chapter 3). For mean, standard deviation, and confidence interval, scores for free recall correct, incorrect, confabulations and ambiguity, see Table 1, and for cued recall correct, partially correct, incorrect, and don't know scores, see Table 2.

Table 1.

Free recall mean, SD, and CI, scores for correct, incorrect, confabulations, and ambiguity, for each condition.

	Condition	*M* (*SD*)	95% CI
Correct	HCL	106.29 (42.50)	[92.31, 120.28]
	MCL	127.09 (40.09)	[115.80, 138.38]
	NCL	127.84 (51.28)	[109.36, 146.33]
Incorrect	HCL	5.94 (3.05)	[4.88, 7.01]
	MCL	6.42 (3.59)	[5.15, 7.70]
	NCL	4.59 (3.17)	[3.45, 5.74]
Confabulation	HCL	4.26 (5.17)	[2.46, 6.07]
	MCL	3.76 (3.87)	[2.38, 5.13]
	NCL	1.59 (2.31)	[.76, 2.42]
Ambiguity	HCL	5.38 (4.12)	[3.95, 6.82]
	MCL	5.21 (3.87)	[3.84, 6.58]
	NCL	4.47 (3.47)	[3.22, 5.72]

Table 2.

Cued-recall mean, SD, and CI, scores for correct, partially correct, incorrect, and don't know, for each condition.

	Condition	*M* (*SD*)	95% CI
Correct	HCL	15.82 (5.55)	[13.87, 17.06]
	MCL	18.73 (5.67)	[16.72, 20.74]
	NCL	22.06 (6.14)	[19.85, 24.28]
Partially correct	HCL	7.47 (2.05)	[6.67, 8.19]
	MCL	7.27 (2.71)	[6.13, 8.23]
	NCL	6.13 (1.83)	[5.47, 6.78]
Incorrect	HCL	8.26 (3.73)	[6.96, 9.57]
	MCL	6.94 (4.01)	[5.52, 8.36]
	NCL	6.31 (3.51)	[5.05, 7.58]
Don't know	HCL	8.29 (3.73)	[6.42, 10.16]
	MCL	6.85 (4.97	[5.09. 8.61]
	NCL	5.41 (3.81)	[4.03, 6.78]

A series of Pearson's correlations were calculated to determine whether the dependent variables of motivation, confidence and task difficulty were correlated with each other. There were significant, but moderate, correlations between the majority of variables (see Table 3).

Table 3.

Pearson correlations, Means and Standard Deviations associated with confidence motivation and task difficulty.

	Confidence in memory accuracy	Motivation to remember the account	Ease of remembering the account	Ease of thinking of questions	Mean	SD
Confidence in memory accuracy	1				4.51	1.11
Motivation to remember the account	.325**	1			5.67	1.15
Ease of remembering the account	-.398**	-.246*	1		4.58	1.08
Ease of thinking of questions	-.247*	-.137	.302**	1	4.96	1.45

* correlation is significant at the 0.5 level; ** correlation is significant as the .01 level

Questionnaire scores.

Scores, on 7-point scale, were obtained for; confidence in memory accuracy; motivation to remember the content of the interview; ease of remembering the content of the child's statement; and ease of coming up with questions. In addition, for participants in the HCL condition, how motivated they were to think about questions whilst listening to the child's statement (see Table 4 for mean, *SD* and *CI* scores for each question).

Table 4.

Questionnaire mean, SD and CI scores for each condition

Question	Condition	*M* (*SD*)	95% CI
Confidence in memory accuracy	HCL	4.24 (1.42)	[3.74, 4.73]
	MCL	4.61 (.93)	[4.28, 4.94]
	NCL	4.69 (.86)	[4.34, 5.00]
Motivation to remember account	HCL	5.59 (1.28)	[5.14, 6.04]
	MCL	5.76 (1.12)	[5.36, 6.15]
	NCL	5.66 (1.07)	[5.27, 6.04]
Ease of remembering account	HCL	4.74 (1.02)	[4.38, 5.09]
	MCL	4.36 (1.14)	[3.96, 4.77]
	NCL	4.63 (1.07)	[4.24, 5.01]
Ease of coming up with questions	HCL	5.03 (1.27)	[4.59, 5.47]
	MCL	4.82 (1.76)	[4.19, 5.44]
	NCL	5.03 (1.31)	[4.56, 5.50]
Motivation to think of questions	HCL	5.24 (1.37)	[4.76, 5.71]

Appendix C: Supplementary materials (Chapter 4: Laboratory study 2)

C.1 Preregistration

C.2 Photographs of the five witnesses

C.3 Matrix for unique details given by each witness

C.4 20 unique detail questions

C.5 Analyses: Confidence, motivation, and task difficulty correlations.

C.1 Preregistration

AS PREDICTED

Wharton CREDIBILITY LAB
University of Pennsylvania

CONFIDENTIAL - FOR PEER-REVIEW ONLY

Cognitive load for interviewers - multiple sources of information. (#34068)

Created: 01/17/2020 08:32 AM (PT)
Shared: 10/28/2020 03:48 PM (PT)

This pre-registration is not yet public. This anonymized copy (without author names) was created by the author(s) to use during peer-review.
A non-anonymized version (containing author names) will become publicly available only if an author makes it public. Until that happens the contents of this pre-registration are confidential.

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
The main goal of this research is to examine the effects of increased cognitive demands on investigative interviewers' perceived cognitive load, their recall of information provided in the accounts of multiple witnesses, and their source monitoring errors. Specifically, we will examine the following questions: What are the effects of increased cognitive demands on the interviewers', i) perceived cognitive load, ii) accuracy of recall of multiple witnesses' accounts, iii) subjective experience of recalling information from multiple witnesses, indicated by remember (recollection), know (familiar), or guess, responses, iv) source monitoring errors when identifying which witness provided the information, and v) subjective experience of recalling which witness provided the information, indicated by remember (recollection), know (familiar), or guess responses? We predict that, in a high cognitive load (HCL) condition, participants, who will be asked to formulate questions in their heads whilst watching multiple interviews, will report increased levels of cognitive load when compared with those in a no cognitive load (NCL) condition, who will be given instructions to merely watch and listen to the accounts. We also predict that participants in the HCL condition will have lower accuracy scores, for their recall of details about the witnesses' accounts, when compared with those in the NCL condition. For interviewers' subjective experience of recall, we predict that in the HCL condition, participants will report i) a lower proportion of remember responses, ii) a lower proportion of know responses, and iii) a higher proportion of guess responses, when compared with those in the NCL condition. Further, participants in the HCL condition will make an increased number of source monitoring errors compared with those in the NCL condition. For interviewers' subjective experience of source monitoring, participants in the HCL condition will also report i) a lower proportion of remember responses, ii) a lower proportion of know responses, and iii) a higher proportion of guess responses, when compared with those in the NCL condition.

3) Describe the key dependent variable(s) specifying how they will be measured.
The dependent variables will be: 1) Perceived cognitive load. The participants' self-reported cognitive load when performing the interview observation task (as measured by the NASA-TLX). 2) Accuracy of answers to 20 forced-choice recognition questions, based on information presented in the witnesses' accounts. Accuracy will be measured as the proportion of correct responses to the questions. 3) Participants' subjective experience of recall will be measured as the proportions of remember, know, and guess (R/K/G), responses for their answers to the 20 forced-choice recognition questions. 4) Accuracy of source monitoring will be measured as the proportion of correct answers to 20 source of information questions. 5) Participants' subjective experience of source monitoring will be measured as the proportions of R/K/G, responses for their answers to the sources of information questions.

4) How many and which conditions will participants be assigned to?
There will be one independent variable of cognitive load, with two levels: no cognitive load (NCL) and high cognitive load (HCL). To assign equal numbers of participants per condition, participants will be pseudo-randomly assigned to one of the conditions.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
To test our hypotheses, independent t-tests will be used. The IV will be cognitive load: NCL vs. HCL. The dependent variables will be: 1) PCL, 2) Accuracy of recall (i.e. the proportion of accurate responses to 20 forced-choice recognition questions), 3) Subjective experience of recall (i.e., the proportion of R/K/G responses for their 20 forced-choice recognition answers), 4) Accuracy of source monitoring (i.e. the proportion of accurate responses to the source of information questions), and 5) Subjective experience of source monitoring (i.e. the proportion of R/K/G responses for their source of information answers).

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
Data will be checked for missing values and outliers. Participants' data will only be excluded if it is clear that the data has been recorded incorrectly (e.g. equipment failure). As outliers may affect some analyses, any data that is an extreme outlier (i.e. beyond 3 SDs from the mean) will be removed. Descriptive statistics will be reported pre and post removal of data.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
102 participants will be recruited. G*power analysis for a one-way t-test, with two independent groups based on alpha of 0.05, power of 0.80 and a medium effect size of .5, gives a desired sample size of 102.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
Nothing else to pre-register.

C.2 Photographs of the five witnesses

WITNESS 1

WITNESS 2

WITNESS 3

WITNESS 4

WITNESS 5

C.3 Matrix for unique details

Unique detail/detail	Witness 1	Witness 2	Witness 3	Witness 4	Witness 5
1	Four friends	Friends	Friends	Friends	Friends
2	Birthday	Megan's birthday	Birthday	Birthday	Birthday
3	Nightclub	Nightclub	Clouds	Nightclub	Nightclub
4	In town	In town	In town	In Brighton	In town
5	Few drinks	Few drinks	Few drinks	Few drinks	Few vodkas
6	Chloe	Friend	Friend	Friend	Friend
7	Shouting	Ollie shouting	Shouting	Shouting	Shouting
8	Short hair	Short hair	Short blonde hair	Short hair	Short hair
9	Sequined top	Sequined top	Sequined top	Red sequined top	Sequined top
10	Punched	Punched	Punched	Punched	Punched in the face
11	Left side	Side	Side	Side	Side
12	Knife	Hunting Knife	Knife	Knife	Knife
13	Exit	Exit	Fire exit	Exit	Exit
14	Jacket	Jacket	Jacket	Black jacket	Jacket
15	Friend	Friend	Friend	Friend	Jenny
16	Sainsburys	Shop	Shop	Shop	Shop
17	Pocket	Jacket pocket	Pocket	Pocket	Pocket
18	Baseball cap	Baseball cap	White baseball cap	Baseball cap	Baseball cap
19	Trousers	Trousers	Trousers	Blue Trousers	Trousers
20	Tattoo	Tattoo	Tattoo	Tattoo	Eagle tattoo

C.4 20 unique detail questions

Q1 How many friends did the witness say she went out with?
3 (1)
4 (2) ✓
5 (3)
6 (4)

Q2 Who was celebrating their birthday?
Chloe (1)
Emily (2)
Maddie (3)
Megan (4) ✓

Q3 What was the name of the nightclub the witness went to?
Clouds (1) ✓
Heaven (2)
Envy (3)
Scandals (4)

Q4 Which town was the nightclub in?
Portsmouth (1)
Southampton (2)
Brighton (3) ✓
Basingstoke (4)

Q5 What did the witness say she had been drinking at the nightclub?
Wine (1)
Gin (2)
Prosecco (3)
Vodka (4) ✓

Q6 Which friend did the witness say she was dancing with?

Jenny (1)
Chloe (2) ✓
Megan (3)
Cassie (4)

Q7 What name did the witness say she heard during the shouting?
Jack (1)
Thomas (2)
Harry (3)
Ollie (4) ✓

Q8 What colour hair did the woman who was attacked have?
Red (1)
Black (2)
Blonde (3) ✓
Brown (4)

Q9 What colour sequined top was the woman who was attacked wearing?
Blue (1)
White (2)
Black (3)
Red (4) ✓

Q10 Where on her body did the man punch the woman?
Leg (1)
Face (2) ✓
Arm (3)
Chest (4)

Q11 The woman was cut and bleeding from an injury to which part of her body?
Left side of head (1) ✓
Right side of head (2)
Right arm (3)
Left arm (4)

Q12 What type of knife did the witness say the man had?
Flick knife (1)
Kitchen knife (2)
Hunting knife (3) ✓
Pen knife (4)

Q13 Which door did the witness say she left the nightclub through?
Fire exit (1) ✓
Rear exit (2)
Front exit (3)
Side exit (4)

Q14 What was the woman holding over her head to protect herself?
Brown jacket (1)
Black jacket (2) ✓
Blue jacket (3)
Pink jacket (4)

Q15 Which friend did the witness say opened the exit door?
Jenny (1) ✓
Maddie (2)
Cassie (3)
Chloe (4)

Q16 Which shop did the witness say the man ran towards?
Tesco's (1)
Co-op (2)
Waitrose (3)
Sainsbury's (4) ✓

Q17 Into which pocket did the witness say the man put the knife as he ran off?
Trouser (1)
Shirt (2)
Jacket (3) ✓
Hoodie (4)

Q18 What colour baseball cap was the man wearing?
White (1) ✓
Blue (2)
Red (3)
Black (4)

Q19 What colour trousers was the man wearing?
Red (1)
Black (2)
Blue (3) ✓
Grey (4)

Q20 What tattoo did the man have on his neck?
Raven (1)
Eagle (2) ✓
Swallow (3)
Hawk (4)

C.5 Analyses: Confidence, motivation, and task difficulty correlations.

Confidence, ease of remembering, and motivation

For experiment 2 (Chapter 4), the post recall questionnaire dependent variables were confidence, ease of remembering the witnesses' accounts, motivation to remember the accounts. A series of Pearson's correlations were conducted to determine whether the dependent variable composite scores were correlated with each other (see Table 1).

Table 1.

Pearson correlations, Means and Standard Deviations associated with confidence motivation and task difficulty.

	Confidence in memory accuracy	Motivation to remember the accounts	Ease of remembering the accounts	Motivation to think of questions	Mean	SD
Confidence in memory accuracy	1				4.58	1.02
Motivation to remember the accounts	.167	1			5.28	.94
Ease of remembering the accounts	-.475**	-.163	1		3.65	1.11
Motivation to think of questions	-.134	-.335*	.179	1	5.56	1.00

* Correlation is significant at the 0.5 level; ** correlation is significant as the .01 level

Appendix D: Supplementary materials (Chapter 5: Laboratory study 3)

D.1 Preregistration

D.2 Examples of Working Memory Capacity tests

D.3 Structured notes template

D.4 Cued recall questions

D.5 Analyses: Descriptive statistics pre-removal of data; AVOVAs for free and cued recall accuracy; Access to notes *t*-tests; Pearson's correlations and MANOVA results for confidence, motivation and ease of remembering the questions.

D.1 Preregistration

CONFIDENTIAL - FOR PEER-REVIEW ONLY

ASPREDICTED

Effects of cognitive load on investigative interviewers' performance. (#19188)

Created: 02/01/2019 08:24 AM (PT)
Shared: 01/07/2020 08:40 AM (PT)

This pre-registration is not yet public. This anonymized copy (without author names) was created by the author(s) to use during peer-review.
A non-anonymized version (containing author names) will become publicly available only if an author makes it public. Until that happens the contents of this pre-registration are confidential.

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
1) What is the effect of note taking on interviewers' perceived cognitive load (PCL) during investigative interviews? 2) What is the effect of note taking, and access to notes, on the amount and accuracy of information recalled by interviewers after an investigative interview? 3) What is the effect of note taking, and access to notes, on interviewers' PCL when asked to recall details of an investigative interview? The following hypotheses will be tested: 1) Participants in the Free Note Taking (FNT) condition will report increased levels of PCL for the interview observation task compared to those in the Structured Note Taking (SNT) condition, who in turn will report increased levels of PCL for the interview observation task than those in the No Note Taking (NNT) condition; 2) Participants in the NNT condition will report fewer details and will have lower percentage accuracy (for free recall and responses to questions) for their recall of the account given by the interviewee than those in the FNT condition, who in turn will report fewer details and will have lower percentage accuracy for their recall of the interviewee's account than those in the SNT condition; 3) Participants in the 'no access to notes' condition will report fewer details and have lower percentage accuracy for their recall of the account given by the interviewee than those in the 'access to notes' condition; 4) Participants in the NNT condition will report increased levels of PCL for the recall tasks compared to those in the FNT condition, who in turn will report increased levels of PCL for the recall tasks than those in the SNT condition 5) Participants in the 'no access to notes' condition will report increased levels of PCL for the recall of the interview task compared to those in the 'access to notes' condition.

3) Describe the key dependent variable(s) specifying how they will be measured.
The DVs will be; 1) Perceived Cognitive Load (PCL): Participants' PCL when performing the interview observation and recall tasks will be measured using the NASA-TLX, which is a subjective workload rating scale. 2) Amount and accuracy of details recalled: Participants will be asked a) to recall as accurately as possible what the witness described and b) 40 questions regarding the witness's account. The participants' a) free recall narrative, will be coded for amount of details and percentage accuracy of details recalled and b) answers to the 40 questions will be coded for percentage accuracy.

4) How many and which conditions will participants be assigned to?
This study will be a 2 x 2 between-participant design. The two IVs will be note-taking and access to notes, there will be a control condition of no note taking in the note-taking condition. Participants will be randomly allocated to one of the five conditions.

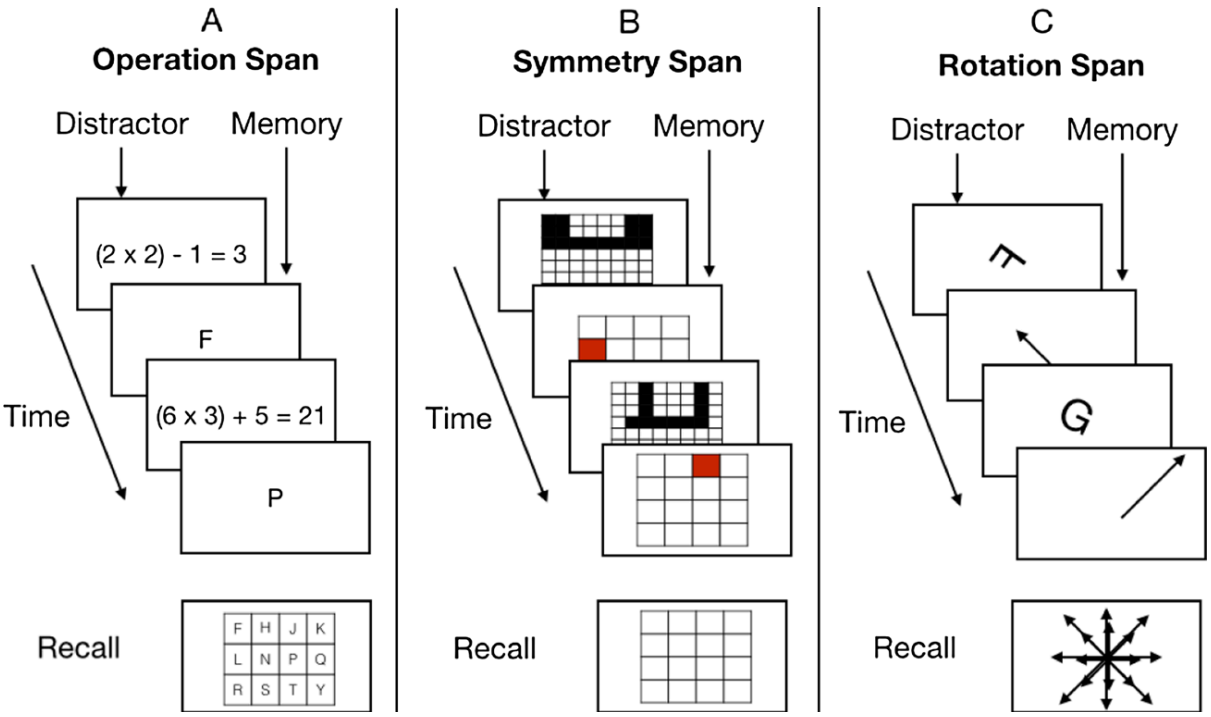
5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
To test our hypotheses, between group ANOVAs will be used. The DVs will be: 1) PCL when performing the interview and recall tasks. 2) Amount and accuracy of details recalled. The IVs will be note taking (NNT x FNT x SNT) and access to notes (access x no access). For the note-taking IV, we will use post-hoc comparisons to determine the nature of any between-group differences.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
Data will be checked for missing values and outliers. Participants' data will only be excluded if it is clear that the data has been recorded incorrectly, e.g. equipment failure. As some analyses may be affected by outliers, any data that is an extreme outlier, i.e. beyond 3 SDs from the mean, will be removed. Descriptive statistics will be reported pre and post removal of data.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
130 participants will be recruited. G*power analysis for a one-way ANOVA, with three groups based on alpha of 0.05, power of 0.80 and a medium effect size of .35, gives a desired sample size of 128 and for a two-way ANOVA with four groups based on alpha of 0.05, power of 0.80 and a medium effect size of.35, a desired sample size of 118.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
We will also investigate if working memory capacity (WMC) is a moderator of; 1) the effects of note taking on PCL during the interview observation task; 2) the effects of note taking on recall performance during the recall task; 3) the effects of access to notes on recall performance during the recall task; 4) the effects of note taking on PCL during the recall task and 5) the effects of access to notes on PCL during the recall task. Participants' WMC will be measured using complex span tasks (OSPAN, RotSPAN, SymSPAN). A series of linear regression models will be estimated to explore if there are any moderator effects. Separate coefficients for each regression equation will be estimated and tested to examine the effect of the IVs on the DVs, when the moderator (WMC) is introduced. To achieve .8 power in the moderation analysis, the sample size will need to be 77 for a medium effect size of .35, with alpha at .05.

D.2 Examples of Working Memory Capacity tests



Examples of the three WMC tests. Participants are presented with a distractor (e.g., for the operation span test, a simple mathematical test to which they respond true or false), they are then presented with the item to be remembered (e.g., a letter). The items to be remembered randomly span three to seven items before recall of the items, in order of presentation, is required (image from Fisher et al, 2015).

D.3 Structured notes template

People, e.g., names of individuals or groups and descriptions		Settings, e.g., descriptions of location, area or space and time
Objects, e.g., weapons, vehicles, inc. description of where and what		Actions, e.g., what happened, what someone doing

D.4 Cued recall questions

No	Question	Answer	Points
1	What time did the witness say he finished work?	5 o'clock	1
2	What was the name of the bar the witness went to for a drink?	The Knights Head	2
3	What was the name of the street the witness mentioned in his interview?	Sankey Street	1
4	What was the name of the train station the witness went to?	Sanford	1
5	What instrument was the busker was playing?	Guitar	1
6	The witness thought the busker was playing a song by which artist?	Ed Sheeran	1
7	What did the witness say the lady outside the train station was wearing?	Dress and coat	2
8	What was the man in the overcoat outside the train station doing?	Smoking a cigarette	2
9	What colour hair did the lady outside the train station have?	Blonde	1
10	What did the witness do while he was sitting and waiting for the train?	Watched the busker perform	2
11	What did the suspect initially shout at the busker?	Some nasty things - "You're not any good"	2
12	What did the witness initially think the suspect was going to do?	Steal the moneybox - take the cash	2
13	What did the suspect do when he first attacked the busker?	Pushed him on to the floor	2
14	What injury did the witness say the busker sustained during the incident?	Banged his head - Bleeding on the side	2
15	What was the suspect waving around while he was shouting?	A knife	1
16	What was the suspect shouting when the busker was on the floor?	"Get away from me"	2

17	How did the witness describe the busker when he was lying on the floor?	Motionless - he was just not moving	2
18	What was on the floor near the busker when he was playing?	A money box	2
19	How many times did the suspect kick the busker?	Twice	1
20	What did the witness say when he shouted at the suspect?	Get away from him or stop	2
21	What did the suspect do when the witness shouted at him?	Turned around and waved knife at him	2
22	Where did the witness tell the suspect the police where?	the police are coming, the police are on their way, they are round the corner	2
23	What did the suspect do when he heard the sirens?	He ran/fled	1
24	What did the suspect throw into a bin?	Big knife	2
25	What did the witness say he was doing while he was having a drink in the bar?	Chatting to people at the bar	2
26	What did the witness do to help the busker after the attack?	Sat him up against wall, rested him up against the wall	2
27	Who else initially arrived at the scene when the witness was helping the busker?	2 men	2
28	What skill did a person who initially came to help the busker have?	First aider	1
29	How long after the attack did the police arrive?	about 5 minutes	1
30	What height did the witness say the suspect was?	5'4"	2
31	How did the witness describe the colour of the suspect's hair?	Short dirty brown	2
32	What did the witness say the suspect was wearing?	Black jacket	2
33	Who did the witness say later came on the scene to check that the busker was OK?	Paramedic	1

34	How many statements did the witness say were obtained by the police?	Five	1
35	For how long did the witness say he was in the pub?	Few hours	2
36	What was the man outside the station wearing under his overcoat?	A suit	1
37	Where was the busker playing?	Outside train station	2
38	What colour was the man's overcoat?	Blue	1
39	Where did the suspect kick the busker?	On the leg	1
40	Why did the witness say he couldn't see the suspect's face?	He was waving knife about in front of his face	2

D.5 Analyses: Descriptive statistics pre-removal of data; AVOVAs for free and cued recall accuracy; Access to notes *t*-tests; Pearson's correlations and MANOVA results for confidence, motivation and ease of remembering the questions.

Data for Experiment 3 (Chapter 5) were checked for missing values and outliers. Any score that was an extreme outlier (i.e., beyond 3 SDs from the mean) was removed from the analysis. One participant's score for PCL (interview observation task) was 3.22 SDs from the mean score. For the 'number of details recalled' measure, one participant's score was 3.46 SDs from the mean and for the 'free recall accuracy' measure, two participants' data were 3.26 and 3.39 SDs from the mean. These four data points were excluded from the analyses reported in Chapter 5. Reported here are descriptive statistics, pre-removal of data, for the PCL 'interview observation' (Table 1) and the amount of free recall details and accuracy of free recall details (Table 2).

Table 1.

Mean and SD scores for PCL for the interview observation task for each condition (N = 130).

	PCL for interview observation task	
Condition	*M* (*SD*)	95% CI
FNT (access)	58.09 (12.10)	[53.20, 62.98]
SNT (access)	60.94 (13.87)	[55.34, 66.54]
FNT (no access)	55.45 (14.62)	[49.54, 61.35]
SNT (no access)	59.95 (11.40)	[55.35, 64.55]
NNT	56.86 (14.72)	[50.91, 62.80]

Table 2.

Mean and SD scores for amount of free recall details and accuracy of free recall details for each condition (N = 130).

	Amount of details		Accuracy of details	
Condition	*M* (*SD*)	95% CI	*M* (*SD*)	95% CI
FNT (access)	110.81 (27.23)	[97.96, 123.65]	.73 (.09)	[.70, .77]
SNT (access)	121.96 (26.88)	[110.97, 133.05]	.76 (.07)	[.73, .79]
FNT (no access)	111.88 (23.86)	[101.80, 121.95]	.70 (.09)	[.68, .75]
SNT (no access)	112.04 (28.52)	[100.27, 123.81]	.71 (.09)	[.68, .75]
NNT	103.42 (27.23)	[92.43, 114.42]	.70 (.08)	[.66, .73]

A one-way between groups ANOVA was conducted to examine differences between note taking conditions (FNT [access to notes] vs. SNT [access to notes] vs FNT [no access to notes] vs SNT [no access to notes] vs NNT) for perceived cognitive load (PCL) during the interview observation and recall tasks. Tukey post-hoc comparisons for accuracy of free recall details are reported (see Table 3) and accuracy of cued recall (see Table 4).

Table 3.

Mean and SD with Tukey post-hoc comparisons between conditions for the proportion of accurate details recalled.

			Control	Access		No Access	
Condition	*M*	*SD*	NNT	FNT	SNT	FNT	SNT
NNT (control)	.70	.08					
FNT (access)	.73	.09	$p = .383$				
SNT (access)	.76	.07	$p = .022$	$p = .700$			
FNT (no access)	.72	.07	$p = .844$	$p = .948$	$p = .279$		
SNT (no access)	.73	.07	$p = .637$	$p = .995$	$p = .464$	$p = .997$	

Table 4.

Mean and SD with Tukey post-hoc comparisons between each condition for the accuracy of cued recall.

			Control	Access		No Access	
Condition	*M*	*SD*	NNT	FNT	SNT	FNT	SNT
NNT (control)	56.25	13.07					
FNT (access)	64.73	15.13	$p = .070$				
SNT (access)	71.25	10.54	$p = .001$	$p = .414$			
FNT (no access)	66.50	13.03	$p = .051$	$p = .989$	$p = .712$		
SNT (no access)	66.05	14.60	$p = .070$	$p = .997$	$p = .637$	$p > .999$	

A series of independent groups *t*-tests were conducted to examine differences between the access to notes conditions (i.e., access to notes and no access to notes) for each of the dependent variables (i.e., PCL 'interview task', PCL 'recall task', total amount of free recall details, accuracy of free recall detail*s* and accuracy of cued recall; see Table 5).

Table 5.

Mean and SD scores, with results of t-tests, for dependent variables of PCL ('recall task'), amount and accuracy of free recall details, and accuracy of free recall details, for access and no access to notes conditions.

	No access to notes		Access to notes				
DV	*N*	*M (SD)*	*N*	*M (SD)*	*t* (*df*)	*p*	*d*
PCL 'recall task'	52	63.40 (12.85)	52	62.47 (10.88)	.40 (102)	.691	.08
FR amount of details	49	111.96 (26.07)	51	116.27 (29.73)	.77 (98)	.443	.16
FR proportion correct details	49	.72 (.07)	51	.75 (.08)	1.69 (98)	.094	.34
CR percentage correct	52	66.27 (13.70)	51	67.92 (13.37)	.62 (101)	.538	.12

Post experiment questionnaire

A series of Pearson's correlations were calculated to determine whether the dependent variables of motivation, confidence and task difficulty were correlated with each other. There were significant, but moderate, correlations between some of the variables (see Table 6).

Table 6.

Mean, SD and Pearson's correlations for confidence, motivation and ease of remembering the questions (N = 130), where scores of 1 = not at all confident and 7 = extremely confident.

	M	*SD*	1	2	3	4	5
1. Confidence in memory	4.82	.92					
2. Motivation to remember	5.77	.98	.357**				
3. Ease of remembering	4.26	1.24	-.434**	-.155			
4. Motivation to think of questions	4.72	1.59	.178*	.116	.905		
5. Ease of generating questions	4.30	1.40	-.055	.017	.249	-.259**	

A one-way between-groups MANOVA was conducted to investigate differences between the conditions (FNT [access to notes] vs. SNT [access to notes] vs FNT [no access to notes] vs SNT [no access to notes] vs NNT) for participants' self-reported motivation, confidence, and ratings of how difficult they found the tasks. There were no significant differences at the univariate level for motivation to remember the witness's account, F (4, 125) = .65, p = .630, η^2_p = .02; ease of remembering the witness's account, F (4, 125) = 1.67, p = .161, η^2_p = .05; motivation to think of questions, F (4, 125) = .70, p = .592, η^2_p = .02; or ease of generating questions, F (4, 125) = .30, p = .875, η^2_p = .01.

Appendix E: Supplementary materials (Chapter 6: Online field survey)

E.1 Chapter 6: Online survey

E.1 Online survey

Study title: Interviewing using ABE guidelines and the PEACE model: Interviewers' perspectives.

University of Portsmouth Ethics Committee Reference Number: SFEC 2020 - 027
Our invitation.
We would like to invite you to take part in our University of Portsmouth research study. Before deciding to participate, we would like you to understand why the research is being conducted and what it will involve for you. Please read the information below and ask us if anything is unclear or if you have any additional questions. We are looking for trained police interviewers who have conducted interviews, as the lead interviewer, with victims of crime using the ABE guidelines and persons suspected of committing crime using the PEACE model.

Study summary.
The study is designed to gain insight and understanding into interviewers' experiences when they conduct ABE and PEACE interviews.

What is the purpose of the study?
The aim of the study is to examine the processes interviewers utilise when they interview victims and suspects. We would like to know more about how these interviews are experienced, from the interviewer's perspective.

Why have I been invited?
There is no specific reason why you have been invited; you may have been invited for a number of reasons. You may have received an email from your force gatekeeper or you may have responded to an advertisement posted on social media.

Do I have to take part?
Taking part in this research is entirely voluntary. The gatekeeper has no role in the research beyond circulation of the invitation email. It is up to you to decide if you want to volunteer for the study. We will describe the study in this information sheet and, if you agree to take part, we will ask you to note your consent below before starting the online survey.

What will happen to me if I take part?
Participation in the research would require you to complete the online questionnaire on the following pages. Initially, you will be asked if you have completed, within the past six months and as the lead interviewer, interviews with, a) victims of crime using the ABE guidelines, and b) suspects of crime using the PEACE model. If you have completed these interviews, you will then complete a questionnaire regarding your experience of conducting these interviews. Lastly, you will be debriefed to learn the purpose of the study. The study should take no more than 20 minutes of your time.

Expenses and payments.
There will be no financial or other reward for taking part in this study.

Are there any possible disadvantages and risks of taking part?
There are no foreseeable risks or disadvantages associated with participating in this study.

What are the possible benefits of taking part?
Participation in this study can give you insight into how psychological research is conducted and you will contribute to the wider community's knowledge and understanding of the experiences of police officers when they conduct witness and suspect interviews.

Will my taking part in the study be kept confidential?
Yes. There will be no way of linking you to your data. The raw data (i.e., the record of your responses to questions) will not identify you. The survey has been designed so that your IP address and geo-location data will not be recorded. The Principal Investigator will keep the downloaded raw data securely on a password-protected University of Portsmouth drive. The anonymous raw data, may be presented to others at scientific meetings, or published as a project report, academic dissertation or scientific paper, or book chapter. The anonymous data may also be used in future research studies approved by an appropriate Research Ethics Committee. Also, the data may be passed to any regulatory authority that have the legal right to access the data for the purposes of conducting an investigation, but only in exceptional cases. In line with current best practices for open access, and as the PI's research is being funded by the UKRI/ESRC, the anonymised data will be archived on the UK Data Storage website on a password-protected account. The anonymised raw data will be retained for a minimum of 10 years.

What will happen if I don't want to carry on with the study?
You will be free to withdraw from the study at any time whilst completing the survey with no negative consequences but, as your responses will be anonymous, you will not be able to withdraw your participation once you have completed the survey and submitted your responses.

What if there is a problem?
If you have a concern about any aspect of this study, you should contact the Principal Investigator in the first instance, if this is appropriate, or her Supervisor (details provided below). If your concern or complaint is not resolved by the Principal Investigator or her Supervisor, you should contact: The Chair of the Ethics Committee, Dr Paul Morris at paul.morris@port.ac.uk or The University Complaints Officer at complaintsadvice@port.ac.uk

Who is funding the research?
The research is being funded by the University of Portsmouth and the ESRC South Coast Doctoral Training Partnership (SCDTP). The researcher will not receive any financial reward for conducting this study, other than her normal bursary as a student of the University.

Who has reviewed the study?
Research involving human participants is reviewed by an ethics committee to ensure that the dignity and wellbeing of participants is respected. This study has been

reviewed by the Science Faculty Ethics Committee at the University of Portsmouth and has been given favourable ethical opinion.

Thank you.
Thank you for taking time to read this information sheet and for considering volunteering for this study. If you are happy to volunteer for this study please complete the consent below.

Principal Investigator: Pamela Hanway - Email: pamela.hanway@myport.ac.uk
Supervisor: Dr Lucy Akehurst - Email: lucy.akehurst@port.ac.uk
University of Portsmouth,
Department of Psychology,
King Henry 1 Street,
Portsmouth, PO1 2DY.

Consent
I confirm that I have read and understood the Participant Information Sheet for this study.
I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason whilst I am completing the survey. I also understand that once I have submitted my responses I will not be able to withdraw my data as the data cannot be linked to me.
I understand that the results of this study may be published and/or presented at meetings or academic conferences. I give my permission for my anonymous data, which does not identify me, to be disseminated in this way.
I understand that data collected during this study could be requested by regulatory authorities. I give my permission to any such regulatory authority, with a right of legal access, to access my anonymised data.
I agree to the data I contribute being retained for open access and for any future research that has been approved by a Research Ethics Committee.
I agree to take part in the above study.

- ○ Yes (23)
- ○ No (24)

ABE Serious
In the past six months have you, **as the lead interviewer**, conducted an interview with the **victim** of a **serious crime,** e.g., child sexual abuse, rape, or wounding, using the ABE guidelines ? (n.b. the interviewee may have been a child, young person, vulnerable, or intimidated witness)

- ○ Yes (1)
- ○ No (2)

ABE Volume
In the past six months have you, **as the lead interviewer**, conducted an interview with the **victim** of a **less serious crime**, e.g., assault, robbery, theft, or burglary, using the ABE guidelines? (n.b. the interviewee may have been a child, young person, vulnerable, or intimidated witness)

- ○ Yes (1)
- ○ No (2)

PEACE Serious
In the past six months have you, **as the lead interviewer**, conducted an interview with a person **suspected** of committing a **serious crime,** e.g., child sexual abuse, rape, or wounding, using the PEACE model ?

- ○ Yes (1)
- ○ No (2)

PEACE Volume
In the past six months have you, **as the lead interviewer**, conducted an interview with a person **suspected** of committing a **less serious crime,** e.g., assault, robbery, theft, or burglary, using the PEACE model?

- ○ Yes (1)
- ○ No (2)

Please take yourself back to the **last interview** you conducted using the **ABE guidelines** with the **victim** of a **serious crime** e.g., child sexual abuse, rape or wounding. Thinking about that interview please answer the following questions

For the **last interview** you conducted using the ABE guidelines with the victim of a serious crime, what offence(s) was being investigated?
_____
_____

For the last interview you conducted using the ABE guidelines with the victim of a serious crime, what age was the interviewee?

- ○ Adult (18 & over) (7)
- ○ Child (Under 18) (8)

For the **last** interview you conducted using the ABE guidelines with the victim of a serious crime, were there any other interviewers present in the interview room?

- ○ Yes (1)
- ○ No (2)

For the last interview you conducted using the ABE guidelines with the victim of a serious crime, did you take notes during the interview?

- ○ Yes (1)
- ○ No (2)

For the last interview you conducted using the ABE guidelines with the victim of a serious crime, please indicate your agreement with the following statements. The interview was part of a complex investigation

- ○ Strongly agree (1)
- ○ Somewhat agree (2)
- ○ Neither agree nor disagree (3)
- ○ Somewhat disagree (4)
- ○ Strongly disagree (5)

For the last interview I conducted using the ABE guidelines with the victim of a serious crime, I felt I had sufficient time to plan and prepare

- ○ Strongly agree (1)
- ○ Somewhat agree (2)
- ○ Neither agree nor disagree (3)
- ○ Somewhat disagree (4)
- ○ Strongly disagree (5)

During the last interview I conducted using the ABE guidelines with the victim of a serious crime, I felt the interviewee was being co-operative

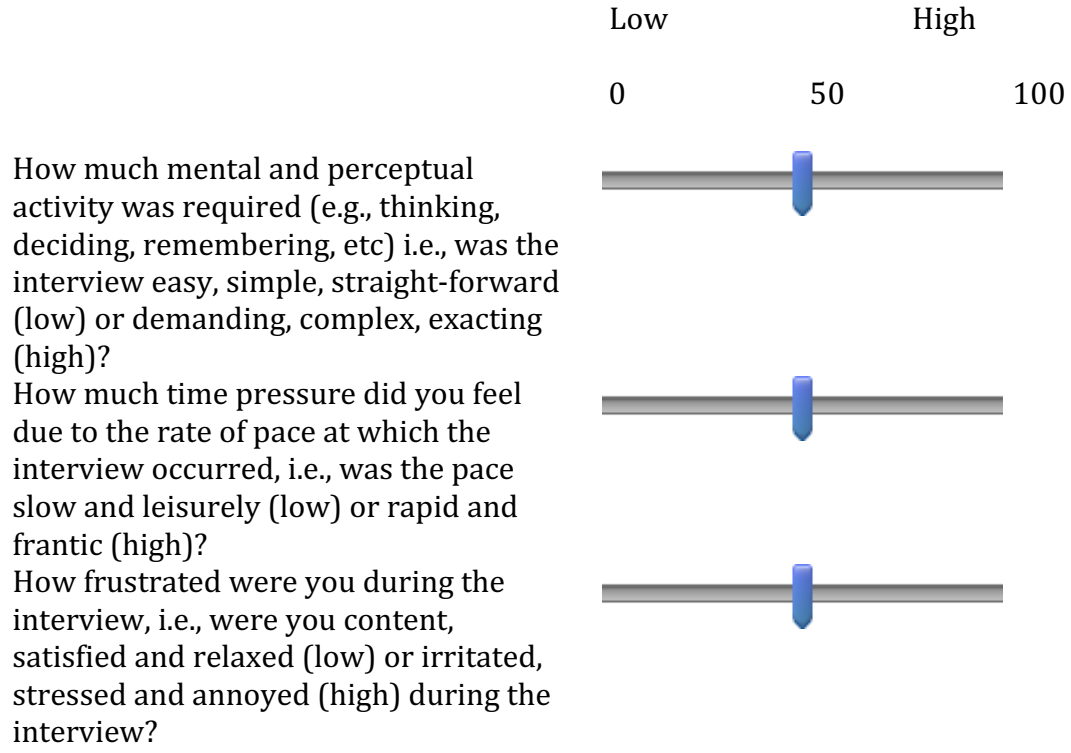
- ○ Strongly agree (1)
- ○ Somewhat agree (2)
- ○ Neither agree nor disagree (3)

○ Somewhat disagree (4)

○ Strongly disagree (5)

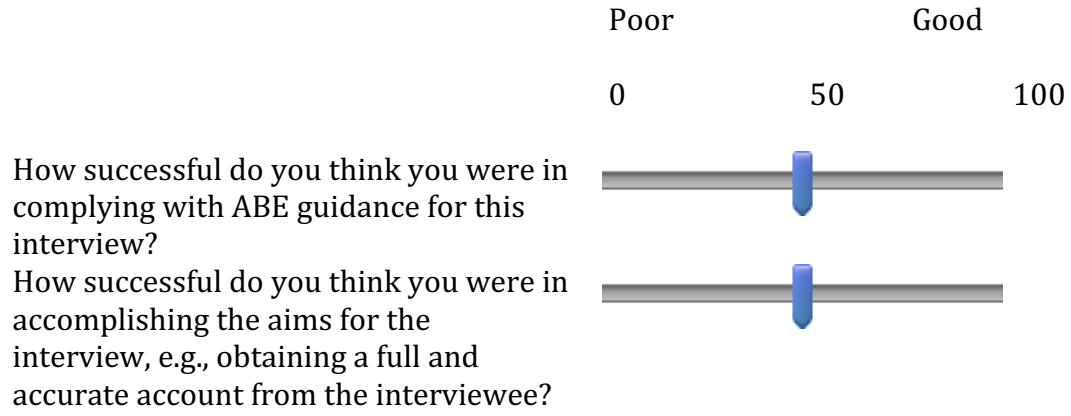
We would now like you to answer three questions about **your experience** of the last interview you conducted using the ABE guidelines with the victim of a serious crime.

Please move the slider between **Low** and **High** to the point which rates **your experience** of this interview.



We would now like you to answer four questions about **your performance** during the last interview you conducted using the ABE guidelines with the victim of a serious crime.

Please move the slider between **Poor** and **Good** to the point which rates **your performance** during this interview.



How satisfied were you with your performance in complying with ABE guidance for this interview?

How satisfied were you with your performance in accomplishing the aims of this interview, e.g., obtaining a full and accurate account from the interviewee?

Linking effort with performance, for the last interview you conducted using the ABE guidelines with the victim of a serious crime.

Please move the slider between Low and High to the point which rates your experience of this interview.

Low High

0 50 100

How hard did you have to work mentally to achieve your level of performance for this interview?

Please take yourself back to the **last interview** you conducted using the **ABE guidelines** with the **victim** of a less serious crime, e.g., assault, robbery, theft, or burglary. Thinking about that interview please answer the following questions

For the last interview you conducted using the ABE guidelines with the victim of a **less** serious crime, what offence(s) was being investigated?

_____

_____

For the last interview you conducted using the ABE guidelines with the victim of a less serious crime, what age was the interviewee?

- ○ Adult (18 & over) (7)
- ○ Child (Under 18) (8)

For the last interview you conducted using the ABE guidelines with the victim of a less serious crime, were there any other interviewers present in the interview room?

- ○ Yes (1)
- ○ No (2)

For the last interview you conducted using the ABE guidelines with the victim of a less serious crime, did you take notes during the interview?

- ○ Yes (1)

- ○ No (2)

For the last interview you conducted using the ABE guidelines with the victim of a less serious crime, please indicate your agreement with the following statements. The interview was part of a complex investigation

- ○ Strongly agree (1)
- ○ Somewhat agree (2)
- ○ Neither agree nor disagree (3)
- ○ Somewhat disagree (4)
- ○ Strongly disagree (5)

For the last interview I conducted using the ABE guidelines with the victim of a less serious crime, I felt I had sufficient time to plan and prepare

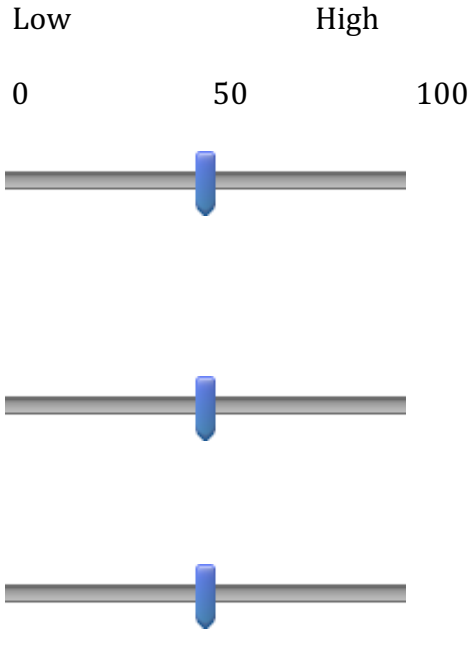
- ○ Strongly agree (1)
- ○ Somewhat agree (2)
- ○ Neither agree nor disagree (3)
- ○ Somewhat disagree (4)
- ○ Strongly disagree (5)

During the last interview I conducted using the ABE guidelines with the victim of a less serious crime, I felt the interviewee was being co-operative

- ○ Strongly agree (1)
- ○ Somewhat agree (2)
- ○ Neither agree nor disagree (3)
- ○ Somewhat disagree (4)
- ○ Strongly disagree (5)

We would now like you to answer three questions about **your experience** of the last interview you conducted using the ABE guidelines with the victim of a less serious crime.

Please move the slider between **Low** and **High** to the point which rates **your experience** of this interview.



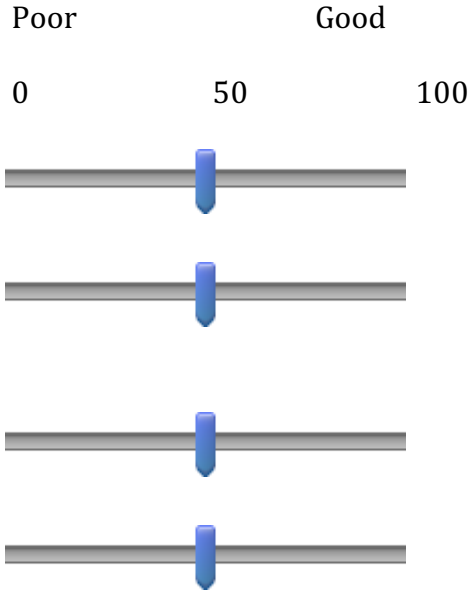
How much mental and perceptual activity was required (e.g., thinking, deciding, remembering, etc) i.e., was the interview easy, simple, straight-forward (low) or demanding, complex, exacting (high)?

How much time pressure did you feel due to the rate of pace at which the interview or elements of it occurred, i.e., was the pace slow and leisurely (low) or rapid and frantic (high)?

How frustrated were you during the interview, i.e., were you content, satisfied and relaxed (low) versus irritated, stressed and annoyed (high) during the interview?

We would now like you to answer four questions about **your performance** during the last interview you conducted using the ABE guidelines with the victim of a less serious crime .

Please move the slider between **Poor** and **Good** to the point which rates **your performance** during this interview.



How successful do you think you were in complying with ABE guidance for this interview?

How successful do you think you were in accomplishing the aims for the interview, e.g., obtaining a full and accurate account from the interviewee?

How satisfied were you with your performance in complying with ABE guidance for this interview?

How satisfied were you with your performance in accomplishing the aims of this interview, e.g., obtaining a full and accurate account from the interviewee?

Linking effort with performance, for the last interview you conducted using the ABE guidelines with the victim of a less serious crime.

Please move the slider between Low and High to the point which rates your experience of this interview.

	Low		High
	0	50	100
How hard did you have to work mentally to achieve your level of performance for this interview?			

Please take yourself back to the **last interview** you conducted using the PEACE model with a person suspected of committing a serious crime e.g., child sexual abuse, rape or wounding. Thinking about that interview please answer the following questions

For the last interview you conducted using the PEACE model with a person suspected of committing a serious crime, what offence was being investigated?

_____

_____

For the last interview you conducted using the PEACE model with a person suspected of committing a serious crime, what age was the interviewee?

- ○ Adult (18 & over) (7)
- ○ Child (Under 18) (8)

For the last interview you conducted using the PEACE model with a person suspected of committing a serious crime, were there any other interviewers present in the interview room?

- ○ Yes (1)
- ○ No (2)

For the last interview you conducted using the PEACE model with a person suspected of committing a serious crime, did you take notes during the interview?

- ○ Yes (1)
- ○ No (2)

For the last interview you conducted using the PEACE model with a person suspected of committing a serious crime, please indicate your agreement with the following

statements.
This interview was part of a complex investigation

- ○ Strongly agree (1)
- ○ Somewhat agree (2)
- ○ Neither agree nor disagree (3)
- ○ Somewhat disagree (4)
- ○ Strongly disagree (5)

For the last interview I conducted using the PEACE model with a person suspected of committing a serious crime, I felt I had sufficient time to plan and prepare

- ○ Strongly agree (1)
- ○ Somewhat agree (2)
- ○ Neither agree nor disagree (3)
- ○ Somewhat disagree (4)
- ○ Strongly disagree (5)

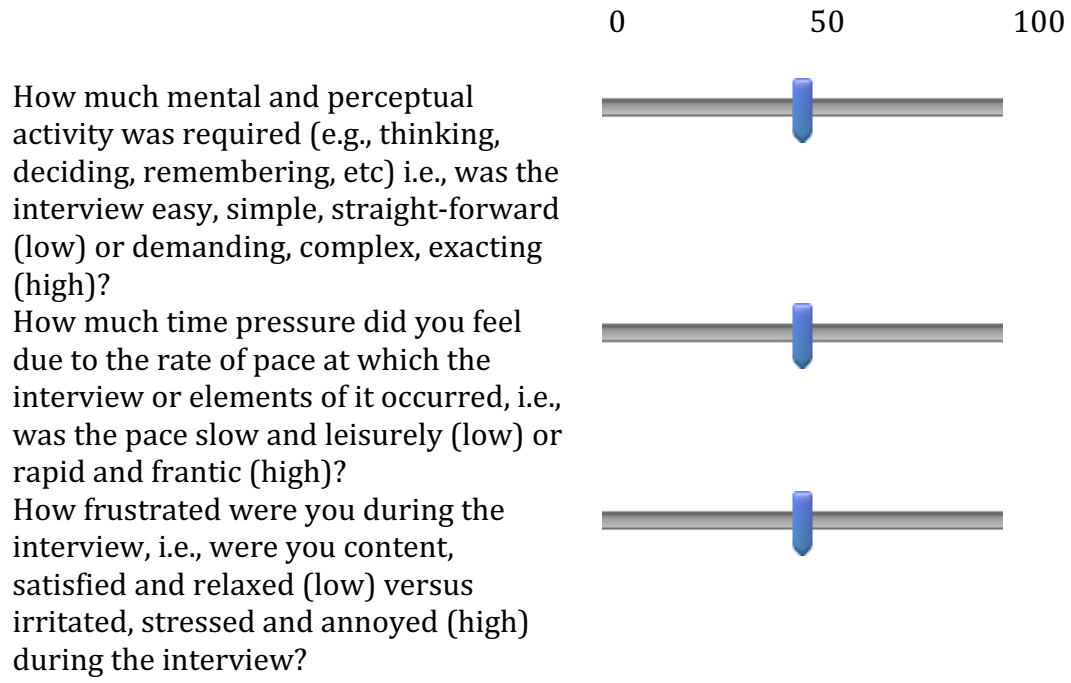
During the last interview I conducted using the PEACE model with a person suspected of committing a serious crime, I felt the interviewee was being co-operative

- ○ Strongly agree (1)
- ○ Somewhat agree (2)
- ○ Neither agree nor disagree (3)
- ○ Somewhat disagree (4)
- ○ Strongly disagree (5)

We would now like you to answer three questions about **your experience** of the last interview you conducted using the PEACE model with a person suspected of committing a serious crime.

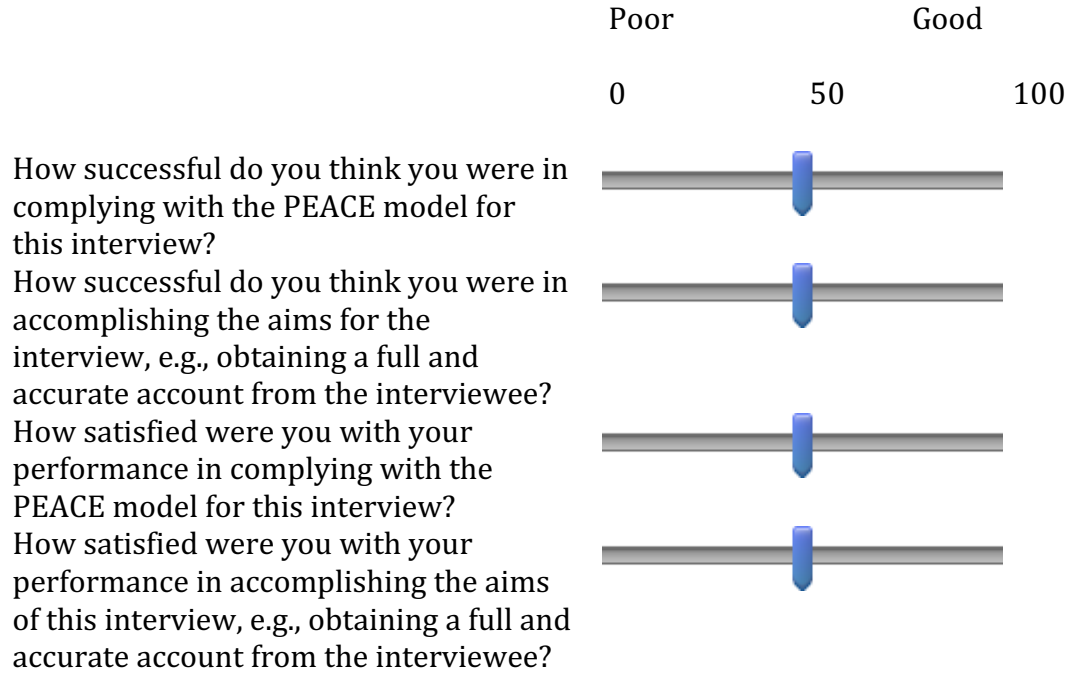
Please move the slider between **Low** and **High** to the point which rates **your experience** of this interview.

Low High



We would now like you to answer four questions about **your performance** during the last interview you conducted using the PEACE model with a person suspected of committing a serious crime.

Please move the slider between **Poor** and **Good** to the point which rates **your performance** during this interview.



Linking effort with performance, for the last interview you conducted using the PEACE model with a person suspected of committing a serious crime.

Please move the slider between Low and High to the point which rates your experience of this interview.

	Low		High
	0	50	100
How hard did you have to work mentally to achieve your level of performance for this interview?			

Please take yourself back to the **last interview** you conducted using the PEACE model with a person suspected of committing a **less** serious crime e.g., assault, robbery, theft or burglary. Thinking about that interview please answer the following questions

For the last interview you conducted using the PEACE model with a person suspected of committing a less serious crime, what offence was being investigated?

_____

_____

For the last interview you conducted using the PEACE model with a person suspected of committing a less serious crime, what age was the interviewee?

- ○ Adult (18 & over) (7)
- ○ Child (Under 18) (8)

For the last interview you conducted using the PEACE model with a person suspected of committing a less serious crime, were there any other interviewers present in the interview room?

- ○ Yes (1)
- ○ No (2)

For the last interview you conducted using the PEACE model with a person suspected of committing a less serious crime, did you take notes during the interview?

- ○ Yes (1)
- ○ No (2)

For the last interview you conducted using the PEACE model with a person suspected of committing a less serious crime, please indicate your agreement with the following statements.
This interview was part of a complex investigation

- ○ Strongly agree (1)

- ○ Somewhat agree (2)
- ○ Neither agree nor disagree (3)
- ○ Somewhat disagree (4)
- ○ Strongly disagree (5)

For the last interview I conducted using the PEACE model with a person suspected of committing a less serious crime, I felt I had sufficient time to plan and prepare

- ○ Strongly agree (1)
- ○ Somewhat agree (2)
- ○ Neither agree nor disagree (3)
- ○ Somewhat disagree (4)
- ○ Strongly disagree (5)

During the last interview I conducted using the PEACE model with a person suspected of committing a less serious crime, I felt the interviewee was being co-operative

- ○ Strongly agree (1)
- ○ Somewhat agree (2)
- ○ Neither agree nor disagree (3)
- ○ Somewhat disagree (4)
- ○ Strongly disagree (5)

We would now like you to answer three questions about **your experience** of the last interview you conducted using the PEACE model with a person suspected of committing a less serious crime.

Please move the slider between **Low** and **High** to the point which rates **your experience** of this interview.

Low		High
0	50	100

How much mental and perceptual activity was required (e.g., thinking, deciding, remembering, etc) i.e., was the interview easy, simple, straight-forward (low) or demanding, complex, exacting (high)?

How much time pressure did you feel due to the rate of pace at which the interview or elements of it occurred, i.e., was the pace slow and leisurely (low) or rapid and frantic (high)?

How frustrated were you during the interview, i.e., were you content, satisfied and relaxed (low) versus irritated, stressed and annoyed (high) during the interview?

We would now like you to answer four questions about **your performance** during the last interview you conducted using the PEACE model with a person suspected of committing a less serious crime.

Please move the slider between **Poor** and **Good** to the point which rates **your performance** during this interview.

Poor Good

0 50 100

How successful do you think you were in complying with the PEACE model for this interview?

How successful do you think you were in accomplishing the aims for the interview, e.g., obtaining a full and accurate account from the interviewee?

How satisfied were you with your performance in complying with the PEACE model for this interview?

How satisfied were you with your performance in accomplishing the aims of this interview, e.g., obtaining a full and accurate account from the interviewee?

Linking effort with performance, for the last interview you conducted using the PEACE model with a person suspected of committing a less serious crime.

Please move the slider between Low and High to the point which rates your experience of this interview.

Low High

0 50 100

How hard did you have to work mentally to achieve your level of performance for this interview?

Finally, please could you answer the following questions.

Please indicate your age

- ○ 18 - 24 (88)
- ○ 25 - 34 (89)
- ○ 35 - 44 (90)
- ○ 45 - 54 (91)
- ○ 55 - 64 (92)
- ○ 65 or older (93)

Gender
Please indicate you gender
_____

Service
How many years/months police service do you have?
_____

Training number
How many interview training courses have you attended? (Please give an estimate if unsure of the exact number)
_____

Training type
Please give details of the interview training courses you have attended
_____
_____

Number ABE
In the **last six months**, how many ABE interviews have you conducted, as the **lead interviewer,** with victims or witnesses of any offence**?** (Please give an estimate if unsure of the exact number)

- ○ 0 - 10 (6)
- ○ 11 - 20 (7)
- ○ 21 or more (8)

Number PEACE
In the **last six months**, how many **PEACE** interviews have you conducted, as the **lead interviewer,** with suspects of **any offence?** (Please give an estimate if unsure of the exact number)

- ○ 0 - 10 (4)
- ○ 11 - 20 (5)
- ○ 21 or more (6)

Region
In which region is your force area?

Region
Please select one region from the below list

- ○ North East, Yorkshire and Humberside (1)
- ○ North West (2)
- ○ East or West Midlands (4)
- ○ West Midlands (5)
- ○ Wales (6)
- ○ East of England (7)
- ○ South East, inc London and BTP (8)
- ○ South West (10)

Debrief De-brief

The survey you have completed was designed to measure your perceived mental workload during the interviews you considered.
We will examine the different aspects involved in these types of interviews that may contribute to interviewers' perceived cognitive load.
The responses you made to the questions will remain anonymous.

If you have any questions or would like further information about the results of this survey please contact Pamela.hanway@port.ac.uk.
If you do make contact with the researcher you will no longer be anonymous, however, the researchers will not be able to identify the response data you have provided.

Appendix F: Ethical approval (Chapters 2, 3, 4, 5, 6 and UPR16 form)

F.1 Chapter 2: Ethical approval

F.2 Chapter 3: Ethical approval

F.3 Chapter 4: Ethical approval

F.4 Chapter 5: Ethical approval

F.5 Chapter 6: Ethical approval

F.6 UPR 16 form

F.1 Chapter 2: Ethical approval.

Pamela Hanway
Department of Psychology
University of Portsmouth

Pamela.Hanway@myport.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

21 November 2017

FAVOURABLE ETHICAL OPINION – WITH CONDITIONS

Study Title: Perceptions of the effects of cognitive load on investigative interviewer performance.

Reference Number: SFEC 2017-111

Date Submitted: 02 November 2017

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)* and the following specific minor conditions. There were also a number of advisory notes that you may wish to take into consideration, as outlined below:

Condition(s)[1]

A. Please ensure that venues intended for use for the interviews are appropriate for private discussion and include this information in the Participant Information Sheet (PIS).

B. Please ensure that participants are aware they need to be careful when providing examples that no confidential information is inappropriately shared with the investigator.

Advisory Note(s)[2]

i. The application states that Interpretive Phenomenological Analysis will used to guide the study and data analysis. The following may be worth consideration, specifically at what point data can no longer be withdrawn with this kind of analysis; whether there may be a need for a second interview; whether a question about 'thought processes' is consistent

[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.

F.2 Chapter 3. Ethical approval.

UNIVERSITY OF PORTSMOUTH

Pamela Hanway
Department of Psychology
University of Portsmouth

Pamela.Hanway@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

13 February 2018

FAVOURABLE ETHICAL OPINION – FOLLOWING RESUBMISSION

Study Title: Examining the effects of cognitive load on interviewers' memory and perception of mental workload during an investigative interview.

Reference Number: SFEC 2018-010

Date Resubmitted: 07 February 2018

Thank you for resubmitting your application to the Science Faculty Ethics Committee (SFEC) for ethical review in accordance with current procedures, for making the requested changes following the first SFEC review, and for the clarifications provided.

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)*.

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

Dr Paul Morris
Vice Chair, Science Faculty Ethics Committee

Annexes

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

Information:

F.3 Chapter 4. Ethical approval.

Pamela Hanway
Department of Psychology
University of Portsmouth

pamela.hanway@port.ac.uk

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

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

18 Nov 2019

FAVOURABLE ETHICAL OPINION – WITH CONDITIONS

Study Title: The effects of increased cognitive demands on investigative interviewers' perceived cognitive load, recollection of multiple accounts, and their source monitoring. (Advertised title – Interviewing multiple witnesses of a crime).

Reference Number: SFEC 2019-107

Date Submitted: 1 Nov 2019

Thank you for submitting your application to the Science and Health 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)*, and the following specific minor conditions:

Condition(s)[1]

1. In section 8, please define WMC (Working Memory Capacity).
2. In section 9.1 and in the "What will happen to me if I take part?" section of the PIS, please communicate the procedure of this experiment in a more clear manner using lay language.
3. In section 9.2 and 12.3, please clarify data retention and storage procedures and ensure they are consistent.
4. In section 11.2, please clarify inclusion criteria. For instance, it states that individuals who identify as male and female are eligible to participate in the study. Does this imply individuals who identify as trans or have a trans history are not eligible to participate in the study? Please also cite the evidence to support the age and language inclusion criteria.
4. Please clarify payments or rewards for non-psychology students.
5. In sections 11.6, 11.9, and PIS, please clarify why participants are not able to withdrawal their data after the experiment. Please ensure this is compliant with GDPR.

[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.

F.4 Chapter 5. Ethical approval.

Pamela Hanway
Department of Psychology
University of Portsmouth

Pamela.Hanway@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

10 December 2018

FAVOURABLE ETHICAL OPINION – WITH CONDITIONS

Study Title: The effects of note taking and access to notes on perceived cognitive load and recall performance for investigative interviewers

Reference Number: SFEC 2018-122

Date Submitted: 20 November 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)*, and the following specific minor conditions:

Conditions[1]

A. Could the PI please submit examples of the WMC and NASA-LTX.

B. Could the PI please give a bit more information in the PIS about what the participant will actually do, e.g. think of questions they would ask the witness and they will have to do the NASA-TLX..

C. Could the PI please include a written version of the debriefing to be used.

D. Could the PI include a statement about anonymisation of participant data in 12.1 Data management.

Advisory Note

These advisory notes are given in good faith and it is hoped they are accepted as such. You do not need to adhere to these comments, or respond to them, unless you wish to.

[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.

F.5 Chapter 6. Ethical approval.

Pamela Hanway
Department of Psychology
University of Portsmouth

pamela.hanway@port.ac.uk

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

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

31 March 2020

FAVOURABLE ETHICAL OPINION

Study Title: The effects of interview type and the seriousness of a crime on investigative interviewers' perceived cognitive load.

Reference Number: SFEC 2020-027

Date Submitted: 12 March 2020

Thank you for submitting your proposal to the Science and Health 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)*.

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.

Dr Paul Morris
Chair, Science and Health Faculty Ethics Committee

Annexes

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

F.6 UPR16 form

FORM UPR16

Research Ethics Review Checklist

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:**	789675
PGRS Name:	Pamela Hanway			
Department:	Psychology	**First Supervisor:**	Dr Lucy Akehurst	
Start Date: (or progression date for Prof Doc students)		**1st October 2017**		
Study Mode and Route:	Part-time ☐ Full-time ☒	MPhil ☐ PhD ☒	MD ☐ Professional Doctorate ☐	

Title of Thesis:	The effects of cognitive load for investigative interviewers
Thesis Word Count: (excluding ancillary data)	

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):	SFEC: 2017-111, 2018-010 2018-122, 2019-107, 2020-027

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)*:	P Hanway	**Date:** 17TH December 2020

UPR16 – April 2018