Chatbots and the Initial Account Interview for Gathering Eyewitness Information: Development, Research and Efficacy of 'ChatCharlie'

Adam, Charlotte

https://doi.org/10.34737/wyy6v

Chatbots and the Initial Account Interview for Gathering Eyewitness Information: Development, Research and Efficacy of 'ChatCharlie'

CHARLOTTE ETAIN ADAM

A thesis submitted in partial fulfilment of the

requirements of the University of Westminster

for the degree of Doctor of Philosophy

May 2024

Abstract

This thesis presents a programme of empirical research centred on promptly eliciting information from witnesses and victims of crime to consolidate memory for later full retrieval. The College of Policing sometimes advocates Initial Account Interviews, yet more research should focus on the prevailing Initial Account interview approach. This novel research considers the nature of episodic memory and highlights a dearth of literature in the domain of initial accounts. In doing so, the potential of a chatbot approach for remotely gathering initial accounts is considered when logistical and personnel challenges preclude eliciting an initial account in person.

This body of work underscores the significance of in-person Initial Account Interviews and suggests the potential of chatbots as a novel interviewing modality referred to as ChatCharlie. The thesis reports the development and first evaluation of ChatCharlie, aligning its design with best interviewing practices and psychological insights into human-computer interaction. Specifically, it uses a mock witness paradigm to explore the impact of typed interviews and technology usage across age groups. The acceptability of chatbots in an eyewitness context was examined, potentially contributing to the 'toolbox' of investigative techniques for frontline investigators.

Study 1 presents the first in a series of empirical investigations of eyewitness memory performance as a function of reporting modality and considers typed modality interviews as one method for potentially expediting the collection of witness information. This study provides insight into the potential costs and benefits of typed versus spoken responses to investigative questioning. Results revealed no difference between conditions for memory performance (correct, incorrect and

percentage accuracy), apart from confabulations. Non-significant differences were consistent globally and as a function of the recall phase.

Study 2 focused on the typed recall of eyewitness information as a function of age. Again, this study uses the mock-witness paradigm to examine episodic recall performance across three different age groups when providing typed information. Results revealed that participants in the 55+ group recalled more correct items than the two younger age groups, with no significant differences across age groups for confabulated or erroneous recall. Despite lower confidence in technology in the 55+ group, here familiarity with technology did not appear to impede the retrieval and reporting process. These findings further suggest that text-based retrieval may have the potential for promptly and accurately capturing witness accounts in some instances and that older witnesses may also benefit.

Study 3 examined the efficacy of ChatCharlie for consolidating eyewitness memory versus an in-person Initial Account Interview. At Time 1, participants were randomly assigned to ChatCharlie, In-person initial account or the no initial account control condition. There were significantly fewer confabulations in the ChatCharlie condition At Time 1, with non-significant differences in correct and incorrect memory performance across the ChatCharlie and In-person conditions. At Time 2, participants in the ChatCharlie and In-person Time 1 conditions were more accurate. They recalled significantly more correct items globally and during both interview phases than the Control group without a concomitant increase in errors and confabulations. No significant differences emerged at Time 2 between ChatCharlie and the In-Person Time 1 conditions.

Study 4 reports perceptions and experiences of ChatCharlie and chatbots in general. Participants in the ChatCharlie condition were asked about their experiences of using a chatbot. In contrast, other participants were asked about their perceptions of chatbots more generally, with an emphasis on chatbots in a criminal justice context. Findings indicate a positive attitude toward chatbots, with participants expressing comfort and willingness to use this technology for initial account interview purposes. Those who engaged with ChatCharlie demonstrated increased ease in admitting memory gaps during subsequent interviews at Time 2. However, privacy and security emerged as paramount concerns, influencing platform preferences and information disclosure behaviours. This thesis offers novel insights relevant to the efficacy and usability of a text-based ChatCharlie chatbot called for collecting Initial Accounts from some witnesses.

Table of Contents

List of Tables

List of Figures

List of Appendices

Publications & Conferences

Parts of the research in this thesis have appeared in the following forms to date:

Journal Articles

Adam, C. & Dando, C. J. (in prep). *Consolidating eyewitness memory using the ChatCharlie chatbot to conduct timely initial accounts interviews: Confidence, accuracy and completeness.* For submission to Nature Scientific reports (special edition: eyewitness memory)

Selected Peer-Reviewed Presentations

Adam, C, & Dando, C.J., (2023) *ChatCharlie: A potential tool to support memory performance at the scene of a crime* [Presentation]. University of Westminster School of Social Sciences PhD Conference. London, United Kingdom.

Adam, C, & Dando, C.J., (2021) *ChatCharlie: Creating a Chatbot to collect brief first accounts remotely* [Presentation]. Criminal Investigative and Forensic Research CIFER Conference. London, United Kingdom.

Adam, C, & Dando, C.J., (2021) *ChatCharlie: Collecting brief first accounts from multiple eyewitnesses remotely* [Presentation]. BPS Psychology Postgraduate Affairs Group (PsyPag) Conference. Online.

Adam, C, & Dando, C.J., (2021) *Initial Account Police Interviewing* [Presentation]. University of Westminster School of Social Sciences PhD Conference. London, United Kingdom.

Adam, C, & Dando, C.J., (2020) *Why are Witnesses so Unreliable?* [Presentation]. University of Westminster School of Social Sciences PhD Conference. London, United Kingdom.

Acknowledgements

Firstly, I would like to say my biggest thank you to my supervisor, Professor Coral Dando, whose ability to trust me as I lost my mind was invaluable. This would not have been possible without you.

I would also like to thank my fellow PhDs for help with everything from the logistics of doing a PhD to the logistics of finding a pub. Thank you, Laura, Kathryn and Giota. Honourable mention to Dave for helping keep spirits up, even in the difficulties of navigating academia.

I would like to thank my friends and family and their community of grandparents for participating and helping to push this research forward. I need to thank Courtney for being there for me every single day, helping me laugh through the tears (and proofreading). I need to thank my incredible partner, Charlie, for loving me through this (and proofreading). I need to thank my Mum and Dad for their endless belief and support and for never asking me when I would be finished. Finally, I want to thank Taylor Swift, who took me from a Blank Space to The Tortured Academics Department.

Unacknowledgements: Covid-19

Author's Declaration

I declare that all the material contained in this thesis is my own work and has not been submitted to any other University.

Charlotte Etain Adam, May 2024

Chapter One: Eyewitness Memory and Gathering Witness Information

1.1. Witness Information

Witness, victim and survivor accounts of their experiences (from here on, the term witness is used to refer to all witnesses, victims and survivors) are a vital element of all criminal investigations; witness information typically forms a significant part of the prosecution case (Granhag et al., 2018; Kebbell & Milne, 1998; Launay et al., 2021; Luknar et al., 2023; Ridley et al., 2012). Witness information often shapes an inquiry from the beginning, with leads established from information gathered as soon as investigators arrive at the crime scene (Chin et al., 2022; College of Policing, 2019; Dando et al., 2009; Gabbert et al., 2015; Launay et al., 2021; Spanoudaki et al., 2019) and throughout the investigative processes that may occur afterwards. Witnesses may be asked to provide information once or on several occasions as the investigation progresses. Later, during any court proceedings that might follow, witnesses often provide vital evidence in court. The importance of witness testimony cannot be understated, and their information is widely recognised as fundamental to criminal justice worldwide.

Despite witness information being crucial to the investigation of crime and criminal justice processes, it is often the case that there is limited understanding of the cognitive demands of remembering and recounting an experienced event. The general public usually believes that memory works similarly to a video recorder and is permanent and accurate (Benton et al., 2006; Henkel et al., 2008; Simons & Chabris, 2011) despite consistent research findings reporting that memory is generally fragile and is typically an incomplete representation of experienced events (Flindall et al., 2016; Henkel, 2017; Wells, 1995; Wells & Loftus, 2003). Indeed, the

fragility of witness memory has been publicly demonstrated using DNA to exonerate wrongly convicted individuals. In the USA, the Innocence Project has found inaccurate eyewitness reporting and misidentification to be critical factors in 69% of the 375 overturned convictions (Innocence Project, 2022). In England and Wales, wrongful convictions have been exemplified by several high-profile cases (C. R. Huff & Naughton, 2017), highlighting the delicate and sometimes inaccurate nature of witness memory (Davies & Griffiths, 2008).

Misidentification is considered an essential element of many miscarriages of justice cases publicised in the USA and UK (C. R. Huff & Naughton, 2017), often changing the direction of an investigation (Kebbell & Milne, 1998; Loftus, 2019). Misidentification often provides, apparently, solid and convincing evidence for both the police and the jury (Gabbert et al., 2015). A plethora of laboratory research supports the idea that an individual who appears to have clearly seen a perpetrator can make an incorrect identification (Blank & Launay, 2014; Bulevich et al., 2022; Dodson et al., 2015; Holliday et al., 2012; Innocence Project, 2022; Wixted & Wells, 2017). It has been widely reported that despite being confident, witnesses can make errors when identifying perpetrators and that the processes and practices involved in running a line-up procedure, including pre and post-line-up experiences, can negatively impact witness memory of the perpetrator (Brewer & Wells, 2006; Milne, Bull, et al., 2008; Wixted et al., 2015, 2016; Wixted & Wells, 2017).

Of more relevance to the programme of research presented here, an extensive body of empirical research reveals that witnesses can remember and report events they have not experienced. Furthermore, witnesses are known to confidently report incorrect information alongside correct event information when providing testimony (Brewer & Douglass, 2019; Bull, 2019; Chan et al., 2009; Dando, 2020; Davis &

Loftus, 2017; Lavis & Brewer, 2017; Loftus, 2005, 2019). Eyewitness errors of this nature can occur at any point when providing information, irrespective of the time between the crime event itself and providing testimony during investigation processes and at court, for example. It has been argued that eyewitnesses can be relied upon to provide good-quality evidence when the encoding and retrieval contexts are uncontaminated (Wixted et al., 2018). While this may well be the case in experimental laboratory research, witnesses rarely, if ever, experience uncontaminated encoding environments or crime events in real life.

Moreover, post-event experiences before the investigative process begins cannot be controlled, so many have argued that encoding should not be a primary concern (Wade et al., 2018). Instead, post-event experiences occurring as part of the investigative process must be considered to mitigate errors and distortions wherever possible. Indeed, with this in mind, many post-event investigative processes are amenable to improvement and have been further developed towards improving memory recall and recognition performance in field settings (Fisher & Geiselman, 1992; Wells et al., 1998, 2015).

1.2. Collecting Witness Information.

Post-event processes, practices or ‘variables’ include how investigators gather witness information after the event. In the UK and many other countries, information is collected from witnesses through an investigative interview. In an investigative context, the term ‘interview’ can range from more conversational question-answer-type interactions at the crime scene to more formal interviews conducted by a professional investigator at a police station or a witness’s home. Cross-examination questioning at court is also a form of interview since the questioner is again seeking

information from the witness. Irrespective of the context, witnesses are asked to recall and verbalise their experiences in detail rather than make a recognition choice from a line-up where the perpetrator may or may not be present. Recalling information from long-term memory is the focus of the thesis presented here. More specifically, how might memory be better supported to maximise memory performance during a post-event interview, be this an interview that occurs quickly after an event, which from hereon is referred to as an initial account interview (College of Policing, 2019) or sometime later during a more in-depth interview, which for the programme of work reported here is referred to as a tier 1 basic PEACE witness interview.

Depending upon the type of crime event itself, the witness status (e.g., person of interest, vulnerable, injured, traumatised, etc.), and the availability of police officers, witness information can be gathered at various time points after the event. Generally, the duration between witnessing a crime and being interviewed can range from within an hour to several years later, depending upon context and environment. Irrespective, all those present who become witnesses participate in an interview, be this an initial account interview, a more in-depth investigative interview, or both. In both instances, the interviewer (typically a police officer) asks questions, and the witness responds. The two types of interviews that are the focus of this thesis are i) the initial account interview, which is structured but more conversational and occurs quickly at the scene of a crime or very soon after, and ii) the more formal and in-depth tier 1 basic PEACE informed investigative interview that occurs sometime later for non-vulnerable witnesses of volume crime and other less severe types of crime.

In England and Wales (and elsewhere), both types of interviews have evolved concerning the theoretical understanding of episodic memory (Brunel et al., 2022;

College of Policing, 2019; Finn et al., 2019; Gabbert et al., 2015; Milne et al., 2019). The interview is a retrieval process, so this must be handled carefully. Conducting an effective witness interview involves extensive training (Zekiroski et al., 2024) and practice because the empirical literature is clear and consistent - a poor interview can interfere with witness memory and the quality of evidence collected (Beek et al., 2021; Brunel et al., 2022; Chin et al., 2022; College of Policing, 2022; Launay et al., 2021; McGaugh, 2000; Spanoudaki et al., 2019; Wells et al., 2020).

Even when informal, early interviewing offers significant investigative advantages by aiding memory consolidation, which supports later recall of information (Finn et al., 2019; Gabbert et al., 2015; Kontogianni et al., 2020). Additionally, early interviewing provides quick access to relevant information for investigation, otherwise unavailable until later. A substantial body of empirical research focuses on optimising witness recall during investigative interviews to ensure the resulting information is suitable for its intended purpose. This guidance and knowledge are crucial for professional practice, emphasising the importance of accuracy, thoroughness, and using non-leading questioning techniques to elicit reliable information from witnesses.

1.3. Witness Memory

When witnesses are asked to provide information about the event or events in question, they are asked to recall personally experienced episodes stored in their long-term memory (Tulving, 1984). This type of memory is widely referred to as ‘episodic memory’. Retrieving episodic information is a reconstructive process involving establishing episodic retrieval mode (Atkinson & Shiffrin, 1968, 1971; Tulving & Craik, 2000). Several cognitive processes are associated with invoking

episodic retrieval mode, a subjective sense of time of being the person who experienced the episode, and autonoetic consciousness, a form of consciousness accompanying the act of remembering (Tulving, 2005; Wheeler et al., 1997).

Rememberers must re-live or re-experience the episode in question by consciously searching for the relevant 'what, where and when' information to reconstruct experiences, a process typically likened to mental time travel (Tulving, 1984, 2002; Wheeler et al., 1997). Episodic memory encompasses actions, environments, people and personal experiences, including the event's context (Baddeley, 2001; Fisher & Geiselman, 1992; Tulving & Thomson, 1973).

The retrieval of stored episodic memories is an effortful, conscious process (Tulving & Craik, 2000) highly dependent on the encoding and retrieval environment, the attention paid to the event itself, and individual differences and perception processes (Hope et al., 2016; Lacy & Stark, 2013; Murphy & Greene, 2016). The amount of event information encoded is determined by what witnesses attend to (Albright, 2017; Chun & Turk-Browne, 2008; Murphy & Greene, 2016) and how much of the event they have seen (Boyce et al., 2008). The time of day (Yarmey, 1986), lighting conditions (Albright, 2017; Bornstein et al., 2012), and the nature of the environment itself all have an impact on encoding, which necessarily carries over to dictate whether or not a witness can recall event information.

For example, chaotic, distracting environments are not uncommon at crime scenes, so memories are often incomplete due to sensory overload, stress and divided attention (Hope et al., 2016; Koriat et al., 2000; Mirandola et al., 2014). Even in ideal conditions, focusing attention on every element of an occurrence is impossible (Craik et al., 2000; Yi & Chun, 2005). Individual witness encoding variables also play a big part in recall performance. For example, age (Aizpurua et

al., 2009; Dodson et al., 2015) and levels of intoxication (Jores et al., 2019; Schreiber Compo et al., 2017; van Oorsouw & Merckelbach, 2012) are known to trigger variable encoding performance. Witness differences and contextual variables impacting encoding are introduced and discussed when relevant to the empirical work presented, centring on post-event retrieval processes.

Events that have been encoded are stored in long-term memory in the form of a memory trace. Long-term memory is believed to have infinite capacity, although episodic memories can degrade and alter over time (Goldsmith et al., 2005; Tuckey & Brewer, 2003), and witness memory is known to be malleable and suggestible (Baxter et al., 2013; Milne & Bull, 2003; Sharman & Powell, 2012; Zaragoza et al., 2013). Indeed, a significant and consistent body of research indicates how post-event experiences can trigger witness memory errors alongside expected degradation (Bulevich et al., 2022; Paterson & Kemp, 2006). Episodic memory traces are believed to consolidate over time and remain in storage until accessed (McGaugh, 2000; Wixted, 2004). However, the time between an event taking place and the event being retrieved can impact recall performance (Tuckey & Brewer, 2003), with the fine details of an event being gradually lost (Goldsmith et al., 2005).

1.3.1. Retrieval-Induced Forgetting

One phenomenon that has significant implications for witness memory is retrieval-induced forgetting (RIF). RIF describes how the selective retrieval of specific memories can impair the recall of related, non-retrieved information (Anderson, Bjork, & Bjork, 1994). While this phenomenon has broad relevance to understanding witness memory, its applicability is nuanced, as RIF may both enhance and impair the reliability of eyewitness accounts depending on the context. RIF highlights the dynamic and reconstructive nature of memory. It illustrates how memory retrieval not

only facilitates access to desired information but also restructures the accessibility of competing memories (Anderson, Bjork, & Bjork, 1994). Early research attributed this phenomenon to cognitive mechanisms like inhibition or interference. However, recent developments in the field have expanded the framework of RIF to include motivational factors, particularly the role of individual goals and the need for cognitive closure (Pica et al., 2018).

RIF is particularly relevant when witnesses must recall specific details about an event. Eyewitness testimony often involves selective retrieval, as witnesses are prompted to remember specific aspects of a crime, such as the perpetrator's appearance, the sequence of events, or the location of particular objects. Research has demonstrated that this process of selective recall can suppress related but unpractised memories, potentially causing witnesses to forget crucial details about the event (Pica et al., 2014).

Moreover, the motivational underpinnings of RIF further complicate its relevance to witness memory. The need for closure (NFC), a motivational tendency to seek cognitive certainty and avoid ambiguity, has been shown to amplify RIF (Pica et al., 2018). High NFC individuals, who often desire quick and precise resolutions, are more likely to inhibit competing memories during recall. In the context of witness testimony, this could lead to the suppression of details that do not align with the witness's current understanding or expectations of the event. For instance, a high NFC witness might focus exclusively on details they perceive as critical, such as the presence of a weapon, while neglecting peripheral but potentially important information, such as the actions of bystanders.

Despite its potential drawbacks, RIF may also serve adaptive functions in the context of eyewitness memory. One of the core benefits of RIF is its ability to reduce

interference, allowing witnesses to provide more focused and coherent accounts of an event (Storm, 2011). Understanding the cognitive and motivational dynamics of RIF is essential for navigating its implications in legal settings. By adopting strategies that minimise the detrimental effects of RIF while leveraging its benefits, legal professionals can improve the accuracy and fairness of eyewitness testimony, ultimately contributing to more just outcomes.

1.4. An Initial Account Interview

Collecting initial accounts from witnesses at the scene of a crime event has increasingly been advocated as an essential part of the investigative process. The rationale for an initial account interview is twofold. First, information gathered at a very early stage can inform the immediate and often time critical early direction of an investigation (Bull, 2019; Gabbert et al., 2015; Loftus, 2019). Second, asking witnesses to recall the event quickly can support memory consolidation, reducing some of the detrimental impacts of delay (Dando et al., 2020; Ebbinghaus, 1913; Goldsmith et al., 2005; Kontogianni et al., 2020), including those arising from the negative effects of post-event information and misinformation. Consolidation has also been associated with the general strengthening of the memory trace so that recall is more robust during an investigative interview that might occur later (Finn et al., 2019; Gabbert et al., 2015; M. J. Huff et al., 2016; Kontogianni et al., 2020; Krix et al., 2014).

An investigative interview with a witness typically takes place sometime after a crime event, ranging from a couple of days to several years, depending upon the event in question, witness factors, and police availability (Gabbert et al., 2009; Hope et al., 2011; Kebbell et al., 1999; Pansky, 2012). Delays in allowing witnesses to

report their experiences can have several negative impacts on witness memory, including reducing the quantity of information recalled (Tuckey & Brewer, 2003) and quality in terms of the number of errors (Goldsmith et al., 2005; Pansky & Nemets, 2012; Shapira & Pansky, 2019). Thus, the initial account interview offers a significant opportunity to quickly document witness experiences and potentially better consolidate and preserve witness memory for later investigative interview opportunities and court proceedings.

The College of Policing defines initial account interviews as immediately necessary initial questioning intended to elicit a brief account of what is alleged to have happened. The primary purpose of these accounts is to protect individuals and secure and preserve evidence, prioritising areas of investigation and assessing risk (College of Policing, 2019). There are ten key guidelines for conducting an in-person initial account interview: rapport building, witness separation, clarifying sources of information, alcohol intoxication, witnesses' own words and open questioning, non-leading approach to questioning, allowing uncertainty, suggesting eye closure, advice on information exposure, identifying and record needs and vulnerabilities. Of these guidelines, eight have a moderate or good evidence base. The remaining two - identifying and recording needs and vulnerabilities and alcohol intoxication – have a practitioner evidence base. In light of the focus of this thesis, only the eight behaviours identified by the guidelines as having a sound evidence base are reviewed here.

1.4.1. Initial Account Interview guidelines: Rapport.

Rapport is believed to be a key element in all witness interviews. Despite the well-documented challenges of operationalising rapport (Gabbert et al., 2021), research has shown that when present, to a lesser or greater degree, it can increase

cooperation (Alison & Alison, 2017), reduce stress, and increase witness recall and accuracy (Dando et al., 2023; Kieckhaefer et al., 2014; Nahouli et al., 2021, 2023). It is also argued to have a protective effect on memory, potentially reducing the influence of misinformation and co-witness contamination (Vallano & Compo, 2011). Rapport builds due to positive mutual attention; it requires a trusting and open atmosphere (Abbe & Brandon, 2013; Tickle-Degnen & Rosenthal, 1990), which involves the engagement of both witness and interviewer, with friendly, effortless interactions leading to feelings of cooperation (K. Collins & Carthy, 2019).

Rapport building is thought to reduce the social demands of an interview (Kieckhaefer et al., 2014; Nahouli et al., 2021), potentially enabling additional cognitive capacity for recalling information (Dando et al., 2023; Dando & Oxburgh, 2016; Fisher & Geiselman, 1992; Nunan et al., 2016) and helping to foster a seamless sharing of information. Rapport building includes providing the witness with a clear understanding of the interview process (Abbe & Brandon, 2013). As with other forms of police interviewing, rapport building is considered an essential element for initial account interviewing. Police must approach witnesses with an empathetic and reflexive attitude (Abbe & Brandon, 2014). Since every witness and situation differs, adapting rapport building to match the witness (Baxter et al., 2003)is crucial. Although this is relevant to all interview types, the first interaction with the police can determine the tone and cooperation of the witness as the investigation moves forward. It is essential to maintain a friendly demeanour as it has a protective effect on memories, which is vital in reducing the effects of suggestibility and increasing witness cooperation (Baxter et al., 2006; McGroarty & Baxter, 2009).

However, questions remain, not about the value of rapport building but the mechanisms of how rapport is built and defined. Building a relationship with a

witness during an initial account interview at a crime scene is challenging, sometimes even impossible. Considerable effort has to be invested in being ‘authentic’ and limiting the perception of false rapport, which may have a detrimental effect (Dando & Oxburgh, 2016; Webster et al., 2021). Small talk and self-disclosure have been shown to increase feelings of rapport, but only if perceived as genuine (Kaski et al., 2018).

It is not only the content of what the interviewer says that influences rapport; active listening and other physical behaviours accompanying spoken words play a role in validating and reassuring witnesses (Bedi et al., 2005). Relaxed body language and a friendly tone of voice have been shown to have differing effects on memory, depending on the interview stage, demonstrating the need for specific training for differing interview types (R. Collins et al., 2002; Zekiroski et al., 2024). The College of Policing highlights these points and emphasises the importance of a genuine, non-judgmental interaction during an initial account interview, using appropriate language and questioning pace (College of Policing, 2019).

Despite the acknowledged value of building and maintaining rapport, it has been reported that it is not always present and that teaching rapport is challenging due to inexact operationalisations of rapport-building behaviours (Brimbal et al., 2021; Clarke & Milne, 2001; Dando & Oxburgh, 2016; Walsh & Bull, 2012). Moreover, at the scene of a crime, it may be impossible or inappropriate to build rapport since one of the many aims of the initial account interview is to assess risk towards the quick apprehension of an offender and to understand ongoing or emerging threats. In such instances, the initial account interview may comprise more of a series of quick-fire questions and answers designed to gather relevant information quickly. As a result, questions have emerged centred on the potential negative impact on memory and

the development of an effective interpersonal interaction when an initial account interview excludes rapport-building behaviours.

1.4.2. Initial Account Interview Guidelines: Witness separation.

Separating witnesses emerges as a critical procedural aspect with multiple implications. The urgency of this practice lies in its potential to mitigate the pervasive effects of co-witness contamination (Gabbert et al., 2003). Co-witness contamination refers to the unintentional influence witnesses can have on each other's recollections, leading to the potential distortion of facts or memories through shared discussions. By prioritising early separation, investigators create a window for witnesses to provide independent, uninfluenced accounts before collaborative discussions occur, thereby safeguarding the integrity of an individual's memories (Jack et al., 2014).

In some cases, the presence of additional individuals has been shown, in some cases, to induce heightened arousal levels, which can be detrimental, particularly in tasks involving complex cognitive processes such as retrieving episodic information. As the number of individuals present increases, witness performance may decline, introducing a noteworthy challenge in obtaining accurate and reliable information (Belletier et al., 2019; Wagstaff, 2008). This decline in performance may be attributed to attentional conflict, wherein the finite cognitive resources available to an individual are spread thin across the demands of the situation, hindering optimal task execution (Baron, 1986; Belletier et al., 2019).

While recognising witness separation’s theoretical and practical importance, real-world challenges may impede this. Instances where uncooperative witnesses outnumber law enforcement officers at a crime scene present a scenario where immediate and comprehensive separation might prove impossible. However,

adhering to the principle of witness separation whenever feasible remains a foundational tenet, as underscored by the initial account practice guidelines (College of Policing, 2019), which emphasises the critical role that separating witnesses plays in ensuring the reliability and accuracy of witness testimonies, thereby enhancing the overall integrity of the investigative process.

1.4.3. Initial Account Interview Guidelines: Clarifying sources of information.

Clarifying sources of information during witness interviews at a crime scene unveils the potential pitfalls associated with source attribution error. This phenomenon underscores the unintentional alteration of witness memories due to the misattribution of information regarding various sources (Gabbert et al., 2003; Zaragoza & Lane, 1994). Recognising and addressing this cognitive vulnerability becomes paramount in maintaining the accuracy of witness testimonies. To mitigate the adverse effects of source attribution error, it proves crucial to inform witnesses about this possibility and underscore the importance of source monitoring. Providing witnesses with an understanding of the intricacies of memory processes empowers them to critically evaluate the origins of their recollections, fostering a more accurate and reliable narrative (Bodner et al., 2009).

However, the practical application of imparting this information has its challenges. Timing and tact become essential considerations, as seeking clarification on the sources of information after initial testimony may inadvertently convey scepticism or disbelief on the interviewer's part (College of Policing, 2019). This potential perception could jeopardise the rapport established with the witness. Therefore, a proactive approach is advocated, wherein the discussion about source attribution and the significance of source monitoring is introduced before the formal interview

process commences. By addressing these aspects upfront, investigators lay the groundwork for collaborative and informed interaction and pre-emptively mitigate any potential damage to the rapport with the witness.

Following practices outlined by the College of Policing in 2019, this proactive engagement ensures that witnesses are aware of the complexities of memory recall and source attribution, fostering a cooperative environment that enhances the reliability of the information provided. Moreover, it reflects a commitment to transparency in the investigative process, fostering trust between the interviewer and the witness. As a result, by navigating the nuances of source clarification with sensitivity and foresight, investigators can optimise the likelihood of obtaining accurate and untainted witness testimonies.

1.4.4. Initial Account Interview Guidelines: Witnesses' own words and open questioning

At the crime scene, there is pressure on investigating officers to find specific information to help immediately clarify what has happened (College of Policing, 2019). This can result in witnesses feeling pressured to provide particular details they may not know (Shepherd & Milne, 1999). To prevent this, investigators should adopt an open questioning style, as it has been shown to increase the quality and quantity of recall compared to closed questions (Brubacher et al., 2020; Oxburgh et al., 2010; Sharman & Powell, 2012). This can be difficult as asking open questions is not, for many people, what they usually do in everyday conversations. Thus, it can be hard to integrate this question style into an interview naturalistically. Asking witnesses to provide a free recall can prevent the accidental input of police bias. It can also give the interviewer information to guide the remainder of the interview, offer a language base, and generally guide the interview process. Using the

witnesses' descriptive language can reduce the input of misinformation. Police interviews have a different focus than other social interactions, with specific motivations focused on solving a crime for prosecution (Chin et al., 2022; Launay et al., 2021; Spanoudaki et al., 2019). Additionally, there is a natural preference for interviewing style that can depend on the individual characteristics of the interviewer (Häkkänen et al., 2009). This can lead to police officers inadvertently using language that can alter the witness's perception of the event (Kleider-Offutt et al., 2015; Kontogianni et al., 2020; Loftus & Palmer, 1974; Luke et al., 2017; V. L. Smith & Ellsworth, 1987).

1.4.5. Initial Account Interview Guidelines: Non-leading approach to questioning.

Navigating the intricacies of an initial account interview demands a focused approach to questioning, recognising the potential pitfalls associated with leading questions. While specific inquiries are crucial for gathering essential details, leading information can inadvertently shape and distort witness recollections (Baxter et al., 2013; Davis & Loftus, 2017; Sharman & Powell, 2012; Thorley & Rushton-Woods, 2013). To counteract this influence, a pivotal strategy involves emphasising neutral language during questioning (Beek et al., 2021; Loftus & Palmer, 1974; Oxburgh et al., 2010). This approach helps mitigate the unintentional sharing of inaccurate information and fosters a more reliable account from the witness.

Despite the negative impact of leading questions, interviewers frequently struggle with information-seeking challenges and confirmation bias (Akca et al., 2021; MacDonald et al., 2017; Memon et al., 1995; Shepherd & Milne, 1999). This inclination toward rigid questioning impedes the openness to alternative explanations, underscoring the critical need for cognitive flexibility (Fahsing & Ask,

2013; Risan et al., 2016). Instead of leading questions, interviewers should move toward using prompts and probes strategically designed to extract information not elicited by open-ended questions (College of Policing, 2019; Kontogianni et al., 2020). These prompts must avoid assumptive details to prevent inadvertently leading the witness, focusing on creating an environment where the witness feels comfortable admitting uncertainty (Dando et al., 2023; Oxburgh et al., 2010; Paulo et al., 2019).

The detrimental impact of misleading questions is exacerbated when witnesses presume the interviewer possesses prior knowledge of the situation or event (V. L. Smith & Ellsworth, 1987). Hence, law enforcement professionals should reassure witnesses that they are not obligated to answer questions if they are uncertain (Evans & Fisher, 2011; Scoboria & Fisico, 2013). Presenting impartial feedback is crucial (Henkel, 2017) since negative feedback can undermine memory confidence and increase susceptibility to misinformation (Leippe et al., 2006).

Beyond verbal cues, physical gestures accompanying communication can influence witness perceptions (Gurney, 2011, 2015). Age-related differences in susceptibility to misinformation are pertinent, with older adults particularly vulnerable (Henkel, 2014) and children struggling with specific or narrow questions (Sheehy & Chapman, 1989). The College of Policing's 2019 guidelines endorses open questions, such as 'What time did this happen,' instead of closed questions, like 'Did it happen before you got home?' to elicit more comprehensive and accurate information.

Recent research, however, exposes deficiencies in interview techniques, with a tendency towards closed questions and a lack of rapport building (Brunel et al., 2022; K. Collins & Carthy, 2019; Dando et al., 2009; Lamb, 2016; Launay et al.,

2021). The persistent dominance of information-seeking over open-question training, leading to the accidental use of closed and leading questions, has been identified as a common pitfall (Akca et al., 2021; Lamb, 2016; Sternberg et al., 2001; Zekiroski et al., 2024).

1.4.6. Initial Account Interview Guidelines: Allowing uncertainty.

Unlike a typical social interaction, a police interview requires specific details that would not be considered relevant or engaging in other conversations. It is, therefore, not unusual for witness accounts to naturally miss this level of detail (Gabbert et al., 2015). Simply informing witnesses that they should provide specific and detailed information can increase the richness of reported information (Koriat et al., 2000). A witness interview is often intimidating for individuals, with an imbalance of power being perceived. This has a detrimental effect on memory, increasing the impact of misinformation and reducing the quantity and quality of recall (Toglia et al., 2017). This can be mediated by letting the witnesses know they have the information, empowering them to feel comfortable and confident during the interview (Fisher & Geiselman, 1992).

Ground rules are recommended to help perceived autonomy (Ali et al., 2020; Fisher et al., 2011). These rules, ‘never guess’, ‘report everything’, ‘say if you do not remember’, and ‘tell me if you do not understand the question’, are all in place to reduce witness error (Blank & Launay, 2014; Goldsmith et al., 2005). Informing witnesses that it is normal for them not to know the answers to questions can minimise guessing (Fisher et al., 2011; Fisher & Geiselman, 1992; Holliday et al., 2012; Scoboria & Fisico, 2013). This means police should reassure and inform witnesses that they do not have to answer questions if they are unsure. This can reduce the quantity of information but ensures a higher accuracy level (Brubacher et

al., 2015; Bull & Milne, 2020; Evans & Fisher, 2011; Pansky, 2012; Paulo et al., 2019; Scoboria & Fisico, 2013).

1.4.7. Initial Account Interview Guidelines: Suggesting eye closure.

In the context of questioning at a crime scene, law enforcement is advised to employ support tools, and one such technique is suggesting eye closure to enhance memory recall, as per the guidelines from the College of Policing in 2019. While evidence indicates that the impact of eye closure is limited in initial account interviews, it proves beneficial in scenarios involving repeat recall (Vredeveldt et al., 2014). The rationale behind suggesting eye closure lies in its potential to alleviate distractions and reduce cognitive load, particularly in the dynamic and visually overwhelming environment of an active crime scene.

The role of eye contact in a memory task is nuanced, and research suggests that maintaining eye contact during questioning can heighten the social demands of the situation (Markson & Paterson, 2009; Vredeveldt et al., 2011). This diversion of cognitive energy away from remembering may, in turn, decrease the amount of information recalled (Perfect et al., 2008). However, it's essential to recognise that eye contact is crucial in building trust and rapport between the interviewer and the witness. Striking a balance between eye closure to reduce distractions and maintaining eye contact for rapport building becomes imperative, necessitating a nuanced approach tailored to the specific demands of the witness and the crime scene environment.

Eye closure proves beneficial in minimising visual input distractions, particularly in the often chaotic setting of an active crime scene, contributing to reduced cognitive load and improved memory (Vredeveldt et al., 2011). Despite the varying effects reported in the literature, the College of Policing recommends using eye closure as

an effective technique, especially during an initial account interview, acknowledging its potential to enhance the witness's ability to recall crucial details, even if it may have a limited impact in the initial stages.

However, it is crucial to approach the suggestion of eye closure with sensitivity and consideration for the witness's comfort and preferences, as emphasised by the College of Policing in 2019. Some witnesses may find this technique conducive to their recall abilities, others may not, and respect for their needs is paramount. The decision to suggest eye closure should be context-specific and made with a thorough understanding of the witness's disposition, ensuring that the chosen approach aligns with their comfort and enhances the overall effectiveness of the interview process.

1.4.8. Initial Account Interview Guidelines: Advice on information exposure.

Following the initial account interview, it is crucial to extend the informational support provided to witnesses by informing them about potential sources of misinformation (College of Policing, 2019). Acknowledging the inherent risks associated with exposure to inaccurate information and its potential impact on personal memories becomes paramount in fostering a more resilient and accurate recollection process. Research supports the notion that explicitly informing witnesses about the possible inaccuracies in the information they might encounter can significantly mitigate the harmful effects of misinformation (Blank & Launay, 2014).

A particular focus should be placed on emphasising the detrimental consequences of discussing the witnessed event in a social setting. Research highlights the vulnerability of memory accuracy when events are casually recounted in social contexts (Gabbert, 2004). Understanding and communicating to the witness

about the potential damage from such discussions are critical in fortifying the witness's memory against distortions.

While it might be impossible to entirely prevent misinformation exposure, creating awareness among witnesses regarding its potential dangers is a proactive strategy to mitigate its impact (Jack et al., 2014). By empowering witnesses to know how external information can shape their memories, law enforcement can contribute to a more vigilant and discerning mindset, enabling witnesses to navigate subsequent encounters with potentially misleading information better.

The advice to witnesses goes beyond cautioning them about misinformation exposure; it extends to actively involving witnesses in safeguarding their memories. Encouraging witnesses to critically approach subsequent information and question the accuracy of new details can further enhance their resilience against misinformation. This collaborative approach between investigators and witnesses aligns with the contemporary understanding of memory dynamics and contributes to witness testimonies' overall accuracy and reliability.

1.5. Initial Account Interview Research

There are several potential theoretical and investigative benefits of conducting an initial account interview quickly at the scene of a witnessed event or soon after. These potential benefits arise from the psychological literature concerning the impact of stress on witness memory, witness confidence, post-event information and misinformation, and memory conformity. Research investigating the use and impact of the initial account interview on witness memory relevant to these topics is sparse. Nonetheless, each topic is introduced, and where it exists, relevant initial account interview research is introduced and evaluated.

1.5.1. Stress.

Although research has extensively investigated witness memory in a laboratory context, one aspect of witness experience that is hard to replicate is the stressful nature of (i) witnessing a crime and (ii) being interviewed by a police officer (Penrod et al., 1995). Generally, however, there is agreement that stress can affect witness memory. Stress has been experimentally examined with military personnel undergoing a survival exercise. High-stress situations reduced eyewitness accuracy and post-event identifications (Morgan et al., 2004). This experimental format has also found that stress increases the vulnerability of memories to be altered by post-event misinformation, with highly stressful situations leading 27% of participants to mistakenly report they were threatened with a gun and 40% to misidentify the perpetrator (Morgan et al., 2013). A meta-analysis confirms these findings, with evidence that negative emotion at the encoding stage impairs subsequent recall (Deffenbacher et al., 2004).

The negative impact of stress is attributed to the allocation of attention in stressful situations (Easterbrook, 1959), resulting in witnesses identifying the source of concern with focused attention at the detriment of peripheral details (Kaplan et al., 2016). This central and peripheral effect often results in accurate recall of specific central details, lacking information about other surrounding verbatim elements (Reisberg & Heuer, 2014). Although peripheral details are essential for police investigations, research shows that witnesses who experience negative emotion during a mock witness paradigm report more accurate descriptive accounts of a perpetrator (Houston et al., 2013). However, they perform poorly when it comes to recognition and identification tasks (Houston et al., 2013). Stress does not appear to render witness memory irrelevant, but understanding the potential effects of stress

on recall memory is a fundamental part of information gathering, albeit interviewing officers cannot control pre-interview stress experiences. That said, it is important to understand the possibility of a stress response during an interview, which can negatively impact memory performance.

An initial account interview can potentially alleviate stress for witnesses through its structured and supportive approach. The initial account interview, conducted by the first responding officer at the crime scene, allows for a timely and systematic information collection. The police officer guiding the witness through the interview with a professional and empathetic demeanour creates a reassuring environment. Moreover, the emphasis on obtaining the initial account promptly helps prevent memory degradation over time, reducing the anxiety associated with memory lapses. By fostering open communication and prioritising witness comfort, a well-conducted police initial account interview not only enhances the accuracy and reliability of the information obtained but also serves as a supportive mechanism that minimises stress and ensures a more constructive experience for witnesses involved in the criminal investigation process.

1.5.2. Confidence.

Examining witness confidence in personal memories involves evaluating and monitoring metamemory (what is known about one's memories) (Metcalfe & Shimamura, 1994). Confidence is not intrinsically representative of memory accuracy but rather represents the perception of memory accuracy. Confidence judgements can be explained using the cue-belief model, where memory judgment is based on the accessibility of memory traces and cues (Leippe et al., 2006). The three cues contributing to the perception of memory accuracy are intrinsic, relating to our knowledge about accurate memories; self-credibility cues, about the personal

perception of memory ability; and extrinsic cues, involving situational elements from the environment or individual feedback.

The general public's perception of memory as an accurate representation of events (83%) (Simons & Chabris, 2011) has implications for an individual's ability to make confidence judgements using self-credibility cues accurately. In turn, this relates to the perception of others' confidence, with mock juries interpreting high confidence as an indicator of reliability (Sauer et al., 2017). This misunderstanding could lead to false confidence in episodic memories that have consistently been shown to be fragile and typically incomplete representations of events (Flindall et al., 2016; Henkel, 2017; Wells, 1995; Wells & Loftus, 2003). In contrast, low confidence is especially present in older adults with negative perceptions of their ageing memory (Wylie et al., 2014), making them more susceptible to misinformation (Roediger & Geraci, 2007). The natural ageing of key brain areas relevant to source monitoring and episodic memory compacts this (Dando, 2013). Furthermore, our confidence levels are very susceptible to positive and negative feedback, with feedback impacting confidence regardless of accuracy (Iida et al., 2020). This highlights how sensitive metamemory is to influence, depending on the perception of intrinsic, self-credibility, and extrinsic cues.

A popular method of metamemory evaluation is asking witnesses to self-report confidence levels using a confidence scale. This has been shown to demonstrate a positive relationship to memory accuracy during a witness interview (Allwood et al., 2005; Luna & Martín-Luengo, 2012). However, it has been argued that using self-report confidence scales would be too time-consuming in a real-world police interview, as the scales would have to be applied to every piece of information the witness reported (Paulo et al., 2016). The relationship between confidence and

accuracy of memories is most reliable when witnesses are asked about confidence immediately after the recall (Brewer & Weber, 2008). This is further demonstrated by research that measured automatic language qualifiers. In an experimental setting, participant confidence was related to accuracy when examining spontaneous verbal utterances of uncertainty or qualitative confidence judgements (I am sure that... it might have been, etc.) (Paulo et al., 2019). This demonstrates that instinctual memory judgements are based upon accurate information.

Unfortunately, confidence does not provide the ability to filter or identify post-event misinformation (Brewin et al., 2020; Wells et al., 2020; Wells & Bradfield, 1998; Wixted et al., 2016; Wixted & Wells, 2017). The possibility of exposure to misinformation increases with time, as does confidence in memories, with misinformation and errors becoming harder to notice (Mudd & Govern, 2004; Wang et al., 2014). Additionally, when confidence decreases over time, it disproportionately affects true memories over false memories, reducing the witness's ability to self-monitor for accuracy (Shapira & Pansky, 2019). In the real world, witnesses will likely experience post-event misinformation (Loftus & Greenspan, 2017). Therefore, it is essential to consider witness confidence as generally representative, understanding that erroneous factors may have impacted metamemory.

An initial account interview serves as a crucial mechanism to address and mediate the multifaceted concerns related to witness confidence, especially regarding memory accuracy and susceptibility to misinformation. By prioritising a timely initial recall, the initial account interview minimises the impact of memory degradation over time and reduces the propensity for post-event misinformation. Additionally, the initial account interview provides an opportunity to gauge witness confidence levels immediately after the recall. This aligns with research highlighting

the heightened reliability of the relationship between confidence and accuracy when assessed promptly (Brewer & Weber, 2008; Paulo et al., 2019). The police officer's skilful use of language during the interview can help elicit qualitative confidence judgments, providing valuable insights into the witness's instinctual memory judgments and contributing to more accurate metamemory evaluations (Paulo et al., 2019).

1.5.3. Post-event information and misinformation.

Information about a crime event that is seen or heard after the crime has been witnessed but not experienced by the witness can be seamlessly subsumed into the original memory trace. This is typically referred to as post-event information and post-event misinformation (Gabbert et al., 2003; Mudd & Govern, 2004). When post-event information is inaccurate and reported in place of original, accurate information, this is known as the misinformation effect (Ayers & Reder, 1998; D. B. Wright & Loftus, 1998). Where the information subsumed is, in fact, correct but was not experienced by the witness, this is referred to as post-event information. Mistakenly assimilating new information into the original memory trace and recalling it as though it was experienced first-hand is known as a source attribution error (Loftus, 2005; Zaragoza & Lane, 1994). There are various ways that post-event information and misinformation can be presented to witnesses in the aftermath of a crime and later during an investigative interview.

Memories are especially vulnerable to misinformation during recall (Chan et al., 2009) and when engaging with other witnesses or social media, for example (Kassin et al., 2007). Although multiple retrievals of the same event can increase the amount of information recalled and strengthen the memory trace and pathways (Roediger & Butler, 2011; Roediger & Karpicke, 2006), memories are susceptible to change,

especially when new information is presented to a witness (Jack et al., 2014; Zaragoza et al., 2013). When conducting a witness interview, it is essential not to introduce new information or influence the witness's memory (Sharman & Powell, 2012). During a police interview, there is an imbalance of power, which can enhance the effect of misinformation (Toglia et al., 2017). Witnesses being interviewed are particularly vulnerable; it has been found that even hand gestures can lead witnesses to alter their recall (Gurney, 2011; Gurney et al., 2016).

Witnesses who experience the same event may have differing memories of what happened. Everyone has different schemas that direct attention to different elements. A witness's schemas can infect the accuracy of memory, particularly when the memory is incomplete or unusual (Davis & Loftus, 2017; Koriat et al., 2000; Tuckey & Brewer, 2003). This often occurs unknowingly, with witnesses struggling with source attribution errors (Lindsay et al., 2004).

Directly after an event, the memory trace has not had time to consolidate and is very vulnerable to suggestion (McGaugh, 2000; Wixted, 2004). Additionally, the act of memory retrieval can alter the memory (Anderson et al., 1994). Thus, one of the first opportunities for exposure to misinformation is immediately after the event. This is particularly significant as 86% of surveyed witnesses reported discussing the event with a co-witness (Paterson & Kemp, 2006). When co-witnesses discuss an event, the individual accounts become more similar (Gabbert, 2004; D. B. Wright et al., 2000), with 71% of experimental co-witnesses assimilating and reporting information they have not witnessed (Gabbert et al., 2003). This collaborative co-witness recall can produce more information than individual recall, with witnesses mistakenly adding additional details they have not experienced. This increase in information could have an advantage in a social context, with close friends and

partners outperforming strangers in recall tasks through mutual memory cueing (Andersson, 2001; Andersson & Rönnberg, 1995; Fussell & Krauss, 1989). However, this leads to increased information but at the cost of inaccurate and false additions (Douglass & Bustamante, 2014; Yaron-Antar & Nachson, 2006). This memory conformity can also remove any outlying correct information and detail from a witness' account, reducing the amount of accurate detail police have about a crime (Douglass & Bustamante, 2014).

In part, memory conformity can be explained through social identity theory, with people deferring to others' accounts to appear agreeable and be perceived as more likeable (Tajfel & Turner, 2004). Known individuals develop unique ways of communicating to reduce cognitive demand (Fleming & Darley, 1991; Fussell & Krauss, 1989), creating a transactional memory system of sharing information (Wegner et al., 1991). Therefore, co-witness contamination is enhanced if the collaborators have a close relationship (Hope et al., 2008), with the strength of social influence increasing the effect of memory conformity (French et al., 2011; C. B. Harris, 2010; Hope et al., 2008). This social influence can affect the perceived context of recalling an event with recalling something formally and conversationally retelling a story requiring different goals and, as a result, different content (Dudukovic et al., 2004). When retelling aims to entertain, elements are exaggerated, whereas when the goal is to inform, perceived irrelevant information is excluded (Marsh & Tversky, 2004). This altered social recall can affect future attempts to recall an event accurately (Hope et al., 2008; Marsh, 2007; Marsh et al., 2005).

Additionally, uncertainty leaves memories vulnerable to change, with witnesses more likely to have a different understanding of an event to fit with another witness

who seems more confident (Gabbert et al., 2003; Sousa & Jaeger, 2022; D. B. Wright et al., 2000). This effect is also apparent when a co-witness is perceived to have an encoding advantage, with witnesses assimilating increased misinformation from co-witnesses when they believed the co-witness had viewed the event for a more extended period (Gabbert et al., 2006). Additionally, co-witness age affects perceived credibility, with younger witnesses being seen as more competent and, therefore, having a more significant influence on co-witness contamination (Kwong See et al., 2001).

An initial account interview has the potential for mitigating the misinformation effect since, in real life, it is difficult to prevent the harmful effects of misinformation when an investigative interview cannot be conducted immediately, as the potential sources are immediate and varied. Separating witnesses at the crime scene has a clear memory protective advantage, with co-witness contamination reducing information gathered. Although it can be hard to separate witnesses, especially if they have a close relationship, it is even more important to do so if this is the case (Gabbert, 2004). The more intimate the witness relationship, the more influential and damaging co-witness contamination can be (French et al., 2011; Hope et al., 2008).

In light of the importance of separating witnesses as soon as possible, further research indicates that giving a witness the opportunity for solo free recall before co-witness discussion can protect against subsequent misinformation (Jack et al., 2014; Sutherland & Hayne, 2001; Wang et al., 2014). However, there are logistical limitations of providing an opportunity for immediate post-event recall at the scene of a crime. One potential solution to this is the use of recall questionnaires. Mock witnesses who received an immediate opportunity to recall, using a questionnaire, post-event were more accurate during a follow-up interview one week later (Wang et

al., 2014). Paper questionnaires such as the Self Administered Interview can improve memory (Dando et al., 2020; Gabbert et al., 2009; Horry et al., 2021) and protect from post-event misinformation (Gabbert et al., 2012).

Informing a witness of the dangers of co-witness contamination and highlighting the importance of source monitoring can reduce the conformity of witness memory (Blank & Launay, 2014; Bodner et al., 2009). It has been demonstrated that warnings are more effective when they precede the presence of misinformation (Bulevich et al., 2022; Echterhoff et al., 2005; D. B. Wright, 1993). However, some research has also found that co-witness contamination is pervasive even with warnings provided (Paterson et al., 2011). This varying effectiveness could be attributed to the modality of the information (written/heard/seen) (Jones et al., 2018).

Regardless, it is considered vital for police to inform witnesses of the dangers of misinformation at the first possible opportunity, generally during an initial account interview at the scene of the crime (College of Policing, 2019). This can reduce co-witness contamination, although it is hard to remove the effect altogether (Jack et al., 2014). Logistically, this becomes difficult when an event has many witnesses, such as a riot. It is, therefore, essential to examine the different possible ways witness memory can be protected as quickly as possible after an event.

1.5.4. Witness separation.

As previously discussed, separating witnesses as soon as possible can help reduce the effects of co-witness contamination (Gabbert et al., 2003). Prioritising this is important as it allows for a solo recall before a co-witness discussion, protecting the memory from future misinformation (Jack et al., 2014). Additionally, the presence of others has been shown to have varying effects on recall. In some cases, the presence of others has been demonstrated to increase arousal, which may have a

detrimental impact on complex tasks. This is so pervasive that witness performance can decrease as the number of individuals present increases (Wagstaff et al., 2008). This could also be attributed to attentional conflict, with fewer cognitive resources available for the task at hand (Baron, 1986). Depending on the circumstances, witness separation may not be possible, for instance, if uncooperative witnesses outnumber officers at a scene; however, it should be done wherever possible (College of Policing, 2019).

1.5.5. Clarifying sources of information.

Source attribution error can unintentionally alter witness memories (Gabbert et al., 2003; Zaragoza & Lane, 1994). Informing witnesses of this possibility and the importance of source monitoring can reduce the adverse effects of source attribution errors in some cases (Bodner et al., 2009), but the practical application of providing this information can be challenging. Asking witnesses to clarify the sources of information after they have provided information could be interpreted as disbelief. Therefore, discussing this with the witness before the interview is a better approach, viewed as essential to invoke and maintain rapport (College of Policing, 2019).

1.6. Investigative Interviews

More formal, in-depth investigative interviews are usually conducted face-to-face in person using the PEACE model. PEACE is an acronym for the stages of an interview (Planning and Preparation, Engage and Explain, Account, Closure, Evaluation. Introduced across England and Wales in 1992, the PEACE model remains the prevailing approach for interviewing cooperative witnesses (College of Policing, 2022; Griffiths & Milne, 2006; Milne, Shaw, et al., 2008; Milne & Bull, 2002;

Milne & Shaw, 1999), and among other techniques, it is a cognitive interview technique for cooperative witnesses.

The cognitive interview (Fisher & Geiselman, 1992; Geiselman et al., 1986) was developed and refined using psychological research with significant practitioner input. In its current form, the cognitive interview combines principles from the psychological understanding of memory (Bower, 1967; Holliday et al., 2012; Westera et al., 2011). It is centred on supporting the retrieval processes involved in episodic recall, which has been likened to a form of time travel whereby the witness has to consciously reexperience the event in question towards accessing the stored memory trace (Kaplan et al., 2016; McGaugh, 2000; Tulving, 1984, 2002; Tulving & Thomson, 1973; Wheeler et al., 1997). Accordingly, the cognitive interview provides interviewers with several mnemonic techniques for improving witness cognition at retrieval, encouraging witnesses to recall as much good-quality information about the event as possible. In addition to the mnemonic components, the cognitive interview includes interpersonal social guidance and good practice processes for managing the social context of the interview to the best effect (Fisher & Geiselman, 1992, 2010; Granhag et al., 2018). Since investigative interviews are necessarily social interactions, it has been widely reported that prosocial and witness-focused interactions support recall since comfortable witnesses have constantly been found to be ‘better’ witnesses regarding reduced erroneous recall (Abbe & Brandon, 2014; Baxter et al., 2006; Dando et al., 2023; McGroarty & Baxter, 2009).

The cognitive interview is widely accepted as one of the most effective methods for improving witness recall without concomitant increase in errors (Bull & Milne, 2020; Memon et al., 2010; Paulo et al., 2013). The cognitive interview is arguably one of the most researched witness recall techniques. Although the PEACE model

countenances the cognitive interview, it is not the focus of the research program reported here. Instead, given the initial account interview (when used) precedes an investigative interview, the two retrieval opportunities are not mutually exclusive since theoretical and applied understanding of witness memory indicates that the former should impact the latter. Accordingly, the latter can be used as a metric to better understand the efficacy of the initial account interview. As such, it is sensible to provide a brief overview of the basic tier 1 PEACE investigative interview taught to police interviewers.

Given the tiered building block approach to training police to conduct investigative interviews in the UK, and since a tier 1 PEACE investigative interview is the one used in this thesis to better understand the efficacy of the initial account interview, the overview that follows is confined to introducing a tier 1 PEACE investigative interview. A basic tier 1 PEACE investigative interview training for police officers includes ‘good’ communication skills and incorporates many of the social and communication elements of the cognitive interview (Griffiths & Milne, 2006) but excludes the more cognitively demanding mnemonic techniques. The PEACE structured approach, which maps onto the cognitive interview technique, helps officers to develop sound basic interviewing skills and supports witnesses who have experienced volume, non-violent, less severe types of crime to explain their experiences. Essential Guideline elements are further explored below.

1.6.1. Rapport building.

The cognitive interview has embedded rapport building into the interview process, with guidelines instructing to be friendly, open, informative, disclose personal information, show interest in the witness and actively listen (Bull & Milne, 2020; Fisher & Geiselman, 1992; Milne et al., 1999). However, guidelines are not specific

in detailing precisely what rapport looks like or how it can be appropriately implemented (Bull & Baker, 2020; Vallano & Schreiber Compo, 2015). Studies looking at the verbal and physical elements of rapport building have found that using an appropriate combination of both results in improved recall (R. Collins et al., 2002; Kieckhaefer et al., 2014; Nahouli et al., 2021).

1.6.2. Witness centred interviewing.

Interviewers are tasked with collecting as much investigation-relevant information during an investigative interview as possible (College of Policing, 2019). This can result in witnesses feeling pressured to provide specific details they may not know or be unable to remember (Shepherd & Milne, 1999). These situational demands can negatively impact the retrieval processes, whether the pressure is real or perceived. Accordingly, a witness-centred or witness-compatible approach to asking questions is countenanced. Investigators are encouraged to adopt an open questioning style as it has been shown to increase the quality and quantity of recall compared to closed questions (Brubacher et al., 2020; Oxburgh et al., 2010; Sharman & Powell, 2012) and use witness-compatible language and terminology to reduce unintentional misinformation and leading questioning. Seminal research demonstrating the power of verb usage in witness questioning, with car crash descriptions (bumped/smashed), resulted in different car speed estimates and even incorrect reports of broken glass (Loftus & Palmer, 1974).

A witness-focused or witness-centred approach to interviewing cooperative witnesses concerns handing social control to the witness and treating the witness as the ‘expert’ (Bull, 2018; Dando et al., 2015; Fisher & Geiselman, 1992). As introduced in section 1.4.1. Rapport building helps to ensure the process is witness-focused and that, as far as possible, the witness is socially comfortable (Fisher et al.,

2011). It is seen as good practice to explain to the interviewee that notes will be taken during the interview and that this is the interviewee's opportunity to describe their experiences. Interviewers should encourage witnesses to voice anything relevant, explaining that there is no time limit for the interview and that as much detail as possible is required, encouraging the interviewee to 'tell everything'. The witness should be reassured that they will not be interrupted. Explaining the interview process and including explanations and instructions to report everything has been found to increase witness-perceived autonomy (Fisher et al., 2011).

1.6.3. Tier 1 PEACE investigative interview.

The tier 1 PEACE investigative interview procedure, as explained by The College of Policing (College of Policing, 2022), constitutes a structured framework to facilitate effective information gathering in investigative interviews. Comprising five distinct stages—Planning and Preparation, Engage and Explain, Account Clarification and Challenge, and Closure and Evaluation—the PEACE model ensures a systematic approach to conducting interviews, which is crucial in law enforcement and investigative contexts.

Planning and Preparation represent the foundational phase of the PEACE model. Considering the complexity of recall and the importance of witness information to investigations, this stage demands meticulous forethought and organisation from investigative interviewers. Central to this stage is formulating a comprehensive plan delineating the interview's objectives, target individuals, and logistical considerations such as appropriate timing. This preparatory phase is essential for optimising the efficacy and focus of the subsequent interview process.

The Engage and Explain stage serves as the initial interaction between the interviewer and interviewee, emphasising the establishment of rapport and providing

a transparent overview of the interview's purpose and objectives. This introductory phase sets the tone for the subsequent exchange, fostering an atmosphere conducive to open communication and cooperation.

Account Clarification and Challenge constitute the core information-gathering phase of the interview, characterised by a structured approach to eliciting relevant details from the interviewee. The process begins with the free recall stage, wherein key points are noted for reference. Subsequently, the cued recall phase takes place, guided by ground rules, and systematically probes the interviewee to elaborate on the previously recalled information. Moreover, prompts and probes are employed to encourage the disclosure of additional pertinent details.

The closure marks the culmination of the interview process, wherein witnesses can supplement or revise their earlier statements. In this stage, interviewers provide a final chance for interviewees to pose any remaining questions or concerns, ensuring a respectful and comprehensive conclusion to the interaction. Evaluation represents the concluding stage of the PEACE model, wherein the information garnered during the interview undergoes critical review and analysis. This assessment involves aligning the gathered data with existing investigative information, potentially informing subsequent investigative strategies and directions. By systematically evaluating the collected information, investigators can refine their focus and decision-making processes, enhancing the overall efficacy of the investigative process.

2. Chapter Two: Gathering an Initial Account. Challenges and Future Directions

2.1. Potential Future Direction

The future of initial account interviewing is poised for significant advancements with the integration of diverse methodologies and technologies. Self-administered interviews are likely to continue evolving, offering witnesses a more accessible and flexible means of providing initial accounts, especially in diverse settings or in scenarios involving multiple witnesses. Remote interviewing, facilitated by video conferencing platforms, may become increasingly prevalent, overcoming geographical barriers and allowing for prompt interactions with witnesses as appropriate. The shift toward video interviewing could enhance the richness of witness statements by capturing recorded non-verbal cues and contextual details. These non-verbal cues are only apparent at the time of interviewing and rarely recorded or accounted for beyond the specific content of the interview. They can provide important information about demeanour and mood that are not immediately apparent and could be relevant to criminal investigations.

Additionally, the ongoing debate between written versus spoken recall may see a nuanced approach, with technology playing a role in adapting to individual preferences and optimising recall accuracy. In this landscape, chatbots may emerge as valuable tools, offering a conversational interface that guides witnesses through the initial account process, providing structure to the interview while allowing for a more natural interaction. These advancements collectively signal a dynamic future for police interviewing, characterised by enhanced accessibility, adaptability, and the integration of technologies catering to a spectrum of witness needs.

2.1.1. The Self Administered Interview.

The self-administered interview was an innovative approach derived from the cognitive interview techniques, allowing for collecting preliminary witness information without needing a trained interviewer (Gabbert et al., 2009). Self-administered interviews are now widely implemented in the UK for gathering accounts, especially in incidents involving multiple witnesses and/or where an interview is unlikely to occur quickly or is unnecessary (Hope et al., 2013). The self-administered interview employs open-ended and probing questions to elicit information. Still, it is detailed and rather extensive, so it is more akin to a Tier 1 style PEACE interview that includes drawing and some mnemonic techniques. A self-administered interview is not akin to an initial account interview, the latter being more straightforward and arguably more basic. A self-administered interview can facilitate the consolidation of memory traces and has been found to contribute to a subsequent increase in accurate information during a more comprehensive cognitive style interview (Dando et al., 2020; Horry et al., 2021; Pfeil, 2018).

One notable advantage of the self-administered interview is that it may have a protective effect in making memories robust to influences from post-event misinformation (Gabbert et al., 2012; Gawrylowicz et al., 2014). By establishing a preliminary account that remains uncontaminated by external influences, witnesses appear better positioned to provide reliable information throughout the investigative process. The effectiveness of the self-administered interview is further enhanced when combined with sketching, catering to the visual processing strengths of individuals (Dando et al., 2020; Matsuo & Miura, 2017).

The self-administered interview has demonstrated particular efficacy in certain demographic groups. For older adults, the self-administered interview has proven

beneficial in memory recall (Dando et al., 2020; Gawrylowicz et al., 2014; Pfeil, 2018). Additionally, in high-stress situations, the self-administered interview has been shown to offer support to witnesses, helping them navigate the challenges of recalling information amid emotional strain (Krix et al., 2016).

However, the self-administered interview has limitations. While the College of Policing (2019) recommends completing the self-administered interview at the scene with an officer present for support, issues such as the suitability of handwritten interviews and potential visual or physical impairments of witnesses need consideration. The recommendation to use the self-administered interview only when taking a spoken initial account is sometimes impossible to achieve and underscores the need for a nuanced approach based on the circumstances and the witness's capabilities.

The potential for digitalising the self-administered interview is a promising avenue for streamlining witness interviews and providing instant responses. Although limited research has been conducted on using technology in conjunction with the self-administered interview, exploring digital options could offer advantages regarding accessibility, efficiency, and data management. In Australia, they have implemented a digital equivalent of the self-administered interview application called ‘iWitnessed’ (Paterson et al., 2018). This tool assists witnesses and victims in immediate recall by allowing them to report details of a crime, including video and images, as soon as the event has occurred. Experimental use of this tool has found that it increases recall compared to participants who took part in a written free-recall condition (Chevroulet et al., 2024). This tool demonstrates the usability and value of providing technological solutions to the criminal justice system. Regardless, the self-administered interview is not a substitute for an initial account interview because of

the immediate need for information gathering at the crime scene and the responsive nature of a conversational interview (College of Policing, 2019).

2.1.2. The use of technology.

Traditionally, witness interviews have followed a face-to-face format (Ministry of Justice, 2022) . However, the landscape is evolving, driven partly by the recent COVID-19 outbreak, necessitating a shift towards remote interactions. The criminal justice system has grappled with the efficacy of remote interviewing, raising relevant questions about its advantages and implications for memory recall (Dando et al., 2020). One notable advantage of remote interviewing, including computer-mediated and video-mediated communication, is the reduction in time and cost associated with witness interviews (Kuivaniemi-Smith et al., 2014). Another critical consideration is the impact of physical presence on memory, as studies show that arousal might increase when another person is present, potentially complicating the recall process (Wagstaff, 2008).

In remote interviewing, computer-mediated communication leverages technology to facilitate social connections and information sharing. Another approach is video-mediated communication, wherein face-to-face interviews occur remotely through electronic screens. The research underscores the effectiveness of video-mediated communication interviews, proving comparable to face-to-face interactions for adults and children in some cases (Nash et al., 2014). Notably, reducing the time between the event and the interview in video-mediated settings has been linked to increased accuracy and detail of information provided (Nash et al., 2014).

Despite concerns about the impact of remote interviewing on communication and rapport building, studies indicate that rapport can often be equally perceived in face-to-face and video-mediated interviews (Jenner & Myers, 2019). Rapport building in

video-mediated interviews is supported by increased comfort and reduced adverse effects of others being present (Weller, 2017). Research suggests that rapport building over video-mediated communication can enhance perception and learning, similar to in-person teaching (Glazier, 2016; Kim & Thayne, 2015). This has been attributed to reduced attentional and cognitively distracting elements of face-to-face interactions (Belletier et al., 2019; Pickard & Roster, 2020).

Moreover, recent exploration into virtual environments offers a unique dimension to remote interviewing, utilising immersive, realistic spaces and avatars (Ahn & Fox, 2017). Research indicates technology has the potential to enhance disclosure, reduce performance anxiety and social pressure, and increase confidence (Baccon et al., 2018; Dando et al., 2023; Herrera et al., 2020; Omarzu, 2000; Rubin, 1975). In the context of mock-witness interviews, virtual environments have shown promise in supporting episodic memory, reducing errors in recall, and fostering an environment where participants feel comfortable admitting uncertainty (Taylor & Dando, 2018). Some studies have found that rapport building more effectively elicits correct episodic memories in a virtual environment than face-to-face interactions (Dando et al., 2023). These developments signify a positive future for police interviewing, embracing technology to optimise efficiency, accessibility, and the quality of witness testimonies.

In addition to the evolving landscape of remote interviewing, the integration of chatbots offers additional advancements. Chatbots, employing conversational interfaces, provide the dynamic and accessible potential for guiding some witnesses through the initial account process. These virtual assistants provide an interview structure, pose relevant questions, and facilitate natural witness interaction. The use of chatbots in police interviewing aligns with the current trend toward leveraging

technology for more efficient investigative practices but also holds the potential for enhancing accessibility and streamlining the initial stages of witness engagement, contributing to the overall effectiveness of the interviewing process. However, more research is needed in this domain.

2.1.3. Modality.

Regarding the impact of modality on episodic witness recall, it is essential to consider the nuances that emerge from using different technologies and interviewing approaches. Traditionally, the superiority of spoken recall over static written accounts has been a consistent finding in memory research, with spoken narratives being deemed more detailed and accurate (Bergmann et al., 2004; Kellogg, 2007). However, a study employing a mock witness paradigm challenges this notion, suggesting that while overall accuracy across modalities might not significantly differ, written recall does tend to reduce the richness of fine details (Sauerland & Sporer, 2011). This reduction is attributed to the increased cognitive load associated with the act of handwriting, where the effortful nature of the process may impact the retrieval of intricate details, in contrast to the more fluent and spontaneous nature of spoken recall. Moreover, research has indicated that participants perceived rapport more positively in in-person interviews regarding attentiveness, trust, respect, and expertise than in spoken interviews (Hoogesteyn et al., 2023).

However, in witness interviewing, particularly within a free recall followed by an open-ended questions format, research indicates that the impact of modality on accuracy is relatively minimal (Sauerland et al., 2014). Recent research has shown that participants interviewed online via chat disclosed comparable amounts of crime-related information and demonstrated accuracy similar to those interviewed in person (Hoogesteyn et al., 2023).

However, a gap in the literature remains concerning comparing text versus spoken recall when the typed recall is embedded in an active conversational reflexive interview setting such as a chatbot. Beyond its role in merely digitising the self-administered Interview, technology opens the door to responsive conversational typed interviewing. This process goes beyond the traditional written account by integrating a conversational dynamic.

While the digitalisation of the self-administered interview addresses practical concerns like time pressures, the potential of reflexive typed interviewing remains an area for exploration. This method acknowledges the changing landscape of societal communication, where social computer-mediated communication has shifted from primarily spoken to increasingly typed conversations facilitated by smartphones and social media. Investigating this evolving modality aligns with contemporary communication norms and offers a nuanced understanding of how technology influences memory recall in witness interviews. The interplay between technology, modality, and memory recall requires further research into optimising witness testimonies.

2.2. Chatbots

The inception of chatbot use is closely linked to the surge in popularity of social media platforms. Notably, within the first year of Facebook's messaging platform, 100,000 chatbots were created (Johnson, 2017). This reflects this technology's rapid adoption and integration into digital communication channels. In the context of computer-mediated communication, the use of chatbots introduces a unique dynamic, as interactions occur without the physical presence of others. The ability of chatbots to create a semblance of human-like conversation opens new possibilities

for engagement within the criminal justice system, from providing information to guiding users through various processes. However, it is crucial to recognise the inherent limitations of chatbots, grounded in their pre-programmed nature, as this has implications for the complexity and depth of interactions they can facilitate within the criminal justice system.

2.2.1. Chatbots: Disclosure of sensitive information.

After witnessing a crime, individuals may encounter various factors influencing their willingness to share information with law enforcement. Among these considerations, apprehensions about divulging sensitive details to a chatbot emerge as an additional hurdle to this disclosure. While existing research has explored the impact of sensitive topics on disclosure to chatbots in healthcare settings (Miles et al., 2021; Nadarzynski et al., 2020), the ramifications of sensitivity in the context of a criminal justice chatbot remain an underexplored area. However, The College of Policing emphasises genuine, non-judgemental interactions during initial account interviews (College of Policing, 2019). Computers are more effective when collecting sensitive information (Lind et al., 2013; Weisband & Kiesler, 1996). This has been attributed to reduced social inhibitions relating to fear of judgement (Joinson et al., 2007).

The non-judgmental nature of chatbots, established in healthcare contexts (Nadarzynski et al., 2021), is noteworthy. The perception of reduced judgment usually fosters increased disclosure, alleviating the cognitive burden associated with managing one's perceived image (Joinson, 2001a; Kang & Gratch, 2010). Furthermore, the absence of a physical presence in virtual interactions typically contributes to decreased arousal levels, possibly positively influencing the quantity and quality of information recalled (Wagstaff et al., 2008).

Considering these cognitive elements, the use of chatbots in initial account interviewing holds the potential to be advantageous. By creating an environment perceived as non-judgmental, chatbots may enhance witnesses' willingness to share information, ultimately improving the richness and accuracy of the details they recall. While the specific implications of disclosing sensitive information to a criminal justice chatbot merit further exploration, existing insights from related fields suggest a promising avenue for using technology to facilitate more practical and comprehensive communication in witness interactions.

2.2.2. Chatbots: Trust.

Rapport building results from positive mutual attention and requires a trusting and open atmosphere (Abbe & Brandon, 2013; Tickle-Degnen & Rosenthal, 1990). Creating this atmosphere requires witness and interviewer engagement, with friendly, effortless interactions leading to feelings of cooperation (K. Collins & Carthy, 2019). Expertise has been shown to influence trust between people (Mayer et al., 1995) and traditionally trust in computers (Fogg & Tseng, 1999). This trust is built through perceptions of competence, accuracy, understanding and knowledge base (Corritore et al., 2003; Nordheim et al., 2019). It has been found that this perception of accuracy and understanding applies to chatbot errors more so than to human errors (Dietvorst et al., 2015; Merritt & Ilgen, 2008). A chatbot is judged more quickly and harshly for making mistakes than a human interviewer. If a chatbot is not perceived as knowledgeable, it directly impacts the users' level of trust in the technology (Nordheim et al., 2019). This underscores a distinctive aspect of chatbot-human interaction dynamics, where the standards for engendering trust in chatbot interactions surpass those applicable to human interactions. Nonetheless, if these elevated standards are met, there is a potential for the consistent cultivation of trust

in chatbot interactions, resulting in a reliability model, interview after interview. This observation highlights the nuanced interplay between perceived knowledge, error tolerance, and the establishment of trust in the context of chatbot-mediated witness initial account interview.

The guidelines for fostering rapport during the initial account interview highlight the importance of appropriate pacing so that the witness feels supported in their recall (College of Policing, 2019). Studies examining the automation of this pacing in a chatbot format have found that the speed of responsiveness is essential in building trust with human users (Nordheim et al., 2019). Concerns arise about the perceived unhuman-like quality associated with swift responses. Addressing this, mirroring the speed of human messaging and breaking up lengthy text segments enhances the chatbot's interaction, rendering it more akin to human-to-human communication (Brandtzaeg & Følstad, 2017). Customer service chatbots are viewed as lacking a natural flow of conversational communication skills (Luger & Sellen, 2016). However, incorporating human-like responsiveness in chatbot interactions can contribute to fostering positive user experiences.

The perception that chatbots are poor at conveying empathy and other human emotions (Nadarzynski et al., 2021) has been exacerbated by reported negative experiences with customer service chatbots (Araujo, 2018). However, users respond congruently when chatbot communication and language are perceived as human-like (Mone, 2016). Providing a chatbot with a human-like name increases the perception of authentic interaction (Araujo, 2018). Additionally, giving a chatbot a human appearance avatar has been found to increase the amount of information participants feel comfortable disclosing, although it does not usually impact the overall accuracy of the memories (Hsu et al., 2023). This highlights how important

the design of a chatbot is as it can influence emotional connection, rapport (Abbe & Brandon, 2013) and, therefore, both acceptability and episodic memory recall (Kieckhaefer et al., 2014).

2.2.3. Chatbots: Risk.

Trust plays a pivotal role in rapport building, and the perception of risk can adversely impact trust in chatbots (Corritore et al., 2003). A study focusing on a health chatbot revealed participant hesitancy in responding to sensitive inquiries due to uncertainties about data usage, privacy rights, and confidentiality (Nadarzynski et al., 2021). Sensitive topics often trigger concerns about others' perceptions and, more significantly, apprehensions regarding the potential repercussions of disclosing sensitive information (Tourangeau & Yan, 2007). Mitigating these concerns involves addressing privacy and confidentiality. Studies indicate heightened privacy and anonymity can enhance individuals' willingness to engage with sensitive questions, mediating disclosure concerns (Joinson et al., 2007). Implementing robust privacy measures and communicating users' rights in chatbot interactions could alleviate perceived risks, fostering trust and facilitating more open and candid engagement.

2.2.4. Chatbots: Artificial intelligence.

A study using an artificial intelligence (AI) chatbot to conduct cognitive interviews to record information after participants witnessed a mock crime video found that the chatbot resulted in more accurate recall than other tools (Minhas et al., 2022). This study compared the artificial intelligence chatbot to a free recall, cognitive interview questionnaire, and basic cognitive interview chatbot. The positive implementation of this functionality in this role is very promising. However, it does not replicate the practical stages of a police investigation. The use of artificial intelligence creates

uncertainty about the outcomes of an interview. Similar issues regarding questioning and biases occur in these self-learning artificial intelligence models.

Additionally, transparency becomes a pronounced concern in the criminal justice system, given the ethical and trust issues that arise when users struggle to comprehend complex artificial intelligence algorithmic and system processes. This lack of understanding hinders the adoption of potentially beneficial technologies in policing environments. Chatbots that do not use artificial intelligence are predictable and do not deviate from the predefined rules they are programmed with. Therefore, creating a chatbot that conducts a simplified replica of the initial account interview would be beneficial as a direct replacement for the current procedure.

2.3. Summary

A broad literature review highlights the current understanding of investigative procedures and practices that can impact witness memory for forensic purposes. In addition, the general nature of episodic memory has been introduced, including why episodic recall is viewed as a demanding cognition. Both positive and negative factors that influence interviewing witnesses after a crime have been examined to highlight the gap in the research. The heightened importance placed on initial account interviews has been discussed, with the current limitations bringing critical areas for improvement into focus. An exploration of existing technology as potential solutions to these limitations has been explored. An encouraging and novel potential for using a chatbot has been highlighted, with the current research being minimal but potentially promising. Thus, the central focus of this thesis is developing and evaluating an additional interviewing modality for gathering initial accounts in the form of a chatbot referred to as ChatCharlie. The design of this tool is carefully

considered in line with both best interviewing practices and the psychological research into human-computer interaction. First, the impact or otherwise of the modality of typed interviews will be explored, as well as the use of technology across age groups. Second, a novel chatbot called ChatCharlie will be developed and empirically investigated using a mock witness paradigm. Finally, the acceptability of chatbots in an eyewitness context will be examined.

Accordingly, the series of broad research questions that naturally emerge are:

1. Concerning episodic memory performance, is a remote-typed interview an effective method for collecting investigative information from crime witnesses versus a remote-spoken interview?

2. Is a remote chatbot interviewer (typed) an effective method for collecting initial account information from witnesses and victims of crime across a wide age range?

3. What are the potential benefits and challenges of integrating chatbots into initial account interviews, and how do they compare to traditional face-to-face interviews?

4. What are the perceptions and experiences of witnesses regarding using chatbots and other technology-assisted interviewing methods in the criminal justice system?

3. Chapter Three: Episodic Memory Performance: A Comparison of Typed and Spoken Modalities

3.1. Brief Introduction

This chapter presents the first in a series of empirical investigations of eyewitness memory performance as a function of reporting modality. Witness testimony is pivotal in criminal investigations, serving as a cornerstone in establishing legal culpability and guiding police investigations. Rapidly collecting an initial brief account from a witness can provide important investigative information promptly and help consolidate memory for the event, making it more robust and stable for later in-depth retrieval (Kontogianni et al., 2020) since memory is fragile and memories degrade over time. The study reported here considers typed modality interviews as one method for potentially expediting the collection of witness information, providing insight into the costs and benefits of typed versus spoken responses to investigative questioning regarding user engagement and cognitive processes. Compared to spoken responding, are eyewitnesses able and willing to provide detailed typed accounts and does typed responding impact the quality and/or quantity of information recalled?

Recent shifts towards remote interactions in criminal justice, prompted by factors such as the COVID-19 pandemic, emphasise the need to examine the efficacy of remote interviewing. Remote interviews offer advantages in reducing time and cost and improving accessibility. Some studies have demonstrated their comparability to face-to-face interviews. However, it has been argued that the absence of an interviewer's physical presence may influence witness recall, underscoring the complex interplay between social dynamics and memory performance. Rapport

building is crucial to successful witness interviews, with positive mutual attention fostering cooperation and memory accuracy. The influence of technology on rapport-building processes is debated, with remote interactions potentially facilitating emotional connections despite physical absence.

The choice of interview modality, whether written or spoken, depends on factors like the severity of the crime and witness significance. Nonetheless, previous research suggests varying effects of different modalities on memory recall, with some studies indicating a spoken superiority effect while others suggest a written superiority effect. However, this is an under-researched area, so this study is one step towards bridging gaps in the literature, thereby contributing to understanding. Further information about the background research into this study can be found in section 1.3. Witness memory, 1.5.4. Witness separation, Rapport Building, 2.1.3. Modality.

3.1.1. The current study.

The current study investigates the differences between typed and spoken recall of an experienced event, examining memory recall as a function of text-based mock witness interviews and traditional spoken video-mediated face-to-face interviews. While previous studies have primarily focused on remote interviewing through video and audio-only modalities, this study uniquely contributes by comparing memory recall across spoken and typed modalities and, as such, sheds some light on the impact of modality in eliciting accurate and reliable witness testimonies. The literature of relevance to this initial study is sparse, and as such, rather than formulating hypotheses, a series of study-specific research questions have been developed, as follows:

1. Does typed retrieval of episodic information impact the quantity of event information recalled versus traditional face-to-face spoken retrieval, as measured by the amount of information recalled and its accuracy?
2. Does typed retrieval of episodic information impact the quality of event information recalled versus traditional face-to-face spoken retrieval as measured by the number of errors and confabulations?

3.2. Methods

3.2.1. Design.

A between-subjects design using the mock witness paradigm (described below) was employed to investigate differences and similarities between spoken and typed recall. The mock witness method is an accepted laboratory method used by psychological scientists since the 1970s for generating data to compare performance metrics (Wells & Bradfield, 1999), offering a significant opportunity to control variables in the initial stages of advancing understanding. The independent variable was the retrieval condition, with two levels (typed and spoken). The dependent variable was memory performance, measured by the number of correct, incorrect, and confabulated items reported and the overall percentage accuracy.

3.2.2. Participants.

Participants were recruited through online advertising and compensated for the time spent taking part through a £10 shopping voucher. A priori power analysis using G*Power 3.1 indicated that a sample size of 60 people would be more than sufficient to detect a large effect size (assuming power = .80 and α = .05). 60 people participated in this study, 8 male, 52 female. There were 30 participants in the spoken condition group with a mean age of 24.57 (SD = 7.56), with 1 male and 29

females. There were 30 participants in the typed condition group, and the mean age was 23.43 (SD = 3.21), with 7 males and 23 females. This is the accepted convention in research of this kind

3.2.3. Procedure.

All potential participants were first emailed a participant information sheet (see Appendix B) and a copy of the consent sheet (see Appendix C). All participants were then called over video at a pre-agreed time, introduced to the interviewer, and given a unique participant number to label their unique contribution to the study. Participants were allowed to ask questions about the study during this initial call before participating. Once the video call had ended, participants were provided with a unique Qualtrics link, which provided access to an online consent process, following which participants watched the video stimulus (individually) before completing the demographic survey (see Appendix D).

Having completed the demographic survey, the Interviewer video-called the participant for a second time and administered the Mini-Mental States Examination (MMSE – see Appendix E). The MMSE was used throughout this PhD research as a filler activity and to indicate a potential cognitive decline in episodic memory outside of the normally expected range since the PhD research programme reported in this thesis concerns the general population (Mean age 40 years). So, some participants were over the age of 50. From age 50, normal cognitive ageing occurs, which can impact episodic memory. While it was important to run inclusive research, it was deemed essential to control for the possibility that some older adults may be experiencing challenges with memory performance outside of that expected due to normal ageing. All participants passed the MMSE and, thus, took part in the

research. There was no variation in participant scores, and it was scored as pass or fail.

Having completed the MMSE procedure, all participants in the typed condition were instructed to end the video call and begin communication with the interviewer through the synchronous text function on MS Teams. Participants in the spoken condition remained on the video call with the interviewer and were informed that the Interview would be audio recorded. Irrespective of retrieval condition (typed or spoken), all participants participated in a basic tier 1 PEACE investigative interview that maps onto a structured PEACE interview by the UK College of Policing. Irrespective of whether the interview was spoken via video call platform or typed using the synchronous text functionality in MS Teams, the interview protocols were the same (see below section 3.3.4.4.). After the interview, all participants completed the post-interview survey (see Appendix F), after which all participants were video-called for a third time and allowed to ask questions. Participant numbers were used to anonymise and label all data collected.

3.2.4. Materials.

3.2.4.1. Stimulus.

Due to the COVID-19 pandemic, all data for this work was collected remotely. Therefore, participants were sent a one-time link to the mock-crime video over email with instructions to watch the video in a quiet room without distractions. All participants were required to use a laptop for all parts of this study. The video was accessed through a linked Qualtrics survey that enabled the researcher to establish the time the video was watched and limit the number of times a participant watched the video to one. The footage used for all studies in this body of work is a non-violent

mock-crime video lasting 1 minute. The video was filmed from a first-person perspective, as though the viewer is a bystander standing on the side of the street.

The video begins with the viewer standing on the pavement, looking towards the left. The view is of a two-way street with cars driving freely along the road. The area has some trees, and a few houses and shops are across the road. The camera/viewer then slowly turns to the right, and you can see more shops across the street and a person walking along the pavement on the other side of the road. The shops are a blue tuck shop, a white laundromat and a Threshers wine shop. Cars continue to travel in both directions. Once the camera/viewer has turned to the right, there is a roundabout with flowers in the centre. Cars are travelling around the roundabout and using all four visible exits. Beyond the roundabout are a few more shops and three-story houses with traditional Tudor beams. Much of the view beyond the roundabout is obscured by trees and vegetation.

The camera/viewer then turns back to the left and zooms in on the corner where the thresher's wine shop is located. From this corner, two people appear, walking side by side. Nearest the street is a white man wearing a hoodie, and beside him is a black woman wearing a baseball cap. The two people walk along the road past the Threshers wine shop and the laundromat and enter the tuck shop. They are in the shop for 10 seconds before they run out back the way they came round the corner and out of view. Directly behind the two people is a third white man running and following them around the corner.

3.2.4.2. Mini-Mental States Examination.

The Mini-Mental States Examination (MMSE) (Crum et al., 1993) is a short questionnaire used extensively in research of this nature to control for cognitive impairments outside of the ‘normal’ range, which might confound the results of the

research (Dando, 2013; Dando et al., 2020). The MMSE is not a diagnostic tool and was used as a filler task in study 1 and study 2 as participants' ages ranged between 18 and 82. All participants passed the MMSE and, thus, took part in the research. There was no variation in participant scores, and it was scored as pass or fail.

3.2.4.3. Ethics and materials.

This research was ethically approved by the University of Westminster research ethics committee: ETH 1920-0542 (See Appendix A). All participants in this research were emailed a participant information sheet before agreeing to participate in any of the experiments (See Appendix B). All participant information sheets contain information unique to the study and the condition to which the participant was allocated. Participants were only aware of other conditions once they received the debrief sheet. Additionally, all participant Information sheets contained background information about the research, requirements for taking part, what participants will be required to do, ethics information, and the researcher's contact details. Additionally, all participants completed an online consent form through a Qualtrics link before beginning the study (See Appendix C).

Participant demographic questions were asked across all studies in this work (See Appendix D). The standardised questions asked participants their age and sex. As this study took place entirely remotely, participants were also asked about their birth country and English language proficiency, which was an exclusionary factor. Additionally, participants were asked to confirm that their eyesight and hearing were not impaired, as this could impact the ability to perceive elements from the video, which could be misinterpreted as poor memory performance.

All participants who participated in any research in this PhD received a debrief sheet (See Appendix G). The debrief sheet contained further information about the

study's aims and purpose. A summary of what the participant had done throughout the study was also provided. A detailed explanation of how the data will be anonymised and instructions for removing data if participants wish to withdraw were given. Information about compensation for taking part and the researcher's contact details was provided.

3.2.4.4. Tier 1 PEACE interview protocol

All studies in this body of work measured memory from transcripts of a tier 1 PEACE investigative interview (See Appendix H). The PEACE model interview procedure outline is taken from The College of Policing (College of Policing, 2022) and, as such, comprises a series of phases, as follows:

Planning and preparation: It is essential to have a clear plan for the interview before it begins; this includes understanding the aims of the interview and who you will be interviewing. Organising appropriate timings for the interview is very important, particularly in Study 3, which involved participants taking part on two separate occasions. It was also essential to clarify that participants had appropriate materials, e.g. a computer with a camera and a microphone.

Engage and Explain: This stage involves building rapport with the interviewee and providing them with an overview of the objectives of the Interview. All participants who took part received a participant information sheet and were allowed to ask questions before taking part. Before the main interview stage, all participants were greeted over video call and introduced to the researcher, following guidelines for building rapport. During this phase, participants were given ground rules, ‘never guess’, ‘report everything’, ‘say if you do not remember’, and ‘tell me if you do not understand the question’.

Account, Clarification, Challenge: This stage is where the information gathering occurs. First, the free recall stage occurs, with participants being asked to recall everything they can from the video. During this stage, notes are taken concerning the key points from the recall. Following this, the cued recall phase of the interview began, with participants being reminded of the ground rules before systematically being asked to provide further information about all the key points noted from the free recall stage. Participants were prompted to give more information after each question.

Closure: After the cued recall phase, participants are asked if they would like to add further information or change anything they have previously shared. Participants are then thanked for their time and allowed to ask questions. After participating, all participants received a debrief sheet with further details about the study and the researcher's contact details.

Evaluation: Once the interview has been transcribed and coded for correct, incorrect and confabulated information, the information is added to the study data set. A summary of the Interview protocol is shown below (Table 1):

Table 1

Tier 1 PEACE investigative interview protocol

PEACE Model	Interview Protocol
Planning and Preparation	Video call participants and introduce the interviewer. Make sure sound, video and text functions are working for the participant and the researcher.
Engage and Explain	Explain what will happen in the Interview. Give the participant an opportunity to ask questions. Start the Interview. Either over synchronous text function or a recorded video interview. Inform participant of the ground rules.
Account, Clarification, Challenge	Ask the participant: Tell me everything you can remember about the video. Take note of the key elements When they have been silent for 5 seconds, ask if they remember anything else (repeat this until the participant indicates there is nothing else). Inform the participant that they will now be asked some more detailed questions about what they remember. Remind the participant of the ground rules. Ask participants to tell you about key elements in as much detail as possible. When they have been silent for 5 seconds, ask if there is anything else they can add. Repeat this until all key elements have been covered.
Closure	Ask if there is anything further the participant would like to add Ask if there is anything the participant would like to change about what they have shared. Thank the participant, end the interview, and stop the recording. Give the participant an opportunity to ask any further questions
Evaluation	Transcribe and code Interviews

3.2.5. Coding

Transcripts from interviews are coded for distinct pieces of information. Information is categorised as correct, incorrect, or confabulated (See Appendix J). The correct information matches an element of the mock crime video. Incorrect information is when there are minor errors or mistakes. Confabulated information is not present in the mock crime video and has been falsely recalled by the participant. Each transcript is coded in the free recall and questioning stages, with each item being coded only once when first verbalised. The percentage accuracy was calculated as a function of the interview stage and overall by dividing correct items by total recall items. The coding throughout this work followed standardised procedures in previous research (Dando et al., 2022).

Below is a fictitious example of how a simple sentence would be coded following the established coding procedure:

This statement is correct:

There was a man walking along the street in a blue t-shirt. A green car goes past, and then the white man goes into a shop.

This would be coded as having 10 correct pieces of information and 100% accuracy. Participants often repeat themselves, but each information item is only coded once. If additional information is provided, then that information is coded. The participant has already said they have seen a man, so in the second sentence, only 'white' is coded as new information. Additionally, the t-shirt and the fact that the t-shirt is blue are coded as separate correct pieces of information.

When incorrect elements are introduced, it would be coded:

There was a man walking along the street in a yellow t-shirt. A blue car goes past, and then the white man goes into a shop.

Note: The t-shirt and car will still be coded as correct information.

This would be coded as 8 correct and two incorrect pieces of information with 80% accuracy.

Confabulated information would look like this:

There was a man on a scooter in a yellow t-shirt. A blue car goes past, and then the white man gets into a car.

This would be coded as 4 correct pieces of information, 2 incorrect pieces, and 2 confabulated pieces of information with 50% accuracy.

Complexities in coding arise when minor errors are present, for instance, saying 'shirt' instead of 't-shirt', as these statements could be perceived as correct or incorrect. Consistency was maintained throughout the coding process as follows. First, all interviews were initially coded by the same researcher to reduce variability. Coding decisions made were documented using a coding 'book' as the coding progressed. Item lists were checked and updated throughout the process.

3.3. Results

3.3.1. Analysis.

A series of one-way between-subjects ANOVAs investigated the effect of modality, typed recall or spoken recall on memory performance. Memory comprised correct, incorrect, confabulated items, and overall accuracy percentage. Items were also analysed within each interview phase (free and cued recall). Where Levine's test for homogeneity of variances revealed the assumption of the homogeneity of variance was violated, a Kruskal-Wallis test was conducted.

3.3.2. Overall memory performance.

Table 2.

Means & standard deviations (in parenthesis) for overall memory performance.

	Condition	
	Typed	Spoken
	Mean (SD)	
Overall Correct	47.63 (16.40)	49.40 (15.61)
Overall Incorrect	4.63 (3.16)	4.53 (2.97)
Overall Confabulations	0.87 (0.92)	0.10 (0.40)
Overall Accuracy %	89.26 (6.18)	91.60 (4.59)

3.3.2.1. Overall correct.

The effect of recall modality on correct memory items was non-significant $F(1,58) = 0.183$ $p = .671$, $\eta p2 = .003$ (see Table 2 for means and SDs).

3.3.2.2. Overall incorrect.

The effect of recall modality on incorrect memory items was non-significant $F(1,58) = 0.16$, $p = .900$, $\eta p2 = .000$ (see Table 2 for means and SDs).

3.3.2.3. Overall confabulations.

Levine's test was significant $F(1,58) = 22.727$, $p < .001$. A Kruskal-Wallis test revealed a significant effect of recall modality on overall confabulation $\chi 2(1, N = 60) = 16.347$, $p < .001$. Participants who took part in a spoken interview recalled significantly fewer confabulated items than participants who took part in a typed interview (see Table 2 for means and SDs).

3.3.2.4. Overall accuracy percentage.

The effect of recall modality on overall accuracy percentage was non-significant $F(1,58) = 2.75$, $p = .103$, $\eta p2 = .045$ (see Table 2 for means and SDs).

3.3.3. Interview phase performance.

3.3.3.1. Free recall correct.

The effect of recall modality on correct memory items during the free recall phase was non-significant $F(1,58) = 0.13$ $p = .911$, $\eta p2 = .000$ (see Table 3 for means and SDs).

3.3.3.2. Free recall incorrect.

The effect of recall modality on incorrect memory items during the free recall phase was non-significant $F(1,58) = 0.46$ $p = .831$, $\eta p2 = .001$ (see Table 3 for means and SDs).

Table 3.

Means & standard deviations (in parenthesis) for interview phase free recall and questioning.

	Condition	
	Typed	Spoken
	Mean (SD)	
Free Recall Correct	27.43 (9.76)	27.70 (8.59)
Free Recall Incorrect	1.33 (1.24)	1.27 (1.17)
Free Recall Confabulations	0.23 (0.43)	0.00 (0.00)
Free Recall Accuracy %	93.84 (5.41)	95.79 (3.91)
Cued Recall Correct	20.23 (8.91)	21.70 (9.17)
Cued Recall Incorrect	3.30 (2.64)	3.27 (2.57)
Cued Recall Confabulations	0.63 (0.81)	0.10 (0.40)
Cued Recall Accuracy %	83.86 (9.63)	86.84 (7.35)

3.3.3.3. Free recall confabulations.

Levene's test was significant $F(1,58) = 72.953$, $p < .001$. A Kruskal-Wallis test revealed a significant effect of recall modality on free recall phase confabulation $\chi2(1, N = 60) = 7.792$, $p = .005$. Participants who took part in a spoken interview

recalled significantly fewer confabulated items than participants who took part in a typed interview in the free recall phase (see Table 3 for means and SDs).

3.3.3.4. Free recall accuracy percentage.

The effect of recall modality on accuracy percentage was non-significant in the free recall phase $F(1,58) = 2.55$, $p = .116$, $\eta p2 = .042$ (see Table 3 for means and SDs).

3.3.3.5. Cued recall phase correct.

The effect of recall modality on correct memory items during the cued recall phase was non-significant $F(1,58) = 0.39$ $p = .532$, $\eta p2 = .007$ (see Table 3 for means and SDs).

3.3.3.6. Cued recall phase incorrect.

The effect of recall modality on incorrect memory items during the cued recall phase was non-significant $F(1,58) = 0.00$ $p = .961$, $\eta p2 = .000$ (see Table 3 for means and SDs).

3.3.3.7. Cued recall phase confabulations.

Again, here, Levene's test was significant $F(1,58) = 23.285$, $p < .001$, and as a result, a Kruskal-Wallis test was conducted. Participants who took part in a spoken interview recalled significantly fewer confabulated items than participants who took part in a typed interview in the cued recall phase, $\chi 2(1, N = 60) = 11.490$, $p < .001$ (see Table 3 for means and SDs).

3.3.3.8. Cued recall phase accuracy percentage.

The effect of recall modality on accuracy percentage was non-significant in the cued recall phase $F(1,58) = 1.81$, $p = .184$, $\eta p2 = .030$ (see Table 3 for means and SDs).

3.4. Discussion

This chapter presents findings from an experimental study examining mock witness memory performance across two distinct recall modalities: via the text chat function or over video-mediated communication. Employing a mock-witness approach, participants were divided into two groups: recalling details through typed and spoken means. The findings of this study were non-significant in all measures (correct, incorrect, and percentage accuracy) across modalities, apart from confabulations. The non-significant differences were consistent globally and as a function of the recall phase. However, turning to confabulations, participants who took part in a spoken interview recalled significantly fewer confabulated items than participants who took part in a typed interview globally and when considering all interview phases. This significant difference was less than 1 information item and did not impact overall accuracy. Nonetheless, this difference is of note for investigative purposes and is worthy of discussion.

This study sheds some light on the potential utility of typed recall for reporting information following some types of witnessed events for some sub-sections of the general population. However, it is worth considering the nature of each of the memorial performance results in turn towards unpicking the findings reported and understanding why, in the context of a witness interview for a volume crime event, modality appears not to impact performance negatively and why typed recall may have triggered small but significant increases in confabulations.

First, participants may have been less cautious when responding by typing. During the free recall and questioning, participants who took part in a spoken interview recalled significantly fewer confabulated items than participants who took part in a typed interview. As is typically the case in mock research, confabulations in

the initial free recall account were very low, zero in the spoken and less than one in the typed. Thus, this phase of the interviews elicited the most accurate recall, which was similarly accurate across retrieval conditions. This finding sits well with the large body of work reporting mock witness recall performance when interviews have been conducted in a witness-compatible, appropriate manner (Dando et al., 2009, 2011; Dodson et al., 2015; Holliday et al., 2012; Memon et al., 2010, 2018; Paulo et al., 2016; A. M. Wright & Holliday, 2007).

Although the absolute number of confabulations was minimal (less than 1) in the free recall of the typed condition, again, as expected, this small error carried over to the cued recall, which is guided by the information provided in the free recall. This carryover effect inflated the confabulations from a mean of .25 to approaching one. It is still being determined why this difference across modalities has emerged. Still, it seems reasonable to suggest that the absence of another person may have resulted in less cautious reporting, possibly akin to online inhibition, (Hollenbaugh & Everett, 2013; Suler, 2004) whereby the absence of another person can trigger cognitive effects that can alter behaviours.

Equally, this increase in confabulated elements in the typed interview may also be linked to perceived levels of rapport during the interview that would typically be present in face-to-face interactions. Recent investigations into rapport building in typed interviews have shown that participants perceive less rapport in this modality than in face-to-face interactions (Hoogesteyn et al., 2023). This finding is important because rapport building has been demonstrated to mitigate the influence of misinformation (Vallano & Compo, 2011) and increase the amount of correct recall (Dando et al., 2023; Gabbert et al., 2021; Nahouli et al., 2021). While no misinformation was presented to participants in this study, the absence of rapport-

building cues might also have had some impact, resulting in participants being less cautious about the information they might otherwise choose not to disclose to an interviewer.

A lack of a physical presence during typed interviews might also have contributed to the observed effects. In some cases research indicates that the presence of another can heighten arousal levels and potentially complicate the recall process, possibly impairing performance in tasks requiring higher cognitive demand (Wagstaff, 2008).

Previous research has questioned the effectiveness of remote spoken interviewing (Bergmann et al., 2004; Kellogg, 2007), with evidence favouring spoken recall (Bekerian & Dennett, 1990; Sauerland & Sporer, 2011), albeit contemporary understanding has suggested this may not necessarily be the case (Dando et al., 2020; Jenner & Myers, 2019; Nash et al., 2014; Taylor & Dando, 2018). Indeed, the current study's findings showed similar accuracy across spoken and typed recall, potentially reducing the cognitive demands emanating from the presence of another person, previously assumed to impact written recall negatively (Kellogg, 2007; Sauerland et al., 2014). This would confirm the previous findings that emphasised the higher cognitive and social pressures that impact spoken recall (Grabowski, 2007). The number of confabulated items reported in the typed condition aligns with research showing that people generally report more information when writing down their memories (Sauerland et al., 2014). However, participants did not report proportionally more in this case. Very recent research examining mechanisms similar to the current research has found similar findings, demonstrating that participants interviewed online via chat disclosed comparable amounts of crime-related information and demonstrated accuracy similar to those interviewed in person

(Hoogesteyn et al., 2023). This is encouraging as it again shows how the findings reported here align with the contemporary understanding.

3.4.1. Limitations.

Individual participant differences, such as cognitive abilities, linguistic proficiency, and familiarity with technology, could have influenced performance. However, the applied nature of this research is such that including diverse participants from the general population enhances the generalisability of the findings by capturing a broader spectrum of cognitive abilities, linguistic proficiency, and familiarity with technology. Thus, this study 'speaks' to the real-world demands of collecting first-hand accounts from witnesses with varying backgrounds who may be involved in providing testimony. Efforts were made to minimise the influence of individual differences through random assignment of participants to the experimental conditions. Randomisation helps distribute any potential confounding variables evenly across the experimental groups, thereby reducing the impact of individual differences on the observed group outcomes.

In the current study, participants had a mean age of 24, contrasting with the general population's mean age of 40. This disparity raises important considerations for generalising the findings, particularly regarding memory performance and cognitive processing differences across age groups. Age-related cognitive differences are well-documented, especially in memory performance and episodic recall. Typically, younger adults tend to show better episodic memory and fewer errors in recall than older adults, partly due to differences in cognitive decline and processing efficiency associated with ageing (Tulving & Thomson, 1973; McGaugh, 2000). Episodic memory stability, influenced by factors like processing speed and attentional capacity, may remain more robust in younger individuals, while the same

tasks may present more challenges for older adults (Goldsmith et al., 2005; Tuckey & Brewer, 2003). This advantage could affect the study's results, primarily if younger participants perform better in recall conditions than an older demographic might.

Given these factors, the findings may not fully generalise to an older, more diverse population. Older adults may differ in their response to various recall modalities (textual or verbal), possibly showing more difficulty with one over the other. Moreover, using digital and remote modalities (such as typed recall) may pose distinct challenges for older individuals who may have less experience with these formats than younger participants, who frequently engage with technology (Kuivaniemi-Smith et al., 2014). Such generational differences suggest that studies aiming for broader applicability should consider including older age groups or adjusting analyses to control for age-related memory factors.

Using a mock witness paradigm in a laboratory setting may lack ecological validity compared to real-world eyewitness scenarios. Efforts were made to simulate a realistic witnessing experience in this initial stage of the research and development process. The stimuli's controlled environment and artificial nature may not fully capture the complexities and emotional aspects involved in actual witness testimonies. Furthermore, relying on a single non-violent mock-crime video as the stimulus may limit the study's ability to generalise findings to diverse eyewitness scenarios with varying levels of complexity and emotional arousal.

Despite these limitations, laboratory studies offer valuable opportunities for rigorous experimental control and manipulation, allowing researchers to isolate specific variables and examine causal relationships more precisely. Additionally, selecting a non-violent scenario was deliberate to ensure ethical considerations and participant safety while eliciting a plausible emotional response. Future research

endeavours would benefit from incorporating a more comprehensive range of stimuli representing different types of crimes and emotional contexts to enhance the ecological validity of the findings.

Moreover, while laboratory studies may lack the richness and complexity of real-world witness testimonies, they offer practical advantages in experimental control, standardised procedures, and the ability to replicate conditions across multiple participants. By carefully designing experimental protocols and minimising extraneous variables, researchers can still glean valuable insights into fundamental cognitive processes underlying eyewitness memory retrieval. The findings obtained from laboratory studies also serve as a foundation for guiding subsequent research conducted in more ecologically valid settings, such as field experiments or retrospective analyses of actual criminal cases, bridging the gap between laboratory research and practical applications.

This research was conducted with a native English-speaking population, and using online recruitment methods may introduce sampling bias, limiting the generalisability of the findings to broader cultural and linguistic contexts. Future research should explore the impact of modality on witness retrieval across diverse populations and languages to enhance the applicability of findings to real-world investigative contexts. While the study may focus on the English-speaking population, the underlying cognitive processes involved in eyewitness memory retrieval likely exhibit cross-cultural universality to some extent. Therefore, while the specific findings reported here may directly relate to English-speaking populations, findings will likely provide valuable insights into memory retrieval across different cultural and linguistic contexts. However, to fully ascertain the generalisability of the

findings, future research should explore the impact of modality across diverse populations and languages.

Conducting the study remotely amid the COVID-19 pandemic may have introduced heightened variability in participants' environments, potentially influencing their focus and engagement during the memory task. Factors such as ambient noise, interruptions, and distractions in participants' home environments could introduce confounding variables. Remote data collection expanded the study's geographic reach and facilitated the inclusion of participants who might otherwise have been excluded due to geographical limitations or mobility constraints. While concerns about noise levels, interruptions, and distractions in participants' home environments are valid, proactive measures were implemented to mitigate potential confounds. For instance, participants received explicit instructions to complete the memory task in a quiet and private setting to minimise distractions.

3.4.2. Conclusion and future directions.

Retrieving episodic memories is a cognitive process that demands significant cognitive effort (Kaplan et al., 2016; McGaugh, 2000; Tulving, 1984, 2002; Tulving & Thomson, 1973; Wheeler et al., 1997). Nevertheless, witnesses are frequently called upon to provide testimonies of critical importance to the criminal justice system (Dando et al., 2020; Kebbell & Milne, 1998; Ridley et al., 2012). Promptness in collecting these statements is imperative due to the need to gather information relevant to the investigation quickly and to consolidate the memory trace as soon as possible to reduce the deterioration of episodic memory quality (Goldsmith et al., 2005). Text-based interviews emerge as a promising solution to address this challenge, offering potentially time-saving benefits by circumventing the necessity for in-person interactions. Text-based interviews have gained traction in various fields

due to their convenience, scalability, and ability to accommodate diverse populations, with much to suggest further research and development is warranted, albeit this approach will not be appropriate for many subsections of the witness population.

In the context of witness testimony, text-based approaches could offer additional benefits, including reducing interviewer variability, managing cognitive demand, and reducing stress and anxiety sometimes associated with traditional face-to-face interviews, thereby potentially enhancing witness cooperation and recall (Dando et al., 2020). Given the promising results of this first study, a clear next step centred on understanding the efficacy of typed retrieval of witness memory is to consider appropriateness and efficacy as a function of various age groups in the general population. Here, the mean age was mid-20s, and as such, this group may be more comfortable than other adult age groups when providing typed responses. Thus, the following chapter reports extending and replicating these findings across three distinct age groups.

4. Chapter Four: Typed Recall of Eyewitness Memory Performance as a Function of Age

4.1. Abstract

Following the results of study 1, this second mock-witness study examines episodic recall performance across three different age groups when providing typed information about a witnessed event. Mirroring the first study's method, participants watched the same non-violent mock crime video and were all subsequently interviewed via text-based chat. Memory performance was analysed across three age groups (18-29, 30-54, 55 and over) for correct, erroneous, and confabulated recall and percentage accuracy globally and across the free and cued recall phases of a basic Tier 1 type interview. Results revealed that participants in the 55 and over group recalled more correct items than the two younger age groups, with no significant differences across age groups for confabulated or erroneous recall. Despite lower confidence in technology in the 55 and over group, familiarity with technology did not appear to impede retrieval. These findings further suggest that text-based retrieval may have the potential for promptly and accurately capturing witness accounts in some instances and that older witnesses may also benefit. Understanding the costs and benefits of technology across wider population groups is crucial for developing investigative tools towards improved access to justice and supporting the investigative process.

4.2. Introduction

This chapter investigates eyewitness memory performance as a function of age when providing typed information. The Office for National Statistics reports an

increasingly ageing population. For example, the number of people aged 65 and over increased from 16.4% in 2011 to 18.6% in 2021 (ONS, 2021). Accordingly, there is an increase in the prevalence of wider ranges of age groups becoming victims and witnesses to crime (Acierno et al., 2010; Biggs et al., 2009). Study 1 of this thesis provides a first indication that typed recall of a volume-type crime event offers promise as an additional 'tool' for investigators in quickly collecting an initial first account and suggests that typed first accounts may be effective for consolidating a memory trace for an experienced event. However, the mean age of participants in Study 1 was mid-20s. Natural ageing impacts numerous complex cognitions such as episodic memory (Eysenck, 2020; Wulff & Thomas, 2021) , whereby many older adults (aged 55 and older), for example, demonstrate reduced performance on episodic memory recall tasks versus teenagers and individuals in their 20s (Balota et al., 2000; Craik & McDowd, 1987; Dando et al., 2020). Hence, it is sensible to consider age-related performance towards investigating the efficacy and suitability of novel approaches, such as using a typed modality for conducting a first initial account.

4.2.1. Age and memory.

As introduced in section 3.2.1., witnesses must access episodic memory to recall a particular event. Episodic memory concerns personally experienced events, including the actions, the environment, the people, and the context within which they occurred (Datta et al., 2022; Tulving, 1984, 2002; Williams et al., 2022). Episodic memory is not an objective replay of an experienced event. Rather, the event must be mentally reconstructed. Consequently, episodic memory is fragile, can degrade and alter over time, and is typically incomplete (Brewer & Douglass, 2019; Bull, 2019; Goldsmith et al., 2005; Tuckey & Brewer, 2003). However, despite this,

witnesses can often be very confident about the things they remember and the accuracy of their recall, resulting in reporting errors sometimes centred on information learned after the event that they did not experience. Conversely, other witnesses can be less confident despite being correct, and it is the case that witnesses generally lose confidence in their memory performance as they move towards middle age and beyond when it is apparent to them that their memory is not as it was. This lack of confidence can also negatively impact performance, causing witnesses to withhold information due to uncertainty or not wishing to make mistakes (Iida et al., 2020; Saraiva et al., 2020).

Many factors contribute to the quality and quality of episodic memory for witnesses of all ages. However, normal deterioration in visual and auditory abilities (Garobbio et al., 2023; Stine & Wingfield, 1987) can reduce the amount of primary information to be encoded in the first instance. Further, perceptual load increases in situations with overwhelming amounts of information, reducing memory accuracy for peripheral details that are not attended to as they might be in situations with lower perceptual loads (Murphy & Greene, 2016). This is particularly true as cognitive processes develop and naturally age (Balota et al., 2000; Craik & McDowd, 1987; Dando, 2013). However, this does not always result in poorer memory performance since appropriate external retrieval support can support equitable performance in some instances (e.g. Dando, 2013; Dando et al., 2020; Thomas et al., 2020).

Encoding transforms perception and attention into a memory trace, including our past experiences, understandings, and feelings about a situation (Kaplan et al., 2016). Memory traces consolidate over time (McGaugh, 2000), with details lost as time passes (Goldsmith et al., 2005). Thus, techniques and tools to help reduce the time between encoding and retrieval could potentially mitigate some differences in

age-related delayed recall performance that are often reported (Aizpurua et al., 2009; Czaja et al., 2001; Roediger & Geraci, 2007). A typed modality Interview could be used in this way, but the implications of this modality across various age groups must be explored to maximise understanding of the utility and efficacy of this approach for age groups other than those in their 20s when cognition and cognitive plasticity in terms of adaption and acceptance of new approaches are at their height (e.g. Anatürk et al., 2021; Staudinger, 2020).

4.2.2. Adult memory credibility and confidence.

Confidence in memory decreases during the natural ageing process. Confidence involves monitoring metamemory and is not an indication of memory accuracy but rather offers perceptions of memory accuracy, which should not be relied upon (Berkowitz et al., 2022; Iida et al., 2020; Metcalfe & Shimamura, 1994). Generally, the public views memory as an accurate representation of events (Simons & Chabris, 2011). However, this view is not robust as individuals naturally age. For example, older adults often negatively perceive their memory (Wylie et al., 2014). Low confidence has implications within the criminal justice system, as mock jurors find high confidence to indicate reliability (Ross et al., 1990; Sauer et al., 2017). Unsurprisingly, police perceive older witnesses as less reliable than younger witnesses (A. M. Wright & Holliday, 2005), a view shared by the general public despite the perception being inaccurate in many instances (Kwong See et al., 2001). Low confidence levels make memory more susceptible to positive and negative feedback, with feedback impacting confidence regardless of accuracy (Iida et al., 2020) and increasing susceptibility to post-event misinformation (Greenspan & Loftus, 2020; Leippe et al., 2006; M. E. Smith et al., 2021).

The misinformation effect is when inaccurate information is reported instead of original, accurate information (Ayers & Reder, 1998; D. B. Wright & Loftus, 1998). This results from source attribution errors when new knowledge is mistakenly recalled as a first-hand experience (Loftus, 2005; Zaragoza & Lane, 1994). Source attribution errors come from errors in source monitoring, errors which increase with age (Dando, 2013).

It has been demonstrated that older adults are particularly susceptible to the effects of negative feedback on recall, deferring more quickly to misinformation than younger adults (Henkel, 2014). Uncertainty leaves memories vulnerable to change, with witnesses more likely to change their understanding of an event to fit with another witness who seems more confident (Gabbert et al., 2003; Sousa & Jaeger, 2022; D. B. Wright et al., 2000). Co-witness age also affects perceived credibility, with younger witnesses being seen as more competent and, therefore, having a greater influence on co-witness contamination (Kwong See et al., 2001). However, informing witnesses about the importance of source monitoring can reduce this effect (Bodner et al., 2009). Thus, from a cognitive perspective, investigating age-related memory performance through typed recalling is critical to advancing understanding of the potential suitability of a remote digital retrieval on two counts. First, this approach offers promise for quickly consolidating memory, possibly supporting adult cognition across cognitive developmental stages when moving towards middle age and beyond. Second, leading on from this, quick initial retrieval may offer some protection from post-event misinformation and possibly boost confidence.

4.2.3. Age, technology, and modality.

It has been shown that memory plays a role in using and completing technology-based activities (Czaja et al., 2001; Sharit et al., 2003). For example, when using a

computer, older adults (e.g., 50 years +) make more errors and take more time than young adults and adolescents (Sayers, 2004). This is compounded by older adults reporting that not knowing how to use technology prevents them from quickly adopting it (M. T. Harris et al., 2022). However, it has been shown that older adults are more likely to use and understand technology if given opportunities to have positive experiences (Mitzner et al., 2019). Older adults benefit from textual support when learning to use technology (Pachman & Ke, 2012) and visual navigation tools (Sayers, 2004). This is an important consideration when developing new technologies to maximise useability.

The preceding Chapter (Chapter 3) explored and reported the differences in memory retrieval performance between spoken and typed recall. Overall, participants engaged in spoken interviews demonstrated a reduction in the reporting of confabulated items compared to those who underwent typed interviews. However, overall accuracy was not compromised because the number of confabulations was so small (< 1). This trend persisted when analysing the distinct phases of the interview process, encompassing both the free recall and cued recall phases. Participants in spoken interviews consistently recall fewer confabulated items than their counterparts in typed interviews during both phases. Nevertheless, the two modalities' percentage accuracy did not differ during either phase.

4.2.4. The current study.

The overarching objective of the study reported in this chapter is to investigate the impact of using typed modalities for collecting witness information across three different age groups. The adult age groupings were selected to map onto known intraindividual cognitive differences across the lifespan. Still, they were dictated by access to age group participants from the general population during the COVID-19

international health emergency. Accordingly, groups of 18 to 29, 30 to 54, and 55 and over emerged. For the study reported here, the 18 to 29 group is referred to as the younger group, 30 to 54 years as the middle group, and 55 years and older is referred to as the older group.

With the widespread adoption of typed communication in everyday social interactions, it is sensible to investigate suitability and efficacy in an investigative context across various age groups since the current literature predominantly focuses on remote interviewing methods, such as video and audio communications with younger adults who are often recruited from student groups. This study contributes to the literature by examining the specific impact of typed witness interviews on memory recall across three distinct age groups from the general population.

Two broad research questions were developed to guide the study reported here centred on age-related differences when providing typed information about an experienced event:

1. Does age impact the quantity and quality of eyewitness memory elicited via typed modality?
2. Is there a relationship between age, familiarity with technology, comfort, and the quality and quantity of event information recalled via typed modality?

4.3. Methods

4.3.1. Design.

A between-subjects design (described below) was employed, using a mock witness paradigm to investigate typed recall across three age groups. Age categories relevant to the cognitive-developmental understanding of episodic memory across the lifespan were used to group participants. The independent

variable of interest was age, with three levels (18 to 29, 30 to 54, and 55 and over). The dependent variable was the quality and quantity of memory performance, measured by the number of correct, incorrect, and confabulated items reported and the percentage accuracy.

4.3.2. Participants.

Participants were recruited through online advertising and compensated for the time spent taking part through a £10 shopping voucher. A total of 82 participants participated in this study, 27 male and 55 female. There were three age groups: 18 to 29, 30 to 54, and 55 and over. There were 29 participants in the 18 to 29 years group with a mean age of 23.21 (SD = 3.02), with 7 males and 22 females. There were 31 participants in the 30 to 54 age group. The mean age was 36.71 (SD = 5.89), with 12 males and 19 females. There were 22 participants in the 55 and over group with a mean age of 62.73 (SD = 6.88), with 8 males and 14 females.

4.3.3. Procedure.

All potential participants were first emailed a participant information sheet (see Appendix B) and a copy of the consent sheet (see Appendix C). All participants were then video-called at a pre-agreed time, introduced to the interviewer, and given a unique participant number to label their unique contribution to the study. Participants were allowed to ask questions about the study during this initial call before participating. Once the video call had ended, participants were provided with a unique Qualtrics link, which provided access to an online consent process, following which participants watched the video stimulus (individually). The Interviewer then video-called the participant for a second time and administered the Mini-Mental States Examination (MMSE – see Appendix E). The MMSE was used throughout this PhD research programme to control for significant cognitive decline in episodic

memory falling outside of the ‘norm’ and was used as a filler task. While the PhD research reported in this thesis concerns the general population, because some participants were over 50, it was deemed appropriate to use the MMSE. All participants passed the MMSE and, thus, took part in the research. There was no variation in participant scores, and it was scored as pass or fail.

Having completed the MMSE procedure, all participants were instructed to end the video call and begin communication with the interviewer through the synchronous text function on MS Teams. All participants participated in a basic information-gathering interview that maps onto a tier 1 PEACE investigative interview (See Appendix I). After the interview, all participants were video-called for a third time. They were allowed to ask questions and respond to two questions regarding how comfortable they were using a keyboard and their familiarity with technology. Participant numbers were used to anonymise and label all data collected.

4.3.4. Materials.

4.3.4.1. Stimulus.

The mock volume crime type stimulus remained consistent across all mock witness studies reported in this PhD research program – see section 3.3.4.1. in Chapter 3 for a description.

4.3.4.2. Mini-Mental States Examination.

The Mini-Mental States Examination (MMSE) (Crum et al., 1993) is a short questionnaire used extensively in research of this nature to control for cognitive impairments outside of the ‘normal’ range, which might confound the research results. The MMSE is not a diagnostic tool and was used as a filler task. Because participant ages ranged between 18 and 82, it was deemed appropriate to control for age-related severe cognitive impairment. No further controls nor assessments were

used throughout this thesis because it was considered important to maximise applicability and application of the findings to the general population, where possible. No participants were found to be outside the normal range for the MMSE (Folstein et al., 1975; Foreman et al., 1996).

4.3.4.3. Ethics Materials.

All participants in this research were emailed a participant information sheet before agreeing to participate in any of the experiments (See Appendix B). All participant information sheets contain information unique to the study and the condition to which the participant was allocated. Further information about all conditions was given in the debrief sheet. Additionally, all participant Information sheets contained background information about the research, requirements for taking part, what participants will be required to do, ethics information, and contact details. Additionally, all participants completed an online consent form through a Qualtrics link before beginning the study (See Appendix C).

Participant demographic questions were asked across all studies in this work (See Appendix D). The standardised questions asked participants their age and sex. As this study took place entirely remotely, participants were also asked about their birth country and English language proficiency, which was an exclusionary factor. Additionally, participants were asked to confirm that their eyesight and hearing were not impaired, as this could impact the ability to perceive elements from the video, which could be misinterpreted as poor memory performance.

All participants received a debrief sheet (See Appendix G). The debrief sheet contained further information about the study's aims and purpose. They were given a summary of what they had done in the study and information about the other conditions. A detailed explanation of how the data will be anonymised and

instructions for removing data if participants wish to withdraw, were given. Information about compensation for taking part and the researcher's contact details were provided.

4.3.4.4. Tier 1 PEACE Investigative Interview.

All studies in this body of work measured memory performance from a transcript of a tier 1 PEACE investigative interview (See Appendix H). The Tier 1 PEACE investigative interview procedure outline is taken from The College of Policing (College of Policing, 2022) – see Chapter 3 Section 3.4.4. and Appendix I for a full description, but in brief, the procedure comprised the following phases:

- Planning and preparation:
- Engage and Explain, including the ground rules i) ‘never guess’, ii) ‘report everything’, iii) ‘say if you do not remember’, and iv) ‘tell me if you do not understand the question’.
- Account, Clarification, plus a reminder of the four ground rules above; question;
- Closure (including the opportunity to add and/or alter anything.

4.3.5. Coding.

See section 3.3.5.

4.4. Results

4.4.1. Analysis approach.

Inferential multivariate statistical techniques compared episodic memory performance across three age groups. Global memory (from the start to the end of the memory retrieval process) was analysed for correct, incorrect, and confabulated recall, and a new variable that combines each of these relevant performance

measures in the first instance. Significant MANOVA results were further investigated concerning univariate findings and post hoc tests as appropriate.

Since an interview comprises a series of phases, each contributing individually and incrementally to global memory performance, a series of ANOVAs were conducted for each of the two recall phases (free recall and cued recall) for correct, incorrect, and confabulated recall. Finally, global and phase percentage accuracy was analysed using a series of ANOVAs to investigate the impact of memory performance on recall accuracy as a function of age. Bonferroni's correction for multiple comparisons is applied throughout, as appropriate.

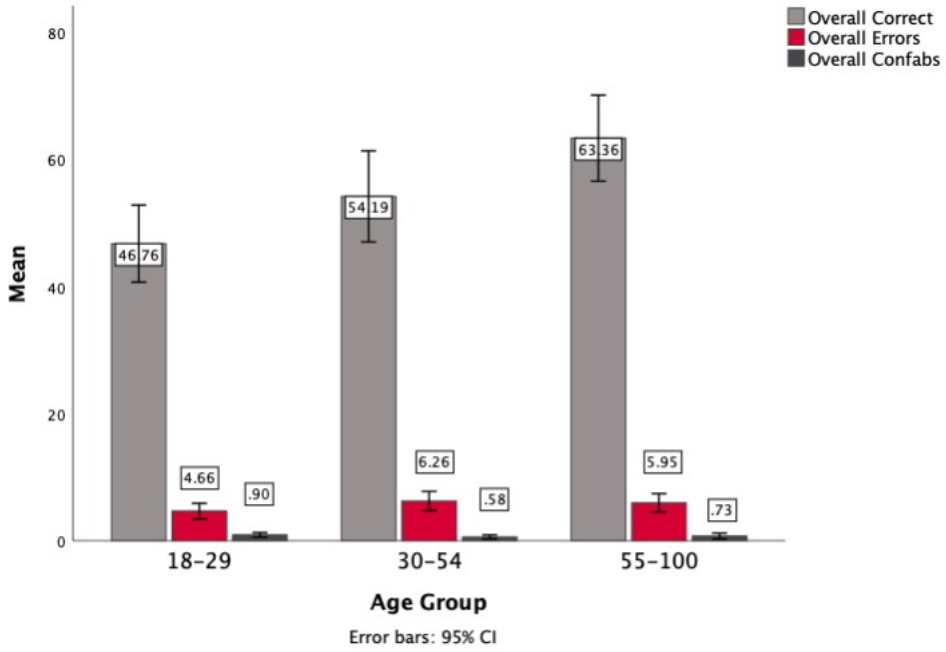
4.4.2. Global memory performance analysis.

Means, standard deviations, and 95% CIs for global memorial performance are displayed in Figure 1 below. The MANOVA was significant, $F(6,154) = 2.62$, $p = 0.019$; Wilks' $\lambda = 0.823$, as a function of age group indicating age-relevant differences in memory performance.

The univariate analysis revealed a significant effect of age group on the number of correct items recalled $F(2,79) = 5.802$, $p = .004$, $\eta p2 = .064$. There was a non-significant effect of age group for the overall number of errors, $F(2,79) = 1.643$, $p = .200$, $\eta p2 = .012$, and confabulations $F(2,79) = 0.861$, $p = .427$, $\eta p2 = .011$. Hence, the locus of the multivariate differences across age groups was the number of correct items recalled. Pairwise comparisons for global correct recall revealed that participants in the over-55 age group recalled more correct items than participants in the 18-29 age group, $p = .001$.

Figure 1

Means & 95% CI for global correct, errors & confabulations as a function of age group (N = 82).



4.4.3. Free recall memory performance.

Means, standard deviations and 95% CIs for free recall memorial performance are displayed in Table 4 below. There was a significant effect of age group on the number of correct Items recalled during the initial free recall phase of the interview $F(2,79) = 5.032$, $p = .009$, $\eta p2 = .11$. Participants in the over 55 age group recalled more correct items than participants in the 18-29 age group, $p = .002$. All other pairwise comparisons were non-significant, all *p*s = .155. There was a non-significant effect of the age group on the number of errors and number of confabulations during the free recall phase of the interview $F(2,79) = 3.389$, $p = .039$, $\eta p2 = .08$, and $F(2,79) = 0.710$, $p = .495$, $\eta p2 = .02$, respectively.

Table 4

Means, standard deviations & 95% confidence intervals for free recall memory performance as a function of age group.

	Age	Mean	SD	95% Confidence Interval	
				Lower	Upper
Correct	18-29	27.10	9.76	23.39	30.82
	30-54	31.13	11.99	26.73	35.53
	Over 55	36.36	8.28	32.73	40.03
Errors	18-29	1.31	1.26	0.83	1.79
	30-54	2.35	1.96	1.64	3.07
	Over 55	2.14	1.50	1.56	2.29
Confabulations	18-29	0.24	0.43	0.08	0.41
	30-54	0.19	0.48	0.02	0.37
	Over 55	0.36	0.66	0.07	0.66

4.4.4. Cued recall memory performance.

Means standard deviations and 95% CIs for cued recall memorial performance are displayed in Table 5 below. There was a significant effect of age group on the number of correct Items recalled during the cued recall phase of the interview $F(2,79) = 4.342$, $p = .016$, $\eta p2 = .10$. Participants in the over 55 age group recalled more correct items in the cued recall than participants in the 18-29 age group, $p = .011$. All other pairwise comparisons were non-significant, all ps = .253. There was a non-significant effect of the age group on the number of errors and confabulations during the cued recall phase of the interview $F(2,79) = 0.354$, $p = .703$, $\eta p2 = .01$, and, $F(2,79) = 1.126$, $p = .329$, $\eta p2 = .03$, respectively.

Table 5

Means, standard deviations & 95% confidence intervals for cued recall memory performance as a function of age group.

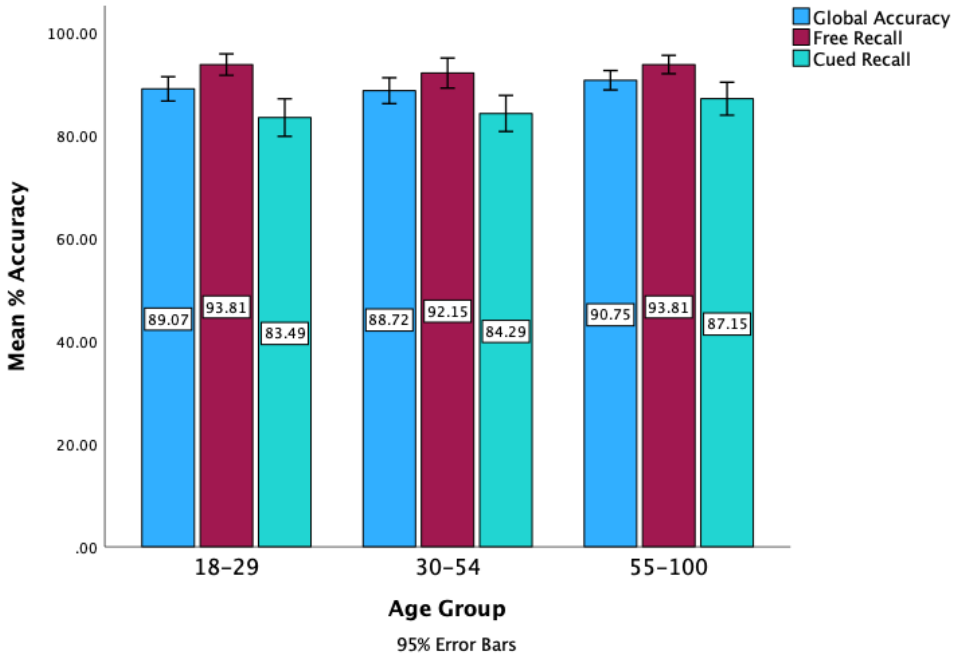
	Age	Mean	SD	95% Confidence Interval	
				Lower	Upper
Correct	18-29	19.69	8.54	16.44	22.94
	30-54	23.06	9.17	19.70	26.43
	Over 55	27.00	8.51	23.23	30.77
Errors	18-29	3.34	2.68	2.33	4.36
	30-54	3.90	2.77	2.89	4.92
	Over 55	3.68	2.69	2.63	5.01
Confabulations	18-29	0.66	0.81	0.35	0.96
	30-54	0.39	0.67	0.14	0.63
	Over 55	0.36	0.95	-0.06	0.79

4.4.5. Percentage accuracy.

There was a non-significant effect of age group on the global percentage accuracy $F(2,79) = 0.794$, $p = .456$, $\eta p2 = .02$, free recall accuracy, $F(2,79) = 0.674$, $p = .513$, $\eta p2 = .02$, and cued recall accuracy, $F(2,79) = 1.109$, $p = .335$, $\eta p2 = .03$ (see Figure 2 below for Means and 95% CIs).

Figure 2

Means standard deviations & 95% CI error bars for free recall, cued recall, and global percentage accuracy (N = 82).



4.4.6. Completeness

The template approach for coding interviews comprised 100 items of event information (See Appendix X). Completeness was the amount of correct recall as a percentage of 100. Analysis of Variance revealed a significant difference across age groups for completeness $F(2, 79) = 5.961$, $p = .004$, η_p^2 .131. Participants in the 18-29 age group were significantly less complete ($M = 46.76$, SD = 3.16, 95% CI 40.47, 53.05) than those in the 55-100 age group, $p < .001$ ($M = 63.36$, SD = 3.63, 95% CI 56.15, 70.58). There was no significant difference between the 30-54 age group and the 18-29 age group, $p = .105$, or the 55-100 age group, $p = .051$.

4.4.7. Additional exploratory analyses.

To explore the relationships between age and memory performance and the directional nature of any emerging relationships, Pearson's correlations (all two-

tailed) were conducted for age and i) correct, errors, confabulated recall, and percentage accuracy globally and as a function of each of the recall phases (free recall and cued recall) and ii) the two questions regarding familiarity/comfort with technology.

4.4.8. Global memory performance.

There was a significant moderate positive correlation between age and overall correct items recalled, r = .345, N = 82, p = .001, explaining 12% of the variation of correct recall. Age and incorrect items were non-significant, r = .106, N = 82, p = .343, likewise confabulations, r = .011 N = 82, p = .925, and overall accuracy, r = .120, N = 82, p = .282. See Table 6.

Table 6

Means, standard deviation, & correlations of overall memory items

Variable	Mean	SD	Age Correlation
Overall Correct	54.02	18.238	.345**
Overall Incorrect	5.61	3.606	.106
Overall Confabulations	0.73	0.930	.011
Overall Accuracy	86.39	6.013	.120

***. Correlation is significant at the 0.01 level (two-tailed)*

4.4.9. Phase memory performance.

Analyses of free recall memory performance showed a significant positive correlation between age and free recall correct items, r = .340, N = 82, p = .002. Again, this was a moderate positive correlation, explaining just 12% of the variation. Thus, the older the participant the more correct information items were recalled in the free recall stage of the interview. Age and the number of free recall errors, r =

.142, N = 82, *p* = .203, and confabulations, *r* = .108 N = 82, *p* = .333, and free recall accuracy, *r* = .031, N = 82, *p* = .781, correlations were all non-significant.

Analyses of cued recall memory performance revealed a significant (but weak) positive correlation between age and cued recall correct items, *r* = .284, N = 82, *p* = .01, two-tailed, explaining just 8% of the variation. Thus, the older the participant the more correct information items were recalled in the cued recall stage of the interview. The age-related correlations for number of cued errors, *r* = .054, N = 82, *p* = .628, two-tailed, cued recall confabulations, *r* = -.057, N = 82, *p* = .609, two-tailed, and cued recall accuracy, *r* = .151, N = 82, *p* = .176, were all non-significant (see Table 7 below)

Table 7

Means, standard deviation, & age group correlations for memory performance as a function of interview phase.

Variable	Mean	SD	Age Correlation (R)
Free Recall Correct	31.11	10.826	.340**
Free Recall Incorrect	1.93	1.661	.142
Free Recall Confabulations	.26	.517	.108
Free Recall Accuracy %	93.182	6.226	.031
Cued Recall Correct	22.93	9.135	.284**
Cued Recall Incorrect	3.68	2.694	.054
Cued Recall Confabulations	.48	.805	-.057
Cued Recall Accuracy %	84.774	9.018	.151

***. Correlation is significant at the 0.01 level (two-tailed)*

4.4.10. Familiarity with technology analysis.

Two questions were asked about participants' familiarity with technology. ‘How comfortable are you with typing with a keyboard on a computer?’ And ‘How comfortable are you with reading from an electronic screen (computer/phone/kindle)’.

Both questions used a 7-point Likert scale, with 1 being extremely comfortable and 7 being extremely uncomfortable.

Familiarity with technology was significantly positively correlated with age and how comfortable participants were typing with a keyboard on a computer, $r = .249$, N = 82, $p = .024$, two-tailed, whereby the older age group appeared more comfortable. However, this was a weak correlation whereby just 6% of the variation is explained. There was a significant positive correlation between the age group and how comfortable participants were reading from an electronic screen, $r = .294$, N = 82, $p = .007$. Again, this weak correlation explained just 9% of the population variation.

4.5. Discussion

This chapter reports the results of a mock-witness study investigating episodic memory performance when providing typed witness information using a keyboard as a function of three age groups: 18 to 29, 30 to 54, and 55 and over. Participants were interviewed via text chat after watching a non-violent video of a mock crime. Several notable and unexpected findings emerged. The global memorial performance revealed that participants in the over-55 age group recalled over 30% more correct items than participants in the youngest of the three age groups (18-29) and performed similarly to the mid-age group. This result appears counterinitiative since the extant literature suggests the opposite result might be more likely. However, this research is novel, so it is challenging to tease apart why this pattern of findings has emerged following one study. However, it seems sensible to suggest that the modality of recall may have offered some benefits for older adults by reducing some of the social demands inherent in witness interviews. Research indicates that older adults are particularly vulnerable to external influence, including

the interviewer's body language and feedback, which can unintentionally cue certain responses (Iida et al., 2020; Leippe et al., 2006). This study's use of typed recall may have mitigated such susceptibility by removing these nonverbal cues, allowing participants to rely more confidently on their own memory (Gurney, 2011). The lack of direct social feedback and time constraints may have afforded older adults the freedom to access and communicate accurate memories without interference from an interviewer, supporting a more self-reliant recall process (Dando, 2013). This may have eased the cognitive load associated with managing or monitoring the social environment, freeing cognitive resources and allowing for more effortful episodic recall. This suggestion seems sensible because there were no age group differences in errors or confabulations.

The findings on completeness add another dimension to our understanding of typed memory recall across age groups. The significant differences in completeness, with older adults (55-100) recalling a greater percentage of event information than younger adults (18-29), challenge assumptions about age and memory performance. Specifically, participants in the older age group achieved an average completeness score of 63.36%, while younger participants achieved only 46.76%. This indicates that older adults may recall more details and a more comprehensive account of the events observed.

These findings are surprising given the common expectation of age-related decline in episodic memory completeness (Craik & McDowd, 1987). One possible explanation for this higher completeness in older adults is the benefit they derive from self-paced recall in the typed modality, allowing them to structure their responses more methodically and potentially avoid memory lapses associated with high-pressure recall settings (Pachman & Ke, 2012). Unlike younger participants,

who may rely more heavily on rapid, spontaneous recall, older adults may engage in deeper, more reflective recall strategies, leading to more complete retrieval of the details available.

This pattern of completeness without a corresponding increase in errors or confabulations (as noted in other sections of this study) reinforces the notion that older adults can provide detailed and accurate accounts when given appropriate recall support. This suggests that memory completeness in older witnesses could be more influenced by the retrieval environment and format than by age alone. These findings support the development of adaptive interviewing strategies, particularly in typed recall settings, that facilitate completeness for older witnesses without risking accuracy.

Participants provided their information textually, and in doing so, the interviewer was not present, physically nor visually, albeit that participant had met the interviewer before providing their memory for the event. As such, they knew this interviewer would be typing the questions. Nonetheless, the physical and visual absence of the interviewer may have reduced real or perceived interviewer demand emanating from feedback, which is easily and often unknowingly communicated through tone of voice, body language, (Gurney, 2011; Gurney et al., 2016) and facial expressions (Gurney, 2015), for example. Contemporary research has reported that remote interviewing in the absence of a human interviewer can improve recall performance in terms of increased correct recall without concomitant increases in errors (Dando et al., 2023; Dickinson et al., 2021; Legg & Song, 2021; Taylor & Dando, 2018), and it appears this may account in part for the significant increase in correct recall reported here. Older adults are known to be susceptible to negative interviewer feedback (e.g. Dando, 2013; Iida et al., 2020; Leippe et al., 2006;

Roediger & Geraci, 2007) , so future research should investigate this aspect of age-related performance and the interaction between external environmental demands emanating from social contexts and cognition.

A similar pattern of results was found when considering each of the two recall phases. Again, participants in the over-55 age group recalled more correct information than participants in the 18-29 age group. There was no difference between confabulated or incorrect information during the free and cued recall phases. Previous research has found that older participants remember fewer details when freely recalling information and are often less accurate (Dodson & Krueger, 2006; List, 1986; Mello & Fisher, 1996; Toglia et al., 2017). However, not always, particularly where appropriate external support is offered (Dando, 2013; Memon et al., 2003, 2013). Indeed, this study provides further insight into the importance of retrieval practices and processes and again suggests that, in some cases, managing the retrieval environment to reduce external stimuli, including actual or perceived social pressures (Henkel, 2014; Leippe et al., 2006) that may interfere with complex cognition, may be beneficial.

Not surprisingly, the cued recall phase of a cognitive interview results in challenges centred on seeking additional information without introducing demand characteristics or confirmation bias (Shepherd & Milne, 1999). Previous research has found that older adults can be half as effective in the cued recall stage of an interview (Dodson et al., 2015), so typed modality recall may remove demand characteristics disproportionately affecting older adults (Henkel, 2014; Pachman & Ke, 2012). An increase in correct information by older adults may be explained by the supportive nature of receiving instruction via text, as others argued (Pachman & Ke, 2012).

Correlation analysis found that as participants' age increased, their familiarity with technology decreased. This aligns with previous research demonstrating that older adults use technology less and feel less confident (M. T. Harris et al., 2022). However, as the results from this study indicate, this lack of confidence appears not to have impacted memorial performance when typing using a computer keyboard. Although the age range of the older adult group for this study was limited, confidence may not necessarily impede the ability to use technology in these circumstances. Low confidence does not necessarily indicate poor memory performance, as others have reported in various contexts (Greener & Wakefield, 2015; Roediger & DeSoto, 2014; Winter et al., 2021).

Unlike traditional verbal or face-to-face recall, typed interviews offer older adults a degree of cognitive control and pacing that is not always possible in spoken formats (Pachman & Ke, 2012). Typing allows older adults to process and structure their memories at a comfortable pace, potentially enhancing recall accuracy by reducing the cognitive load associated with real-time spoken interaction (Sauerland & Sporer, 2011). This contrasts with findings in traditional eyewitness studies where older adults typically recall fewer details and are more susceptible to confabulations in free recall (List, 1986). In a typed format, older adults might benefit from the ability to carefully retrieve and articulate memories without the pressure or demand characteristics that can arise in face-to-face settings (Henkel, 2014).

Contrary to the common belief that cognitive decline uniformly affects older adults, research suggests that certain cognitive strengths, like life experience and semantic memory, can support episodic recall (Craik & McDowd, 1987). Older adults might employ more robust retrieval strategies developed over time, such as categorization or chronological sequencing, which facilitate accurate recall. In the typed interview,

these strategies could be especially advantageous, as older adults can construct a narrative without the pressure to respond immediately, unlike in a live verbal interview (Evans & Fisher, 2011). This adaptability in recalling personal narratives might explain the higher accuracy observed in older adults when using typed recall, aligning with literature suggesting that age-related memory challenges are context-dependent and can be alleviated through adaptive recall methods (Tulving & Thomson, 1973).

4.5.1. Limitations.

The study's sample may not fully represent the general population's diversity despite remote participant recruitment. Participants had to be able to access a computer or tablet, access the internet, and feel confident enough to sign up for the research. This would undoubtedly have closed down opportunities for some groups. Further, the requirement of English language proficiency and access to technology per se could have favoured individuals with higher socioeconomic status or educational attainment. Consequently, the findings may not apply to individuals from non-English-speaking backgrounds or those with limited access to technology. Conversely, remote recruitment methods can broaden the participant pool beyond geographical constraints and widen access.

Access to technology was necessary for participation in the study. Technology usage has become increasingly pervasive, particularly over the COVID-19 pandemic period when many adults of all ages, including adults in the older age ranges, were ‘forced’ to use technology more often and were encouraged to understand better a broader range of communication methods falling outside of their pre-COVID experiences to mitigate social isolation, for example, (Haase et al., 2021). Thus, while there may be some bias towards individuals with greater access to technology,

this limitation may not be as pronounced as before the international public health emergency.

While participants demonstrated proficient use of a computer keyboard during the text chat interview, the study revealed a negative correlation between age and reported familiarity with technology. This technology familiarity disparity could influence feelings of comfort and thus impact engagement during the memory task, although no negative impact on memory performance was observed here. However, this discrepancy underscores the need to consider and better understand technological literacy as a potential confounding factor in future studies. The negative correlation between age and technology familiarity may not fully capture the nuances of participants' technological skills and experiences. It is worth noting that younger and middle-aged participants in this study may not necessarily exhibit higher levels of technological literacy if they have limited exposure to specific technologies. Typing skills are often more closely associated with everyday computer use, regardless of age, and participants' proficiency in this particular task may have contributed to consistent performance levels across age groups.

Categorising participants into broad age groups (18-29, 30-54, 55 and over) may oversimplify the complex relationship between age and memory performance. For instance, in the older adult category, significant variability in cognitive functioning and memory abilities may exist due to health status and cognitive reserve. Subgroup analyses based on additional demographic and cognitive variables could provide more nuanced insights into age-related differences in memory recall. However, the approach employed for this research study was dictated by access and balancing the practical considerations for applying findings to the general population who are not screened when coming into contact with the criminal justice system.

Subdividing the older adult category into smaller age brackets may capture more nuanced variations in cognitive functioning and memory abilities, and future research should consider this. Factors such as educational background, health status, and cognitive reserve can significantly influence memory performance within this demographic (Beyer et al., 2021). Age-related memory impairments were filtered for by integrating the MMSE into the study (Crum et al., 1993). This was done to control for cognitive impairments outside of the ‘normal’ range, which might confound the research results and as a filler (Dando, 2013; Dando et al., 2020).

Conducting the study remotely due to the COVID-19 pandemic may have introduced additional variability in participants' environments, potentially impacting their focus and engagement during the memory task. Factors such as noise levels, interruptions, and distractions in participants' home environments could have affected their performance and responses. Although participants were instructed to complete the memory task in a quiet and private location to minimise distractions, standardised instructions and procedures were also used to ensure consistency across participants and reduce the influence of environmental factors on study outcomes. Without a controlled laboratory setting, monitoring participants’ behaviour during the memory task was impossible. Future research might wish to consider this, albeit in a quieter, more controlled environment, performance would likely improve since controlled environments are known to improve complex cognition.

Finally, a low-impact volume crime video was used for this research. To minimize extraneous variables, a short video was used, although this may limit generalizability. High accuracy rates, consistent with similar research (Dando et al., 2009a: 2009b; Deffenbacher et al., 2004; Evans & Fisher, 2011; Köhnken et al., 1999), were likely due to structured interview protocols emphasizing precise recall.

These rates suggest effective consolidation and retrieval manipulations, but future studies could explore varied, more complex stimuli to deepen understanding of generalizability and accuracy outcomes.

Where events are traumatic and or more serious, and for vulnerable witnesses, typed interviews may be inappropriate for monitoring for a trauma response or monitoring for understanding by the interviewee, for example. Consideration should be given in future research to the type of crime and the witness status, alongside looking to mirror unintentional encoding environments as often happens in the real world.

4.5.2. Conclusion.

The study suggests that text interviews are a promising avenue for expediting the initial account witness testimony process in some instances. Text-based interviews offer a practical solution that saves time and resources by eliminating face-to-face interactions. Participants can provide detailed accounts of events through typed responses, allowing for efficient data collection without the logistical constraints associated with traditional interview methods. This shift towards text-based modalities not only streamlines the interview process but also minimises potential biases introduced by interviewer influence or non-verbal cues, thus enhancing the objectivity of the information obtained. The implications of these findings extend beyond academic research, with potential applications for witnesses of all ages being called upon to provide accounts of their experiences.

5. Chapter Five: ChatCharlie for Gathering Initial Accounts from Eyewitnesses: Development and Evaluation.

5.1. Introduction

This chapter reports an empirical evaluation of a novel approach to collecting initial accounts from witnesses immediately following a crime event. A hybrid 'chatbot' was developed, hereon referred to as 'ChatCharlie'. This empirical study examines the efficacy of ChatCharlie for consolidating memory for an experienced event. Consolidation is known to stabilise a memory trace, thus potentially improving memory performance when asked to recall the event in more detail at a later time). Using the same mock witness paradigm reported in previous chapters, 90 participants took part in a two-stage research process. At Time 1, participants were randomly assigned to ChatCharlie, In-person initial account or the no initial account control condition. Within 15 minutes of viewing the stimulus event, participants provided an initial account according to the condition (or did not provide an initial account in the Control condition). At Time 2, seven days later, all participants were interviewed in person face-to-face using a basic Tier 1 type investigative interview. Analyses included memory performance at Time 1 (as appropriate) and Time 2. Global memory, interview phase performance, and percentage accuracy were coded. At Time 1, there were significantly fewer confabulations in the ChatCharlie condition and non-significant differences in correct and incorrect memory performance between participants in the ChatCharlie and In-person conditions. At Time 2, participants in the ChatCharlie and In-person Time 1 conditions were more accurate. They recalled significantly more correct globally and during both interview phases than the Control group without a concomitant increase in errors and

confabulations. No significant differences emerged at Time 2 between ChatCharlie and the In-Person Time 1 conditions.

5.1.1. The current study.

In light of a shortage of research in the domain of initial account interviewing, as advocated by the College of Policing, this chapter reports a study describing the development and an initial empirical investigation of ChatCharlie, a hybrid chatbot, for gathering an initial account immediately after witnessing a mock crime. Integrating contemporary technology with established principles of investigative practice that are largely scripted and prescriptive, the research seeks to assess whether ChatCharlie might be effective for consolidating witness memory similar to the initial account in-person interview, thus comparatively improving recall 7 days later versus a control, without increasing incorrect or confabulated information.

5.1.2. Development and application of ChatCharlie.

ChatCharlie merges contemporary technology with professional investigative practices. Some potential advantages to remote collection of initial accounts include reducing the time and the cost of interviewing witnesses in person, widening access to larger numbers of eyewitnesses and enabling human resources to be strategically directed to where they are most needed (Kuivaniemi-Smith et al., 2014). In the case of a mass event, for example, it would be impossible to collect initial accounts from all witnesses promptly. Delays in gathering information affect the quality and quantity of information collected. Thus, ChatCharlie offers a potential tool to assist police investigations in some circumstances. ChatCharlie would allow witnesses to quickly recall and record their initial account experiences remotely using a smartphone, tablet or laptop computer, thus collecting and preserving initial eyewitness information that could be immediately accessed and time and date ‘stamped’ in a

similar manner to digital photographs, for example, in a way that would otherwise be impossible.

A chatbot is a computer program designed to simulate conversation with human users. Although conversations can be similar to those with a person to some degree, chatbots have a limited set of responses to questions and a limited number of questions that, in turn, can be asked. Therefore, chatbot questions and replies are guided by the platform or program that hosts and directs the chatbot interaction. One study investigated using an artificial intelligence chatbot to conduct cognitive interviews and collect information after participants witnessed a mock crime video. The study compared the artificial intelligence (AI) chatbot to a free recall cognitive interview style questionnaire and a basic cognitive interview via a non-AI chatbot. The AI chatbot resulted in the most accurate and descriptive recall (Minhas et al., 2022). However, the process did not replicate the practical initial stages of a police investigation in terms of mapping the process of gathering an initial account interview. Instead, the researchers were more concerned with developing a potential alternative to an in-depth, more detailed, full investigative interview.

An initial account collected via ChatCharlie may, however, trigger concerns over sharing sensitive information with a chatbot, which may be a barrier to disclosure. Although research has examined the impact of sensitive topics on disclosure in the context of healthcare (Miles et al., 2021; Nadarzynski et al., 2020), suggesting that in some cases, sharing with a non-human is ‘easier’ since such interactions are non-judgemental, potentially increasing disclosure, which is one of the guidance topics countenanced by the College of Policing when collecting an initial account (College of Policing, 2019). Indeed, computer-mediated communication is beneficial in some instances, including for detecting deception (Dando et al., 2023) and reducing errors

in interviews with eyewitnesses because interviewees found it easier to admit a lack of memory than in a face-to-face interview (Dando et al., 2023; Taylor & Dando, 2018) and revealing sensitive information in some circumstances (Lind et al., 2013; Weisband & Kiesler, 1996), attributed in part to reduced social inhibitions relating to fear of judgment (Joinson et al., 2007).

It has recently been reported that chatbots are perceived as being non-judgemental in a health context when revealing personal health information (Nadarzynski et al., 2021). Reducing fear of judgment increases disclosure by reducing the cognitive load of managing how one is being perceived (Joinson, 2001b; Kang & Gratch, 2010). The absence of a physical presence may also reduce physiological arousal, potentially increasing the quantity and quality of episodic information recalled (Belletier et al., 2019; Dando et al., 2023; Wagstaff, 2008). Thus, using chatbots may benefit some witnesses by increasing their willingness to share information and improving the quality of the information they recall in the first instance, alongside the cognitive benefits of consolidation.

A novel chatbot prototype, ‘ChatCharlie,’ was designed for this study to collect real-time initial account ‘best evidence’ from witnesses. The interview protocol presented to witnesses via ChatCharlie is modelled on the current initial account interview procedure (College of Policing, 2019). The guidelines for conducting initial account interviews have been interpreted and modified for this thesis according to the evidence base to create an appropriate comparative but functional tool. The ten key guidelines for conducting an in-person initial account interview are outlined in Chapter 1.4. To briefly recap, the guidelines included by the College of Policing but excluded in this first empirical evaluation of ChatCharlie are consideration of witness

separation, alcohol intoxication, eye-closure, advice on information exposure, and identifying and recording needs and vulnerabilities.

Witness separation and alcohol intoxication were not included as participants took part individually and were required not to be under the influence of any substance. Eye closure was not included as the chatbot modality is typed, requiring participants to look at the interview questions they are being asked. These features may change if ChatCharlie is further developed into a spoken modality. Participants were not advised about information exposure or misinformation, as a mock crime video was used, meaning there would be no exposure to others' views on the content. However, participants were advised not to discuss the content of the video with anyone before taking part in the second interview. Finally, no records of participants' needs or vulnerabilities were taken since this controlled study was undertaken with pre-screened participants in the general population. All participants were required to wear glasses as needed to correct their vision, and all participants were required to speak fluent English. There is a clear case for the ease of translating the chatbot to accommodate all languages to support the general population's diversity, and it would be possible to record needs and vulnerabilities to triage witnesses as the development process evolves. All the other guidelines were included, with careful consideration regarding creating and integrating each element, each detailed below.

5.1.2.1. Rapport building.

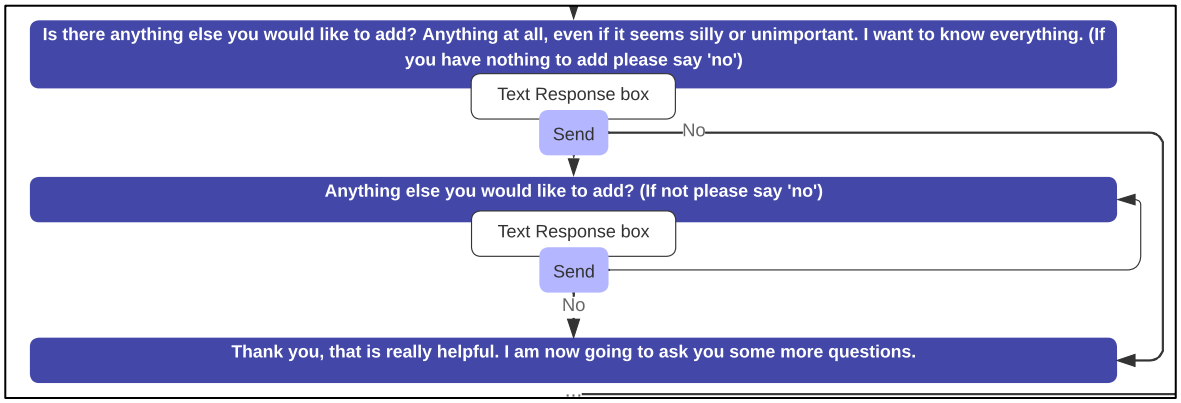
The literature surrounding rapport building for initial account interviews and the literature on chatbot usability and acceptance reveals that rapport results from positive mutual attention and requires a trusting and open atmosphere (Abbe & Brandon, 2013; Tickle-Degnen & Rosenthal, 1990). Creating this atmosphere involves the engagement of both witness and interviewer, with friendly, effortless

interactions leading to feelings of cooperation (K. Collins & Carthy, 2019). Expertise has been shown to influence trust between people (Mayer et al., 1995) and, traditionally, trust in computers (Fogg & Tseng, 1999). This trust is built through perceptions of competence, accuracy, understanding, and knowledge base (Corritore et al., 2003; Nordheim et al., 2019). It has been found that the perception of accuracy and understanding applies disproportionately to chatbot errors, with people being more likely to forgive or understand human errors (Dietvorst et al., 2015; Merritt & Ilgen, 2008) but becoming annoyed and irritated by chatbot errors. Hence, when a chatbot makes a mistake, users judge it much more harshly than when a human makes a mistake. If a chatbot is not perceived as knowledgeable, this directly impacts the user's level of trust in the technology (Nordheim et al., 2019).

ChatCharlie was extensively tested to ensure users would have a 'good' experience. Testing and trailing ChatCharlie involved user interface considerations, ensuring all the buttons, icons, and input fields worked as intended. User experience testing was implemented to ensure the flow of ChatCharlie was logical, and the questions were understandable and straightforward. That said, the questions and flow mapped onto the current College of Policing initial account procedure, and it was expected that the questions had already been the subject of some review and, therefore, would not be problematic for the general population. Limit testing was also conducted to establish chatbot responses for unforeseen scenarios. This was particularly important when different text responses resulted in different outcomes (See Figure 3).

Figure 3

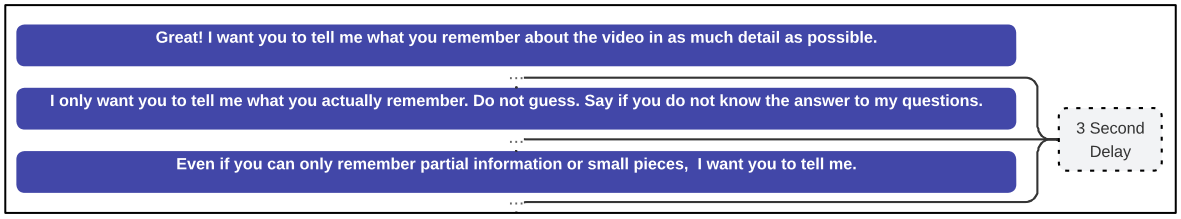
Example of ChatCharlie outcome reliant on text input



The guidelines for fostering rapport during the initial account interview highlight the importance of appropriate pacing during the interview (College of Policing, 2019). Studies examining the automation of this pacing in a chatbot format have found that the speed of responsiveness is essential in building trust with human users (Nordheim et al., 2019). However, the speed of response can also be perceived as unhuman-like, reducing trust, so a careful balance must be struck in this regard. Mirroring the speed of human messaging and breaking up long chunks of text can provide a more human-to-human-like interaction (Brandtzaeg & Følstad, 2017). Customer service chatbots are negatively viewed as lacking a natural flow of conversational communication skills (Luger & Sellen, 2016). Therefore, messages from ChatCharlie were timed to appear after the user had time to read the previous message, where a response was not required but where information was being presented. This design would prevent the chatbot from appearing unresponsive without overwhelming the reader with information (Figure 4).

Figure 4

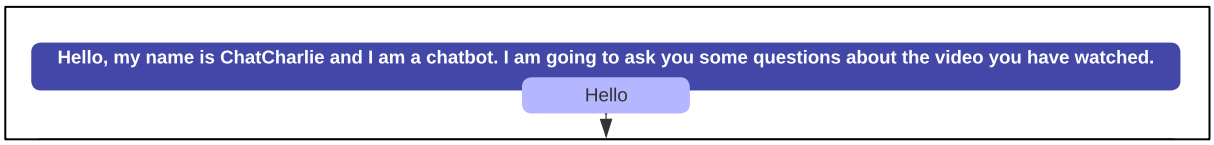
Example of Time delay messages from ChatCharlie.



Other features were added to develop a more human-like interaction. Users respond more positively when chatbot communication and language are perceived as human-like (Mone, 2016). Even small details, such as providing a chatbot with a human-like name, increase the perception of authentic interaction (Araujo, 2018). The perception that chatbots are poor at conveying empathy and other human emotions (Nadarzynski et al., 2021) has been exacerbated by reported negative experiences with customer service chatbots (Araujo, 2018). This highlights the importance of the design of a chatbot as it can influence emotional connection and rapport and, thus, acceptability of the chatbot and improved episodic memory recall (Kieckhaefer et al., 2014). Therefore, giving the chatbot a name and a friendly human-like demeanour is a vital design consideration (Figure 5).

Figure 5

Friendly and personable introduction from ChatCharlie



Trust is a crucial element of building rapport. However, trust in chatbots can be negatively affected by the perception of risk (Corritore et al., 2003). In a study looking at perceptions of a health chatbot, participants were hesitant to answer sensitive questions as they were unclear about how the data could be used and their

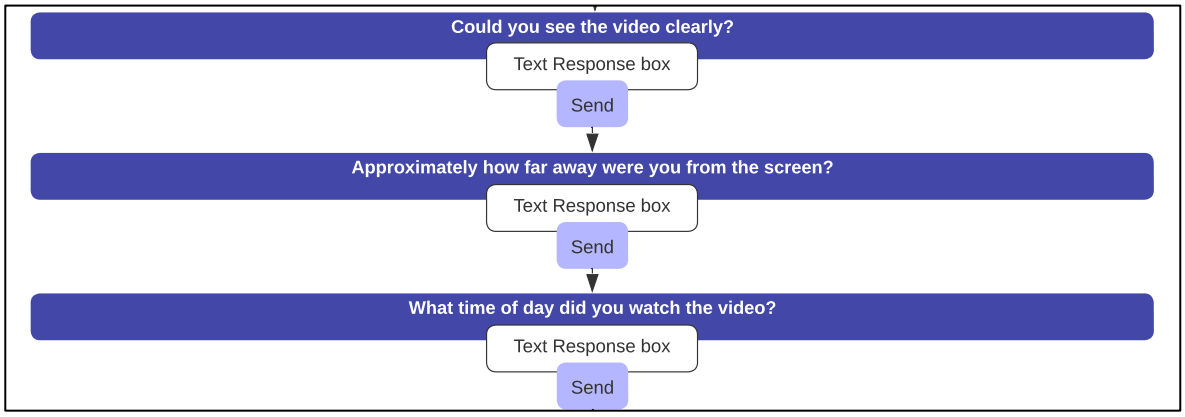
rights regarding privacy and confidentiality (Nadarzynski et al., 2021). Sensitive topics raise concerns in individuals about how others will perceive them and, more seriously, fears surrounding the repercussions of disclosure (Tourangeau & Yan, 2007). Increasing levels of privacy and anonymity have been shown to increase willingness to engage with sensitive questions (Joinson et al., 2007). Regarding an initial account interview chatbot, reassurances about information security would be paramount in fostering trust, particularly as anonymity would not be able to be consistently considered in real-world utilisation of this technology. Confidentiality and anonymity could be guaranteed for experimental purposes per appropriate ethical guidelines.

5.1.2.2. Clarifying sources of information.

For experimental purposes, participants were not asked how they knew the information they were sharing with the chatbot. This question was not included so as not to compromise the conversational elements of the prototype chatbot that were designed to build rapport. However, participants were asked the following questions to simulate the core questions recommended to clarify the source of information recorded (College of Policing, 2019). The questions were ‘could you see the video clearly?’, ‘Approximately how far away were you from the screen?’ and ‘What time did you watch the video?’ (Figure 6).

Figure 6

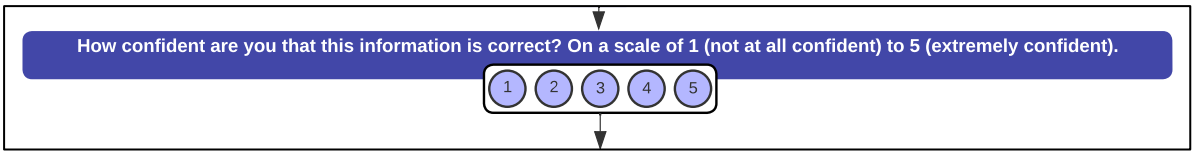
Questions to clarify the source of information from ChatCharlie.



Generally, confidence in memory is a positive indicator of an accurate account (Paulo et al., 2019; Wixted & Wells, 2017), although not always (Brewin et al., 2020; Shapira & Pansky, 2019; Wells et al., 2020; Wells & Bradfield, 1998). Generally, asking witnesses to self-report confidence levels has demonstrated a positive relationship to memory accuracy in both the cued recall phase of an interview (Caso et al., 2024; Luna & Martín-Luengo, 2012) and the free recall phase (Allwood et al., 2005), but not always. Furthermore, the relationship between confidence and accuracy of memories is most reliable when witnesses are asked about their confidence immediately after the recall (Brewer & Weber, 2008). Thus, confidence questions were introduced after every phase of the chatbot's initial account interview (See Figure 7). The inclusion of this question is in line with initial account interview guidelines and encourages participants to consider their perception of accuracy (College of Policing, 2019).

Figure 7

Confidence questions asked by ChatCharlie

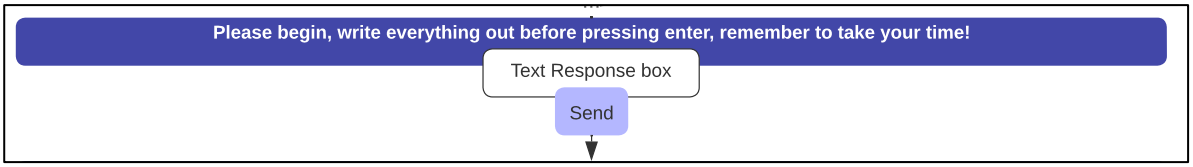


5.1.2.3. Witnesses' own words and non-leading open questioning.

During an initial account interview, the focus is on summarising what has happened (College of Policing, 2019). Although specific information must be gathered at this stage, investigators should adopt an open questioning style as it has been shown to increase the quality and quantity of recall compared to closed questions (Brubacher et al., 2020; Oxburgh et al., 2010; Sharman & Powell, 2012). The language should be neutral, as assumptions can be taken from leading words and cause incorrect information to be shared (Aharoni et al., 2020; Beek et al., 2021; Kleider-Offutt et al., 2015; Loftus & Palmer, 1974). Accordingly, ChatCharlie begins with an invitation to the participant to provide a free recall account of what they remember. The open format allows for the provision of an uninterrupted first-hand account. To prevent the inclusion of additional information, ChatCharlie only asked specific questions about simple concepts (e.g. How many people did you see?). Therefore, ChatCharlie cannot include further details or accidental red herrings, unlike human interviewers, who can struggle with information-seeking and confirmation bias (Akca et al., 2021; MacDonald et al., 2017; Memon et al., 1995; Shepherd & Milne, 1999). Furthermore, open questions typically elicit more accurate accounts of what has been witnessed (Brubacher et al., 2020; Oxburgh et al., 2010; Sharman & Powell, 2012). ChatCharlie is designed with open-text answer boxes, allowing complete and descriptive responses to further questioning (Figure 8).

Figure 8

Example of an open question asked by ChatCharlie.

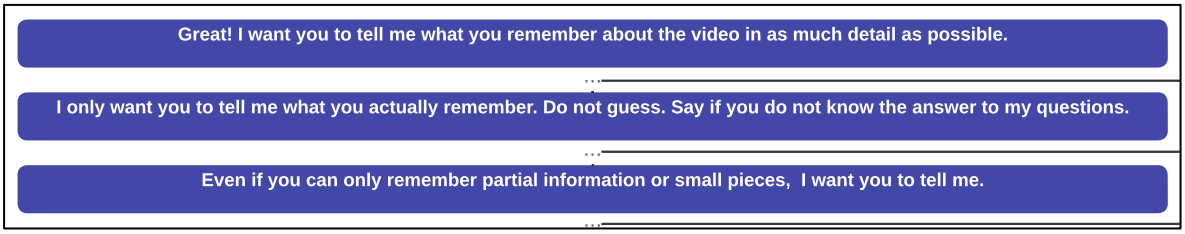


5.1.2.4. Allowing uncertainty.

Empowering witnesses to feel comfortable saying they do not know the answer to a question is important (Dando et al., 2023; Oxburgh et al., 2010; Paulo et al., 2019). Being able to say “I do not know” can reduce the quantity of information while ensuring higher accuracy of responses (Bull & Milne, 2020; Dando et al., 2023; Evans & Fisher, 2011; Holliday et al., 2012; Scoboria & Fisico, 2013). Police are thus instructed to reassure and inform witnesses that they do not have to answer questions if they are unsure. Ground rules are recommended to help develop ‘witnesses-perceived autonomy’ (Fisher et al., 2011). These rules: ‘never guess’, ‘report everything’, ‘say if you do not remember’, and ‘tell me if you do not understand the question’, are all in place to reduce witness error (Blank & Launay, 2014; Goldsmith et al., 2005). Informing witnesses that it is normal for them not to know the answers to questions can also minimise guessing (Bull & Milne, 2020; Fisher et al., 2011; Fisher & Geiselman, 1992; Holliday et al., 2012; Scoboria & Fisico, 2013). These ground rules are included at the beginning of the ChatCharlie initial account interview (Figure 9).

Figure 9

The ground rules within ChatCharlie.



5.1.2.5. In-Person Initial Account protocol.

The in-person initial account interview followed the guidelines from the College of Policing and mapped onto the ChatCharlie initial account interview (College of Policing, 2019) (See Appendix K). Thus, five of the ten key guidelines were implemented for this experimental study. The excluded guidelines were witness separation, alcohol intoxication, eye-closure, advice on information exposure, and identifying and recording needs and vulnerabilities. Details on why these were excluded can be found in section 1.4. The in-person initial account interview included the following guidelines: rapport building, clarifying sources of information, witnesses' own words, non-leading questions, and allowing uncertainty.

5.1.3. The Current Study Research Questions

A series of research questions were developed to guide the analysis of this novel research, which are as follows:

1. Is ChatCharlie an effective method for collecting an initial account to consolidate a memory trace for improved recall during later, more in-depth interviews? To investigate this question, participant recall performance at Time 2 was compared as a function of the three Time 1 conditions.
2. How does ChatCharlie compare at Time 1 to an in-person initial account regarding the quantity (amount) and quality (errors) of information elicited?

3. Does the Time 1 retrieval condition impact self-reported confidence and perceived accuracy at Time 2?

5.2. Method

5.2.1. Design.

A between-subjects mock eyewitness paradigm was employed to investigate the efficacy of ChatCharlie. The design had two distinct time points, Time 1 and Time 2. At Time 1, participants were randomly allocated to one of three Time 1 conditions: ChatCharlie In-person initial account or Control (no initial account). Having viewed the stimulus video, participants' initial accounts were collected within 15 minutes as a function of the Time 1 condition. One week later, at Time 2, all participants were interviewed about the film using a basic Tier 1 investigative interview. The dependent variable was Time 2 memory performance measured by the number of correct, incorrect, and confabulated items reported and percentage accuracy. Time 1 memory was also analysed across the ChatCharlie and In-person conditions. Confidence and self-report accuracy data were also analysed and reported.

5.2.2. Participants.

A total of 90 participants took part in this study: 28 male (31%), 60 female (67%), and 2 non-binary/third gender (2%). The mean age was 29.26 (SD = 8.01), ranging from 18 to 54 years. There was a non-significant difference in age across each of the three conditions, $F(2, 87) = 2.128$, $p = .125$ ($M_{ChatCharlie} = 30.83$; $M_{Control} = 30.10$; $M_{In\text{-}Person} = 26.83$). Given the results of the previous study, this study focused on participants under 55 to focus on the impact of using a chatbot. Participants were recruited through online advertising, snowball sampling, and word of mouth and were compensated for the time spent taking part through a £10 shopping voucher.

5.2.3. Procedure.

As this study was conducted remotely, all recruitment and communication took place over email. All potential participants were first emailed a participant information sheet (see Appendix M and Appendix N) and a copy of the consent sheet (see Appendix C). Participants committed to taking part in two elements precisely one week apart. Although all aspects were remote, the researcher was available to help with any difficulties in real-time.

Time 1: All participants received an email with a link to complete the consent form, watch a one-minute video of a non-violent mock crime and complete a short demographic survey. Once completed, participants in the ChatCharlie condition were linked to the designated website to participate in an initial account interview with ChatCharlie or were interviewed in person by one of three interviewers (referred to as interviewers A, B, and C), using the College of Policing guidance for initial account interviews, following an interview protocol, verbatim (see Appendix K). Finally, participants in these two conditions completed a short online survey on their experiences and perceptions of ChatCharlie and In-person interviews. Participants in the control condition had no further tasks at Time 1. Interviewer A conducted 12 interviews, Interviewer B conducted 14 interviews, and Interviewer C conducted 4. A Kruskal-Wallis test for participant memory performance at Time 2 as a function of Time 1 interviewer revealed a non-significant differences for global memory performance on all measures as a function of Time 1 interviewer, overall correct, $H(2) = 1.015$, $p = .602$, overall Errors $H(2) = .391$, $p = .882$, overall confabulations, $H(2) = 2.030$, $p = .362$, and overall accuracy, $H(2) = .829$, $p = .629$.

Time 2: One week later, all participants took part in a basic tier 1 type PEACE investigative interview conducted by the same interviewer using an interview protocol

(see Appendix H and I). Interviews took place face-to-face via video call platform and were digitally audio and video recorded. After the interview, all participants received a debriefing sheet (See Appendix O).

5.2.4. Materials.

5.2.4.1. Stimulus.

All participants viewed the same mock-crime video using the procedure described in Chapters 3 and 4 of this thesis (see sections 3.3.4.1.).

5.2.4.2. Ethics materials.

This research was ethically approved by the University of Westminster research ethics committee: ETH 2122-2235 (See Appendix L). All participants were emailed a participant information sheet before agreeing to participate in any of the experiments (See Appendices M and N). All participant information sheets contain information unique to this study and the condition to which the participant was allocated. Participants who did not take part in an interview with a chatbot were not informed about the use of this technology before taking part in the study. All participant information sheets contained participation requirements, what participants would be asked to do, ethics information, and contact details.

Additionally, all participants completed an online consent form through a Qualtrics link before beginning the study (See Appendix C). Participant demographic questions were asked across all studies in this work (See Appendix D and Section 3.3.4.3.). All participants received a debrief sheet (See Appendix O) with further information about the relevant study's aims and purpose. All participants, regardless of condition, were provided with details about chatbots and the purpose of the research. A summary of what they had done in the study and additional information about the other conditions was given. A detailed explanation of how the data will be

anonymised and instructions for removing data if participants wish to withdraw, were given. Information about compensation for taking part and the researcher's contact details were provided.

5.2.4.3. Time 1 Initial Account in-person interview.

The primary focus of this PhD thesis is the initial account interview as described by the College of Policing (2019). Here, at Time 1, participants took part in an initial account interview with a chatbot (See Appendix K) or an in-person initial account as defined as immediately necessary initial questioning to elicit a brief account of what is alleged to have happened (See Appendix K) (College of Policing, 2019).

5.2.4.4. Time 1 ChatCharlie interview.

The interview protocol presented to participants via ChatCharlie was modelled on the current initial account interview (College of Policing, 2019) (See Appendix K). The guidelines for conducting initial account interviews have been interpreted and modified accordingly to create an appropriate comparative tool for this functionality (See Figure 12 below).

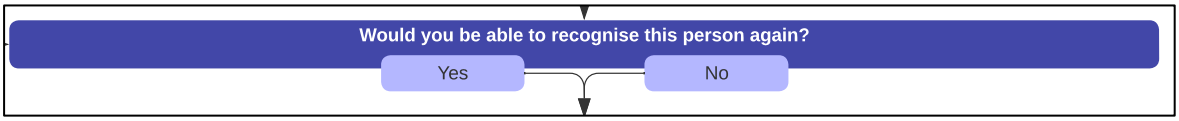
Several chatbot platforms were considered, as some chatbot creation services require social media to communicate with the chatbot. It was concluded that utilising any social media platform would compromise experimental ethics, so chatbots limited to Facebook Messenger, Instagram, and Twitter were eliminated. The chatbot would be created with a synchronous text functionality, eliminating all voice-activated services such as Amazon Alexa, Google Assistant, and iOS voice applications. It was found that additional costs would apply to text messaging services and WhatsApp messages, so these services were eliminated. Some chatbot services were designed to be embedded in user websites, which would not be appropriate for

this study, so only chatbot platforms with dedicated hosting websites were considered.

Question features were considered as it was important for the chatbot service to offer question features and types that could replicate an initial account interview. There was a requirement for unlimited open-text responses to enable a total free recall without limiting the number or length of messages (See Figure 12 below). Multiple-choice questions were also required to allow for yes or no questions to be asked (Figure 10).

Figure 10

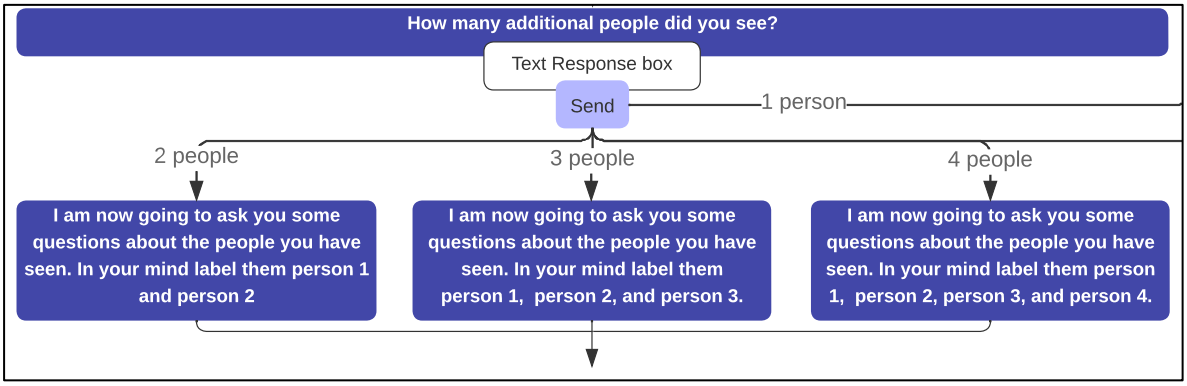
Example of a yes or no question asked by ChatCharlie



Additionally, a certain level of question logic would be required to ensure the chatbot could respond accurately. An example of this is the ability of the chatbot to understand how many people the participant reported so that subsequent questions ask about the correct number of people (Figure 11).

Figure 11

An example of ChatCharlie using question logic to ask accurate questions



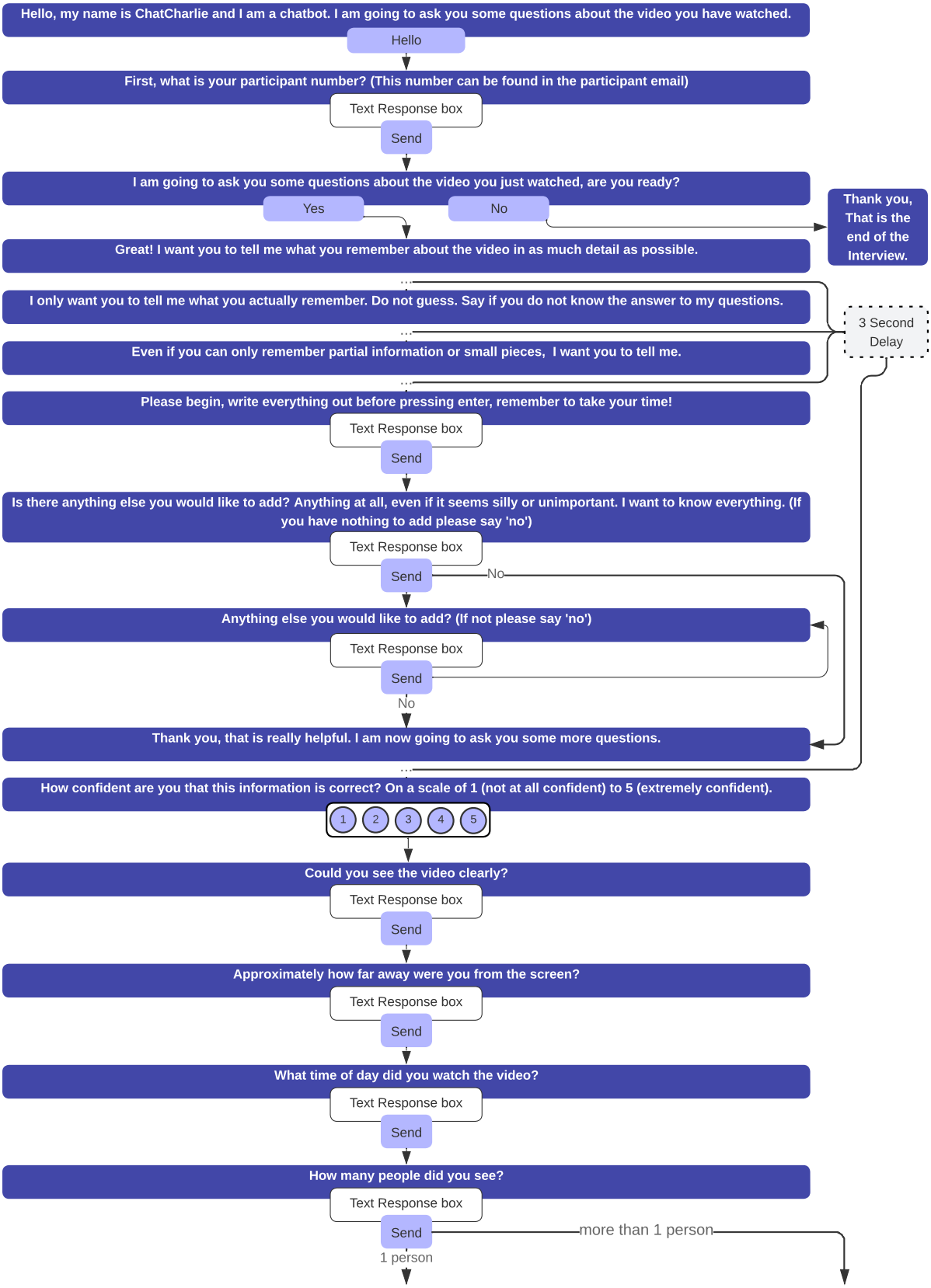
Some chatbot services limit the amount of respondent data that could be downloaded. Once identified, these services were removed from the list as full access to respondent data was needed. Complete data was necessary for accuracy and enabled the findings from this work to be collated for future analysis. Additionally, to qualify, chatbot websites had to comply with GDPR to ensure that participant information would not breach the ethical requirements for this body of work.

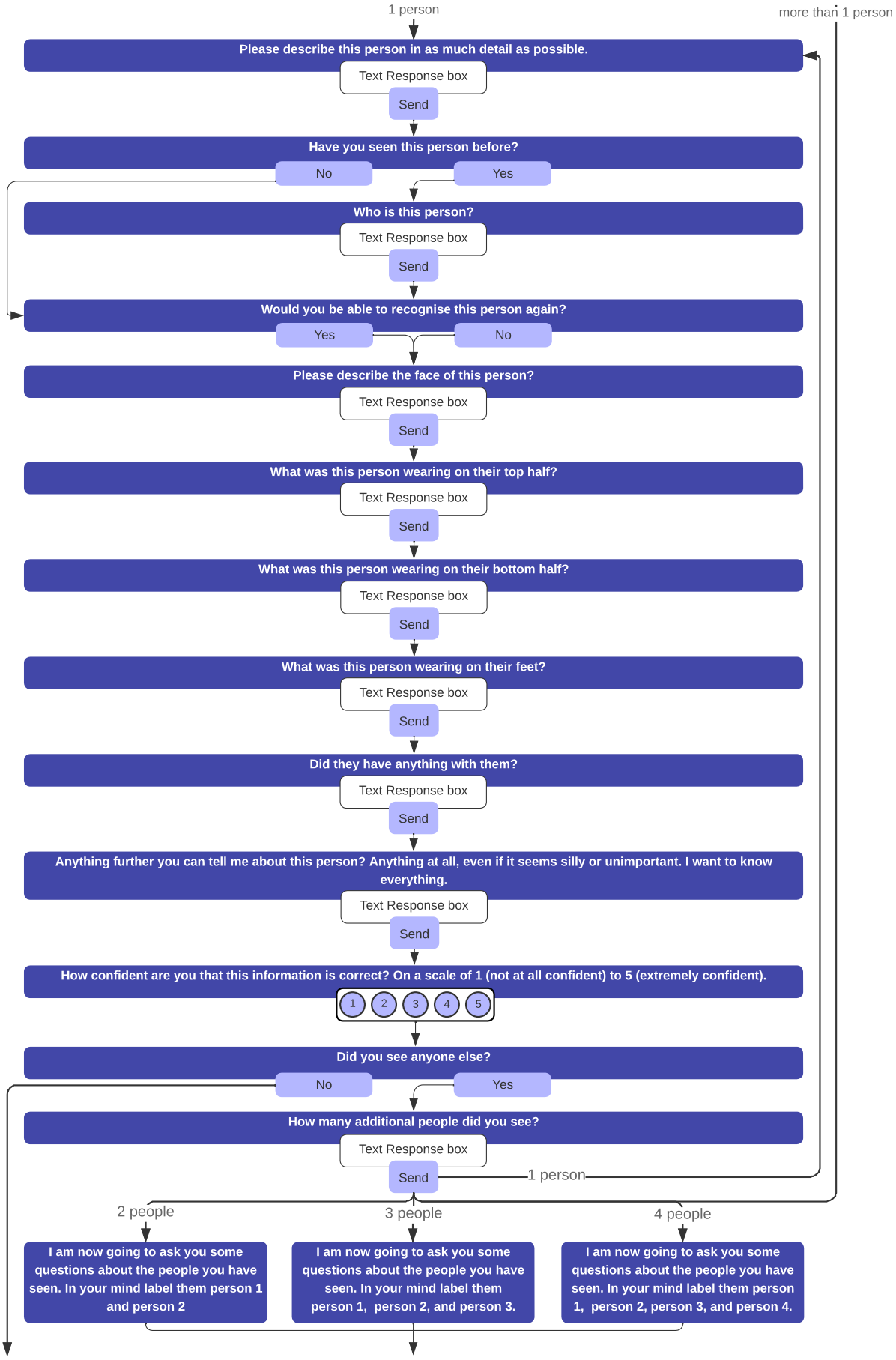
Many elements were considered when exploring the potential cost of creating the chatbot. While some chatbot services offer a free trial period, most offer tiered costing ranging from basic to premium levels. We aimed to find a chatbot website that could enable the required question features, provide a dedicated hosting website, and allow an appropriate number of respondents whose data could be fully exported.

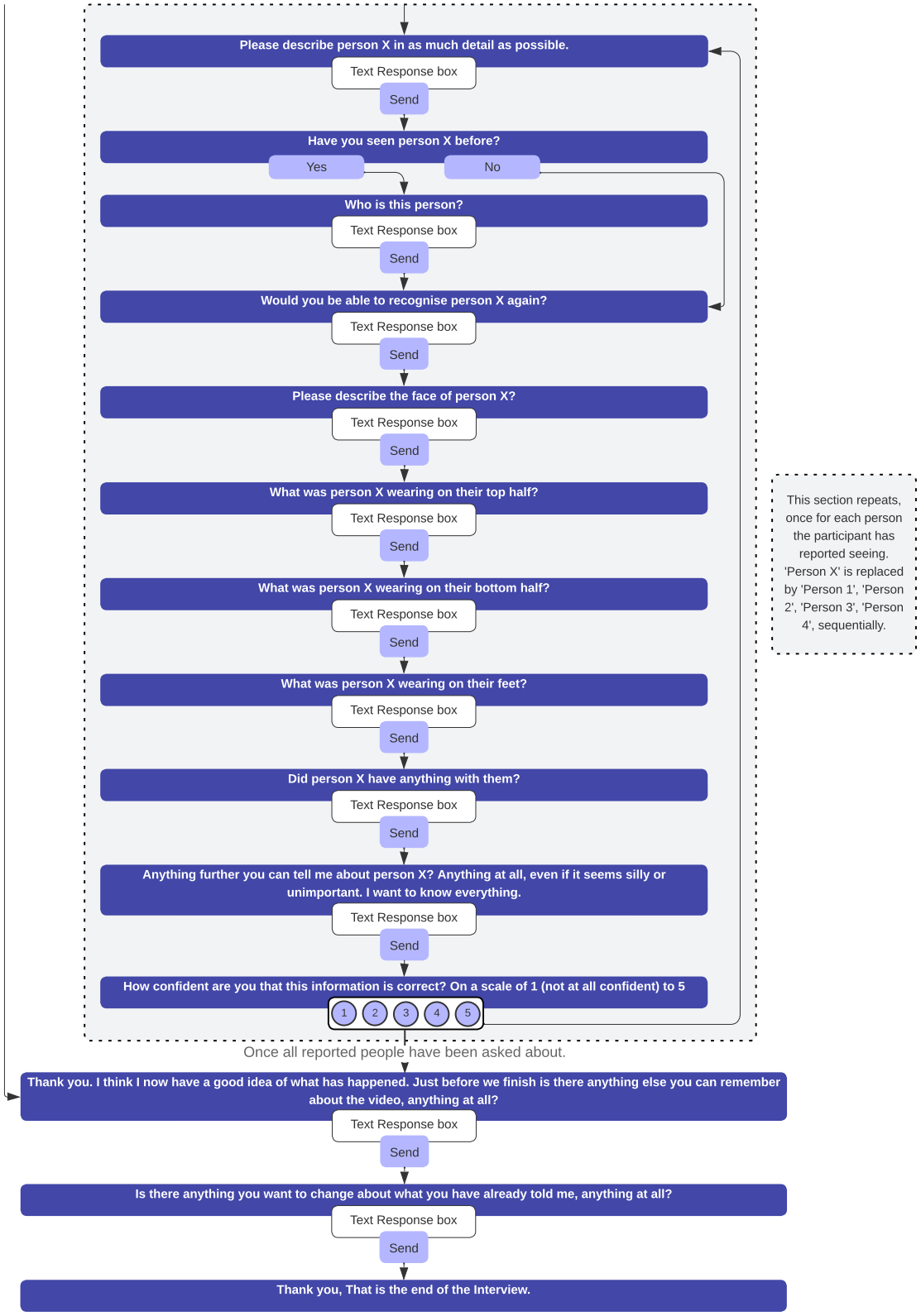
Survey Sparrow hosted the chatbot platform that was most appropriate for this project. It offered all the required functionality within a basic tier pricing and GDPR compliance. Subsequently, a chatbot named ChatCharlie was built using this platform.

Figure 12

ChatCharlie Conversation flow




1 person
more than 1 person
Please describe this person in as much detail as possible.
Text Response box
Send
Have you seen this person before?
No
Yes
Who is this person?
Text Response box
Send
Would you be able to recognise this person again?
Yes
No
Please describe the face of this person?
Text Response box
Send
What was this person wearing on their top half?
Text Response box
Send
What was this person wearing on their bottom half?
Text Response box
Send
What was this person wearing on their feet?
Text Response box
Send
Did they have anything with them?
Text Response box
Send
Anything further you can tell me about this person? Anything at all, even if it seems silly or unimportant. I want to know everything.
Text Response box
Send
How confident are you that this information is correct? On a scale of 1 (not at all confident) to 5 (extremely confident).
1
2
3
4
5
Did you see anyone else?
No
Yes
How many additional people did you see?
Text Response box
Send
1 person
2 people
3 people
4 people
I am now going to ask you some questions about the people you have seen. In your mind label them person 1 and person 2
I am now going to ask you some questions about the people you have seen. In your mind label them person 1, person 2, and person 3.
I am now going to ask you some questions about the people you have seen. In your mind label them person 1, person 2, person 3, and person 4.


Please describe person X in as much detail as possible.
Text Response box
Send
Have you seen person X before?
Yes
No
Who is this person?
Text Response box
Send
Would you be able to recognise person X again?
Text Response box
Send
Please describe the face of person X?
Text Response box
Send
What was person X wearing on their top half?
Text Response box
Send
What was person X wearing on their bottom half?
Text Response box
Send
What was person X wearing on their feet?
Text Response box
Send
Did person X have anything with them?
Text Response box
Send
Anything further you can tell me about person X? Anything at all, even if it seems silly or unimportant. I want to know everything.
Text Response box
Send
How confident are you that this information is correct? On a scale of 1 (not at all confident) to 5
1
2
3
4
5
This section repeats, once for each person the participant has reported seeing. 'Person X' is replaced by 'Person 1', 'Person 2', 'Person 3', 'Person 4', sequentially.
Once all reported people have been asked about.
Thank you. I think I now have a good idea of what has happened. Just before we finish is there anything else you can remember about the video, anything at all?
Text Response box
Send
Is there anything you want to change about what you have already told me, anything at all?
Text Response box
Send
Thank you, That is the end of the Interview.

5.2.4.5. Time 2 basic tier 1 PEACE type investigative interview.

All studies in this body of work measured memory performance through a tier 1 PEACE investigative interview (See Appendix H). The Tier 1 PEACE investigative interview procedure outline is taken from The College of Policing (College of Policing, 2022) – see Chapter 3 Section 3.4.4. and Appendix I for a full description, but in brief, the procedure comprised the following phases:

- Planning and preparation:
- Engage and Explain, including the ground rules i) 'never guess', ii) 'report everything', iii) 'say if you do not remember', and iv) 'tell me if you do not understand the question'.
- Account, Clarification, plus a reminder of the four ground rules above; question;
- Closure (including the opportunity to add and/or alter anything.

5.2.5. Coding.

Transcripts from Time 1 and Time 2 interviews were coded for percentage accuracy and recall of distinct pieces of information. Information is categorised as correct, incorrect, or confabulated (See section 3.3.5. and Appendix J).

5.3. Results

5.3.1. Analysis approach.

To investigate the research questions formulated for this study, a series of one-way between-subjects ANOVAs were conducted, applying Bonforroni's correction as appropriate. Analyses investigated Time 1 conditions on Time 2 memory performance globally, that is, across the entire interview. Next, performance in each recall phase (free and cued recall) was analysed individually. A comparison of

memory performance at Time 1 was conducted comparing ChatCharlie and In-person Time 1 interview performance. Finally, repeated measures analyses were performed to investigate the potential carryover and consolidation effects of Time 1 (using the above memorial measures) on performance at Time 2.

5.3.2. Time 1 Initial Account Interview.

A series of independent samples t-tests for Time 1 memory performance revealed a significant difference for the number of confabulations $t(60) = 8.297$, $p = .006$ ($M_{ChatCharlie} = .57$, SD = .68; $M_{InPerson} = 1.00$, SD = 1.23). Participants in ChatCharlie confabulated less at Time 1 than in the In-Person condition. There were non-significant differences between the ChatCharlie and In-person Time 1 retrieval groups for the amount of correct information reported at Time 1, $t(60) = 3.404$, $p = .007$ ($M_{ChatCharlie} = 37.03$, SD = 9.05; $M_{InPerson} = 39.50$, SD = 13.08) and number of errors, $t(60) = 2.007$, $p = .155$ ($M_{ChatCharlie} = 3.93$, SD = 2.23; $M_{InPerson} = 4.73$, SD = 2.84).

5.3.3. Time 2 interview global recall performance.

5.3.3.1. Correct recall.

There was a significant main effect of the Time 1 condition for correct recall at Time 2, $F(2, 87) = 13.935$, $p < .001$, η_p^2 .24. Participants in the ChatCharlie and In-Person Time 1 conditions recalled more correct information at Time 2 than those in the Control condition, all *p*s < .001. There was a non-significant difference between the former two conditions, $p = .256$ (see Figure 13 below)

5.3.3.2. Erroneous recall.

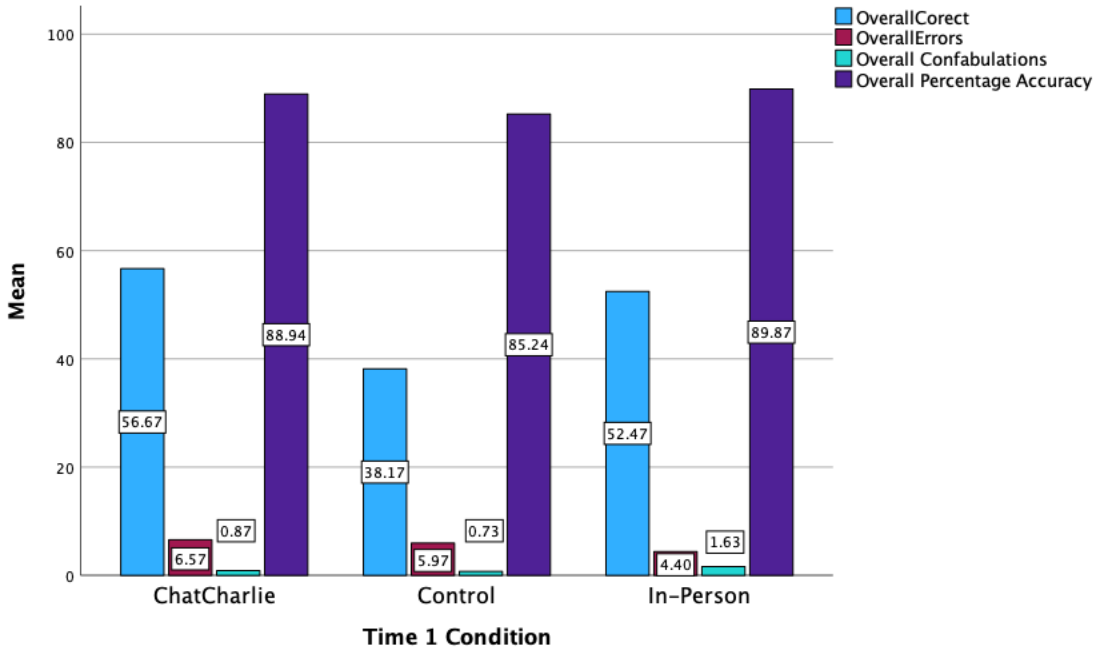
There was a non-significant main effect of Time 1 condition for the amount of erroneous recall at Time 2, $F(2, 87) = 3.791$, $p = .026$, η_p^2 .08 (see Figure 13 below).

5.3.3.3. Confabulations.

There was a significant main effect of Time 1 condition for the number of confabulations at Time 2, $F(2, 87) = 6.019$, $p = .004$, η_p^2 .12. Participants in both the In-person and ChatCharlie conditions reported fewer confabulations at Time 2 than those in the Control, $p = .002$ and $p = .007$, respectively. There was a non-significant difference between the former two conditions, $p = .635$ (see Figure 13 below).

Figure 13.

Time 2 mean global memory performance & percentage accuracy as a function of Time 1 retrieval condition (N = 90).



5.3.3.4. Global percentage accuracy.

Analysis of variance revealed a significant main effect of Time 1 condition for percentage accuracy in Time 2 interviews, $F(2, 87) = 180.126$, $p = .005$, η_p^2 .12 (see Figure 8 for means). Participants in the Control condition were less accurate (*SD* = 7.11, 95% CI 82.58, 87.89) at Time 2 than those in both the In-person (*SD* = 4.70,

95% CI 88.11, 91.62) and ChatCharlie (*SD* = 4.61, 95% CI 87.89, 90.66) conditions, *p* = .006 and *p* = .037, respectively, with a non-significant difference between the Chatbot and In-Person conditions, *p* = .718

5.3.4. Time 2 interview phase memory performance.

5.3.4.1. Free recall.

There was a significant main effect of Time 1 condition for correct recall in the free recall phase of Time 2 interviews, $F(2, 87) = 20.786$, $p < .001$, η_p^2 .32. Participants in the Control condition recalled significantly fewer correct information items than those in both ChatCharlie and In-person Time 1 conditions, all *p*s < .001, with a non-significant difference between the latter two conditions, *p* = .718.

There were non-significant differences for errors and confabulations, $F(2, 87) = .932$, p = .397, η_p^2 .02, and, $F(2, 87) = 3.053$, p = .052, η_p^2 .07, respectively. There was a significant difference of the Time 1 condition for percentage accuracy, $F(2, 87) = 7.572$, $p < .001$, η_p^2 .15. Participants in the ChatCharlie condition were more accurate than those in both in-person and Control conditions, p = .003 and In-person conditions, *p* = .013, with non-significant differences between the latter two conditions, *p* = .937 (see Table 8)

5.3.4.2. Cued recall.

There was a significant main effect of Time 1 condition for the amount of correct information items recalled in the cued recall phase at Time 2, $F(2, 87) = 12.215$, p < .001, η_p^2 .22. Participants in ChatCharlie recalled significantly more correct information than participants in both the Control and In-person conditions, *p* =.009 and *p* < .001, respectively, with no difference between the latter two conditions, *p* = .204.

There was a significant difference of Time 1 condition for errors, F(2, 87) = 7.641, p < .001, η_p^2 .22. Participants in the In-person condition recalled significantly fewer errors than Control and ChatCharlie, *p* =.003 and *p* < .001, respectively, with a non-significant difference between the latter two conditions, *p* = .993.

There was a significant difference across the Time 1 condition for confabulations in the cued recall phase of Time 2 interviews, *F*(2, 87) = 5.635, *p* = .005, η_p^2 .11. Participants in the In-person Time 1 condition confabulated more than participants in the ChatCharlie condition, *p* = .004. All other differences were non-significant, all *p*s >.061. There was a significant difference in percentage accuracy in the cued recall as a function of the Time 1 condition, *F*(2, 87) = 6.581, *p* = .002, η_p^2 .13. Participants in ChatCharlie were significantly more accurate than those in the In-Person Time 1 condition, *p* = .005. All other differences were non-significant, all *p*s > .0.59.

Table 8

Means & standard deviations (in parenthesis) for interview phase free recall & cued question recall.

	Condition Mean (SD)		
	ChatCharlie	In-Person	Control
Free Recall Correct	34.37 (10.64)	38.60 (12.23)	21.57 (8.81)
Free Recall Incorrect	2.50 (2.89)	2.13 (1.17)	1.77 (1.79)
Free Recall Confabulations	0.47 (0.68)	0.60 (0.72)	0.20 (0.48)
Free Recall Accuracy %	92.66 (5.73)	85.97 (8.90)	92.37 (6.43)
Cued Recall Correct	22.13 (7.45)	13.87 (5.77)	16.60 (6.46)
Cued Recall Incorrect	4.07 (2.25)	2.27 (1.14)	4.00 (2.42)
Cued Recall Confabulations	0.40 (0.62)	1.03 (0.81)	0.53 (0.86)
Cued Recall Accuracy %	84.11 (6.96)	71.14 (19.79)	78.27 (11.70)

5.3.4.3. Memory confidence at Time 2.

There was a significant main effect for how confident participants were that they had recalled a lot of information at Time 2, $F(2, 87) = 18.286$, $p < .00$, η_p^2 .29. Participants in the In-Person Time 1 condition ($M = 1.23$, SD = .50) were significantly less confident at Time 2 than those in both the ChatCharlie ($M = 2.67$, SD = 1.24) and Control conditions ($M = 2.57$, SD = 1.16), all *p*s < .001. There was a non-significant difference between the latter two conditions, $p = .945$.

There was also a significant main effect for how confident participants were that they had not made any errors when recalling the event at Time 2, $F(2, 87) = 11.694$, $p < .00$, η_p^2 .21. Participants in the In-person Time 1 condition ($M = 3.33$, SD = .88) were more confident they had not made any errors at Time 2 than those in both the ChatCharlie ($M = 2.50$, SD = 1.00) and Control Time 1 conditions ($M = 2.22$, SD = .92), $p = .003$ and $p < .001$, respectively. There was a non-significant difference between the ChatCharlie and the Control conditions, $p = .458$. Confidence at Time 2 was a non-significant predictor of percentage accuracy at Time 2, accounting for just 3% (adjusted $r^2 = 0.32$) of the variance, $F(2, 87) = 2.447$, $p = .092$.

5.3.4.4. Memory confidence Time 1 to Time 2

Repeated measures analysis revealed a main effect of Time for self-reported confidence for correct information, $F(1, 58) = 66.906$, $p = < .001$, η_p^2 .54, and a significant Time and Condition mixed interaction, $F(1, 58) = 52.678$, $p = < .001$, η_p^2 .48. Irrespective of Time 1 condition (ChatCharlie; In-person), confidence in the amount of correct information recalled dropped from Time 1 (M = 3.13, SD =1.08, 95% CI, 2.866, 3.033) to Time 2 ($M = 1.95$, SD = .86, 95% CI, 2.051, 2.649) at Time 2. Participants' confidence in the In-person Time 1 condition dropped significantly

from Time 1 to Time 2, $p < .001$, whereas the confidence of participants in ChatCharlie remained constant from Time 1 to Time 2, $p = .517$.

There was a non-significant repeated measure main effect for confidence in errors from Time 1 ($M = 3.08$, SD = 1.17) to Time 2 ($M = 2.92$, SD = 1.02) and a non-significant Time X Condition mixed interaction, $Fs < 1.495$, all $ps > .226$ ($M_{\text{ChatCharlie T1}} = 2.50$, SD = 1.14; $M_{\text{ChatCharlie T2}} = 2.55$, SD = 1.00; $M_{\text{In-person T1}} = 3.67$, SD = .88; $M_{\text{In-person T 2}} = 3.33$, SD = .84).

5.4. Discussion

This chapter reports the results of an empirical investigation of the efficacy of ChatCharlie as a method of collecting initial account information after witnessing a mock crime. The findings previously reported are now discussed, centred on the research questions developed for this chapter. First, considering overall or global memory performance at Time 2, results show that participants who provided an initial account via ChatCharlie or in-person recalled 47% and 38% (respectively) more correct event information at Time 2 than those in the control condition.

Furthermore, this significantly improved performance did not emerge accompanied by any significant increase in errors or confabulations. Thus, overall percentage accuracy was also improved. The benefits of providing a quick retrieval were found to be robust across both interview phases, with a similar pattern of results in terms of improved recall versus the Control condition.

This pattern of results is in line with previous research demonstrating that recalling an event again promptly can consolidate and strengthen the memory trace towards improving recall at a later date (Dando et al., 2020; Gabbert et al., 2009; Gawrylowicz et al., 2014; Roediger & Butler, 2011) and theoretical understanding of

the benefits of memory consolidation (Naji et al., 2019; Tucker et al., 2020). One of the challenges for applied cognitive research such as this is that a balance has to be struck between increasing the quantity of information an eyewitness provides without compromising quality (accuracy and errors) (Goldsmith et al., 2005; Mudd & Govern, 2004; Tuckey & Brewer, 2003; Weber & Brewer, 2008), and it seems as if ChatCharlie offers the potential for doing so, even when In-Person first accounts may be impossible.

While previous research has reported the benefits of providing a timely written and spoken initial recall, this is the first research to design and use a targeted chatbot. Accordingly, this study demonstrates that allowing participants to provide an initial account offers recall-strengthening benefits, whether via an in-person first account or ChatCharlie. The ChatCharlie questioning style helped preserve additional information that might otherwise have been lost over time, and the absence of a human interviewer is known to reduce demand characteristics. While the differences between ChatCharlie and In-Person conditions on memory performance at Time 2 were not statistically significant, differences were approaching significance, and the percentage differences in correct recall were notable (approaching 10% in favour of ChatCharlie). The reason for this is not entirely apparent. Still, ChatCharlie combines open questions with specific information-seeking questions, as described in the initial account guidance provided by the College of Policing (2019). This approach is based on the psychological understanding of eliciting good quality information. It balances the collection of investigative information while preserving episodic memory in terms of not ‘triggering’ errors by using inappropriate questions that can often emerge in a high-pressure human-to-human environment (Paulo et al., 2021). ChatCharlie inherently removes interruptions, including leading and multiple questions, and

reduces the accidental intrusion of assumptions. However, arguably more importantly, the social demands of conversing with a human that witnesses often report concerning memory performance are absent (Dando et al., 2023; Taylor & Dando, 2018).

Turning to memory confidence, some suggest that confidence is a helpful predictor of accuracy, although others would argue with this suggestion in part (Iida et al., 2020; Saraiva et al., 2020). Here, the findings indicate that even under the controlled conditions typical of mock witness studies, confidence assessments made shortly after the event at Time 1 and Time 2 did not prove to be reliable performance indicators. These findings suggest that in applied settings, recall confidence should be cautiously approached, especially following face-to-face interviews. In-person interviews focus on the witness, creating a supportive environment that may enhance the witness's comfort. This social comfort could increase self-assessed performance confidence (Newcomb et al., 2021; Tahan & Sminkey, 2012). Although further research should be done to confirm, our In-person initial account protocol incorporated rapport-building efforts, which appeared to result in notably higher confidence levels at Time 1, which then significantly decreased from Time 1 to Time 2. Conversely, ChatCharlie participants reported lower initial confidence levels, which remained consistent across both time points.

Previous studies have documented a decline in confidence under conditions involving actual stimuli and repeated confidence checks (Jondani et al., 2023). While our study did not include repeated confidence checks within the same interview, we did collect confidence data at both Time 1 and Time 2 for participants in the In-person and ChatCharlie conditions. The reason for the observed decrease in confidence across time points among In-person participants, as opposed to the

stability of confidence levels in the ChatCharlie group, remains unclear. This outcome highlights the need for further research in applied settings, particularly regarding confidence, accuracy, and comparisons between digital and human interview agents.

Participants performed equally as well as those in the ChatCharlie condition for correct recall, and they performed better than those in the control group but confabulated more. This invites further examination of underlying factors that may contribute to this apparent anomaly. In face-to-face interviews, complex social dynamics, such as nonverbal cues and perceived authority, can influence a witness's responses. Social interactions can create implicit expectations, leading individuals to alter their memories inadvertently. Research suggests these social dynamics can shape recall by encouraging witnesses to "fill in" memory gaps to meet perceived expectations (Dando et al., 2009; Hollenbaugh & Everett, 2013). This is less likely in the structured, predictable environment of an automated system like ChatCharlie.

Another contributing factor is rapport. While rapport is essential for building trust, it may also increase the tendency to over-report as witnesses strive to cooperate, leading to a greater risk of confabulation. Face-to-face interactions are generally more effective at establishing rapport than virtual or automated formats. This rapport may have encouraged participants to provide more extensive information, aligning with findings that rapport can enhance recall, though sometimes at the expense of accuracy (Vallano & Compo, 2011). ChatCharlie's standardised responses lack these interpersonal dynamics, potentially reducing the urge to elaborate under social pressure.

Cognitive load associated with processing social cues during in-person interviews may further impact recall accuracy. Studies show that social interactions impose

cognitive demands, detracting from an individual's focus on accurate memory retrieval. Witnesses may divide attention between the memory task and managing the social dynamics of the interview, leading to reduced recall fidelity (Wagstaff, 2008). In contrast, ChatCharlie's text-based, neutral interface may alleviate cognitive burden, allowing witnesses to concentrate on memory recall without social distractions.

Finally, feedback mechanisms in human interviews may also play a role in the observed anomaly. Human interviewers can, often unconsciously, provide verbal or nonverbal feedback that affects a witness's confidence and recall accuracy. ChatCharlie's neutral, non-interactive responses may mitigate these effects by offering a consistent, controlled environment free of reinforcing or discouraging cues (College of Policing, 2019). These combined factors—social influence, cognitive load, rapport, and feedback effects—may help explain why confabulations were less frequent in ChatCharlie interactions than in-person interviews.

Interestingly, confidence was not a predictor of percentage accuracy, again lending support to questions centred on the utility of confidence indicators in interviews for investigative purposes (Caso et al., 2024). Previous research has found that confidence in recall accuracy has had variable results in eyewitness tasks (Sauer et al., 2010; Sporer et al., 1995).

Given this set of results, ChatCharlie appears to have applied potential as an effective method for collecting an initial account to consolidate a memory trace for improved recall during later, more in-depth interviews since ChatCharlie's results were comparable to an in-person initial account. It is also worth noting that, as far as can be ascertained, this is the first empirical study of the College of Policing Initial

Account protocol. While sufficient theoretical literature supports the collection of an in-person initial account, this study now provides some empirical support.

Including ground rules is a primary element in the design of ChatCharlie. Previous research has found that while ground rules reduce the quantity of information, they ensure higher accuracy of responses (Evans & Fisher, 2011; Scoboria & Fisico, 2013). This is in line with the findings of this study. ChatCharlie use increased accuracy and correct elements while not increasing the incorrect and confabulated elements. However, future research should include an exclusion condition to isolate and further explore the specific effect of ground rules in using this modality.

Rapport building was considered when designing ChatCharlie since rapport building in the traditional sense was deemed impossible, and to try to mimic it would seem inappropriate. Considerations included giving the chatbot a human-like name (Araujo, 2018), creating a conversational pace to the interview (Nordheim et al., 2019), and maintaining a clear and error-free interaction (Dietvorst et al., 2015; Merritt & Ilgen, 2008; Nordheim et al., 2019) were included, however. However, it is unclear what effects and to what extent each element specifically had on recall and how the combination was received. This could be examined in future studies by utilising variations on the current chatbot format for comparison. Rapport is an integral part of the initial account interview (College of Policing, 2019); it increases cooperation (Alison & Alison, 2017), reduces stress, and increases witness accuracy (Kieckhaefer et al., 2014), so research in this regard is important in the future.

One advantage of ChatCharlie is that it was designed to have small elements, which lend themselves to being altered without affecting other parts of the initial account interview structure. ChatCharlie can thus explore the specific effects of interview interactions within the initial account interview structure. Future research

should be conducted to understand the impact of the ground rules on the protection and accuracy of episodic memories within this format and the rapport-building elements. Introducing different rapport-building elements could be explored in further detail. Not only would this help improve the functionality of ChatCharlie, but it is uniquely positioned to remove the human variables contributing to rapport building in traditional contexts – it may be that rapport building *per se* is unnecessary since witnesses would not expect to build rapport with ChatCharlie (Dando et al., 2023; Dando & Oxburgh, 2016; Nahouli et al., 2021, 2023). Thus, ChatCharlie could potentially provide a more accurate understanding of the broader literature on rapport building.

It would also be useful to explore the specific effects of language usage and question phrasing on memory accuracy using ChatCharlie. ChatCharlie offers a format where interview elements and variables can be controlled. Removing the inevitable variation with human interviewers would allow the identification of preferred and specific language use for initial account interviews.

The design of ChatCharlie included providing clear instructions so that the tool could be used in isolation instead of with another person. Previous research has found that witness performance decreases as the number of other people present increases (Wagstaff et al., 2008). Furthermore, this tool could be translated into the witness's first language, removing the presence of a translator and reducing misunderstandings with language barriers. Memory performance could be attributed to the removal of the presence of others. Future research should compare the use of ChatCharlie alone compared to using ChatCharlie with another person present. This would help further research examining the passive presence of others on episodic memory recall.

5.4.1. Limitations.

This study provides valuable first-step insights into the potential of ChatCharlie as a tool for enhancing witness memory recall accuracy for some witnesses in some contexts. However, several limitations must be acknowledged at this stage. First, the sample size of 90 participants may limit the generalizability of the findings, although power analysis did indicate that this was enough for experiments of this kind. Effect size measures were reported to provide insights into the magnitude of observed effects, mitigating some limitations associated with sample sizes. Efforts were made to recruit a diverse group of participants, but for experimental control, all were adults from the general population with English as their native language. More research is required to extrapolate to broader populations.

These data were collected during the COVID-19 health emergency and access to adults in the older age brackets was challenging for the following reasons: i) older adults may have been less confident in using remote video platforms at that point and communicating remotely in general. Social distancing resulted in a lack of access to younger generations and others who may have been able to assist. This is an explicit limitation, but as the research progressed during this challenging period, it became apparent that this limitation could not be countered. It was deemed unethical to exclude these data. Future research must consider older adults and the impact of access to technology and technology awareness more generally since reduced access is likely to lower confidence, which will negate the efficacy of ChatCharlie moving forward. Indeed, it has been reported that over 3% of people aged 55 to 64 in the UK have never used the internet for personal use or have no access to the internet, which rises to 25% aged 65 years and older (Age UK, 2024). Outside of pandemic conditions, researchers should consider employing digital champions to

travel to assist older adults in participating in research of this nature. This approach was considered here but was truncated by social distancing rules and older adults' susceptibility to COVID-19.

Efforts were made to incorporate ground rules and rapport-building elements into ChatCharlie's design. The specific effects of these components on memory recall accuracy were not examined here. The psychological literature concerning the impact of exclusion and inclusion of these elements sits firmly in the in-person, face-to-face literature. So more research is needed in this regard, maybe isolating each element using remote chatbot platforms. Future research endeavours could benefit from a more granular exploration of the individual contributions of ground rules and rapport-building elements to memory recall accuracy. Researchers could examine their respective effects on witness memory recall by systematically varying and manipulating these components within the ChatCharlie interface.

There is always a trade-off between experimentally controlling extraneous variables such as potential interviewer effects and seeking to maximise ecological validity. Here, given the serious and significant constraints of conducting experimental research during an international public health emergency, combined with the desire to maintain interviewer consistency in this initial proof of concept programme of PhD research, it was decided to keep the interviewer consistent. However, this is a potential limitation, so future research should vary interviewers to improve generalizability. It is not unusual for the interviewer to be kept consistent in the first instance, moving to vary this variable as empirical validation progresses.

Finally, the study focused on memory recall accuracy within a controlled laboratory using the traditional mock witness paradigm, which does not fully capture the complexities of real-world contexts. The benefits are that researchers can control

variables and confounding factors, thereby enhancing the internal validity of the findings and facilitating a systematic investigation into the specific effects of ChatCharlie on memory recall accuracy in the first instance. More research is necessary in conditions that map onto ecologically relevant contexts and across various populations that might encompass stress, anxiety, leading questioning, and social pressure.

5.4.2. Conclusion.

ChatCharlie offers a potential alternative method for conducting initial account interviews for some witnesses in some circumstances. This aligns with the growing interest in using technology to enhance investigative (Ministry of Justice, 2022). The primary findings of this study are very promising in terms of improved quantity and quality of information both at Time 1 and at Time 2, performing at least equally to an In-Person initial account and often better. ChatCharlie's adherence to evidence-based questioning techniques and its impartial nature possibly minimises the risk of memory distortion, a longstanding and enduring challenge in investigative interviewing (Weber & Brewer, 2008). However, more research must be undertaken, and public acceptability must be investigated (see Chapter 6).

The ChatCharlie effect was robust and enduring from Time 1 to Time 2 and across interview phases at Time 2, again highlighting the importance of memory consolidation. In summary, the present study underscores the potential of ChatCharlie as a tool in the investigative process, offering law enforcement agencies a reliable and standardised method for initial witness account retrieval. Future research should aim to further validate ChatCharlie's effectiveness across diverse populations and real-world crime scenarios while also exploring its integration with existing investigative protocols to maximise its use in criminal investigations.

6. Chapter 6: Acceptability of Chatbots: Gathering an Initial Account, Experiences and Perceptions

6.1. Introduction

The timely gathering of initial accounts from witnesses, particularly in the aftermath of mass events, presents a logistical challenge. Traditional face-to-face initial account interviews may not be possible, but gathering an initial account as soon as possible is desirable for investigative and memory consolidation purposes. Previous chapters of this thesis have reported a series of empirical laboratory studies concerned with investigating the potential of ChatCharlie for gathering an initial account from some people, in some circumstances, focusing on the potential cognitive benefits and mapping out the design approach. This chapter reports perceptions and experiences of ChatCharlie and chatbots in general. Participants in the ChatCharlie condition were asked about their chatbot experiences. In contrast, all other participants were asked about their perceptions of chatbots more generally, emphasising chatbots in a criminal justice context. Findings indicate a positive attitude to chatbots in general, with participants expressing comfort and willingness to use this technology for initial account interview purposes, and those who engaged with ChatCharlie demonstrated increased ease in admitting memory gaps during subsequent interviews. However, privacy and security emerged as paramount concerns, influencing platform preferences and information disclosure behaviours. While ChatCharlie may offer cognitive and investigative benefits in some instances by way of prompt and non-judgmental interactions that may improve witness cooperation and information sharing, addressing privacy concerns and ensuring transparent communication regarding data usage is crucial for maintaining public

trust. Further information about this chapter can be found in the following section of the thesis: Section 1.4. Initial Account Interviews, 1.5.2. Confidence, 2.1.3. Modality, 2.2. Chatbots.

6.1.1. Application to Initial Account Interviews.

As outlined in the trust model proposed by Nordheim et al. (2019), the perception and reputation of the service provider constitute crucial environmental factors influencing trust dynamics. In scenarios involving a customer service chatbot, the positively viewed brands inspire confidence in customers, illustrating how brands can effectively establish trust with their clientele. The inverse is also true, with negatively viewed brands creating distrust. Transposing this to the criminal justice system, the notion of brand takes on a different dimension, representing the public's perception of law enforcement agencies. At any given time, societal perception of the police will influence how trustworthy the public finds initial account chatbots. This extends to the individual; trust in a police chatbot will be influenced by social background and prior experiences of police engagement with their community.

6.1.2. The current study.

This chapter explores perceptions and acceptability of chatbots in investigative contexts, including factors influencing user acceptance and trust in chatbots, concerns about privacy and security, interview pacing, trustworthiness, accuracy, handling of sensitive information, and privacy concerns. The research addresses gaps in understanding how chatbots might be accepted or otherwise for initial account interviews within the criminal justice system, thus contributing to a better understanding of chatbot use and the development of chatbot technology in a way that enhances its utility and effectiveness in investigative contexts. The research questions for this study were:

1. What are the perceptions and experiences of using ChatCharlie, a chatbot, for an initial account interview?
2. What is the impact of being interviewed by ChatCharlie on individuals' perspectives of their memory?
3. How might privacy concerns dictate the development and use of a chatbot to collect initial account interviews?

6.2. Method

6.2.1. Design.

This study collected post-interview data from participants who completed a ChatCharlie initial account at Time 1 of the mock witness study reported in Chapter 5 and the control group participants. Participants were randomly assigned to no initial account interview (control), an In-Person initial account or the ChatCharlie condition. One week later, all participants underwent a face-to-face interview. Perception and acceptability data was collected from ChatCharlie participants, focusing on detailed feedback regarding experiences and perceptions of the acceptability of chatbots for initial account interview purposes. More general data was collected from the control participants regarding their general impressions of chatbots. Participants in the In-Person condition did not participate in this study since the in-person experience may have confounded the response to the chatbot survey. As others have reported, perceptions and acceptability of computer-mediated communication, such as chatbots, are typically negative following in-person interactions for reasons linked to a lack of emotion and humanness (Mygland et al., 2021; Rapp et al., 2021).

6.2.2. Participants.

A total of 60 participant datasets are reported in this chapter: 18 male, 41 female and one non-binary/third gender, randomly allocated to the ChatCharlie or Control conditions. The mean age was 30.47 years (SD: 8.52) and ranged from 19 to 54 years. Participants were recruited through online advertising and compensated for the time spent taking part through a £10 shopping voucher. There was no significant difference in age between groups (control: 30.1 years, SD: 9.51; ChatCharlie: 30.83 years, SD: 7.55), $t = 0.33$, df = 58, $p = 0.742$.

6.2.3. Procedure.

As this study was conducted remotely, all recruitment and communication took place over email. All potential participants were first emailed a participant information sheet (see Appendix M and N) and a copy of the consent sheet (see Appendix C). Participants committed to taking part in two elements exactly one week apart. Although all elements were remote, the researcher was on hand to help with any difficulties in real time.

Time 1: All participants received an email with a link to complete the consent form, watch a one-minute video of a non-violent mock crime and complete a short demographic survey (See Appendix D). Once completed, participants in the ChatCharlie condition were linked to a designated website to participate in an initial account interview with ChatCharlie. Finally, these participants were also asked to complete a short online survey on their experiences and perceptions of ChatCharlie. Participants in the control condition had no further tasks at time 1.

Time 2: One week later, all participants took part in a tier 1 type basic PEACE investigative interview. Interviews took place via face-to-face video call and were digitally recorded. Immediately following the interview, participants were asked to

complete a survey regarding their interview experience. Participants who did not participate in the chatbot interview were then provided with further information about chatbots and asked about their general perceptions of this as a potential tool for information gathering during criminal investigations.

Data collection procedures relevant to this chapter commenced after the Time 2 interviews (see Chapter 3 Section 3.4.4. and Appendix I for a full description). They comprised two groups: participants who took part in the Control condition at Time 1 and participants who took part in the ChatCharlie interview at Time 1. They were asked to read this material and reflect on chatbots and their experiences with chatbots in general. They were then asked to complete an acceptability survey (see Appendix Q) detailing their perceptions of chatbots in general and as a potential tool for information gathering during criminal investigations. Chatbot participants were asked to complete comparison surveys. The comparison survey asked participants about their experiences and acceptability of ChatCharlie and their perceptions of chatbots more generally (see Appendix R).

6.2.4. Materials.

6.2.4.1. Ethics materials.

All participants were emailed a participant information sheet before agreeing to participate in any experiments (See Appendix N). All participant information sheets contain information unique to the study and the condition to which the participant was allocated. Participants who did not take part in an interview with a chatbot were not informed about the use of this technology before taking part in the study. All participant information sheets detailed requirements for taking part, what participants would be required to do, ethics information, and contact details. Additionally, all participants completed an online consent form through a Qualtrics link before

beginning the study (See Appendix C). Participant demographic questions were asked across all studies in this body of work (See appendix D) (See section 3.3.4.3.).

All participants received a debrief sheet (See Appendix O). The debrief sheet contained further information about the study's aims and purpose. All participants, regardless of condition, were provided with further information about chatbots and the purpose of the research. A summary of what they had done in the study and additional information about the other conditions was given. A detailed explanation of data anonymisation and withdrawal protocols was given. Information about compensation for taking part and the researcher's contact details were provided.

6.2.4.2. Chatbot post-interview survey questions.

This survey was altered from the original post-interview survey to apply to an interview done by a chatbot rather than a person. This survey was administered to all participants who participated in a ChatCharlie interview. The focus of these questions concerned participants' perception of their memory performance and any previous experience with reporting a crime as a witness. Questions were also asked about how easily participants found remembering and how confident they were in their memory performance (See Appendix Q). However, instead of asking participants about their perceptions of the interviewer, they were asked about their perceptions of being interviewed by a chatbot called ChatCharlie (See Appendix P). For example: ‘How easy did you find it to communicate with ChatCharlie?’.

Additionally, participants were asked questions about their user experience with this technology. For example: ‘How satisfied were you with ChatCharlie?’. Participants were also asked to judge alternative interview modality options (See Appendix P). For example: ‘Which format of messages did you prefer the most?;

Free text, open-ended questions, Yes or No buttons, Scaled 1-5 button questions, A mixture of ‘free text’ and buttons, Not sure.’.

6.2.4.3. Post-interview survey questions: Acceptability.

This survey was designed to collect information about the utilisation of chatbots in the context of a police interview (See Appendix R). The questions addressed concerns surrounding the potential messaging platforms available for chatbot hosting. For example, ‘How comfortable would you be disclosing information to a police chatbot about a real crime you witnessed using one of these platforms?; Facebook Messenger, WhatsApp, Text message, Email, Dedicated website’. Additionally, questions were asked about sharing personal information or information about a crime with a non-human chatbot interviewer. For example: ‘How comfortable would you be disclosing the following information to a chatbot?; Your name, age, email address, telephone number, current level of education, information about any physical disability, and any learning difficulties. These questions aimed to establish if privacy and trust would be barriers to utilising this technology and if experiencing a chatbot interview would increase or decrease these concerns.

6.2.4.4. Post-interview survey questions: Comparison.

All participants in a chatbot interview were asked questions about their memory performance a week later. These questions were designed to establish if participants perceived the use of ChatCharlie positively impacting their memory. It also asked them to compare their experience with both interview modalities and their opinion of utilising a chatbot in this functionality (See Appendix R). For example, a 5-point Likert scale (with a range of definitely agree, somewhat agree, neither agree nor disagree, somewhat disagree, definitely disagree) with the question: ‘I remembered more in the chatbot interview than the second interview.’.

6.3. Results

6.3.1. Analysis.

A series of one-way between-subjects ANOVAs investigated the impact of using a chatbot conducting initial account interviews on participant perception of this technology's use in this functionality. Where Levine's test for homogeneity of Variance revealed the assumption that the homogeneity of variance was violated, a Kruskal-Wallis test was conducted.

6.3.2. Time 2 Memory Perception and Confidence

When asked about memory perception and confidence, there was no significant difference between conditions. Means and SD are in Table 9 below.

Table 9

Means & standard deviations (in parenthesis) Memory perception & confidence.

	Condition	
	ChatCharlie	Control
	Mean (SD)	
Overall, how easy did you find it to remember what happened in the video today?	2.80 (1.06)	3.17 (1.09)
Overall, how difficult did you find it to remember what happened in the video?	2.77 (1.01)	2.63 (1.00)
I am confident that I remembered a lot of what I saw.	2.67 (1.24)	2.57 (1.17)
I am confident I did not make any errors.	2.50 (1.01)	2.20 (0.92)

Overall, how easy did you find it to remember what happened in the video today? (1 – very easy, 2 – somewhat easy, 3 – neither easy nor difficult, 4 – somewhat difficult, 5 – very difficult). $F(1,58) = 1.747$, $p = .191$, $N_p^2 = .029$.

Overall, how difficult did you find it to remember what happened in the video? Not reversed (1 – very difficult, 2 – somewhat difficult, 3 neither easy nor difficult, 4 – somewhat easy, 5 – very easy) $F(1,58) = 0.265$, $p = .609$, $N_p^2 = .005$.

I am confident that I remembered a lot of what I saw (1 – not at all confident, 2 - somewhat confident, 3 – undecided, 4 – quite confident, 5 – completely confident) $F(1,58) = 0.104$, $p = .749$, $N_p^2 = .002$.

I am confident I did not make any errors (1 – not at all confident, 2 - somewhat confident, 3 – undecided, 4 – quite confident, 5 – completely confident) $F(1,58) = 1.442$, $p = .235$, $N_p^2 = .024$.

6.3.3. Time 2 interviewer questions.

There was a significant difference between the conditions when the interviewer communication was questioned. Means and SD are in Table 10 below.

Table 10.

Means & standard deviations (in parenthesis) for Interviewer questions

	Condition	
	ChatCharlie	Control
	Mean (SD)	
How easy was it to tell the interviewer that you did not know the answer to a question?	1.37 (0.76)	2.23 (1.25)
How difficult was it to tell the interviewer when you could not remember?	4.50 (0.73)	3.70 (1.15)
How difficult was it to tell the interviewer that you did not know the answer to a question?	4.53 (0.73)	3.73 (1.17)

How easy was it to tell the interviewer that you did not know the answer to a question? (1 – very easy, 2 – somewhat easy, 3 – neither easy nor difficult, 4 – somewhat difficult, 5 – very difficult) $F(1,58) = 1.483$, $p = .002$, $N_p^2 = .153$. Levine's

test was significant $F(1,58) = 13.061$, $p < .001$. A Kruskal-Wallis test revealed a significant effect of recall modality on overall confabulation $\chi2(1, N = 30) = 9.273$, $p = .002$. Participants in the ChatCharlie condition found it easier to tell the interviewer at time 2 that they did not know the answer to a question.

How difficult was it to tell the interviewer when you could not remember? (1 – very difficult, 2 – somewhat difficult, 3 neither easy nor difficult, 4 – somewhat easy, 5 – very easy) $F(1,58) = 10.349$, $p = .002$, $N_p^2 = .151$. Levine's test was significant $F(1,58) = 8.855$, $p = .004$. A Kruskal-Wallis test revealed a significant effect of recall modality on overall confabulation $\chi2(1, N = 30) = 8.285$, $p = .004$. Again, participants in the ChatCharlie condition found it easier to tell the interviewer at time 2 that they could not remember.

How difficult was it to tell the interviewer that you did not know the answer to a question? Not reversed (1 – very difficult, 2 – somewhat difficult, 3 neither easy nor difficult, 4 – somewhat easy, 5 – very easy) $F(1,58) = 10.063$, $p = .002$, $N_p^2 = .148$. Levine's test was significant $F(1,58) = 8.700$, $p = .005$. A Kruskal-Wallis test revealed a significant effect of recall modality on overall confabulation $\chi2(1, N = 30) = 8.265$, $p = .004$. Participants in the ChatCharlie condition found it easier to tell the interviewer at time 2 that they did not know the answer to a question.

6.3.4. Chatbot acceptability.

There was a significant difference between the conditions with the following questions regarding the chatbot's acceptability. The questions used a 5-point Likert scale (1 - Very comfortable, 2 - Somewhat comfortable, 3 - Neither comfortable or uncomfortable, 4 - Somewhat uncomfortable, 5 - Very uncomfortable). Means and SD are shown in Table 11 below.

Table 11

Means & standard deviations (in parenthesis) for Chatbot Acceptability.

	Condition	
	ChatCharlie	Control
	Mean (SD)	
How comfortable would you be typing with a chatbot using a computer to quickly provide some initial basic information about a crime event before being interviewed face-to-face?	1.53 (0.90)	2.07 (1.11)
How comfortable would you be speaking to a chatbot about a crime but with no option to speak to a person later?	3.17 (1.60)	4.13 (0.82)
How comfortable would you be texting with a chatbot using your phone to quickly provide some initial basic information about a crime to the police before being interviewed face-to-face.	1.70 (1.09)	2.20 (1.13)

How comfortable would you be typing with a chatbot using a computer to quickly provide some initial basic information about a crime event before being interviewed face-to-face, for example? $F(1,58) = 4.171$, $p = .046$, $N_p^2 = .067$. Here, Levine's test was non-significant $F(1,58) = 0.534$, $p = .468$. Participants in the ChatCharlie condition were more comfortable with this scenario.

How comfortable would you be speaking to a chatbot about a crime but with no option to speak to a person later? $F(1,58) = 8.682$, $p = .005$, $N_p^2 = .130$. Levine's test was significant $F(1,58) = 25.842$, $p < .001$. A Kruskal-Wallis test revealed a significant effect of recall modality on overall confabulation $\chi 2(1, N = 30) = 4.788$, $p = .029$. Again, participants in the ChatCharlie condition were more comfortable with this scenario.

The following questions about comfort using a chatbot were non-significant: How comfortable would you be texting with a chatbot using your phone to quickly provide some initial basic information about a crime to the police before being interviewed face-to-face, for example? $F(1,58) = 3.059$, $p = .086$, $N_p^2 = .050$.

6.3.5. Chatbot platform.

Questions were asked about participants' comfort in disclosing information to a police chatbot using different modalities (Facebook Messenger, WhatsApp, Text message, Email, Dedicated website). The questions used a 5-point Likert scale (1 - Very comfortable, 2 - Somewhat comfortable, 3 - Neither comfortable or uncomfortable, 4 - Somewhat uncomfortable, 5 - Very uncomfortable). When asked about different chatbot modalities, there was no significant difference between conditions. Means and SD are in Table 12 below.

Table 12

Means & standard deviations (in parenthesis) for Chatbot platform preference

How comfortable would you be disclosing information to a police chatbot about a real crime you had witnessed that used one of these platforms?	Condition	
	ChatCharlie	Control
	Mean (SD)	
Facebook Messenger	3.57 (1.52)	3.67 (1.29)
WhatsApp	2.77 (1.55)	3.27 (1.31)
Text message	2.30 (1.44)	2.73 (1.20)
Email	2.17 (1.31)	2.57 (1.28)
Dedicated Website	1.63 (1.13)	2.07 (1.34)

How comfortable would you be disclosing information to a police chatbot about a real crime you had witnessed that used one of these platforms? - Facebook Messenger, $F(1,58) = 0.075$, $p = .785$, $N_p^2 = .001$; WhatsApp, $F(1,58) = 1.824$, $p =$

.182, N_p^2 = .030; Text, $F(1,58) = 1.599$, $p = .211$, N_p^2 = .027; Email ($F(1,58) = 1.427$, $p = .237$, N_p^2 = .024; Dedicated website e.g.(https://www.eviebot.com/en/), $F(1,58) = 1.839$, $p = .180$, N_p^2 = .031. Overall, participants preferred a Dedicated Website over other platforms, including social media platforms, with means and SD shown in Table 13 below.

Table 13

Overall Means & standard deviations (in parenthesis) Chatbot platform preference.

	Mean (SD)
Facebook messenger	3.62 (1.40)
WhatsApp	3.02 (1.44)
Text message	2.52 (1.33)
Email	2.37 (1.30)
Dedicated Website	1.85 (1.25)

6.3.6. Chatbot personal information disclosure.

Questions were asked about participants' comfort in disclosing personal information to a police chatbot (Name, age, email, telephone, level of education, physical disabilities, learning difficulties). The questions used a 5-point Likert scale (1 - Very comfortable, 2 - Somewhat comfortable, 3 - Neither comfortable or uncomfortable, 4 - Somewhat uncomfortable, 5 - Very uncomfortable). Means and SD are shown in Table 14 below.

Table 14

Means & standard deviations (in parenthesis) for personal disclosure questions.

How comfortable would you be disclosing the following information to a chatbot?	Condition	
	ChatCharlie	Control
	Mean (SD)	
Name	2.37 (1.33)	2.33 (1.21)
Age	2.13 (1.28)	2.23 (1.19)
Email	2.53 (1.43)	2.50 (1.22)
Telephone	2.63 (1.37)	2.87 (1.41)
Level of Education	2.00 (1.11)	2.47 (1.19)
Physical disabilities	2.57 (1.28)	2.87 (1.43)
Learning difficulties	2.60 (1.38)	2.76 (1.43)

There was no difference between groups regarding disclosure of personal information to a chatbot, all *F*s < 2.446, all *p*s > .040:

The overall means for personal disclosure are shown in Table 15 below.

Table 15

Overall Means & standard deviations (in parenthesis) for personal disclosure questions.

	Mean (SD)
Name	2.35 (1.26)
Age	2.18 (1.23)
Email	2.52 (1.32)
Telephone	2.75 (1.38)
Level of Education	2.23 (1.17)
physical disabilities	2.72 (1.35)
learning difficulties	2.68 (1.39)

6.3.7. Chatbot in Initial Account Interview functionality.

Questions were asked about participants' comfort in disclosing information about different aspects of a crime to a police chatbot. The questions used a 5-point Likert scale (1 - Very comfortable, 2 - Somewhat comfortable, 3 - Neither comfortable or uncomfortable, 4 - Somewhat uncomfortable, 5 - Very uncomfortable)

There was no difference between groups when disclosing information about a crime to a chatbot, all *F*s < 3.384, all *p*s >. 031. Means and SD are shown in Table 16 below.

Table 16

Means & standard deviations (in parenthesis) for Initial Account Interview functionality questions.

How comfortable would you be disclosing basic information about a real crime you had just witnessed to a chatbot, that would be quickly accessible to police?	Condition	
	ChatCharlie	Control
	Mean (SD)	
A description of what happened.	1.50 (0.97)	1.90 (1.09)
Information about crime location.	1.37 (0.81)	1.70 (1.05)
Information about the people involved in the crime.	1.63 (0.96)	2.17 (1.26)
How confident you were in your memories of the crime.	1.87 (1.07)	2.40 (1.40)
If you were under the influence of alcohol or drugs when you witnessed the crime.	2.50 (1.43)	2.93 (1.41)

The overall means (SD) for disclosing a crime to a chatbot are shown in Table 17 below.

Table 17

Overall Means & standard deviations (in parenthesis) for Initial Account Interview functionality questions.

How comfortable would you be disclosing basic information about a real crime you had just witnessed to a chatbot that would be quickly accessible to police?	Mean (SD)
A description of what happened	1.70 (1.05)
Information about crime location	1.53 (0.95)
Information about the people involved in the crime	1.90 (1.14)
How confident you were in your memories of the crime	2.13 (1.27)
If you were under the influence of alcohol or drugs when you witnessed the crime	2.72 (1.43)

6.4. Discussion

Exploring alternative methods for collecting initial accounts from witnesses, such as utilising chatbots, must include understanding perceptions and experiences, potentially hindering efficacy. Positive perceptions of chatbots indicate a willingness to embrace technology to support law enforcement agencies in their investigative efforts, and experiential feedback is essential for maximising cognitive utility. Understanding the factors influencing the acceptability of chatbots and their impact on witness memory can inform the development of additional interviewing techniques, particularly when the process is prescriptive, such as initial accounts. These results reported in this chapter will be further examined alongside existing research on the acceptability of chatbots and the reliability of witness memory.

6.4.1. Time 2 confidence questions.

There was no significant difference between the two conditions when asked about memory perception and confidence. However, Chapter 5 demonstrated that participants who had an interview with ChatCharlie recalled more correct information

than participants in the Control condition and performed equally as well as the In-Person condition. This implies that participants did not think an additional and immediate recall opportunity with a chatbot increased their memory recall, irrespective of their actual performance. This aligns with the research indicating confidence is not an indicator of memory accuracy (Flindall et al., 2016; Henkel, 2017; Wells, 1995; Wells & Loftus, 2003). While participants did not perceive a boost in memory recall from the ChatCharlie interaction, recall quality and quantity were markedly improved (See Chapter 5).

Despite not reporting increased confidence in their memories, participants who had an interview with ChatCharlie at time 1 found it easier to tell the interviewer at time 2 that they did not know the answer to a question and could not remember. This implies that ChatCharlie increased confidence and certainty in understanding when recall was incomplete or limited and feeling empowered to say so. One plausible interpretation is that the initial interview with ChatCharlie preserved memories and fortified trust in their cognitive recollections. This phenomenon could be linked to the provision of a non-judgmental initial account interview associated with the absence of a human (Dando et al., 2023; Taylor & Dando, 2018). Historically, computers have proven exceptionally adept at soliciting sensitive information, a trait well-documented in research (Lind et al., 2013; Weisband & Kiesler, 1996). This effectiveness is often attributed to reduced social inhibitions, particularly those stemming from the fear of being judged (Joinson et al., 2007). This diminished fear of judgment may have persisted into the subsequent interview, fostering an environment where participants felt more at ease admitting gaps in their memory to the interviewer (i.e. a carryover effect).

6.4.2. Chatbot acceptability.

Participants interviewed with ChatCharlie were more comfortable typing with a chatbot using a computer to quickly provide some initial basic information about a crime event before being interviewed face-to-face. This demonstrates that positive experiences have a positive impact on perceptions, maybe. This has implications surrounding previous chatbot use, with reports suggesting that one-third of people have interacted with chatbots unknowingly (Elsner, 2017). Making people more aware of using chatbots may support improved favourable perceptions. However, it is sensible to suggest that a negative experience with a chatbot will bias perceptions of the capability of chatbots, which may be enduring meaningful information (Nadarzynski et al., 2021). More concerningly, negative chatbot experiences may alter the perception of the company or organisation the chatbot represents (Araujo, 2018).

Positively, both groups felt comfortable with the idea of using a chatbot in this modality, with the ChatCharlie group reporting they felt very comfortable with this option, compared to the control group, who felt somewhat comfortable. Furthermore, participants who had experienced ChatCharlie felt more favourably about speaking to a chatbot about a crime but with the option to speak to someone later. The ChatCharlie group rated this option neutrally compared to people who had not used ChatCharlie and were somewhat uncomfortable with this option. This has implications surrounding the modality of the chatbot and previous experiences. There was no difference between groups when asked how comfortable they would be texting with a chatbot using their phone to quickly provide some initial basic information about a crime to the police before being interviewed face-to-face. Both groups feel somewhat comfortable with this option.

6.4.3. Chatbot platform.

Regardless of conditions, participants preferred disclosing information to a police chatbot through a dedicated website, with both groups reporting being very/somewhat comfortable with this platform option. Participants in both groups felt somewhat comfortable disclosing information to a police chatbot through email or text messages, likely stemming from repeated experiences with both methods. Participants felt more neutral about using WhatsApp to disclose information to a police chatbot. Participants felt less comfortable using Facebook Messenger to disclose information, reflecting potential privacy concerns when reporting a crime.

The discomfort surrounding the Facebook Messenger modality warrants attention, particularly in light of the robust initial adoption of chatbots within the platform (Johnson, 2017). This disparity requires consideration and might indicate broader societal apprehensions regarding privacy and security, topics prevalent in discussions surrounding trust in social media platforms (Office for National Statistics, 2020). Indeed, the perceived risks associated with divulging personal information can significantly erode trust and inhibit disclosure (Corritore et al., 2003). The unease surrounding sharing sensitive details is further compounded by anxieties about potential social judgment and, more critically, fears of potential repercussions stemming from such disclosures (Tourangeau & Yan, 2007). These concerns surfaced in the present study, wherein participants preferred platforms that offered more significant privacy safeguards.

In contemplating the potential application of an initial account interview chatbot within the criminal justice system, ensuring and communicating information security emerges as a paramount concern. Such assurances could be pivotal in cultivating trust, particularly given the inherent challenges in consistently guaranteeing

anonymity in digital interactions. As such, robust measures to safeguard user data and ensure confidentiality become imperative considerations in deploying such technology effectively.

6.4.4. Chatbot personal information disclosure.

Across the various experimental conditions, no discernible differences emerged in the extent to which participants were willing to divulge information to a chatbot about personal details. Individuals from both groups exhibited a comparable level of comfort when sharing details. This parity in disclosure is encouraging, as it underscores a general openness among individuals to explore novel and more convenient avenues for communicating such personal information, especially with law enforcement agencies. The willingness displayed by the public is reminiscent of trends observed in healthcare settings, where chatbots have been leveraged effectively to share pertinent information swiftly, offering a viable alternative to traditional human support systems (Nadarzynski et al., 2021; Nordheim et al., 2019).

6.4.5. Chatbot in Initial Account Interview functionality.

There was no difference between conditions when participants were asked about disclosing information about a crime to a chatbot. Interestingly, participants felt very/somewhat comfortable revealing a description of what, the location of the crime, and information about the people involved, but only somewhat comfortable reporting their memory confidence. Participants felt fairly neutral about telling a chatbot if they were under the influence of any substances. This may reflect the trend of people feeling more comfortable giving information that does not directly relate to themselves. Moreover, it seems linked to concerns about privacy and security (Office for National Statistics, 2020), with personal disclosures being seen as less favourable. It is possible this could be combatted with reassurances about

information use. Chatbots have been used to collect health information, demonstrating that the technology can be successfully implemented when these privacy elements are addressed (Nadarzynski et al., 2021; Nordheim et al., 2019).

Chatbots have emerged as a promising substitute for conventional human-mediated online communication channels with the capacity to simulate human conversational dynamics, including prompt responses and natural pacing (Brandtzaeg & Følstad, 2017). This resonates strongly with established guidelines on conducting initial account interviews, which underscore the importance of interview pace in fostering effective communication dynamics (College of Policing, 2019). This alignment between the attributes of chatbot-mediated interactions and the qualities for successful interview processes underscores the potential efficacy of this technological modality within investigative contexts. The positive reception observed among participants not only shows their willingness to embrace technological advancement but also suggests a recognition of the practical advantages that chatbots offer, including streamlining interviewing logistics and preserving the integrity of recollections.

From a theoretical standpoint, the favourable reception of chatbot-mediated initial account interviews underscores broader shifts in societal attitudes towards human-computer interfaces. As perceptions surrounding the capabilities and limitations of automated conversational agents continue to evolve, stakeholders within the criminal justice system are motivated to capitalise on the advantages afforded by emerging technologies. This necessitates a nuanced exploration of the interplay between technological innovation, procedural efficacy, and user acceptance.

6.4.6. Limitations.

While this novel data sheds light on the potential of chatbots as tools for conducting Initial Account Interviews in investigative contexts, several limitations should be acknowledged. The study recruited participants through online advertising, which may have introduced selection bias. The sample primarily consisted of individuals comfortable with remote communication methods, potentially limiting the generalizability of the findings to the broader population. Further, all participants were adults, and the event in question was intentionally encoded, not traumatic nor real, which can impact memory and, most likely, the usability of ChatCharlie.

The study employed a between-subjects design, randomly assigning participants to either a control group or a group that interacted with ChatCharlie. While this design allows for group comparisons, it may overlook individual differences that could influence responses. For example, the study did not account for factors like prior experience with chatbots or technology literacy. However, randomisation helps mitigate the influence of individual differences by evenly distributing known and unknown variables across groups, potentially limiting confounding effects. Additionally, while individual differences may exist in participants' prior experience with chatbots or technology literacy, the study's focus on perceptions and acceptability of chatbots in a criminal justice context may mitigate the impact of these factors. Attitudes towards chatbots in law enforcement are likely influenced by broader societal perceptions and experiences rather than individual technological proficiency alone.

The study highlighted a preference for platforms that prioritise privacy when interacting with chatbots. However, the study did not directly assess concerns about data security or their trust in the confidentiality of chatbot interactions. However,

findings revealed discomfort with disclosing sensitive information via specific platforms, such as Facebook Messenger, suggesting a level of apprehension regarding data security. This discomfort likely stems from broader societal concerns about privacy and data security on social media platforms, as evidenced by existing research (ONS, 2020). Participants implicitly express their concerns about data security in chatbot interactions by preferring platforms with more significant privacy safeguards.

A focus on acceptability and perceptions of chatbots within a law enforcement context inherently encompasses privacy and data security considerations. Participants' willingness to engage with chatbots for initial account interviews reflects their trust in the confidentiality of these interactions, as such interviews involve disclosing potentially sensitive information related to criminal incidents. Thus, while the study did not directly measure participants' trust in chatbot confidentiality, their acceptance of chatbots for law enforcement purposes suggests confidence in the security measures. Regardless, future research should employ measures to evaluate participants' perceptions of privacy and data security when engaging with chatbots in law enforcement contexts.

Participants in the study demonstrated a willingness to disclose personal information to chatbots. However, the study did not explore the understanding of how chatbot systems would use or store their personal information, which may impact their decision to share sensitive details in real life, and willingness to disclose personal information in research may not fully reflect their behaviour in real-world interactions with law enforcement chatbots. Acknowledging the complexity surrounding participants' understanding of how their data would be used or stored

and the potential disparities between research settings and real-world interactions with law enforcement chatbots (Nadarzynski et al., 2021; Nordheim et al., 2019).

6.4.7. Conclusion.

The research on chatbots for initial account interviewing presents significant insights into the potential of technology to address challenges inherent in traditional investigative methods. Historically reliant on face-to-face interactions, witness interviews face logistical constraints. Hence, the findings support the potential of chatbots for information gathering, offering a viable alternative to traditional face-to-face interviews in some circumstances. This conclusion aligns with existing research highlighting the limitations of conventional interviewing methods, particularly in scenarios involving mass incidents (Dando et al., 2023; Howe & Knott, 2015; Taylor & Dando, 2018; Vredeveldt et al., 2011). The positive reception of chatbots suggests general acceptability and aligns with previous research findings that appropriately designed chatbots are generally received positively (Brandtzaeg & Følstad, 2017). This acceptance suggests a shift in societal attitudes towards integrating digital solutions in investigative procedures, echoing broader trends in technology adoption across various domains (Johnson, 2017).

Moreover, the study revealed participants' preference for platforms prioritising privacy in interactions with chatbots. This aligns with existing literature highlighting the significance of privacy and data security in fostering user trust and technology acceptance (Corritore et al., 2003; ONS, 2020). The emphasis on privacy underscores the importance of transparent communication and robust measures to safeguard user data, especially in sensitive contexts like law enforcement. Enhancing privacy and anonymity has been associated with increased willingness by participants to engage with sensitive questions (Joinson et al., 2007). This aligns

with research demonstrating the successful implementation of chatbots for data collection in healthcare settings (Nordheim et al., 2019; Nadarzynski et al., 2021).

In light of these findings, there are several potential implications for law enforcement agencies. Integrating chatbots into initial witness interviews could enhance the efficiency and effectiveness of information-gathering processes, particularly in time-sensitive situations. However, agencies must prioritise privacy, ensure transparent communication, and address user concerns surrounding data security to maintain public trust and acceptance of chatbot technology (College of Policing, 2019).

Overall, the findings demonstrate the potential of chatbots to be practical tools in conducting initial account interviews. Chatbots offer prompt and non-judgmental interactions that may enhance witness cooperation and information sharing. However, it is imperative to address privacy concerns and ensure transparent communication regarding data usage to maintain public trust. By leveraging the benefits of chatbot technology while addressing privacy considerations, law enforcement agencies might enhance their investigative capabilities and improve timely information gathering.

7. Chapter 7: Discussion

7.1. Aims of the programme of PhD research

The research programme reported in this thesis investigates the efficacy of a novel approach for gathering initial accounts from witnesses via a chatbot called ChatCharlie, especially when in-person interviews are impossible. This alternative interview method aimed to align with established interviewing standards and insights from psychological research on human-computer interaction.

A novel tool was created to address existing constraints, capturing initial witness testimonies following criminal incidents for certain individuals and contexts. This solution shows promise in enhancing the efficiency and accuracy of information gathering during early investigations, helping to consolidate memory. The thesis also critically appraises the feasibility of integrating chatbots into practice, contributing to the emerging remote interviewing literature.

The research questions addressed in this thesis were:

1. Concerning episodic memory performance, is a remote-typed interview an effective method for collecting investigative information from crime witnesses versus a remote-spoken interview?

2. Is a remote chatbot interviewer (typed) an effective method for collecting initial account information from witnesses and victims of crime across a wide age range?

3. What are the potential benefits and challenges of integrating chatbots into initial account interviews, and how do they compare to traditional face-to-face interviews?

4. What are the perceptions and experiences of witnesses regarding using chatbots and other technology-assisted interviewing methods in the criminal justice system?

What follows is a general summary of each of the studies.

Study 1: Episodic Memory Performance: A Comparison of Typed and Spoken Modalities

Study 1 explored differences in witness recall between typed versus spoken interviews. The aim was to see if typing, without verbal cues, could be as effective as spoken interviews in the accuracy and completeness of information. Results revealed no difference between conditions for memory items (correct, incorrect and percentage accuracy), apart from confabulations. The non-significant differences were consistent globally and as a function of the recall phase. This study is one step towards bridging gaps in the literature, thereby contributing to understanding. This study answers the following research questions:

- Does typed retrieval of episodic information impact the quantity of event information recalled versus traditional face-to-face spoken retrieval, as measured by the amount of information recalled and its accuracy?
- Does typed retrieval of episodic information impact the quality of event information recalled versus traditional face-to-face spoken retrieval as measured by the number of errors and confabulations?

Study 2: Typed Recall of Eyewitness Memory Performance as a Function of Age

Study 2 examined the episodic recall across three age groups when typing. Results revealed that older participants (55 +) recalled significantly more correct items than younger groups despite lower confidence in technology use, both in the interview's free recall and cued recall phases. These findings suggest that text-based retrieval may have the potential for promptly and accurately capturing witness

accounts in some instances and that older witnesses may also benefit. These results answer the following research questions:

- Does age impact the quantity and quality of eyewitness memory elicited via a typed modality?
- Is there a relationship between age, familiarity with technology, comfort, and the quality and quantity of event information recalled via a typed modality?

Study 3: ChatCharlie for Gathering Initial Accounts from Eyewitnesses: Development and Evaluation.

This study examined the efficacy of ChatCharlie for memory consolidation using an initial account interview. Participants were randomly assigned to ChatCharlie, In-person, or no initial account condition. At Time 1, ChatCharlie had fewer confabulations, with no significant difference in correct or incorrect recall between ChatCharlie and in-person conditions. At Time 2, ChatCharlie and In-person conditions were more accurate and recalled more correct items than the Control group without an increase in errors and confabulations. There were no significant differences between ChatCharlie and the In-Person interviews. This study answers the following research questions:

- Is ChatCharlie an effective method for collecting an initial account to consolidate a memory trace for improved recall during later, more in-depth interviews?
- How does ChatCharlie compare at Time 1 to the in-person initial account interview regarding the quantity (amount) and quality (errors) of information elicited?

- Does the Time 1 retrieval condition impact self-reported confidence and perceived accuracy at Time 2?

Study 4: Acceptability of Chatbots: Gathering an Initial Account, Experiences and Perceptions

This chapter reports perceptions and experiences of ChatCharlie and chatbots. ChatCharlie participants were asked about their chatbot experiences. All other participants were asked about chatbots more generally, emphasising a criminal justice context. Findings indicate a positive attitude toward chatbots in general, with comfort and willingness to use chatbots for initial account interviews. Those who engaged with ChatCharlie had increased ease in admitting memory gaps later. Privacy and security concerns influenced platform preferences and information sharing. This study answers the following research questions:

- What are the perceptions and experiences of using ChatCharlie, a chatbot, for an initial account interview?
- What is the impact of being interviewed by ChatCharlie on individuals' perspectives of their memory?
- How might privacy concerns dictate the development and use of a chatbot to collect initial account interviews?

7.2. How each of the research questions has been answered

7.2.1. Research question 1.

Concerning episodic memory performance, is a remote-typed interview an effective method for collecting investigative information from crime witnesses versus a remote-spoken interview?

Study 1 (See Chapter 3) explores using typed recall to collect witness information after an event. Findings indicate that text-based recall does not harm memory performance but leads to a slight increase in confabulations. Although not surpassing spoken interviews, the text-based condition proves to be comparably effective. The rise in confabulations may be due to reduced vigilance in typed responses, similar to online inhibition (Hollenbaugh & Everett, 2013; Suler, 2004). The absence of a physical presence during typed interviews may further contribute to the observed effects, with heightened arousal levels potentially complicating the recall process (Wagstaff, 2008).

Additionally, lower perceived rapport in typed interviews may contribute to increased confabulation. Research suggests that participants perceive lower levels of rapport in typed interviews compared to face-to-face interactions (Hoogesteyn et al., 2023). This finding carries interest as rapport building has been demonstrated to mitigate the impact of misinformation (Vallano & Compo, 2011) and enhance the accuracy of recall (Dando et al., 2023; Gabbert et al., 2021; Nahouli et al., 2021).

However, overall accuracy percentages were unaffected, consistent with existing literature on witness recall when interviews align with cognitive processes (Dando et al., 2009, 2011; Dodson et al., 2015; Holliday et al., 2012; Memon et al., 2010, 2018; Paulo et al., 2016; A. M. Wright & Holliday, 2007). This thesis answers the research question and finds that a remote-typed interview is an effective method for collecting investigative information from witnesses of a crime compared to a remote spoken interview.

7.2.2. Research question 2.

Is a remote chatbot interviewer (typed) an effective method for collecting initial account information from witnesses and victims of crime across a wide age range?

Study 1 (See Chapter 3) found typed responses did not harm memory performance, though they increased confabulations. These results underscore the potential of employing typed modalities to elicit investigative details from witnesses. Although they do not surpass spoken interviews, the text-based condition proves to be comparably effective, making it suitable for chatbot interviews.

Study 2 (See Chapter 4) further explored age effects. Participants over 55 recalled over 30% more correct items than those aged 18-29 and performed similarly to the mid-age group. Notably, participants over 55 also had no differences in the number of confabulated or incorrect items. This suggests that a text-based chatbot could be beneficial for collecting witness information from witnesses of all ages. However further research should be conducted examining the use of the text-based chatbot with this age group.

The reason for this finding could be attributed to the absence of a physical interviewer, which likely reduced perceived demand effects (Gurney, 2011; Gurney et al., 2016; Gurney, 2015), particularly for younger participants who showed less vigilance, akin to online inhibition (Hollenbaugh & Everett, 2013; Suler, 2004). Older adults, often affected by negative interviewer feedback (Dando, 2013; Iida et al., 2020; Leippe et al., 2006), may benefit from reduced social cues, leading to improved recall (Henkel, 2014; Pachman & Ke, 2012). Text-based guidance supports older adults by minimizing distractions, aiding accurate recall (Henkel, 2014; Dando et al., 2020; Leippe et al., 2006; Roediger & Geraci, 2007).

Study 3 speaks to the Chatbot element of this research question (See Chapter 5). Participants who provided an initial account via a chatbot or in-person recalled 47% and 38% (respectively) more correct information at Time 2 than those in the control condition. Furthermore, this improvement came without significant increase in errors

or confabulations. While the differences between ChatCharlie and In-Person conditions were non statistically significant, ChatCharlie showed a notable 10% higher correct recall. This may be due to strict adherence to best practices, avoiding inappropriate questions (Paulo et al., 2021), and reducing social demands (Dando et al., 2023; Taylor & Dando, 2018).

7.2.3. Research question 3.

What are the potential benefits and challenges of integrating chatbots into initial account interviews, and how do they compare to traditional face-to-face interviews?

The potential benefits of integrating chatbots into initial account interviews can be seen in Study 3 (See Chapter 5). Participants using ChatCharlie or in-person modalities demonstrated a 47% and 38% (respectively) increase in recall of accurate details at Time 2 compared to the control group, without a significant rise in errors or false recollections.

This improvement may stem from ChatCharlie's adherence to best practices for initial interviews (College of Policing, 2019), including clear ground rules to reduce errors and improve accuracy (Blank & Launay, 2014; Goldsmith et al., 2005; Evans & Fisher, 2011; Scoboria & Fisico, 2013). Unlike human interviewers, the chatbot consistently applied these guidelines, though it lacked flexibility.

As discussed in Chapter 5, ChatCharlie's structured, conversational style, paired with human-like features (Araujo, 2018), controlled pacing (Nordheim et al., 2019), and error reduction (Dietvorst et al., 2015; Merritt & Ilgen, 2008), may explain its effectiveness compared to in-person interviews. Additionally, utilising ChatCharlie to Introduce different rapport-building elements could be explored in further detail. Not only would this help improve the functionality of ChatCharlie, but it is uniquely positioned to remove the human variables contributing to rapport building in

traditional contexts. ChatCharlie offers a format where interview elements and variables can be controlled but at the expense of flexibility, which is a significant limitation for practice. Removing the inevitable variation that occurs with human interviewers in the field would allow the identification of preferred and specific language use for initial account interviews. Findings from Study 1 support this; using a text-based format did not harm recall accuracy, though confabulations increased. This issue was not seen with ChatCharlie, suggesting the chatbot's design mitigated these confounds. ChatCharlie also provided advantages by offering clear instructions and the possibility of independent use, reducing performance issues linked to the presence of others (Wagstaff et al., 2008). Moreover, this tool can be translated into the witness's native language, eliminating the need for a translator and reducing potential misunderstandings due to language barriers.

As shown in Study 2, the use of a text-based interview format could provide advantages and protective measures for different age groups, previously thought to be vulnerable, both for memory recall (Dodson & Krueger, 2006; List, 1986; Mello & Fisher, 1996; Toglia et al., 2017), and the use of technology (M. T. Harris et al., 2022). Of note is the potential for this tool to use a spoken modality, which could be more appropriate for some witnesses who might be unable to type their responses.

The potential challenges associated with integrating chatbots into initial account interviews are far more logistical. Witnesses would need access to technology to enable the integration of chatbots in initial account interviews. However, reflecting the growing use of technology, particularly since the COVID-19 pandemic, accelerated this trend, prompting individuals of various age groups, including older adults, to engage more frequently with technology. Consequently, while a bias might

exist favouring those with better technology access, this logistical difficulty may not be as significant as before the global health crisis.

Although ChatCharlie shows promise for its inclusionary abilities, further research should examine the exact impact its use would have on vulnerable witnesses, such as the impact of using a spoken chatbot to support those who cannot type. While the improvement in memory performance could be associated with the absence of others (See Study 1 and Study 3), future research could compare the efficacy of using ChatCharlie alone versus in the presence of another person.

Overall, this research addresses the question: ‘What are the potential benefits and challenges of integrating chatbots into initial interviews, and how do they compare to face-to-face interviews?’ Further investigation, including stakeholder perspectives, such as those of police officers, is needed to fully understand the practical implications.

7.2.4. Research question 4.

What are the perceptions and experiences of witnesses regarding using chatbots and other technology-assisted interviewing methods in the criminal justice system?

Study 4 (Chapter 6) explored witnesses' perceptions of chatbots and technology-assisted interviews. Participants in in-person interviews reported higher confidence in their memory performance at Time 2 but were less confident about the correctness of their recalled information. Despite this, they performed equally well as those in the ChatCharlie condition for correct recall, though they confabulated more (See Chapters 5 and 6).

Confidence was not a predictor of percentage accuracy, in line with research raising concerns about using confidence as an indicator for investigative purposes (Caso et al., 2024). This discrepancy raises concerns about the reliability of

confidence as a standalone indicator, echoing findings in memory research (Sauer et al., 2010; Sporer et al., 1995). In human interactions, implicit or explicit feedback can influence confidence without impacting recall accuracy. ChatCharlie's controlled interface lacks such feedback, potentially stabilising confidence assessments.

In Chapter 6 participants who had not used the chatbot reported lower comfort, whereas participants who had used the chatbot reported feeling more comfortable using the technology. This finding supports the integration of new technologies regardless of potential apprehension (Elsner, 2017). Negative experiences can, however, alter these perceptions (Nadarzynsky et al., 2021; Araujo, 2018).

Interestingly, participants from Studies 3 and 4, regardless of condition, preferred using a chatbot in text modality rather than spoken modality. This allignes with Study 1, findings on the benefits of text-based recall, linked to online inhibition (Hollenbaugh & Everett, 2013; Suler, 2004) and the absence of a physical interviewer (Wagstaff, 2008). It could be that typing is becoming more common as a form of communication and, thus, a preferred modality for sharing memory information. This is illustrated by the favourable reaction of participants in both conditions (Study 4) when asked how comfortable they would be communicating with a chatbot using their phone to quickly provide some initial basic information about a crime to the police before being interviewed face-to-face.

Participants viewed a dedicated website for the chatbot positively, similar to email or text messages, but were wary of platforms like WhatsApp or Facebook Messenger. Ensuring information security is critical for deploying such technology (Tourangeau & Yan, 2007). Study 4 participants showed a preference for platforms with strong privacy safeguards, feeling comfortable sharing information with a text chatbot on a dedicated website. These concerns are also highlighted in Studies 1

and 2, which discuss the impact of online communication on inhibitions and arousal related to the presence of others. Encouragingly, in Study 4, participants felt comfortable sharing personal information with a text chatbot with its own dedicated website.

Speaking to the specific use of a chatbot in the criminal justice system, participants in study 4 reported being very/somewhat comfortable disclosing details of a crime. The positive reception of chatbot initial account interviews reflects broader societal shifts in attitudes toward human-computer interfaces from a theoretical perspective. As perceptions regarding the capabilities and limitations of chatbots evolve, stakeholders within the criminal justice system should be increasingly inclined to leverage the benefits offered by emerging technologies.

Overall, participants who used ChatCharlie perceived their memory performance as lower, however, they had increased positive perceptions of the technology. Participants in Study 4 preferred typing with a chatbot rather than speaking to a chatbot, aligning with Study 1 and Study 2, showing the benefit of typed recall. The most favourable chatbot platform was a dedicated website or email/text messages. Thus, the research question ‘What are the perceptions and experiences of witnesses regarding the use of chatbots and other technology-assisted interviewing methods in the criminal justice system?’ has been answered within the scope of the current investigations. However, further exploring the perspectives of those working within the criminal justice system could illuminate potential benefits and/or limitations of using a chatbot to conduct initial account interviews.

7.3. Limitations of this Research

Variations among participants, such as cognitive capabilities, language skills, and technology familiarity, will undoubtedly have impacted the outcomes across all experiments reported in this thesis. However, given the practical focus of this work, including a diverse range of participants from the general populace enhances the applicability of the findings. By encompassing a broader spectrum of cognitive abilities, linguistic aptitudes, and technological familiarity, this research better reflects the real-world complexities of gathering first-hand accounts from witnesses with diverse backgrounds. Measures were taken to mitigate individual differences by randomly allocating participants to experimental conditions. Randomisation aids in evenly distributing potential confounding factors among the experimental groups, thereby minimising the influence of individual disparities on the observed results.

This research was carried out with participants who are native English speakers, and using online recruitment methods might mean there is a skewed sample, which could restrict how widely the results can be applied to different cultural and linguistic backgrounds. Even though this research looked at English speakers, to some degree, the basic cognitive processes involved in remembering what we have seen are similar across different cultures.

Despite all participants being recruited remotely, the study's sample might not adequately represent the diversity found in the general population. Criteria such as access to a computer or tablet, internet connectivity, and confidence in signing up for the research may have excluded certain demographic groups. Additionally, the requirement for English proficiency and technological access could have inadvertently favoured individuals with higher socioeconomic status or educational attainment. Consequently, the study's findings may not be directly applicable to

individuals from non-English-speaking backgrounds or those with limited access to technology. Regarding technology access, it is worth considering the broader context. The COVID-19 pandemic prompted widespread technology adoption, even among older adults, to mitigate social isolation. Thus, while some bias towards those with greater technological access may exist, it might not be as pronounced as before the pandemic.

Using a mock witness paradigm in a laboratory setting lacks ecological validity compared to real-world eyewitness scenarios. Using one non-violent simulated crime video as a trigger limits how the current findings might drive forward research since real-world situations vary widely in complexity and emotional intensity. However, lab studies allow us to control and manipulate variables carefully and offer a valuable first step. Using a non-violent scenario ensured ethical requirements were met and maintained participants' safety while provoking genuine emotional responses. Although experimental studies may lack the depth of real-life scenarios, they bring practical benefits like experimental control, standardised procedures, and the ability to replicate conditions across different participants. The insights gained from lab studies lay a solid foundation for further research in more realistic settings, like field experiments or analysing actual criminal cases.

Keeping the stimulus consistent controls for an effect of materials in terms of the stimulus leveraging changes in performance rather than experimental manipulations. There is always a trade-off between controlling potentially extraneous variables such as materials effects and making research more generalisable. Here, the decision was to control the stimulus by keeping it consistent across studies. Future research should consider altering the to-be-remembered event across studies since this would

allow a better understanding of the generalisability of the interventions in a manner that has not been possible here.

Turning to consistently high percentage accuracy as has been found here, this is not uncommon in laboratory eyewitness research of this nature (see Dando et al., 2009a: 2009b; Deffenbacher et al., 2004; Evans & Fisher, 2011; Köhnken et al., 1999). The percentage accuracy of what has been recalled was high and consistent across studies, whereby most of the information recalled was accurate. This high accuracy rate may have emanated from the 'gold standard' interview protocols used, which included the four ground rules, of which three focus on improving accuracy, namely 'tell me everything', 'only tell me what you remember' and 'tell me if you do not know the answer or cannot remember'. These instructions, followed by a procedure that asks for a free recall followed by probing questions that are witness recall centred, are known to significantly improve the quality of information recalled, as has been widely reported by others worldwide (see Dando et al., 2009a: 2009b; Memon et al., 2010). Equally, paradigms that use intentional encoding typically report high accuracy rates that are unlikely to be seen in the real world, where intentional encoding is less common.

Percentage accuracy was constant across studies, offering insights into the potential efficacy of the consolidation and retrieval experimental manipulations. Nonetheless, future research should consider alternative types of stimuli since this may be a stimulus effect, or it may not be given high percentage accuracy rates often reported by others. The stimulus lasted 1 minute and was non-violent. Nonetheless, it was clear that 'something' significant had occurred, and the stimulus used was an event explicitly produced for police training purposes and was

information-rich. However, a longer and more complex violent event could be used to improve understanding as research in this domain moves forward.

A limitation of this thesis is the decision not to control for alcohol intoxication among participants despite its relevance to memory accuracy and completeness in real-world crime contexts. Crime frequently occurs in environments where alcohol consumption is typical, and witnesses are often still intoxicated at the time of initial interviews (Block & Block, 1995; Palmer et al., 2013). Research suggests that police and the public perceive intoxicated witnesses as less credible, potentially influencing legal outcomes and interview dynamics (Evans et al., 2009; Benton et al., 2006; Evans & Compo, 2010). Despite these perceptions, prior studies indicate that prompt interviews can yield more accurate recall from intoxicated individuals, as the effects of delay on memory can be more damaging than intoxication itself (Altman et al., 2018; Hildebrand Karlén et al., 2017; La Rooy et al., 2013).

Research on alcohol's impact on memory is mixed, with some studies finding reduced recall accuracy and detail (Calhoun et al., 2004; Mintzer & Griffiths, 2001) and others showing minimal impact (Hildebrand Karlén et al., 2015; Schreiber Compo et al., 2011), controlling for intoxication levels would add complexity without necessarily improving the relevance of the findings to a broader witness population (van Oorsouw & Merckelbach, 2012). Moreover, field and lab-based studies have indicated that high levels of intoxication are more likely to impair recall than moderate levels, making it challenging to generalise findings based on a controlled alcohol level alone (Crossland et al., 2016; Jores et al., 2019). Hence, this thesis excludes alcohol as a variable, acknowledging that real-world witness memory may be impacted by varied and unregulated levels of intoxication that are difficult to replicate or control in research.

Amid the COVID-19 pandemic, conducting research remotely may have increased variability in participants' surroundings, potentially impacting their concentration and involvement in the memory task. Factors such as background noise, interruptions, and distractions in participants' homes could introduce additional variables affecting the study's outcomes. However, remote data collection extended the study's reach geographically and facilitated the inclusion of participants who might otherwise have been excluded due to geographical constraints or limited mobility. While concerns regarding environmental factors such as noise and interruptions are valid, proactive measures were taken to address potential confounding variables. For example, participants were given clear instructions to complete the memory task in a quiet, secluded space to minimise potential distractions.

Regarding the creation of ChatCharlie, it is essential to acknowledge some significant limitations. Primarily, the integration of such a tool relies heavily on the input and contribution of individuals working in the criminal justice system. This study did not address this technology's practical or logistical limitations, such as the perspectives of the police officers who conduct initial account interviews. Gathering this perspective could provide essential insights to inform the design and use of a tool such as ChatCharlie.

The research also underscored a preference for platforms that prioritise privacy when engaging with chatbots. However, direct assessments regarding concerns about data security or trust in the confidentiality of chatbot interactions were not conducted. The research did not explore their comprehension of how chatbot systems would utilise or store personal data, which could influence decisions to share sensitive details in real-life scenarios. Moreover, willingness to disclose

personal information in research may not fully mirror behaviour in real-world interactions with law enforcement chatbots.

Finally, ChatCharlie may be usable and practical in a laboratory. However, it does not offer human flexibility in understanding the complexities of individual witnesses and various contexts and being able to adapt, respond, and/or alter behaviour accordingly. This is undoubtedly the biggest limitation, but the arguments and research programme presented in this thesis are not aimed at suggesting ChatCharlie could ever replace highly skilled interviewers. Instead, ChatCharlie may support investigators by offering a ‘tool’ whereby something is better than nothing, perhaps. Memory consolidation is known to improve a memory trace's robustness, so ChatCharlie undoubtedly offers potential for consolidation.

7.3.1. Practical limitations impacting this thesis.

The COVID-19 health emergency has had significant practical implications for this thesis, which, with hindsight, have offered some benefits regarding supporting suggestions that remote interviewing needs more research. Conversely, the challenges have resulted in a thesis that sits outside what was expected at the start of this PhD journey. Investigative interviewing is traditionally all about human-to-human social interactions, so moving online just 6 months after starting this body of work has been impactful in terms of methodological changes introduced at very short notice and challenges centred on recruitment and supervision arrangements.

7.4. Recommendations for future research.

Primarily, creating a chatbot that might conduct initial account interviews provides an exciting platform to explore many different avenues and variables, which, until now, were impossible to isolate and examine with human interviewers. While efforts

were made to embed ground rules and rapport-building elements into ChatCharlie's design, the specific effects of these components on memory recall accuracy were not examined in this research. Existing psychological literature on the influence of including or excluding these elements predominantly resides within face-to-face, in-person interactions, necessitating further research in this area, possibly by isolating each element using remote chatbot platforms. Future research endeavours could gain from a more detailed exploration of the individual impacts of ground rules and rapport-building elements on memory recall accuracy. Researchers could investigate their respective effects on witness memory recall by systematically varying and manipulating these components within the ChatCharlie interface.

Future research could consider incorporating a broader array of stimuli representing various types of crimes and emotional contexts to enhance the ecological validity of the findings. To fully ascertain the generalisability of the findings, future research should explore the impact of modality across diverse populations and languages. It would also be prudent for future research to consider that performance is likely to improve in quieter, more controlled environments, as such environments are known to enhance complex cognition. Future research must investigate how different methods of presenting information affect individuals' ability to remember things, particularly across diverse groups of people and languages. Such investigations will enhance the relevance of the findings to real-world investigative contexts.

7.5. Conclusion

The thesis focuses on developing and evaluating the initial account interview using a chatbot called ChatCharlie. It aligns with established interviewing standards

and psychological research on human-computer interaction. The research explores typed interviews and the intersection of technology with age-related factors. ChatCharlie is designed to improve the capture of initial witness testimonies following criminal incidents, aiming to enhance efficiency and accuracy during investigations. The thesis critically evaluates the acceptability and feasibility of integrating chatbots into investigative interviewing, contributing significantly to the literature in this area.

This thesis answered the research question, ‘Concerning episodic memory performance, is a remote-typed interview an effective method for collecting investigative information from witnesses of a crime versus a remote spoken interview?’. The findings from this thesis suggest that while text-based interviews do not impair memory performance, they lead to more confabulated items. This indicates that text is a viable means for eliciting details, though not surpassing spoken interviews. The rise in confabulations may stem from reduced vigilance in typed communication, potentially influenced by the absence of a physical interviewer and lower perceived rapport compared to face-to-face interactions. Despite this, accuracy remains unaffected, aligning with previous research on witness recall performance.

The research question ‘Is a remote chatbot interviewer (typed) an effective method for collecting initial account information from witnesses and victims of crime across a wide age range?’ was answered in this thesis. Study 1 investigates the effectiveness of remote typed interviews for collecting investigative information, finding that text-based interviews are comparable to spoken ones, suggesting a suitable format for chatbot interviews. Study 2 addresses concerns about age-appropriateness, showing that participants over 55 recall more correct information

and perform similarly to mid-aged participants, with no age group differences in errors or confabulations. Study 3 focuses on the Chatbot element, revealing that ChatCharlie and in-person interviews significantly improve memory performance without increasing errors or confabulations. Together, these studies demonstrate that a remote-typed chatbot interviewer is effective across a wide age range for collecting initial account information from witnesses and victims of crime.

The third research question addressed and answered by this thesis is ‘What are the potential benefits and challenges associated with integrating chatbots into initial account interviews, and how do they compare to traditional face-to-face interviews?’. Study 3 examines the benefits of integrating chatbots into initial account interviews, focusing on participants' memory performance using ChatCharlie. Findings show significant increases in recalling accurate event details compared to the control group, without a rise in errors or false recollections. The chatbot's design captures the benefits of conversational interviewing while minimising the negative impacts of human-to-human interviews. Additionally, ChatCharlie offers advantages such as clear instructions, potential for language translation, and protection for different age groups and technology users. Overall, this research demonstrates the potential benefits and logistical challenges of integrating chatbots into initial account interviews, answering the research question effectively.

This body of work answers the final research question: ‘ What are the perceptions and experiences of witnesses regarding using chatbots and other technology-assisted interviewing methods in the criminal justice system?’. Study 4 demonstrates that confidence was not found to predict accuracy, highlighting the utility questions regarding confidence indicators in interviews. Apprehensions about technology use were evident. However, those who had used the chatbot felt more comfortable

supporting integrating new technologies. Interestingly, participants across conditions preferred using a chatbot in text modality over spoken modality, possibly due to the benefits of textural recall and increased online inhibition. The perception of a chatbot for the criminal justice system was favourable, especially when accessed through a dedicated website, email, or text messages, which offered more significant privacy safeguards.

Overall, participants using ChatCharlie perceived their memory performance lower than those who did not use it, but using the chatbot increased positive perceptions of the technology. The preference for typing with a chatbot aligns with the benefits of typed recall. Thus, the research question regarding the perceptions and experiences of witnesses has been addressed.

Although future research should be done to fully test different attributes of ChatCharlie, the creation of this tool opens up the door to isolating and examining previously un-tested elements of human interviewing. The perspectives of stakeholders, such as the police or investigative interviewers, should be considered. As it stands, ChatCharlie appears to be a promising tool to assist with collecting witness evidence at the scene of a crime.

References

Abbe, A., & Brandon, S. E. (2013). The Role of Rapport in Investigative Interviewing: A Review. *Journal of Investigative Psychology and Offender Profiling*, *10*(3), 237–249. https://doi.org/10.1002/jip.1386

Abbe, A., & Brandon, S. E. (2014). Building and maintaining rapport in investigative interviews. *Police Practice and Research*, *15*(3), 207–220. https://doi.org/10.1080/15614263.2013.827835

Acierno, R., Hernandez, M. A., Amstadter, A. B., Resnick, H. S., Steve, K., Muzzy, W., & Kilpatrick, D. G. (2010). Prevalence and Correlates of Emotional, Physical, Sexual, and Financial Abuse and Potential Neglect in the United States: The National Elder Mistreatment Study. *American Journal of Public Health*, *100*(2), 292–297.

Aharoni, E., Kleider-Offutt, H. M., Brosnan, S. F., & Fernandes, S. (2020). Slippery scales: Cost prompts, but not benefit prompts, modulate sentencing recommendations in laypeople. *PLOS ONE*, *15*(7), e0236764. https://doi.org/10.1371/journal.pone.0236764

Ahn, S. J. (Grace), & Fox, J. (2017, June 28). *Immersive Virtual Environments, Avatars, and Agents for Health*. Oxford Research Encyclopedia of Communication. https://doi.org/10.1093/acrefore/9780190228613.013.325

Aizpurua, A., Garcia-Bajos, E., & Migueles, M. (2009). Memory for actions of an event: Older and younger adults compared. *Journal of General Psychology*, *136*(4), 428–441. https://doi.org/10.1080/00221300903269816

Akca, D., Larivière, C. D., & Eastwood, J. (2021). Assessing the efficacy of investigative interviewing training courses: A systematic review. *International*

Journal of Police Science & Management, *23*(1), 73–84. https://doi.org/10.1177/14613557211008470

Albright, T. D. (2017). Why eyewitnesses fail. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(30), 7758–7764.

Ali, M. M., Brubacher, S. P., Earhart, B., Powell, M. B., & Westera, N. J. (2020). The utility of ground rule instructions with younger and older adult witnesses. *Applied Cognitive Psychology*, *34*(3), 664–677. https://doi.org/10.1002/acp.3648

Alison, L., & Alison, E. (2017). Revenge versus rapport: Interrogation, terrorism, and torture. *American Psychologist*, *72*(3), 266–277. https://doi.org/10.1037/amp0000064

Allwood, C. M., Ask, K., & Granhag, P. A. (2005). The Cognitive Interview: Effects on the realism in witnesses' confidence in their free recall. *Psychology, Crime and Law*, *11*(2), 183–198. https://doi.org/10.1080/10683160512331329943

Altman, C. M., Compo, N. S., McQuiston, D., Hagsand, A. V., & Cervera, J. (2018). Witnesses' memory for events and faces under elevated levels of intoxication. *Memory*, *26*(7), 946–959. https://doi.org/10.1080/09658211.2018.1445758

Anatürk, M., Kaufmann, T., Cole, J. H., Suri, S., Griffanti, L., Zsoldos, E., Filippini, N., Singh-Manoux, A., Kivimäki, M., Westlye, L. T., Ebmeier, K. P., & De Lange, A. G. (2021). Prediction of brain age and cognitive age: Quantifying brain and cognitive maintenance in aging. *Human Brain Mapping*, *42*(6), 1626–1640. https://doi.org/10.1002/hbm.25316

Anderson, M. C. (2003). Rethinking interference theory: executive control and the mechanisms of forgetting. *Journal of Memory and Language*. 49, 415–445. doi: 10.1016/j.jml.2003. 08.006

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(5), 1063–1087. https://doi.org/10.1037/0278-7393.20.5.1063

Andersson, J. (2001). Net effect of memory collaboration: How is collaboration affected by factors such as friendship, gender and age? *Scandinavian Journal of Psychology*, *42*(4), 367–375. https://doi.org/10.1111/1467-9450.00248

Andersson, J., & Rönnberg, J. (1995). Recall suffers from collaboration: Joint recall effects of friendship and task complexity. *Applied Cognitive Psychology*, *9*(3), 199–211. https://doi.org/10.1002/acp.2350090303

Aquila, M. S. H. (2017). *Building the Personal: Instructors' Perspectives of Rapport in Online and Face-to-Face Classes* [George Mason University]

Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*, *85*, 183–189. https://doi.org/10.1016/J.CHB.2018.03.051

Atkinson, R. C., & Shiffrin, R. M. (1968). Human Memory: A Proposed System and its Control Processes1. In K. W. Spence & J. T. Spence (Eds.), *Psychology of Learning and Motivation* (Vol. 2, pp. 89–195). Academic Press. https://doi.org/10.1016/S0079-7421(08)60422-3

Atkinson, R. C., & Shiffrin, R. M. (1971). The Control of Short-Term Memory. *Scientific American*, *225*(2), 82–91.

Augustinova, M., & Ferrand, L. (2012). The influence of mere social presence on Stroop interference: New evidence from the semantically-based Stroop task.

Journal of Experimental Social Psychology, *48*(5), 1213–1216. https://doi.org/10.1016/j.jesp.2012.04.014

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

Baccon, L., Elodie, C., & MacDougall, H. (2018). Virtual Reality for Teletherapy: Avatars May Combine the Benefits of Face-to-Face Communication with the Anonymity of Online Text-Based Communication. *Cyberpsychology, Behavior, and Social Networking*, *22*. https://doi.org/10.1089/cyber.2018.0247

Baddeley, A. D. (2001). Is working memory still working? *American Psychologist*, *56*(11), 851–864. https://doi.org/10.1037/0003-066X.56.11.851

Balota, D. A., Dolan, P. O., & Duchek, J. M. (2000). Memory changes in healthy older adults. In *The Oxford handbook of memory* (pp. 395–409). Oxford University Press.

Baron, R. S. (1986). Distraction-conflict theory: Progress and problems. *Advances in Experimental Social Psychology*, *19*(C), 1–40. https://doi.org/10.1016/S0065-2601(08)60211-7

Baxter, J. S., Boon, J. C. W., & Marley, C. (2006). Interrogative pressure and responses to minimally leading questions. *Personality and Individual Differences*, *40*(1), 87–98. https://doi.org/10.1016/j.paid.2005.06.017

Baxter, J. S., Charles, K., Martin, M., & McGroarty, A. (2013). The relative influence of leading questions and negative feedback on response change on the Gudjonsson Suggestibility Scale (2): Implications for forensic interviewing. *Psychology, Crime and Law*, *19*(3), 277–285. https://doi.org/10.1080/1068316X.2011.631538

Baxter, J. S., Jackson, M., & Bain, S. A. (2003). Interrogative suggestibility: Interactions between interviewees' self-esteem and interviewer style. *Personality and Individual Differences*, *35*(6), 1285–1292. https://doi.org/10.1016/S0191-8869(02)00349-5

Bedi, R. P., Davis, M. D., & Williams, M. (2005). Critical incidents in the formation of the therapeutic alliance from the client's perspective. *Psychotherapy*, *42*(3), 311–323. https://doi.org/10.1037/0033-3204.42.3.311

Beek, M. van, Bull, R., & Chen, M. (2021). When the Evidence Is Incorrect: An Exploration of What Happens When Interviewers Unwittingly Present Inaccurate Information in Interviews with Suspects. *Journal of Police and Criminal Psychology*, *36*(4), 769–782. https://doi.org/10.1007/s11896-021-09494-3

Bekerian, D. A., & Dennett, J. L. (1990). Spoken and written recall of visual narratives. *Applied Cognitive Psychology*, *4*(3), 175–187. https://doi.org/10.1002/acp.2350040303

Belletier, C., Normand, A., & Huguet, P. (2019). Social-Facilitation-and-Impairment Effects: From Motivation to Cognition and the Social Brain. *Current Directions in Psychological Science*, *28*(3), 260–265. https://doi.org/10.1177/0963721419829699

Benton, T. R., Ross, D. F., Bradshaw, E., Thomas, W. N., & Bradshaw, G. S. (2006). Eyewitness memory is still not common sense: Comparing jurors, judges and law enforcement to eyewitness experts. *Applied Cognitive Psychology*, *20*(1), 115–129. https://doi.org/10.1002/acp.1171

Bergmann, M. M., Jacobs, E. J., Hoffmann, K., & Boeing, H. (2004). Agreement of Self-Reported Medical History: Comparison of an In-Person Interview with a

Self-Administered Questionnaire. *European Journal of Epidemiology*, *19*(5), 411–416. https://doi.org/10.1023/B:EJEP.0000027350.85974.47

Berkowitz, S. R., Garrett, B. L., Fenn, K. M., & Loftus, E. F. (2022). Convicting with confidence? Why we should not over-rely on eyewitness confidence. *Memory*, *30*(1), 10–15. https://doi.org/10.1080/09658211.2020.1849308

Beyer, L., Meyer-Wilmes, J., Schönecker, S., Schnabel, J., Sauerbeck, J., Scheifele, M., Prix, C., Unterrainer, M., Catak, C., Pogarell, O., Palleis, C., Perneczky, R., Danek, A., Buerger, K., Bartenstein, P., Levin, J., Rominger, A., Ewers, M., & Brendel, M. (2021). Cognitive reserve hypothesis in frontotemporal dementia: A FDG-PET study. *NeuroImage: Clinical*, *29*, 102535. https://doi.org/10.1016/j.nicl.2020.102535

Biggs, S., Erens, B., Doyle, M., Hall, J., & Sanchez, M. (2009). *Abuse and neglect of older people: Secondary analysis of UK prevalence study*.

Blank, H., & Launay, C. (2014). How to protect eyewitness memory against the misinformation effect: A meta-analysis of post-warning studies. *Journal of Applied Research in Memory and Cognition*, *3*(2), 77–88. https://doi.org/10.1037/h0101798

Block, R. R. L., & Block, C. R. (1995). Space, place, and crime: Hot spot areas and hot places of liquor-related crime. *Crime and Place*, 145–183.

Bodner, G. E., Musch, E., & Azad, T. (2009). Reevaluating the potency of the memory conformity effect. *Memory and Cognition*, *37*(8), 1069–1076. https://doi.org/10.3758/MC.37.8.1069

Bornstein, B. H., Deffenbacher, K. A., Penrod, S. D., & McGorty, E. K. (2012). Effects of exposure time and cognitive operations on facial identification accuracy: A meta-analysis of two variables associated with initial memory

strength. *Psychology, Crime and Law*, *18*(5), 473–490. https://doi.org/10.1080/1068316X.2010.508458

Bower, G. (1967). A Multicomponent Theory of The Memory Trace. *Psychology of Learning and Motivation - Advances in Research and Theory*, *1*(C), 229–325. https://doi.org/10.1016/S0079-7421(08)60515-0

Boyce, M. A., Lindsay, D. S., & Brimacombe, C. A. E. (2008). Investigating investigators: Examining the impact of eyewitness identification evidence on student-investigators. *Law and Human Behavior*, *32*(5), 439–453. https://doi.org/10.1007/s10979-007-9125-5

Brandtzaeg, P. B., & Følstad, A. (2017). Why people use chatbots. *Lecture Notes in Computer Science: 4th International Conference, INSCI 2017, Thessaloniki, Greece, November 22-24, 2017 Proceedings 4 (pp. 377-392).* Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-70284-1_30

Brewer, N., & Douglass, A. B. (2019). Psychological science and the law. In *Psychological science and the law.*

Brewer, N., & Weber, N. (2008). Eyewitness confidence and latency: Indices of memory processes not just markers of accuracy. *Applied Cognitive Psychology*, *22*(6), 827–840. https://doi.org/10.1002/acp.1486

Brewer, N., & Wells, G. L. (2006). The confidence-accuracy relationship in eyewitness identification: Effects of lineup instructions, foil similarity, and target-absent base rates. *Journal of Experimental Psychology: Applied*, *12*(1), 11–30. https://doi.org/10.1037/1076-898X.12.1.11

Brewin, C. R., Andrews, B., & Mickes, L. (2020). Regaining Consensus on the Reliability of Memory. *Current Directions in Psychological Science*, *29*(2), 121–125. https://doi.org/10.1177/0963721419898122

Brimbal, L., Dianiska, R. E., Swanner, J. K., & Meissner, C. A. (2019). Enhancing cooperation and disclosure by manipulating affiliation and developing rapport in investigative interviews. *Psychology, Public Policy, and Law*, *25*(2), 107–115. https://doi.org/10.1037/law0000193

Brimbal, L., Meissner, C. A., Kleinman, S. M., Phillips, E. L., Atkinson, D. J., Dianiska, R. E., Rothweiler, J. N., Oleszkiewicz, S., & Jones, M. S. (2021). Evaluating the benefits of a rapport-based approach to investigative interviews: A training study with law enforcement investigators. *Law and Human Behavior*, *45*(1), 55–67. https://doi.org/10.1037/lhb0000437

Brubacher, S. P., Benson, M. S., Powell, M. B., Goodman-Delahunty, J., & Westera, N. J. (2020). Chapter Twenty-Two—An overview of best practice investigative interviewing of child witnesses of sexual assault. In I. Bryce & W. Petherick (Eds.), *Child Sexual Abuse* (pp. 445–466). Academic Press. https://doi.org/10.1016/B978-0-12-819434-8.00022-2

Brubacher, S. P., Poole, D. A., & Dickinson, J. J. (2015). The use of ground rules in investigative interviews with children: A synthesis and call for research. *Developmental Review*, *36*, 15–33. https://doi.org/10.1016/j.dr.2015.01.001

Brunel, M., Launay, C., Hermant, M., & Py, J. (2022). Perception of Acceptability and Usability of a Modified Cognitive Interview in the Evaluation of Police Training in France. *Journal of Police and Criminal Psychology*, *37*(1), 38–48. https://doi.org/10.1007/s11896-020-09416-9

Bulevich, J. B., Gordon, L. T., Hughes, G. I., & Thomas, A. K. (2022). Are witnesses able to avoid highly accessible misinformation? Examining the efficacy of different warnings for high and low accessibility postevent misinformation.

Memory & Cognition, *50*(1), 45–58. https://doi.org/10.3758/s13421-021-01255-1

Bull, R. (2018). What is 'believed' or actually 'known' about characteristics that may contribute to being a good/effective interviewer? *Investigating the Truth*, 265–281. https://doi.org/10.4324/9781315169910-16

Bull, R. (2019). Roar or "PEACE": Is it a "tall story?" In *The Routledge International Handbook of Legal and Investigative Psychology*. Routledge.

Bull, R., & Baker, B. (2020). *Obtaining Valid Discourse from Suspects PEACE-fully: What Role for Rapport and Empathy? In. M. Mason & F. Rock (Eds.) 'The Discourse of Police Interviews'.*

Bull, R., & Milne, B. (2020). Recommendations for Collecting Event Memory Evidence. In *Memory and Sexual Misconduct*. Routledge.

Calhoun, V. D., Altschul, D., McGinty, V., Shih, R., Scott, D., Sears, E., & Pearlson, G. D. (2004). Alcohol Intoxication Effects on Visual Perception: An fMRI Study. *Human Brain Mapping*, *21*(1), 15–26. https://doi.org/10.1002/hbm.10145

Caso, A., Gabbert, F., & Dando C. J. (2024) Eyewitness confidence in the interviewing context: Understanding the impact of question type and order. *Applied Cognitive Psychology*, ISSN 0888-4080 [Article] (In Press)

Chan, J. C. K., Thomas, A. K., & Bulevich, J. B. (2009). Recalling a witnessed event increases eyewitness suggestibility: The reversed testing effect. *Psychological Science*, *20*(1), 66–73. https://doi.org/10.1111/j.1467-9280.2008.02245.x

Chevroulet, C., Paterson, H. M., Kemp, R. I., & van Golde, C. (2024). Technology and eyewitness memory: evaluating the efficacy of a novel digital cued recall tool. *Psychology, Crime & Law*, 30(6), 538-555.

Chin, J., Milne, B., & Bull, R. (2022). Fuelling an investigative mindset: The importance of pre-interview planning in police interviews with suspects. *Psychology, Crime & Law*, *0*(0), 1–25. https://doi.org/10.1080/1068316X.2022.2139829

Chun, M. M., & Turk-Browne, N. B. (2008). Associative Learning Mechanisms in Vision. In S. J. Luck & A. Hollingworth (Eds.), *Visual Memory* (pp. 209–246). Oxford University Press.

Clarke, C., & Milne, B. (2001). A National Evaluation of the PEACE Investigative Interviewing Course. (p. 187). London: Home office.

College of Policing. (2019). *Guidelines for first responders Obtaining initial accounts from victims and witnesses*. https://www.college.police.uk/app/investigation/investigative-interviewing/investigative-interviewing

College of Policing. (2022). *Investigative interviewing*. College of Policing. https://www.college.police.uk/app/investigation/investigative-interviewing/investigative-interviewing

Collins, K., & Carthy, N. (2019). No rapport, no comment: The relationship between rapport and communication during investigative interviews with suspects. *Journal of Investigative Psychology and Offender Profiling*, *16*(1), 18–31. https://doi.org/10.1002/jip.1517

Collins, R., Lincoln, R., & Frank, M. G. (2002). The effect of rapport in forensic interviewing. *Psychiatry, Psychology and Law*, *9*(1), 69–78. https://doi.org/10.1375/pplt.2002.9.1.69

Cooper, P., Dando, C. J., Ormerod, T. C., Mattison, M., Marchant, R., Milne, R., & Bull, R. (2018). *One step forward and two steps back? The '20 Principles' for*

questioning vulnerable witnesses and the lack of an evidence-based approach. https://doi.org/10.1177/1365712718793435

Corritore, C. L., Kracher, B., & Wiedenbeck, S. (2003). On-line trust: Concepts, evolving themes, a model. *International Journal of Human-Computer Studies*, *58*(6), 737–758. https://doi.org/10.1016/S1071-5819(03)00041-7

Craik, F. I. M., & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(3), 474–479. https://doi.org/10.1037/0278-7393.13.3.474

Craik, F. I. M., Naveh-Benjamin, M., Ishaik, G., & Anderson, N. D. (2000). Divided Attention during Encoding and Retrieval: Differential Control Effects? *Journal of Experimental Psychology: Learning Memory and Cognition*, *26*(6), 1744–1749. https://doi.org/10.1037/0278-7393.26.6.1744

Crossland, D., Kneller, W., & Wilcock, R. (2016). Intoxicated Witnesses: Testing the Validity of the Alcohol Myopia Theory. *Applied Cognitive Psychology*, *30*(2), 270–281. https://doi.org/10.1002/acp.3209

Crum, R. M., Anthony, J. C., Bassett, S. S., & Folstein, M. F. (1993). Population-Based Norms for the Mini-Mental State Examination by Age and Educational Level. *JAMA: The Journal of the American Medical Association*, *269*(18), 2386–2391. https://doi.org/10.1001/jama.1993.03500180078038

Czaja, S. J., Sharit, J., Ownby, R., Roth, D. L., & Nair, S. (2001). Examining age differences in performance of a complex information search and retrieval task. *Psychology and Aging*, *16*(4), 564–579. https://doi.org/10.1037/0882-7974.16.4.564

Dando, C. J. (2013). Drawing to Remember: External Support of Older Adults' Eyewitness Performance. *PLoS ONE*, *8*(7). https://doi.org/10.1371/journal.pone.0069937

Dando, C. J. (2020). *Is Interviewing Witnesses and Victims Remotely a Good Idea?* Psychology Today. https://www.psychologytoday.com/gb/blog/forensic-insights/202006/is-interviewing-witnesses-and-victims-remotely-good-idea

Dando, C. J., Gabbert, F., & Hope, L. (2020). *Supporting older eyewitnesses' episodic memory: The Self-Administered Interview and Sketch Reinstatement of Context*. https://doi.org/10.1080/09658211.2020.1757718

Dando, C. J., Geiselman, R. E., MacLeod, N., & Griffiths, A. (2015). Interviewing Adult Witnesses and Victims. In G. E. Oxburgh, T. Myklebust, T. Grant, & R. Milne (Eds.), *Communication in Investigative and Legal Contexts* (1st ed., pp. 79–106). Wiley. https://doi.org/10.1002/9781118769133.ch5

Dando, C. J., Ormerod, T. C., Cooper, P., Marchant, R., Mattison, M., Milne, R., & Bull, R. (2018). No Evidence Against Sketch Reinstatement of Context, Verbal Labels or the Use of Registered Intermediaries for Children with Autism Spectrum Disorder: Response to Henry et al. (2017). *Journal of Autism and Developmental Disorders*, *48*(7), 2593–2596. https://doi.org/10.1007/s10803-018-3479-z

Dando, C. J., Ormerod, T., Wilcock, R., & Milne, B. (2011). When help becomes hindrance: Unexpected errors of omission and commission in eyewitness memory resulting from change temporal order at retrieval? *Cognition*, *121*(3), 416–421. https://doi.org/10.1016/j.cognition.2011.06.015

Dando, C. J., & Oxburgh, G. E. (2016). Empathy in the field: Towards a taxonomy of empathic communication in information gathering interviews with suspected

sex offenders. *European Journal of Psychology Applied to Legal Context*, *8*(1), 27–33. https://doi.org/10.1016/j.ejpal.2015.10.001

Dando, C. J., Taylor, D., Caso, A., Nahouli, Z., & Adam, C. (2023). Interviewing in virtual environments: Towards understanding the impact of rapport building behaviours and retrieval context on eyewitness memory. *Memory & Cognition*.

Dando, C. J., Wilcock, R., & Milne, B. (2009). The cognitive interview: Novice police officers' witness/victim interviewing practices. *Psychology, Crime and Law*, *15*(8), 679–696. https://doi.org/10.1080/10683160802203963

Dando, C., Wilcock, R., & Milne, R. (2009a). The cognitive interview: The efficacy of a modified mental reinstatement of context procedure for frontline police investigators. Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 23(1), 138-147.

Dando, C., Wilcock, R., Milne, R., & Henry, L. (2009b). A modified cognitive interview procedure for frontline police investigators. Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 23(5), 698-716.

Datta, S., Dharur, S., Cartillier, V., Desai, R., Khanna, M., Batra, D., & Parikh, D. (2022). *Episodic Memory Question Answering*. 19119–19128. DOI:10.1109/CVPR52688.2022.01853

Davies, G., & Griffiths, L. (2008). Eyewitness identification and the English courts: A century of trial and error. *Psychiatry, Psychology and Law*, *15*(3), 435–449. https://doi.org/10.1080/13218710802101605

Davis, D., & Loftus, E. (2017). Internal and external sources of misinformation in adult witness memory. In *The Handbook of Eyewitness Psychology: Volume I: Memory for Events* (pp. 195–238).

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A Meta-Analytic Review of the Effects of High Stress on Eyewitness Memory. *Law and Human Behavior*, *28*(6), 687–706. https://doi.org/10.1007/s10979-004-0565-x

Dickinson, J. J., Lytle, N. E., & Poole, D. A. (2021). Tele-forensic interviewing can be a reasonable alternative to face-to-face interviewing of child witnesses. *Law and Human Behavior*, *45*(2), 97–111. https://doi.org/10.1037/lhb0000443

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, *144*(1), 114–126. https://doi.org/10.1037/xge0000033

Dilevski, N., Cullen, H. J., van Golde, C., Flowe, H. D., Paterson, H. M., Takarangi, M. K. T., & Monds, L. A. (2024). Juror Perceptions of Bystander and Victim Intoxication by Different Substances. *Criminal Justice and Behavior*, *51*(4), 569–587. https://doi.org/10.1177/00938548241227543

Dodson, C. S., & Krueger, L. E. (2006). I misremember it well: Why older adults are unreliable eyewitnesses. *Psychonomic Bulletin & Review*, *13*(5), 770–775. https://doi.org/10.3758/BF03193995

Dodson, C. S., Powers, E., & Lytell, M. (2015). Aging, confidence, and misinformation: Recalling information with the cognitive interview. *Psychology and Aging*, *30*(1), 46–61. https://doi.org/10.1037/a0038492

Doherty-Sneddon, G., & McAuley, S. (2000). Influence of Video-Mediation on Adult-Child Interviews: Implications for the Use of the Live Link with Child Witnesses. *Applied Cognitive Psychology*, *14*(4), 379–392. https://doi.org/10.1002/1099-0720(200007/08)

Douglass, A. B., & Bustamante, L. (2014). Social influences on eyewitness memory. In *The SAGE handbook of applied memory* (pp. 614–632).

Dudukovic, N. M., Marsh, E. J., & Tversky, B. (2004). Telling a story or telling it straight: The effects of entertaining versus accurate retellings on memory. *Applied Cognitive Psychology*, *18*(2), 125–143. https://doi.org/10.1002/acp.953

Duke, M. C., Wood, J. M., Bollin, B., Scullin, M., & LaBianca, J. (2018). Development of the Rapport Scales for Investigative Interviews and Interrogations (RS3i), Interviewee Version. *Psychology, Public Policy, and Law*, *24*(1), 64–79. https://doi.org/10.1037/law0000147

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, *66*(3), 183–201. https://doi.org/10.1037/h0047707

Ebbinghaus, H. (1913). Memory: A contribution to experimental psychology. *New York: Teachers College, Columbia University. (Original Work Published 1885)*.

Echterhoff, G., Hirst, W., & Hussy, W. (2005). How eyewitnesses resist misinformation: Social postwarnings and the monitoring of memory characteristics. *Memory & Cognition*, *33*(5), 770–782. https://doi.org/10.3758/BF03193073

Elsner, N. (2017, May 15). *KAYAK Mobile Travel Report | KAYAK*. KAYAK Blog UK. https://www.kayak.co.uk/news/mobile-travel-report-2017/

Evans, J. R., & Compo, N. S. (2010). Mock jurors' perceptions of identifications made by intoxicated eyewitnesses. *Psychology, Crime and Law*, *16*(3), 191–210. https://doi.org/10.1080/10683160802612890

Evans, J. R., Compo, N. S., & Russano, M. B. (2009). Intoxicated Witnesses and Suspects: Procedures and Prevalence According to Law Enforcement. *Psychology, Public Policy, and Law*, *15*(3), 194–221. https://doi.org/10.1037/a0016837

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

Fahsing, I., & Ask, K. (2013). Decision Making and Decisional Tipping Points in Homicide Investigations: An Interview Study of British and Norwegian Detectives. *Journal of Investigative Psychology and Offender Profiling*, *10*(2), 155–165. https://doi.org/10.1002/jip.1384

Fahsing, I., & Ask, K. (2016). The making of an expert detective: The role of experience in English and Norwegian police officers' investigative decision-making. *Psychology, Crime and Law*, *22*(3), 203–223. https://doi.org/10.1080/1068316X.2015.1077249

Finn, W., Mcneill, A., & Mclean, F. (2019). *Obtaining initial accounts from victims and witnesses: A Rapid Evidence Assessment to support the development of College guidelines on obtaining initial accounts from victims and witnesses*. http://whatworks.college.police.uk/Research/Pages/Published.aspx

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

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. https://doi.org/10.1016/j.ijlp.2010.09.004

Fisher, R. P., Milne, B., & Bull, R. (2011). Interviewing Cooperative Witnesses. *Https://Doi.Org/10.1177/0963721410396826*, *20*(1), 16–19.

Fiske, S. T., & Taylor, S. E. (2013). Representation in memory. *Social Cognition: From Brains to Culture (2nd Ed.). London: SAGE Publishing*.

Fleming, J. H., & Darley, J. M. (1991). Mixed Messages: The Multiple Audience Problem and Strategic Communication. *Social Cognition*, *9*(1), 25–46. https://doi.org/10.1521/soco.1991.9.1.25

Flindall, I. R., Leff, D. R., Pucks, N., Sugden, C., & Darzi, A. (2016). The Preservation of Cued Recall in the Acute Mentally Fatigued State: A Randomised Crossover Study. *World Journal of Surgery*, *40*(1), 56–65. https://doi.org/10.1007/s00268-015-3317-9

Flowe, H. D., & Schreiber Compo, N. (2021). The lack of robust evidence for the effects of alcohol on false memory. *Neuroscience & Biobehavioral Reviews*, *127*, 332–333. https://doi.org/10.1016/j.neubiorev.2021.04.029

Fogg, B. J., & Tseng, H. (1999). The elements of computer credibility. *Conference on Human Factors in Computing Systems - Proceedings*, 80–87. https://doi.org/10.1145/302979.303001

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”. *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Foreman, M. D., Fletcher, K., Mion, L. C., Simon, L., & Niche Faculty. (1996). Assessing cognitive function: The complexities of assessment of an individual’s cognitive status are important in making an accurate and

comprehensive evaluation. *Geriatric Nursing*, *17*(5), 228–232. https://doi.org/10.1016/S0197-4572(96)80210-2

French, L., Garry, M., & Mori, K. (2011). Relative—Not absolute—Judgments of credibility affect susceptibility to misinformation conveyed during discussion. *Acta Psychologica*, *136*(1), 119–128. https://doi.org/10.1016/j.actpsy.2010.10.009

Fussell, S. R., & Krauss, R. M. (1989). Understanding friends and strangers: The effects of audience design on message comprehension. *European Journal of Social Psychology*, *19*(6), 509–525. https://doi.org/10.1002/ejsp.2420190603

Gabbert, F. (2004). *Use with Reviewing the Literature in OneNote: Memory conformity between co-witnesses: The effects of discussion on subsequent memory accuracy*. (University of Aberdeen).

Gabbert, F., Hope, L., Carter, E., Boon, R., & Fisher, R. (2015). The Role of Initial Witness Accounts within the Investigative Process. In G. E. Oxburgh, T. Myklebust, T. Grant, & B. Milne (Eds.), *Communication in Investigative and Legal Contexts* (pp. 107–131). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118769133.ch6

Gabbert, F., Hope, L., & Fisher, R. P. (2009). Protecting eyewitness evidence: Examining the efficacy of a self-administered interview tool. *Law and Human Behavior*, *33*(4), 298–307. https://doi.org/10.1007/s10979-008-9146-8

Gabbert, F., Hope, L., Fisher, R. P., & Jamieson, K. (2012). Protecting Against Misleading Post-event Information with a Self-Administered Interview. *Applied Cognitive Psychology*, *26*(4), 568–575. https://doi.org/10.1002/acp.2828

Gabbert, F., Hope, L., Luther, K., Wright, G., Ng, M., & Oxburgh, G. E. (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

Gabbert, F., Memon, A., & Allan, K. (2003). Memory Conformity: Can Eyewitnesses Influence Each Other's Memories for an Event? *Applied Cognitive Psychology*, *17*(5), 533–543. https://doi.org/10.1002/acp.885

Gabbert, F., Memon, A., & Wright, D. B. (2006). Memory conformity: Disentangling the steps toward influence during a discussion. *Psychonomic Bulletin & Review*, *13*(3), 480–485. https://doi.org/10.3758/BF03193873

Gabbert, F., Tamonyte, D., Apps, J., Caso, A., Woolnough, P., Hope, L., Handscomb, M., & Waterworth, G. (2020). Examining the efficacy of a self-administered report form in missing person investigations. *Legal and Criminological Psychology*, *25*(1), 1–16. https://doi.org/10.1111/lcrp.12163

Garcia-Bajos, E., & Migueles, M. (2009). When stereotype knowledge prevents retrieval-induced forgetting. *Acta Psychologica*, 131(1), 63–71.

Garobbio, S., Pilz, K. S., Kunchulia, M., & Herzog, M. H. (2023). No Common Factor Underlying Decline of Visual Abilities in Mild Cognitive Impairment. *Experimental Aging Research*, *49*(3), 183–200. https://doi.org/10.1080/0361073X.2022.2094660

Gawrylowicz, J., Memon, A., & Scoboria, A. (2014). Equipping witnesses with transferable skills: The Self-Administered Interview. *Psychology, Crime and Law*, *20*(4), 315–325. https://doi.org/10.1080/1068316X.2013.777961

Geiselman, R. E., Fisher, R. P., MacKinnon, D. P., & Holland, H. L. (1986). Enhancement of Eyewitness Memory with the Cognitive Interview. *The American Journal of Psychology*, *99*(3), 385–401. https://doi.org/10.2307/1422492

Glazier, R. A. (2016). Building Rapport to Improve Retention and Success in Online Classes. *Journal of Political Science Education*, *12*(4), 437–456. https://doi.org/10.1080/15512169.2016.1155994

Goldsmith, M., Koriat, A., & Pansky, A. (2005). Strategic regulation of grain size in memory reporting over time. *Journal of Memory and Language*, *52*(4), 505–525. https://doi.org/10.1016/j.jml.2005.01.010

Grabowski, J. (2007). The Writing Superiority Effect in the Verbal Recall of Knowledge: Sources and Determinants. In *Writing and Cognition* (pp. 163–179). Brill. https://doi.org/10.1163/9781849508223_012

Granhag, P. A., Fisher, R. P., Ross, S. J., & Cahill, B. S. (2018). Interviewing witnesses and victims. In *Forensic Psychology in Context* (pp. 56–74).

Greener, S., & Wakefield, C. (2015). Developing Confidence in the Use of Digital Tools in Teaching. *Electronic Journal of E-Learning*, *13*(4), 260–267.

Greenspan, R. L., & Loftus, E. F. (2020). Eyewitness confidence malleability: Misinformation as post-identification feedback. *Law and Human Behavior*, *44*(3), 194–208. https://doi.org/10.1037/lhb0000369

Griffiths, A., & Milne, B. (2006). Will it all end in tiers? Police interviews with suspects in Britain. *Investigative Interviewing: Rights, Research, Regulation*.

Gurney, D. J. (2011). *The Misleading Potential of Communicative Hand Gestures in a Forensic Interview*. http://uhra.herts.ac.uk/handle/2299/6003

Gurney, D. J. (2015). What's Left Unsaid_ How Nonverbal Influence Compares with Verbal Influence. *Psychiatry, Psychology and Law*, *22*(3), 465–473. https://doi.org/10.1080/13218719.2014.985624

Gurney, D. J., Ellis, L. R., & Vardon-Hynard, E. (2016). The saliency of gestural misinformation in the perception of a violent crime. *Psychology, Crime and Law*, *22*(7), 651–665. https://doi.org/10.1080/1068316X.2016.1174860

Haase, K. R., Cosco, T., Kervin, L., Riadi, I., & O'Connell, M. E. (2021). Older Adults' Experiences With Using Technology for Socialization During the COVID-19 Pandemic: Cross-sectional Survey Study. *JMIR Aging*, *4*(2), e28010. https://doi.org/10.2196/28010

Hagsand, A. V., Hjelmsäter, E. R. af, Granhag, P. A., Fahlke, C., & Gordh, A. S. (2017). Witnesses stumbling down memory lane: The effects of alcohol intoxication, retention interval, and repeated interviewing. *Memory*, *25*(4), 531–543. https://doi.org/10.1080/09658211.2016.1191652

Hagsand, A. V., Roos-af-Hjelmsäter, E., Granhag, P. A., Fahlke, C., & Söderpalm-Gordh, A. (2013). Do sober eyewitnesses outperform alcohol intoxicated eyewitnesses in a lineup? *European Journal of Psychology Applied to Legal Context*, *5*(1), 23–47.

Häkkänen, H., Ask, K., Kebbell, M. R., Alison, L., & Granhag, P. A. (2009). Police officers' views of effective interview tactics with suspects: The effects of weight of case evidence and discomfort with ambiguity. *Applied Cognitive Psychology*, *23*(4), 468–481. https://doi.org/10.1002/acp.1491

Harris, C. B. (2010). *Collaborative Remembering: When Can Remembering With Others Be Beneficial?* 131–134. https://doi.org/10.5096/ascs200921

Harris, M. T., Blocker, K. A., & Rogers, W. A. (2022). Older Adults and Smart Technology: Facilitators and Barriers to Use. *Frontiers in Computer Science*, *4*. 835927

Henderson, S., & Gilding, M. (2004). 'I've Never Clicked this Much with Anyone in My Life': Trust and Hyperpersonal Communication in Online Friendships. *New Media & Society*, *6*(4), 487–506. https://doi.org/10.1177/146144804044331

Henkel, L. A. (2014). Do older adults change their eyewitness reports when re-questioned? *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, *69*(3), 356–365. https://doi.org/10.1093/geronb/gbt071

Henkel, L. A. (2017). Inconsistencies across repeated eyewitness interviews: Supportive negative feedback can make witnesses change their memory reports. *Psychology, Crime and Law*, *23*(2), 97–117. https://doi.org/10.1080/1068316X.2016.1225051

Henkel, L. A., Coffman, K. A. J., & Dailey, E. M. (2008). A survey of people's attitudes and beliefs about false confessions. *Behavioral Sciences and the Law*, *26*(5), 555–584. https://doi.org/10.1002/bsl.826

Henry, L., Crane, L., Nash, G., Hobson, Z., Kirke-Smith, M., & Wilcock, R. (2017). Verbal, Visual, and Intermediary Support for Child Witnesses with Autism During Investigative Interviews. *Journal of Autism and Developmental Disorders*, *47*. https://doi.org/10.1007/s10803-017-3142-0

Herrera, F., Oh, S. Y., & Bailenson, J. N. (2020). Effect of Behavioral Realism on Social Interactions Inside Collaborative Virtual Environments. *Presence*, *27*(2), 163–182. https://doi.org/10.1162/pres_a_00324

Hildebrand Karlén, M., Hjelmsäter, E. R. af, Fahlke, C., Granhag, P. A., & Gordh, A. S. (2015). Alcohol intoxicated eyewitnesses' memory of intimate partner violence. *Psychology, Crime and Law*, *21*(2), 156–171. https://doi.org/10.1080/1068316X.2014.951644

Hildebrand Karlén, M., Hjelmsäter, E. R. A., Fahlke, C., Granhag, P. A., & Söderpalm-Gordh, A. (2017). To wait or not to wait? Improving results when interviewing intoxicated witnesses to violence. *Scandinavian Journal of Psychology*, *58*(1), 15–22. https://doi.org/10.1111/sjop.12345

Hollenbaugh, E. E., & Everett, M. K. (2013). The Effects of Anonymity on Self-Disclosure in Blogs: An Application of the Online Disinhibition Effect: Anonymity and self-disclosure. *Journal of Computer-Mediated Communication*, *18*(3), 283–302. https://doi.org/10.1111/jcc4.12008

Holliday, R. E., Humphries, J. E., Milne, B., Memon, A., Houlder, L., Lyons, A., & Bull, R. (2012). Reducing misinformation effects in older adults with cognitive interview mnemonics. *Psychology and Aging*, *27*(4), 1191–1203. https://doi.org/10.1037/a0022031

Hoogesteyn, K., Verigin, B. L., Finnick, D., & Meijer, E. H. (2023). Rapport-building: Chat versus in-person witness interviews. *Journal of Investigative Psychology and Offender Profiling*, *20*(2), 162–178. https://doi.org/10.1002/jip.1609

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

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

Hope, L., Gabbert, F., Heation-Armstrong, A., & Wolchover, D. (2013). Self administered witness interviews–part IV. *Criminal Law & Justice Weekly*, *177*(4).

Hope, L., Ost, J., Gabbert, F., Healey, S., & Lenton, E. (2008). 'With a little help from my friends...': The role of co-witness relationship in susceptibility to misinformation. *Acta Psychologica*, *127*(2), 476–484. https://doi.org/10.1016/j.actpsy.2007.08.010

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

Houston, K. A., Hope, L., Memon, A., & Read, J. D. (2013). Expert testimony on eyewitness evidence: In search of common sense. *Behavioral Sciences and the Law*, *31*(5), 637–651. https://doi.org/10.1002/bsl.2080

Howe, M. L., & Knott, L. M. (2015). The fallibility of memory in judicial processes: Lessons from the past and their modern consequences. *Memory*, *23*(5), 633–656. https://doi.org/10.1080/09658211.2015.1010709

Hsu, C.-W., Gross, J., Colombo, M., & Hayne, H. (2023). Look into my eyes: A "faceless" avatar interviewer lowers reporting threshold for adult eyewitnesses. *Memory & Cognition*, *51*(8), 1761–1773. https://doi.org/10.3758/s13421-023-01424-4

Huff, C. R., & Naughton, M. (2017). Wrongful conviction reforms in the U.S. and UK: Taking Stock. In E. Plywaczewski (Ed.), *Current Problems of the Penal Law and Criminology. 7th Edition* (pp. 482–505).

Huff, M. J., Weinsheimer, C. C., & Bodner, G. E. (2016). Reducing the Misinformation Effect Through Initial Testing: Take Two Tests and Recall Me in the Morning? *Applied Cognitive Psychology*, *30*(1), 61–69. https://doi.org/10.1002/acp.3167

Iida, R., Itsukusima, Y., & Mah, E. Y. (2020). How do we judge our confidence? Differential effects of meta-memory feedback on eyewitness accuracy and confidence. *Applied Cognitive Psychology*, *34*(2), 397–408. https://doi.org/10.1002/acp.3625

Innocence Project. (2022). *Eyewitness Identification Reform*. Innocence Project. https://innocenceproject.org/eyewitness-identification-reform/

Jack, F., Zydervelt, S., & Zajac, R. (2014). Are co-witnesses special? Comparing the influence of co-witness and interviewer misinformation on eyewitness reports. *Memory*, *22*(3), 243–255. https://doi.org/10.1080/09658211.2013.778291

Jenner, B. M., & Myers, K. C. (2019). Intimacy, rapport, and exceptional disclosure: A comparison of in-person and mediated interview contexts. *International Journal of Social Research Methodology*, *22*(2), 165–177. https://doi.org/10.1080/13645579.2018.1512694

Johnson, K. (2017, April 18). Facebook Messenger hits 100,000 bots. *VentureBeat*. https://venturebeat.com/ai/facebook-messenger-hits-100000-bots/

Joinson, A. N. (2001a). Knowing me, knowing you: Reciprocal self-disclosure in Internet-based surveys. *Cyberpsychology and Behavior*, *4*(5), 587–591. https://doi.org/10.1089/109493101753235179

Joinson, A. N. (2001b). Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity. *European Journal of Social Psychology*, *31*(2), 177–192. https://doi.org/10.1002/ejsp.36

Joinson, A. N., Woodley, A., & Reips, U. D. (2007). Personalization, authentication and self-disclosure in self-administered Internet surveys. *Computers in Human Behavior*, *23*(1), 275–285. https://doi.org/10.1016/j.chb.2004.10.012

Jones, K. A., Crozier, W. E., & Strange, D. (2018). Objectivity is a Myth for you but not for me or police: A bias blind spot for viewing and remembering criminal events. *Psychology, Public Policy, and Law*, *24*(2), 259–270. https://doi.org/10.1037/law0000168

Jores, T., Colloff, M. F., Kloft, L., Smailes, H., & Flowe, H. D. (2019). A meta-analysis of the effects of acute alcohol intoxication on witness recall. *Applied Cognitive Psychology*, *33*(3), 334–343. https://doi.org/10.1002/acp.3533

Kang, S. H., & Gratch, J. (2010). Virtual humans elicit socially anxious interactants' verbal self-disclosure. *Computer Animation and Virtual Worlds*, *21*(3–4), 473–482. https://doi.org/10.1002/cav.345

Kaplan, R. L., Damme, I. V., Levine, L. J., & Loftus, E. F. (2016). Emotion and False Memory. *Emotion Review*, *8*(1), 8–13. https://doi.org/10.1177/1754073915601228

Kaski, T., Niemi, J., & Pullins, E. (2018). Rapport building in authentic B2B sales interaction. *Industrial Marketing Management*, *69*, 235–252. https://doi.org/10.1016/j.indmarman.2017.08.019

Kassin, S. M., Leo, R. A., Meissner, C. A., Richman, K. D., Colwell, L. H., Leach, A. M., & Fon, D. L. (2007). Police interviewing and interrogation: A self-report survey of police practices and beliefs. *Law and Human Behavior*, *31*(4), 381–400. https://doi.org/10.1007/s10979-006-9073-5

Kebbell, M. R., & Milne, B. (1998). Police officers' perception of eyewitness factors in forensic investigations. *Journal of Social Psychology, 128*, 323–330.

Kebbell, M. R., Milne, B., & Wagstaff, G. F. (1999). The cognitive interview: A survey of its forensic effectiveness. *Psychology, Crime and Law*, *5*(1–2), 101–115. https://doi.org/10.1080/10683169908414996

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

Kieckhaefer, J. M., Vallano, J. P., & Compo, N. S. (2014). Examining the positive effects of rapport building: When and why does rapport building benefit adult eyewitness memory? *Memory*, *22*(8), 1010–1023. https://doi.org/10.1080/09658211.2013.864313

Kim, Y., & Thayne, J. (2015). Effects of learner–instructor relationship-building strategies in online video instruction. *Distance Education*, *36*(1), 100–114. https://doi.org/10.1080/01587919.2015.1019965

Kleider-Offutt, H. M., Cavrak, S. E., & Knuycky, L. R. (2015). Do Police Officers' Beliefs About Emotional Witnesses Influence the Questions They Ask? *Applied Cognitive Psychology*, *29*(2), 314–319. https://doi.org/10.1002/acp.3111

Köhnken, G., Milne, R., Memon, A., & Bull, R. (1999). The cognitive interview: A meta-analysis. Psychology, crime and law, 5(1-2), 3-27.

Kontogianni, F., Hope, L., Taylor, P. J., Vrij, A., & Gabbert, F. (2020). "Tell me more about this…": An examination of the efficacy of follow-up open questions following an initial account. *Applied Cognitive Psychology*, *34*(5), 972–983. https://doi.org/10.1002/acp.3675

Koriat, A., Goldsmith, M., & Pansky, A. (2000). Toward a Psychology of Memory Accuracy. *Annual Review of Psychology*, *51*(Volume 51, 2000), 481–537. https://doi.org/10.1146/annurev.psych.51.1.481

Krix, A. C., Sauerland, M., Gabbert, F., & Hope, L. (2014). Providing eyewitnesses with initial retrieval support: What works at immediate and subsequent recall? *Psychology, Crime and Law*, *20*(10), 1005–1027. https://doi.org/10.1080/1068316X.2014.902456

Krix, A. C., Sauerland, M., Raymaekers, L. H. C., Memon, A., Quaedflieg, C. W. E. M., & Smeets, T. (2016). Eyewitness Evidence Obtained with the Self-Administered Interview© Is Unaffected by Stress. *Applied Cognitive Psychology*, *30*(1), 103–112. https://doi.org/10.1002/acp.3173

Kruglanski, A. W. (2004). *The Psychology of Closed Mindedness*. Psychology Press.

Kuivaniemi-Smith, H. J., Nash, R. A., Brodie, E. R., Mahoney, G., & Rynn, C. (2014). Producing facial composite sketches in remote Cognitive Interviews: A preliminary investigation. *Psychology, Crime and Law*, *20*(4), 389–406. https://doi.org/10.1080/1068316X.2013.793339

Kwong See, S. T., Hoffman, H. G., & Wood, T. L. (2001). Perceptions of an old female eyewitness: Is the older eyewitness believable? *Psychology and Aging*, *16*(2), 346–350. https://doi.org/10.1037/0882-7974.16.2.346

La Rooy, D. J., Nicol, A., & Terry, P. (2013). Intoxicated Eyewitnesses: The Effects of Alcohol on Eyewitness Recall across Repeated Interviews. *Open Journal of Medical Psychology*, *02*(03), 107–114. https://doi.org/10.4236/ojmp.2013.23017

Lacy, J. W., & Stark, C. E. L. (2013). The neuroscience of memory: Implications for the courtroom. *Nature Reviews Neuroscience*, *14*(9), 649–658. https://doi.org/10.1038/nrn3563

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), 710–718. https://doi.org/10.1037/amp0000039

Launay, C., Brunel, M., & Bull, R. (2021). What Do Police Interview Eyewitnesses for? A Review. *Policing: A Journal of Policy and Practice*, *16*. https://doi.org/10.1093/police/paab081

Lavis, T., & Brewer, N. (2017). Effects of a proven error on evaluations of witness testimony. *Law and Human Behavior*, *41*(3), 314–323. https://doi.org/10.1037/lhb0000217

Legg, M., & Song, A. (2021). The courts, the remote hearing and the pandemic: From action to reflection. *University of New South Wales Law Journal, The*. https://search.informit.org/doi/abs/10.3316/informit.688022482335487

Leippe, M. R., Eisenstadt, D., Rauch, S. M., & Stambush, M. A. (2006). Effects of social-comparative memory feedback on eyewitnesses' identification confidence, suggestibility, and retrospective memory reports. *Basic and Applied Social Psychology*, *28*(3), 201–220. https://doi.org/10.1207/s15324834basp2803_1

Lind, L. H., Schober, M. F., Conrad, F. G., & Reichert, H. (2013). Why do survey respondents disclose more when computers ask the questions? *Public Opinion Quarterly*, *77*(4), 888–935. https://doi.org/10.1093/poq/nft038

Lindsay, D. S., Allen, B. P., Chan, J. C. K., & Dahl, L. C. (2004). Eyewitness suggestibility and source similarity: Intrusions of details from one event into memory reports of another event. *Journal of Memory and Language*, *50*(1), 96–111. https://doi.org/10.1016/j.jml.2003.08.007

List, J. A. (1986). Age and schematic differences in the reliability of eyewitness testimony. *Developmental Psychology*, *22*(1), 50–57. https://doi.org/10.1037/0012-1649.22.1.50

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *12*(4), 361–366. https://doi.org/10.1101/lm.94705

Loftus, E. F. (2019). Eyewitness testimony. *Applied Cognitive Psychology*, *33*(4), 498–503. https://doi.org/10.1002/acp.3542

Loftus, E. F., & Greenspan, R. L. (2017). If I'm Certain, Is It True? Accuracy and Confidence in Eyewitness Memory. *Psychological Science in the Public Interest*, *18*(1), 1–2. https://doi.org/10.1177/1529100617699241

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, *13*(5), 585–589. https://doi.org/10.1016/S0022-5371(74)80011-3

Luger, E., & Sellen, A. (2016). 'Like having a really bad pa': The gulf between user expectation and experience of conversational agents. *Conference on Human Factors in Computing Systems - Proceedings*, 5286–5297. https://doi.org/10.1145/2858036.2858288

Luke, T. J., Crozier, W. E., & Strange, D. (2017). Memory errors in police interviews: The bait question as a source of misinformation. *Journal of Applied Research in Memory and Cognition*, *6*(3), 260–273. https://doi.org/10.1016/j.jarmac.2017.01.011

Luknar, I., Bull, R., & Stanojević, P. (2023). Improving the preparation of the investigative interviewing of suspects: A multidisciplinary perspective. *Socioloski Pregled*, *57*, 1235–1256. https://doi.org/10.5937/socpreg57-46448

Luna, K., & Martín-Luengo, B. (2012). Improving the accuracy of eyewitnesses in the presence of misinformation with the plurality option. *Applied Cognitive Psychology*, *26*(5), 687–693. https://doi.org/10.1002/acp.2845

Lynch, K. R., Wasarhaley, N. E., Golding, J. M., & Simcic, T. (2013). Who Bought the Drinks? Juror Perceptions of Intoxication in a Rape Trial. *Journal of Interpersonal Violence*, *28*(16), 3205–3222. https://doi.org/10.1177/0886260513496900

MacDonald, S., Snook, B., & Milne, B. (2017). Witness Interview Training: A Field Evaluation. *Journal of Police and Criminal Psychology*, *32*(1), 77–84. https://doi.org/10.1007/s11896-016-9197-6

Markson, L., & Paterson, K. B. (2009). Effects of gaze-aversion on visual-spatial imagination. *British Journal of Psychology*, *100*(3), 553–563. https://doi.org/10.1348/000712608X371762

Marsh, E. J. (2007). Retelling is not the same as recalling: Implications for memory. *Current Directions in Psychological Science*, *16*(1). https://doi.org/10.1111/j.1467-8721.2007.00467.x

Marsh, E. J., & Tversky, B. (2004). Spinning the stories of our lives. *Applied Cognitive Psychology*, *18*(5), 491–503. https://doi.org/10.1002/acp.1001

Marsh, E. J., Tversky, B., & Hutson, M. (2005). How eyewitnesses talk about events: Implications for memory. *Applied Cognitive Psychology*, *19*(5), 531–544. https://doi.org/10.1002/acp.1095

Matsuo, K., & Miura, H. (2017). Effectiveness of the Self-Administered Interview and Drawing Pictures for Eliciting Eyewitness Memories. *Psychiatry, Psychology and Law*, *24*(5), 643–654. https://doi.org/10.1080/13218719.2016.1254587

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *The Academy of Management Review*, *20*(3), 709. https://doi.org/10.2307/258792

McGaugh, J. L. (2000). Memory-a century of consolidation. *Science*, *287*(5451), 248–251.

McGroarty, A., & Baxter, J. S. (2009). Interviewer behaviour, interviewee self-esteem and response change in simulated forensic interviews. *Personality and Individual Differences*, *47*(6), 642–646. https://doi.org/10.1016/j.paid.2009.05.024

Mello, E. W., & Fisher, R. P. (1996). Enhancing older adult eyewitness memory with the cognitive interview. *Applied Cognitive Psychology*, *10*(5), 403–417.

Memon, A., Bartlett, J., Rose, R., & Gray, C. D. (2003). The ageing eyewitness: The effects of face-age and delay upon younger and older observers. *Gerontology*, *58*, 338–345.

Memon, A., Bull, R., & Smith, M. (1995). Improving the quality of the police interview: Can training in the use of cognitive techniques help? *Policing and Society*, *5*(1), 53–68. https://doi.org/10.1080/10439463.1995.9964710

Memon, A., Gabbert, F., & Hope, L. (2013). The ageing eyewitness. In *Forensic psychology* (pp. 96–112). Willan.

Memon, A., Meissner, C. A., & Fraser, J. (2010). The cognitive interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, and Law*, *16*(4), 340. https://doi.org/10.1037/a0020518

Memon, A., Wark, L., Bull, R., & Koehnken, G. (2018). *Isolating the effects of the cognitive interview techniques: Selected Works of Ray Bull* (pp. 36–55). https://doi.org/10.4324/9781315169910-4

Merritt, S. M., & Ilgen, D. R. (2008). Not all trust is created equal: Dispositional and history-based trust in human-automation interactions. *Human Factors*, *50*(2), 194–210. https://doi.org/10.1518/001872008X288574

Metcalfe, J., & Shimamura, A. P. (Eds.). (1994). *Metacognition: Knowing about Knowing*. The MIT Press. https://doi.org/10.7551/mitpress/4561.001.0001

Miles, O., West, R., & Nadarzynski, T. (2021). Health chatbots acceptability moderated by perceived stigma and severity: A cross-sectional survey. *Digital Health*, *7*. https://doi.org/10.1177/20552076211063012

Milne, B., & Bull, R. (2002). Back to basics: A componential analysis of the original cognitive interview with three age groups. *Applied Cognitive Psychology*, *16*(7), 743–753. https://doi.org/10.1002/acp.825

Milne, B., & Bull, R. (2003). Does the cognitive interview help children to resist the effects of suggestive questioning? *Legal and Criminological Psychology*, *8*(1), 21–38. https://doi.org/10.1348/135532503762871219

Milne, B., Bull, R., & Wilcock, R. (2008). *Criminal identification by witnesses: Psychology and practice*. Oxford University Press.

Milne, B., Clare, I. C. H., & Bull, R. (1999). Interviewing adults with learning disability with the cognitive interview. *Psychology, Crime and Law*, *5*, 81–100.

Milne, B., Griffiths, A., Clarke, C., & Dando, C. J. (2019). The Cognitive Interview: A Tiered Approach in the Real World. In *Evidence-based Investigative Interviewing*. Routledge.

Milne, B., & Shaw, G. (1999). Obtaining witness statements: The psychology, best practice and proposals for innovation. *Medicine, Science and the Law*, *39*(2), 127–138. https://doi.org/10.1177/002580249903900207

Milne, B., Shaw, G., & Bull, R. (2008). *Investigative Interviewing: The Role of Research* (pp. 65–80). https://doi.org/10.1002/9780470713068.ch4

Minhas, R., Elphick, C., & Shaw, J. (2022). *Protecting victim and witness statement: Examining the effectiveness of a chatbot that uses artificial intelligence and a cognitive interview*. *37*, 265–281. https://doi.org/10.1007/s00146-021-01165-5

Ministry of Justice. (2022). *Achieving Best Evidence in Criminal Proceedings*. 246.

Mintzer, M. Z., & Griffiths, R. R. (2001). Alcohol and false recognition: A dose-effect study. *Psychopharmacology*, *159*(1), 51–57. https://doi.org/10.1007/s002130100893

Mirandola, C., Toffalini, E., Grassano, M., Cornoldi, C., & Melinder, A. (2014). Inferential false memories of events: Negative consequences protect from distortions when the events are free from further elaboration. *Memory*, *22*(5), 451–461. https://doi.org/10.1080/09658211.2013.795976

Mitzner, T. L., Savla, J., Boot, W. R., Sharit, J., Charness, N., Czaja, S. J., & Rogers, W. A. (2019). Technology Adoption by Older Adults: Findings From the PRISM Trial. *The Gerontologist*, *59*(1), 34–44. https://doi.org/10.1093/geront/gny113

Mone, G. (2016). The Edge of the Uncanny Scientists are learning more about what makes robots and chatbots engaging. *Communications of the ACM*, *59*(9), 17–19. https://doi.org/10.1145/2967977

Morgan, C. A., Hazlett, G., Doran, A., Garrett, S., Hoyt, G., Thomas, P., Baranoski, M., & Southwick, S. M. (2004). Accuracy of eyewitness memory for persons

encountered during exposure to highly intense stress. *International Journal of Law and Psychiatry*, *27*(3), 265–279. https://doi.org/10.1016/j.ijlp.2004.03.004

Morgan, C. A., Southwick, S., Steffian, G., Hazlett, G. A., & Loftus, E. F. (2013). Misinformation can influence memory for recently experienced, highly stressful events. *International Journal of Law and Psychiatry*, *36*(1), 11–17. https://doi.org/10.1016/j.ijlp.2012.11.002

Mudd, K., & Govern, J. M. (2004). Conformity to Misinformation and Time Delay Negatively Affect Eyewitness Confidence and Accuracy. In *North American Journal of Psychology* (Vol. 6, Issue 2, pp. 227–238).

Murphy, G., & Greene, C. M. (2016). Perceptual load affects eyewitness accuracy and susceptibility to leading questions. *Frontiers in Psychology*, *7*(AUG). https://doi.org/10.3389/fpsyg.2016.01322

Mygland, M. J., Schibbye, M., Pappas, I. O., & Vassilakopoulou, P. (2021). Affordances in Human-Chatbot Interaction: A Review of the Literature. In D. Dennehy, A. Griva, N. Pouloudi, Y. K. Dwivedi, I. Pappas, & M. Mäntymäki (Eds.), *Responsible AI and Analytics for an Ethical and Inclusive Digitized Society* (pp. 3–17). Springer International Publishing. https://doi.org/10.1007/978-3-030-85447-8_1

Nadarzynski, T., Llewellyn, J., Kidsley, C., & Graham, S. (2020). *Acceptability of artificial intelligence (AI)-enabled chatbots, video consultations and live webchats as online platforms for sexual health advice*. https://doi.org/10.1136/bmjsrh-2018-200271

Nadarzynski, T., Puentes, V., Pawlak, I., Mendes, T., Montgomery, I., Bayley, J., & Ridge, D. (2021). Barriers and facilitators to engagement with artificial intelligence (AI)-based chatbots for sexual and reproductive health advice: A

qualitative analysis. *Sexual Health*, *18*(5), 385–393. https://doi.org/10.1071/SH21123

Nahouli, Z., Dando, C. J., Mackenzie, J.-M., & Aresti, A. (2021). Rapport building and witness memory: Actions may 'speak' louder than words. *PLOS ONE*, *16*(8), e0256084. https://doi.org/10.1371/JOURNAL.PONE.0256084

Nahouli, Z., Mackenzie, J.-M., Aresti, A., & Dando, C. J. (2023). Rapport building with offenders in probation supervision: The views of English probation practitioners. *Probation Journal*, *70*, 026455052211374.

Naji, M., Krishnan, G. P., McDevitt, E. A., Bazhenov, M., & Mednick, S. C. (2019). Coupling of autonomic and central events during sleep benefits declarative memory consolidation. *Neurobiology of Learning and Memory*, *157*, 139–150. https://doi.org/10.1016/j.nlm.2018.12.008

Nash, R. A., Houston, K. A., Ryan, K., & Woodger, N. (2014). Remembering remotely: Would video-mediation impair witnesses' memory reports? *Psychology, Crime and Law*, *20*(8), 756–768. https://doi.org/10.1080/1068316X.2013.857669

Nordheim, C. B., Følstad, A., & Bjørkli, C. A. (2019). An Initial Model of Trust in Chatbots for Customer Service—Findings from a Questionnaire Study. *Interacting with Computers*, *31*(3), 317–335. https://doi.org/10.1093/IWC/IWZ022

Nunan, J., Milne, B., Shawyer, A., Wakefield, A., & Dalton, G. (2016). Don't break the chain: From effective communication to successful decision-making. *Emergency Services Times*, *17*(6), 50–52.

Office for National Statistics. (2020). *Local authority ageing statistics, population projections for older people*. https://www.ons.gov.uk/datasets/ageing-population-projections/editions/time-series/versions/1

Omarzu, J. (2000). A Disclosure Decision Model: Determining How and When Individuals Will Self-Disclose. *Personality & Social Psychology Review (Lawrence Erlbaum Associates)*, *4*(2), 174–185. https://doi.org/10.1207/S15327957PSPR0402_05

ONS. (2020). *Internet access – households and individuals, Great Britain: 2020*. https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/bulletins/internetaccesshouseholdsandindividuals/2020

ONS. (2021). *Profile of the older population living in England and Wales in 2021 and changes since 2011—Office for National Statistics*. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing/articles/profileoftheolderpopulationlivinginenglandandwalesin2021andchangessince2011/2023-04-03#population-ageing

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 and the Law*, *17*(1), 45–66. https://doi.org/10.1558/ijsll.v17i1.45

Pachman, M., & Ke, F. (2012). Environmental support hypothesis in designing multimedia training for older adults: Is less always more? *Computers & Education*, *58*(1), 100–110. https://doi.org/10.1016/j.compedu.2011.08.011

Palmer, F. T., Flowe, H. D., Takarangi, M. K. T., & Humphries, J. E. (2013). Intoxicated witnesses and suspects: An archival analysis of their involvement

in criminal case processing. *Law and Human Behavior*, *37*(1), 54–59. https://doi.org/10.1037/lhb0000010

Pansky, A. (2012). Inoculation against forgetting: Advantages of immediate versus delayed initial testing due to superior verbatim accessibility. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *38*(6), 1792–1800. https://doi.org/10.1037/a0028460

Pansky, A., & Nemets, E. (2012). Enhancing the quantity and accuracy of eyewitness memory via initial memory testing. *Journal of Applied Research in Memory and Cognition*, *1*(1), 2–10. https://doi.org/10.1016/j.jarmac.2011.06.001

Paterson, H. M., & Kemp, R. I. (2006). Co-witnesses talk: A survey of eyewitness discussion. *Psychology, Crime & Law*, *12*(2), 181–191. https://doi.org/10.1080/10683160512331316334

Paterson, H. M., Kemp, R. I., & Ng, J. R. (2011). Combating Co-witness contamination: Attempting to decrease the negative effects of discussion on eyewitness memory. *Applied Cognitive Psychology*, *25*(1), 43–52. https://doi.org/10.1002/acp.1640

Paterson, H. M., Van Golde, C., Devery, C., Cowdery, N., & Kemp, R. (2018). iWitnessed: Capturing contemporaneous accounts to enhance witness evidence. *Current Issues in Criminal Justice*, 29(3), 273-281.

Paulo, R. M., Albuquerque, P. B., & Bull, R. (2016). The enhanced cognitive interview: Expressions of uncertainty, motivation and its relation with report accuracy. *Psychology, Crime & Law*, *22*(4), 366–381. https://doi.org/10.1080/1068316X.2015.1109089

Paulo, R. M., Albuquerque, P. B., & Bull, R. (2019). Witnesses' Verbal Evaluation of Certainty and Uncertainty During Investigative Interviews: Relationship with Report Accuracy. *Journal of Police and Criminal Psychology*, *34*(4), 341–350. https://doi.org/10.1007/s11896-019-09333-6

Paulo, R. M., Albuquerque, P., & Bull, R. (2013). The Enhanced Cognitive Interview: Towards a Better Use and Understanding of This Procedure. *International Journal of Police Science & Management*, *15*, 190–199. https://doi.org/10.1350/ijps.2013.15.3.311

Paulo, R. M., Jones, E., & Mendes, R. (2021). Testing two retrieval strategies to enhance eyewitness memory: Category and location clustering recall. *Psychology, Public Policy, and Law*, *27*(1), 140–149. https://doi.org/10.1037/law0000281

Penrod, S. D., Fulero, S. M., & Cutler, B. L. (1995). Expert psychological testimony on eyewitness reliability before and after Daubert: The state of the law and the science. *Behavioral Sciences & the Law*, *13*(2), 229–259. https://doi.org/10.1002/bsl.2370130206

Perfect, T. J., Wagstaff, G. F., Moore, D., Andrews, B., Cleveland, V., Newcombe, S., Brisbane, K. A., & Brown, L. (2008). How can we help witnesses to remember more? It's an (eyes) open and shut case. *Law and Human Behavior*, *32*(4), 314–324. https://doi.org/10.1007/s10979-007-9109-5

Pfeil, K. (2018). *The Effectiveness of the Self-Administered Interview©—A Meta-Analytic Review and Empirical Study with Older Adult Witnesses* [University of Cambridge].

Pica, G., Chernikova, M., Pierro, A., Giannini, A. M., & Kruglanski, A. W. (2018). Retrieval-induced forgetting as motivated cognition. *Frontiers in Psychology*, 9, Article 2030.

Pickard, M. D., & Roster, C. A. (2020). Using computer automated systems to conduct personal interviews: Does the mere presence of a human face inhibit disclosure? *Computers in Human Behavior*, *105*. https://doi.org/10.1016/j.chb.2019.106197

Rapp, A., Curti, L., & Boldi, A. (2021). The human side of human-chatbot interaction: A systematic literature review of ten years of research on text-based chatbots. *International Journal of Human-Computer Studies*, *151*, 102630. https://doi.org/10.1016/j.ijhcs.2021.102630

Reisberg, D., & Heuer, F. (2014). The influence of emotion on memory in forensic settings. In *Handbook Of Eyewitness Psychology 2 Volume Set* (pp. 81–116). Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9781315805535-3/influence-emotion-memory-forensic-settings-daniel-reisberg-friderike-heuer

Ridley, A. M., Gabbert, F., & La Rooy, D. J. (2012). Suggestibility in Legal Contexts: Psychological Research and Forensic Implications. In *Suggestibility in Legal Contexts: Psychological Research and Forensic Implications*. https://doi.org/10.1002/9781118432907

Risan, P., Binder, P. E., & Milne, B. (2016). Emotional Intelligence in Police Interviews—Approach, Training and the Usefulness of the Concept. *Journal of Forensic Psychology Practice*, *16*(5), 410–424. https://doi.org/10.1080/15228932.2016.1234143

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*(1), 20–27. https://doi.org/10.1016/j.tics.2010.09.003

Roediger, H. L., & DeSoto, K. A. (2014). Confidence and memory: Assessing positive and negative correlations. *Memory*, *22*(1), 76–91. https://doi.org/10.1080/09658211.2013.795974

Roediger, H. L., & Geraci, L. (2007). Aging and the misinformation effect: A neuropsychological analysis. *Journal of Experimental Psychology: Learning Memory and Cognition*, *33*(2), 321–334. https://doi.org/10.1037/0278-7393.33.2.321

Roediger, H. L., & Karpicke, J. D. (2006). The Power of Testing Memory: Basic Research and Implications for Educational Practice. *Perspectives on Psychological Science*, *1*(3), 181–210. https://doi.org/10.1111/j.1745-6916.2006.00012.x

Ross, D. F., Dunning, D., Toglia, M. P., & Ceci, S. J. (1990). The child in the eyes of the jury: Assessing mock jurors' perceptions of the child witness. *Law and Human Behavior*, *14*(1), 5–23. https://doi.org/10.1007/BF01055786

Rubin, Z. (1975). Disclosing oneself to a stranger: Reciprocity and its limits. *Journal of Experimental Social Psychology*, *11*(3), 233–260. https://doi.org/10.1016/S0022-1031(75)80025-4

Saraiva, R. B., Hope, L., Horselenberg, R., Ost, J., Sauer, J. D., & van Koppen, P. J. (2020). Using metamemory measures and memory tests to estimate eyewitness free recall performance. *Memory*, *28*(1), 94–106. https://doi.org/10.1080/09658211.2019.1688835

Sauer, J. D., Brewer, N., Zweck, T., & Weber, N. (2010). The effect of retention interval on the confidence-accuracy relationship for eyewitness identification. *Law and Human Behavior*, *34*(4), 337–347. https://doi.org/10.1007/s10979-009-9192-x

Sauer, J. D., Palmer, M. A., & Brewer, N. (2017). Mock-juror evaluations of traditional and ratings-based eyewitness identification evidence. *Law and Human Behavior*, *41*(4), 375–384. https://doi.org/10.1037/lhb0000235

Sauerland, M., Krix, A. C., Kan, N. van, Glunz, S., & Sak, A. (2014). Speaking is silver, writing is golden? The role of cognitive and social factors in written versus spoken witness accounts. *Memory and Cognition*, *42*(6), 978–992. https://doi.org/10.3758/s13421-014-0401-6

Sauerland, M., & Sporer, S. L. (2011). Written vs. Spoken Eyewitness Accounts: Does Modality of Testing Matter Research Report. *Behavioral Sciences & the Law*, *29*(6), 846–857.

Sayers, H. (2004). Desktop virtual environments: A study of navigation and age. *Interacting with Computers*, *16*(5), 939–956. https://doi.org/10.1016/j.intcom.2004.05.003

Schreiber Compo, N., Carol, R. N., Evans, J. R., Pimentel, P., Holness, H., Nichols-Lopez, K., Rose, S., & Furton, K. G. (2017). Witness memory and alcohol: The effects of state-dependent recall. *Law and Human Behavior*, *41*(2), 202–215. https://doi.org/10.1037/lhb0000224

Schreiber Compo, N., Evans, J. R., Carol, R. N., Kemp, D., Villalba, D., Ham, L. S., & Rose, S. (2011). Alcohol intoxication and memory for events: A snapshot of alcohol myopia in a real-world drinking scenario. *Memory*, *19*(2), 202–210. https://doi.org/10.1080/09658211.2010.546802

Scoboria, A., & Fisico, S. (2013). Encouraging and clarifying 'Don't Know' responses enhances interview quality. *Journal of Experimental Psychology: Applied*, *19*(1), 72–82. https://doi.org/10.1037/a0032067

Shapira, A. A., & Pansky, A. (2019). Cognitive and metacognitive determinants of eyewitness memory accuracy over time. *Metacognition and Learning*, *14*(3), 437–461. https://doi.org/10.1007/s11409-019-09206-7

Sharit, J., Czaja, S. J., Nair, S., & Lee, C. C. (2003). Effects of age, speech rate, and environmental support in using telephone voice menu systems. *Human Factors*, *45*(2), 234–251.

Sharman, S. J., & Powell, M. B. (2012). A Comparison of Adult Witnesses' Suggestibility Across Various Types of Leading Questions. *Applied Cognitive Psychology*, *26*(1), 48–53. https://doi.org/10.1002/acp.1793

Sheehy, N. P., & Chapman, A. J. (1989). Assessing the veracity of children's testimony. *Medicine and Law*, *8*(3), 311–315.

Shepherd, E., & Milne, B. (1999). Full and faithful: Ensuring quality practice and integrity of outcome in witness interviews. In *Analysing witness testimony: A guide for legal practitioners and other professionals* (pp. 124–145).

Simons, D. J., & Chabris, C. F. (2011). What people believe about how memory works: A representative survey of the U.S. population. *PLoS ONE*, *6*(8). https://doi.org/10.1371/journal.pone.0022757

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

Smith, M. E., Loschky, L. C., & Bailey, H. R. (2021). Knowledge guides attention to goal-relevant information in older adults. *Cognitive Research: Principles and Implications*, *6*(1), 56. https://doi.org/10.1186/s41235-021-00321-1

Smith, V. L., & Ellsworth, P. C. (1987). The Social Psychology of Eyewitness Accuracy: Misleading Questions and Communicator Expertise. *Journal of Applied Psychology*, *72*(2), 294–300. https://doi.org/10.1037/0021-9010.72.2.294

Sousa, W. S., & Jaeger, A. (2022). Memory conformity for high-confidence recognition of faces. *Memory & Cognition*, *50*(6), 1147–1156. https://doi.org/10.3758/s13421-022-01325-y

Spanoudaki, E., Ioannou, M., Synnott, J., Tzani-Pepelasi, C., & Pylarinou, N. R. (2019). Investigative decision-making: Interviews with detectives. *Journal of Criminal Psychology*, *9*(2), 88–107. https://doi.org/10.1108/JCP-02-2018-0005

Sporer, S. L., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, Confidence, and Accuracy: A Meta-Analysis of the Confidence-Accuracy Relation in Eyewitness Identification Studies. *Psychological Bulletin*, *118*(3), 315–327. https://doi.org/10.1037/0033-2909.118.3.315

Staudinger, U. M. (2020). The positive plasticity of adult development: Potential for the 21st century. *American Psychologist*, *75*(4), 540–553. https://doi.org/10.1037/amp0000612

Sternberg, K. J., Lamb, M. E., Davies, G. M., & Westcott, H. L. (2001). The Memorandum of Good Practice: Theory versus application. *Child Abuse and Neglect*, *25*(5), 669–681. https://doi.org/10.1016/S0145-2134(01)00232-0

Stine, E. L., & Wingfield, A. (1987). Process and strategy in memory for speech among younger and older adults. *Psychology and Aging*, *2*(3), 272–279. https://doi.org/10.1037/0882-7974.2.3.272

Storm, B. C. (2011). The benefit of forgetting in thinking and remembering. *Current Directions in Psychological Science*, 20(5), 291–295.

Storm, B. C., Angello, G., Buchli, D. R., Koppel, R. H., Little, J. L., & Nestojko, J. F. (2015). A review of retrieval-induced forgetting in the contexts of learning, eyewitness memory, social cognition, autobiographical memory, and creative cognition. *Psychology of Learning and Motivation*, 62, 141–194.

Storm, B. C., & Jobe, T. A. (2012). Retrieval-induced forgetting predicts failure to recall negative autobiographical memories. *Psychological Science*, 23(11), 1356–1363.

St-Yves, M. (2005). The psychology of rapport: FIve basic rules. In *Investigative Interviewing*. Willan.

Suler, J. (2004). The Online Disinhibition Effect. *CyberPsychology & Behavior*, *7*(3), 321–326. https://doi.org/10.1089/1094931041291295

Sutherland, R., & Hayne, H. (2001). The Effect of Postevent Information on Adults' Eyewitness Reports. *Applied Cognitive Psychology*, *15*(3), 249–263. https://doi.org/10.1002/acp.700

Tajfel, H., & Turner, J. (2004). The social identity theory of intergroup behavior. In *Political psychology* (pp. 276–293). Psychology Press.

Taylor, D., & Dando, C. J. (2018). Eyewitness memory in face-to-face and immersive avatar-to-avatar contexts. *Frontiers in Psychology*, *9*(APR). https://doi.org/10.3389/fpsyg.2018.00507

Thomas, A. K., Smith, A. M., & Mazerolle, M. (2020). The unexpected relationship between retrieval demands and memory performance when older adults are faced with age-related stereotypes. *The Journals of Gerontology: Series B*, *75*(2), 241–250.

Thorley, C., & Rushton-Woods, J. (2013). Blame conformity: Leading eyewitness statements can influence attributions of blame for an accident. *Applied Cognitive Psychology*, *27*(3), 291–296. https://doi.org/10.1002/acp.2906

Tickle-Degnen, L., & Rosenthal, R. (1990). The Nature of Rapport and Its Nonverbal Correlates. *Psychological Inquiry*, *1*(4), 285–293. https://doi.org/10.1207/s15327965pli0104_1

Toglia, M. P., Read, J. D., Ross, D. F., & Lindsay, R. C. L. (2017). The handbook of eyewitness psychology: Volume I: Memory for events. In *The Handbook of Eyewitness Psychology: Volume I: Memory for Events*. https://doi.org/10.4324/9781315086309

Tourangeau, R., & Yan, T. (2007). Sensitive Questions in Surveys. *Psychological Bulletin*, *133*(5), 859–883. https://doi.org/10.1037/0033-2909.133.5.859

Tucker, M. A., Humiston, G. B., Summer, T., & Wamsley, E. (2020). Comparing the Effects of Sleep and Rest on Memory Consolidation. *Nature and Science of Sleep*, *12*, 79–91. https://doi.org/10.2147/NSS.S223917

Tuckey, M. R., & Brewer, N. (2003). How schemas affect eyewitness memory over repeated retrieval attempts. *Applied Cognitive Psychology*, *17*(7), 785–800. https://doi.org/10.1002/acp.906

Tulving, E. (1984). Précis of Elements of episodic memory. *Behavioral and Brain Sciences*, *7*(2), 223–238. https://doi.org/10.1017/S0140525X0004440X

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

Tulving, E. (2005). Episodic memory and autonoesis: Uniquely human. *The Missing Link in Cognition: Origins of Self-Reflective Consciousness*, 3–56.

Tulving, E., & Craik, F. I. M. (2000). The Oxford handbook of memory. *Oxford University Press*.

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

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

Vallano, J. P., & Schreiber Compo, N. (2015). Rapport-building with cooperative witnesses and criminal suspects: A theoretical and empirical review. *Psychology, Public Policy, and Law*, *21*(1), 85–99. https://doi.org/10.1037/law0000035

van Oorsouw, K., & Merckelbach, H. (2012). The Effects of Alcohol on Crime-related Memories: A Field Study: Intoxication and memory for crime. *Applied Cognitive Psychology*, *26*(1), 82–90. https://doi.org/10.1002/acp.1799

Vredeveldt, A., Baddeley, A. D., & Hitch, G. J. (2014). The effectiveness of eye-closure in repeated interviews. *Legal and Criminological Psychology*, *19*(2), 282–295. https://doi.org/10.1111/lcrp.12013

Vredeveldt, A., Hitch, G. J., & Baddeley, A. D. (2011). Eyeclosure helps memory by reducing cognitive load and enhancing visualisation. *Memory and Cognition*, *39*(7), 1253–1263. https://doi.org/10.3758/s13421-011-0098-8

Wade, K., Nash, R., & Lindsay, D. (2018). Reasons to Doubt the Reliability of Eyewitness Memory: Commentary on Wixted, Mickes, and Fisher (2018). *Perspectives on Psychological Science*, *13*, 339–342. https://doi.org/10.1177/1745691618758261

Wagstaff, G. F. (2008). Hypnosis and the law: Examining the stereotypes. *Criminal Justice and Behavior*, *35*(10), 1277–1294. https://doi.org/10.1177/0093854808321669

Wagstaff, G. F., Wheatcroft, J., Cole, J. C., Brunas-Wagstaff, J., Blackmore, V., & Pilkington, A. (2008). Some cognitive and neuropsychological aspects of social inhibition and facilitation. *European Journal of Cognitive Psychology*, *20*(4), 828–846. https://doi.org/10.1080/09541440701469749

Walsh, D., & Bull, R. (2010). What really is effective in interviews with suspects? A study comparing interviewing skills against interviewing outcomes. *Legal and Criminological Psychology*, *15*(2), 305–321. https://doi.org/10.1348/135532509X463356

Wang, E., Paterson, H., & Kemp, R. (2014). The effects of immediate recall on eyewitness accuracy and susceptibility to misinformation. *Psychology, Crime and Law*, *20*(7), 619–634. https://doi.org/10.1080/1068316X.2013.854788

Waterhouse, G. F., Ridley, A. M., Bull, R., Satchell, L., & Wilcock, R. (2023). Rapport-building in multiple interviews of children. *Applied Cognitive Psychology*, *37*(6), 1210–1222. https://doi.org/10.1002/acp.4116

Weber, N., & Brewer, N. (2008). Eyewitness Recall: Regulation of Grain Size and the Role of Confidence. *Journal of Experimental Psychology: Applied*, *14*(1). https://doi.org/10.1037/1076-898X.14.1.50

Webster, W. S., Oxburgh, G. E., & Dando, C. J. (2021). The use and efficacy of question type and an attentive interviewing style in adult rape interviews. *Psychology, Crime & Law*, *27*(7), 656–677. https://doi.org/10.1080/1068316X.2020.1849694

Wegner, D. M., Erber, R., & Raymond, P. (1991). Transactive memory in close relationships. *Journal of Personality and Social Psychology*, *61*(6), 923–929. https://doi.org/10.1037/0022-3514.61.6.923

Weisband, S., & Kiesler, S. (1996). *Self disclosure on computer forms*. 3–10. https://doi.org/10.1145/238386.238387

Weller, S. (2017). Using internet video calls in qualitative (longitudinal) interviews: Some implications for rapport. *International Journal of Social Research Methodology*, *20*(6), 613–625. https://doi.org/10.1080/13645579.2016.1269505

Wells, G. L. (1995). Scientific Study of Witness Memory: Implications for Public and Legal Policy. *Psychology, Public Policy, and Law*, *1*(4), 726–731. https://doi.org/10.1037/h0092727

Wells, G. L., & Bradfield, A. L. (1998). 'Good, you identified the suspect': Feedback to eyewitnesses distorts their reports of the witnessing experience. *Journal of Applied Psychology*, *83*(3), 360–376. https://doi.org/10.1037/0021-9010.83.3.360

Wells, G. L., & Bradfield, A. L. (1999). Distortions in Eyewitnesses' Recollections: Can the Postidentification-Feedback Effect Be Moderated? *Psychological Science*, *10*(2), 138–144. https://doi.org/10.1111/1467-9280.00121

Wells, G. L., Kovera, M. B., Douglass, A. B., Brewer, N., Meissner, C. A., & Wixted, J. T. (2020). Policy and procedure recommendations for the collection and preservation of eyewitness identification evidence. *Law and Human Behavior*, *44*(1). https://doi.org/10.1037/lhb0000359

Wells, G. L., & Loftus, E. F. (2003). Eyewitness Memory for People and Events. In I. B. Weiner (Ed.), *Handbook of Psychology* (p. wei1109). John Wiley & Sons, Inc. https://doi.org/10.1002/0471264385.wei1109

Wells, G. L., Small, M., Penrod, S., Malpass, R., Fulero, S., & Brimacombe, C. (1998). Eyewitness Identification Procedures: Recommendations for Lineups and Photospreads. *Law and Human Behavior*, *22*. https://doi.org/10.1023/A:1025750605807

Wells, G. L., Steblay, N., & Dysart, J. (2015). Double-Blind Photo Lineups Using Actual Eyewitnesses: An Experimental Test of a Sequential Versus Simultaneous Lineup Procedure. *Law and Human Behavior*, *39*. https://doi.org/10.1037/lhb0000096

Westera, N. J., Kebbell, M. R., & Milne, B. (2011). Interviewing witnesses: Do investigative and evidential requirements concur? *British Journal of Forensic Practice*, *13*(2), 103–113. https://doi.org/10.1108/14636641111134341

Westera, N. J., Kebbell, M. R., Milne, B., & Green, T. (2016a). The prospective detective: Developing the effective detective of the future. *Policing and Society*, *26*(2), 197–209. https://doi.org/10.1080/10439463.2014.942845

Westera, N. J., Kebbell, M. R., Milne, B., & Green, T. (2016b). Towards a more effective detective. *Policing and Society*, *26*(1), 1–17. https://doi.org/10.1080/10439463.2014.912647

Wheatcroft, J. M., Wagstaff, G. F., & Kebbell, M. R. (2004). The influence of courtroom questioning style on actual and perceived eyewitness confidence and accuracy. *Legal and Criminological Psychology*, *9*(1), 83–101. https://doi.org/10.1348/135532504322776870

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychological Bulletin*, *121*(3), 331–354. https://doi.org/10.1037/0033-2909.121.3.331

Williams, S. E., Ford, J. H., & Kensinger, E. A. (2022). The power of negative and positive episodic memories. *Cognitive, Affective, & Behavioral Neuroscience*, *22*(5), 869–903. https://doi.org/10.3758/s13415-022-01013-z

Winter, E., Costello, A., O'Brien, M., & Hickey, G. (2021). Teachers' use of technology and the impact of Covid-19. *Irish Educational Studies*, *40*(2), 235–246. https://doi.org/10.1080/03323315.2021.1916559

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology*, *55*, 235–269. https://doi.org/10.1146/annurev.psych.55.090902.141555

Wixted, J. T., Mickes, L., Clark, S. E., Gronlund, S. D., & Roediger, H. L. (2015). Initial eyewitness confidence reliably predicts eyewitness identification accuracy. *American Psychologist*, *70*(6), 515–526. https://doi.org/10.1037/a0039510

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

Wixted, J. T., Read, J. D., & Lindsay, D. S. (2016). The Effect of Retention Interval on the Eyewitness Identification Confidence–Accuracy Relationship. *Journal of Applied Research in Memory and Cognition*, *5*(2), 192–203. https://doi.org/10.1016/j.jarmac.2016.04.006

Wixted, J. T., & Wells, G. L. (2017). The Relationship Between Eyewitness Confidence and Identification Accuracy: A New Synthesis. *Psychological Science in the Public Interest*, *18*(1), 10–65. https://doi.org/10.1177/1529100616686966

Wright, A. M., & Holliday, R. E. (2005). Police officers' perceptions of older eyewitnesses. *Legal and Criminological Psychology*, *10*, 211–223. https://doi.org/10.1348/135532505X37001

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. https://doi.org/10.1002/acp.1260

Wright, D. B. (1993). Misinformation and Warnings in Eyewitness Testimony: A New Testing Procedure to Differentiate Explanations. *Memory*, *1*(2), 153–166. https://doi.org/10.1080/09658219308258229

Wright, D. B., & Loftus, E. F. (1998). How Misinformation Alters Memories. *Journal of Experimental Child Psychology*, *71*(2), 155–164. https://doi.org/10.1006/jecp.1998.2467

Wright, D. B., Self, G., & Justice, C. (2000). Memory conformity: Exploring misinformation effects when presented by another person. *British Journal of Psychology*, *91*(2), 189–202. https://doi.org/10.1348/000712600161781

Wulff, A. N., & Thomas, A. K. (2021). The Dynamic and Fragile Nature of Eyewitness Memory Formation: Considering Stress and Attention. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.666724

Wylie, L. E., Patihis, L., McCuller, L. L., Davis, D., Brank, E. M., Loftus, E. F., & Bornstein, B. H. (2014). Misinformation effect in older versus younger adults: A meta-analysis and review. In *The Elderly Eyewitness in Court*. https://doi.org/10.4324/9781315813936

Yarmey, A. D. (1986). Verbal, Visual, and Voice Identification of a Rape Suspect Under Different Levels of Illumination. *Journal of Applied Psychology*, *71*(3), 363–370. https://doi.org/10.1037/0021-9010.71.3.363

Yaron-Antar, A., & Nachson, I. (2006). Collaborative remembering of emotional events: The case of Rabin's assassination. *Memory*, *14*(1), 46–56. https://doi.org/10.1080/09658210444000502

Yi, D. J., & Chun, M. M. (2005). Attentional modulation of learning-related repetition attenuation effects in human parahippocampal cortex. *Journal of Neuroscience*, *25*(14), 3593–3600. https://doi.org/10.1523/JNEUROSCI.4677-04.2005

Yuille, J. C., & Tollestrup, P. A. (1990). Some Effects of Alcohol on Eyewitness Memory. *Journal of Applied Psychology*, *75*(3), 268–273. https://doi.org/10.1037/0021-9010.75.3.268

Zaragoza, M. S., Belli, R. F., & Payment, K. E. (2013). Misinformation effects and the suggestibility of eyewitness memory. In *Do Justice and Let the Sky Fall:*

Elizabeth F. Loftus and her Contributions to Science, Law, and Academic Freedom. https://doi.org/10.4324/9780203774861

Zaragoza, M. S., & Lane, S. M. (1994). Source Misattributions and the Suggestibility of Eyewitness Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 934–945. https://doi.org/10.1037/0278-7393.20.4.934

Zekiroski, H., Powell, M. B., Brubacher, S. P., & Curtis, P. (2024). Evaluation of a police training programme designed to enhance open-ended questions with adult witnesses. *International Journal of Police Science & Management*, 14613557241233857. https://doi.org/10.1177/14613557241233857

Appendices

Appendix A

Ethics Application Decision Letter for Study 1 and Study 2

UNIVERSITY OF
FORWARD
THINKING
WESTMINSTER

Project title: Doctoral Research Project

Application ID: ETH1920-0542

Date: 11 Dec 2019

Dear Charlotte

I am writing to inform you that your application was considered by the Psychology Ethics Committee.

The proposal was approved.

Yours,

Prof. Coral Dando

Psychology Ethics Committee

I am advised by the Committee to remind you of the following points:

Your responsibility to notify the Research Ethics Committee immediately of any information received by you, or of which you become aware, which would cast doubt upon, or alter, any information contained in the original application, or a later amendment, submitted to the Research Ethics Committee and/or which would raise questions about the safety and/or continued conduct of the research.

The need to comply with the Data Protection Act 2018 and General Data Protection Regulation (GDPR) 2018.

The need to comply, throughout the conduct of the study, with good research practice standards.

The need to refer proposed amendments to the protocol to the Research Ethics Committee for further review and to obtain Research Ethics Committee approval thereto prior to implementation (except only in cases of emergency when the welfare of the subject is paramount).

The desirability of including full details of the consent form in an appendix to your research, and of addressing specifically ethical issues in your methodological discussion.

The requirement to furnish the Research Ethics Committee with details of the conclusion and outcome of the project, and to inform the Research Ethics Committee should the research be discontinued. The Committee would prefer a concise summary of the conclusion and outcome of the project, which would fit no more than one side of A4 paper, please.

Appendix B

Participation Information Sheet for Study 1 and Study 2

Text and spoken language collection from witness Interviews.

You are being invited to take part in a research project collecting language data from witness interviews. This project is being conducted by Charlotte Adam, supervised by Prof. Coral Dando, at the University of Westminster.

What will I be asked to do?

If you would like to take part in this research you will be asked to do the following:

- View a short one minute film of a non-violent mock crime.
- You will then complete a short memory task and a survey collecting demographic data.
- Then an interview will take place where you will be asked to explain what you can remember about the film. These interviews typically last between 30 to 40 minutes. You will be interviewed ether over the text or video function on skype. The interviews will be audio-recorded to allow us to analyse your memory of the event.
- Immediately following the interview you will then be asked to complete a short online survey on your experience of being interviewed (duration 5 minutes approximately).

The findings of this research will allow us to better understand the variety of language used by witnesses when being interviewed about a crime.

This research is being conducted in accordance with the University Of Westminster Code Of Ethical Conduct, and the BPS Code of ethics. These documents are available online:

https://www.bps.org.uk/news-and-policy/bps-code-ethics-and-conduct

https://www.westminster.ac.uk/research/research-framework/research-ethics

Throughout these processes your data will be labelled with your individual identification number, you will not be identifiable.

Your participation in this research is on an entirely voluntary basis, and you are able to withdraw without providing any reason, at any time up until the research has been published, or submitted in any form of a report (e.g., conference presentation, dissertation, etc.)

NOTE: We will not store any personal identifying data. Should you wish to withdraw at any time (until publication) simply refer to this document and contact us so that we can remove your contribution.

We will not be able to give feedback on individual performance, but we can provide all participants with a summary of the overall findings if requested

If you would like to take part in this research, in the first instance please contact:

Charlotte Adam – c.adam1@westminster.ac.uk

We will be very happy to provide further details, and to arrange a suitable date/time for you to participate.

Appendix C

Consent Sheet

Participant No. _____

I agree to participate in the research '**Spoken and text language collection from witness Interviews.'.** The research has been explained to my satisfaction, and I am aware that:

- My participation in this research is on an entirely voluntary basis
- I am able to stop at any point during data collection
- Once I have taken part, I am still able to withdraw **my data** at any point until the research has been published/submitted as part of my research project.
- In order to withdraw my data I will need to contact the researchers and quote my individual ID number, which has been provided to me.
- My data will be anonymised, and all identifying features will be removed so that my contribution will not be identifiable when reporting this research.
- My data will be securely stored, and destroyed in accordance with the General Data Protection Regulation 2018 and Data Protection Act 2018 (UK based Research).
- My identity and contact details will not be stored by the research team.
- I agree that my data from this study may be used for future research, and may undergo secondary analysis. Future research may be related or unrelated to the goals of this study.
- My anonymised data will be used as part of a larger data set by members of the research team.

Signed _____

Date _____

Participant No _____

Researcher _____

Appendix D

Demographic Survey for all Studies in this PhD

Q1. How old are you (years)?

Q2. What is the highest level of education you have completed?

Q3. In which country where you born?

Q4. Is English your first language?

Q5. How long have you spoken English for?

Q6. What is your sex?

Q7. Is your gender the same as the sex you were registered at birth?

Q8. If no, write in gender (optional)

Q9. Do you have eyesight problems not corrected with glasses?

Q10. Do you have hearing problems not corrected with a hearing aid?

Q11. How comfortable are you with typing with a keyboard on a computer?

Q12. How comfortable are you with reading from a electronic screen?

Appendix E

Mini-Mental States Examination

Standardized Mini-Mental State Examination (SMMSE)

DIRECTIONS FOR ADMINISTRATION OF THE SMMSE

1. Before the questionnaire is administered, try to get the person to sit down facing you. Assess the peron's ability to hear and understand very simple conversation, e.g. What is your name? If the person uses hearing or visual aids, provide these before starting.
2. Introduce yourself and try to get the person's confidence. Before you begin, get the person's permission to ask questions, e.g. Would it be all right to ask you same questions about your memory? This helps to avoid catastrophic reactions.
3. Ask each question a maximum of three times. If the subject does not respond, score 0.
4. If the person answers incorrectly, score 0. Accept that answer and do not ask the question again, hint, or provide any physical clues such as head shaking, etc.
5. The following equipment is required to administer the instrument: A watch, a pencil, Page 2 of this SMMSE with CLOSE YOUR EYES written in large letters and two five-sided figures intersecting to make a four-sided figure, and Page 3, a blank piece of paper.
6. If the person answers What did you say?, do not explain or engage in conversation. Merely repeat the same directions a maximum of three times.
7. If the person interrupts (e.g. What is this for?), just reply: I will explain in a few minutes, when we are finished. Now if we could proceed please... we are almost finished.

I am going to ask you some questions and give you same problems to solve. Please try to answer as best you can.

1. (Allow 10 seconds for each reply)

a) What year is this? (accept exact answer only) ..... / 1
b) What season is this? (during the last week of the old season or first week of a new season, accept either) ..... / 1
c) What month is this? (on the first day of a new month or the last day of the previous month, accept either) ..... / 1
d) What is today's date? (accept previous or next date)..... / 1
e) What day of the week is this? (accept exact answer only) ..... / 1

2. (Allow 10 seconds for each reply)

a) What country are we in? (accept exact answer only) ..... / 1
b) What province or state are we in? (accept exact answer only)..... / 1
c) What city/town are we in? (accept exact answer only) ..... / 1
d) (In home) What is the street address of this house?(accept street name and house number or equivalent in rural areas) / 1
(In facility) What is the name of this buliding? (accept exact name of institution only)..... / 1
e) (In home) What room are we in?(accept exact answer only)
(In facility) What floor of the buliding are we on?(accept exact answer only).....

3. Say: I am going to name three objects. When I am finished, I want you to repeat them. Remember what they are because I am going to ask you to name them again in a few minutes. (say slowly at approximatly one-second intervals)

Ball Car Man

For repeated use: Bell, jar, fan; Bill, tar, can; Bull, bar, pan
Please repeat the three items for me. (score one point for each correct reply on the first attempt) ..... / 3
Allow 20 seconds for reply; if the person did not repeat all three, repeat until they are learned or up to a maximum of five times. (but only score first attempt)

4. Spell the word WORLD. (you may help the person to spell the word correctly) Say: Now spell it backwards please.. / 5
(allow 30 seconds; if the subject cannot spell world even with assistance, score 0) Refer to Page 2 for scoring instructions

5. Say: *Now what were the three objects I asked you to remember?* ..... / 3
(score one point for each correct answer regardless of order; allow 10 seconds)

6. Show wristwatch. *Ask: What is this called?* ..... / 1
(score one point for correct response; accept "wristwatch" or "watch"; do not accept "clock" or "time", etc.; allow 10 seconds)

7. Show pencil. *Ask: What is this called?* ..... / 1
(score one point for correct response; accept "pencil" only; score 0 for pen; allow 10 seconds for reply)

8. Say: *I would like you to repeat a phrase after me: No ifs, ands, or buts*..... / 1
(allow 10 seconds for response. Score one point for a correct repetition. Must be exact, e.g. no ifs or buts, score 0)

9. Say: Read the words on this page and then do what it says..... / 1
Then, hand the person the sheet with CLOSE YOUR EYES on it. If the subject just reads and does not close eyes, you may repeat: Read the words on this page and then do what it says, (a maximum of three times. This is covered in #3 directions section above). Allow 10 seconds, score one point only if the subject closes eyes. The subject does not have to read aloud.

10. Hand the person a pencil and paper (Page 3). Say: Write any complete sentence on that piece of paper. Allow 30 seconds. Score one point. The sentence must make sense. Ignore spelling errors.

11. Place design, pencil, eraser and paper in front of the person. Say: Copy this design please. Allow multiple tries. Wait until the person is finished and hands it back. Score one point for a correctly copied diagram. The person must have drawn a four-sided figure between two five-sided figures. Maximum time: One minute. / 1

12. Ask the person if he is right or left handed. Take a piece of paper, hold it up in front of the person and say the following: Take this paper in your right/left hand (whichever is non-dominant), fold the paper in half once with both hands and put the paper down on the floor. / 1
Allow 30 seconds. Score one point for each instruction executed correctly.

Takes paper in correct hand..... / 1
Folds it in half..... / 1
Puts it on the floor..... / 1

TOTAL TEST SCORE: / 30
ADJUSTED SCORE : /

Molloy DW, Alemayehu E, Roberts R. Reliability of a standardized Mini-Mental State Examination compared with the traditional Mini-Mental state Examination. American Journal of Psychiatry, Vol. 14, 1991a, pp.102-105.

 Name :

Appendix F

Post Interview Survey Study 1 and Study 2

Q1. Overall, how easy did you find it to remember what happened in the video? (Extremely easy, Somewhat easy, Neither easy nor difficult, Somewhat difficult, Extremely difficult)

Q2. How easy was it to inform the interviewer when you could not remember? (Extremely easy, Somewhat easy, Neither easy nor difficult, Somewhat difficult, Extremely difficult)

Q3. How easy was it to inform the interviewer that you did not know the answer to a question? (Extremely easy, Somewhat easy, Neither easy nor difficult, Somewhat difficult, Extremely difficult)

Q4. Overall, how difficult did you find it to remember what happened in the video? (Extremely easy, Somewhat easy, Neither easy nor difficult, Somewhat difficult, Extremely difficult)

Q5. How difficult was it to inform the interviewer when you could not remember? (Extremely easy, Somewhat easy, Neither easy nor difficult, Somewhat difficult, Extremely difficult)

Q6. How difficult was it to inform the interviewer that you did not know the answer to a question? (Extremely easy, Somewhat easy, Neither easy nor difficult, Somewhat difficult, Extremely difficult)

Q7. Did you find the interviewer friendly? (Definitely yes, Probably yes, Might or might not, Probably not, Definitely not)

Q8. Did you find the interviewer easy to communicate with? (Definitely yes, Probably yes, Might or might not, Probably not, Definitely not)

Q9. Did you like the environment that you were interviewed in? (Definitely yes, Probably yes, Might or might not, Probably not, Definitely not)

Q10. I am confident that I remembered a lot of what I saw. (Not at all confident, Somewhat confident, Undecided, Quite confident, Completely confident)

Q11. I am confident that the information I gave at interview was correct. (Not at all confident, Somewhat confident, Undecided, Quite confident, Completely confident)

Q12. I am confident that I did not make any errors. (Not at all confident, Somewhat confident, Undecided, Quite confident, Completely confident)

Q13. Have you witnessed a crime before? (Yes, No)

Q14. Have you ever had a witness interview with a police officer? (Yes, No)

Appendix G

Debrief Sheet Study 1 and Study 2

Thank you for taking part in this research project collecting language data from witness interviews. You watched a short film and took part in an interviewed about what you could remember about the film. Face-to-face interviews were audio recorded to allow me to code and analyse your memory of the event, text interviews were downloaded from skype and removed from the service.

The findings of this research will give a better understanding the type of language used when people share memories from an event they have seen. It will also demonstrate the difference in language used when people have a face-to-face interview or when they are asked to remember over text interviews. The language data will be used to inform the future development of an automated witness interview system.

Throughout the interviews you were only referred to by your first name and/or your unique participant number. I will not store any personal identifying data, and so please keep this document safe because your personal participation code/number is included at the end. Should you wish to withdraw at any time (until publication) simply refer to this document and contact us so that we can remove your contribution stating this number.

If you do not have this number I will NOT be able to locate your data because it is completely anonymous.

I will not be able to give feedback on individual performance, but I can provide you with a summary of the overall findings if requested – please email (see below)

Charlotte Adam w1624134@my.westminster.ac.uk

Unique ID Number _____

Appendix H

Tier 1 PEACE Investigative Interview Script

“Hello, This is Charlotte Adam the researcher for Online Forensic Witness interviews. I will be calling you shortly, if that still works for you?”

- Video call participant make sure the sound/picture and text functions are working on skype.
- Introduce yourself and explain what is going to happen in the study
- allocate participant number
- End video call and message them the pre interview link with instructions to message when it is completed.

“Your participant number is XXXX. Please click on the link below and complete the consent form, watch the video and answer the following questions. If you have any questions send me a message and I will be happy to help. Please only watch the video once and don't take any notes.

https://westminsterpsych.az1.qualtrics.com/jfe/form/SV_6MdfqcAnDBjcXHv

When you have finished please message me back here and I will call you back.”

- Video Call Participant
- Digitally go through the MMSE
- Explain the interview will take place over text/spoken and hang up.
- Conduct Interview

Script:

To begin I am going to ask you some questions about the video you just watched. Do you have any questions before I start?

Ok, great. Before I start there are just a few very important ground rules that I want you to remember as you answer my questions.

I want you to tell me what you remember in as much detail as possible. I have not seen the film, so I do not know what it is about. I only want you to tell me what you actually remember. Do not guess. Say if you do not know the answer to my questions. Even if you can only remember partial information, or small pieces I want you to tell me. If you do not understand what I am asking then please say so, and I will rephrase it.
Does that make sense?

To start, I want you to type out everything you can remember about the video, in as much detail as possible. Whatever you can remember. Don't worry about taking your time writing everything out, just concentrate on remembering. OK, please begin.

When there has been no typing for 5 seconds type:
Is there anything else you can remember?

Thank you. I am now going to ask you some more detailed questions about the things you have just told me about. Before I start, I just want to remind you of the ground rules. I want you to tell me what you remember in as much detail as possible. I have not seen the film, so I do not know what it is about. I only want you to tell me what you actually remember. Do not guess. Say if you do not know the answer to my questions. Even if you can only remember partial information, or small pieces I want you to tell me. If you do not understand what I am asking then please say so, and I will rephrase it. When you are ready let me know and I will begin

Script Cued Recall

The first thing that you mentioned was XXXXXX

Think back to the film and tell me about that in as much detail as you can.

The next thing that you mentioned was XXXXXXX

Think back to the film and tell me about that in as much detail as you can.

Then you said XXXXXXXXX

try and remember the video, and add to that as much as you can?

End:

Thank you. I think I now have a good idea about what has happened. Just before we finish is there anything else you want to add or alter about what you have told me, anything at all?

Thank you that is the end of the interview. Please complete this post interview survey, let me know when you are done, and I will answer any questions you might have.

https://westminsterpsych.az1.qualtrics.com/jfe/form/SV_5clP6xmFNj5dXQp

Thank you for completing the survey and taking part in my study, do you have any questions before we finish? Would you like me to call you back to go over anything?

Great! thank you so much for taking part that is really kind of you,

I have noted your email and I will be sending off for your voucher, let me know if you don't receive it in the next seven days. Feel free to send me an email if you do have any questions.

Appendix I

Tier 1 PEACE Investigative Interview Protocol

PEACE Model	Interview Protocol
Planning and Preparation	- Video call participants and introduce the interviewer. - Make sure sound, video and text functions are working for the participant and the researcher.
Engage and Explain	- Explain what will happen in the Interview. - Give the participant an opportunity to ask questions. - Start the Interview. Either over synchronous text function or a recorded video interview. - Inform participant of the ground rules.
Account, Clarification, Challenge	- Ask the participant: Tell me everything you can remember about the video. - Take note of the key elements - When they have been silent for 5 seconds, ask if they remember anything else (repeat this until the participant indicates there is nothing else). - Inform the participant that they will now be asked some more detailed questions about what they remember. - Remind the participant of the ground rules. - Ask participants to tell you about key elements in as much detail as possible. - When they have been silent for 5 seconds, ask if there is anything else they can add. Repeat this until all key elements have been covered.
Closure	- Ask if there is anything further the participant would like to add - Ask if there is anything the participant would like to change about what they have shared. - Thank the participant, end the interview, and stop the recording. - Give the participant an opportunity to ask any further questions
Evaluation	- Transcribe and code Interviews

Appendix J

Excerpt From a coded Transcript from Study 1

I: To start, I want you to tell me everything you can remember about the video, in as much detail as possible. Whatever you can remember. Don't worry about taking your time writing everything out, just concentrate on remembering. OK, please begin.

TP006: The incident happened near a busy roundabout. There was a small blue shop that was selling newspaper. At some point a couple emerged from around the corner, a white male and a black female, they were around the same height. The female had long braids that were tied in a ponytail. They entered the shop and after around 15-20 seconds they came outside running and a white male in his 40s came out running after them. All three of them went around the corner, where the couple initially came from.

I: Is there anything else you can remember?

TP006: There was also a heavy traffic at the roundabout and at one point a police car went by. But that's it :)

I: Thank you. I am now going to ask you some more detailed questions about the things you have just told me about. The first thing that you mentioned was a busy roundabout. Think back to the film and tell me about that in as much detail as you can.

TP006: I can't remember much apart from the two cars that caught my attention: one was yellow and the other was a police car

I: ok great, The next thing that you mentioned was a small blue shop. Think back to the film and tell me about that in as much detail as you can.

TP006: It had a blue tent over it saying which newspapers they were selling. the name of the shop was written on either the tent or on the window of the shop, and was yellow in colour

I: You said there was a couple, a male and a female, can you tell me about the female in as much detail as possible?

TP006: She wore jeans and a jacket. She was almost the same height as the male. She had dark hair which was in long braids and tied up in a ponytail.

I: great! Now tell me about the male in as much detail as possible

TP006: He was a white slim male, with dark hair but I can't remember what clothes he was wearing.

I: thats ok, you said they came from around the corner, If you think back to the video is there anything else you can add to that?

TP006: No

I: The next thing that you mentioned was they entered the shop. Think back to the film and tell me about that in as much detail as you can.

TP006: They were talking to each other and were walking quite quickly, not just taking a walk. They didn't stay there long and after 15-20 seconds left the shop running

I: you said they came outside running If you think back to the video and tell me anything you can add to that?

TP006: There's nothing I can add to that, apart from that another white male, probably the owner of the shop, went out too chasing them

I: yes, can you tell me about him in as much detail as you can?

TP006: I can't remember much, only that he looked like he was in his 40s

I: The last thing that you mentioned was heavy traffic, think back to the film and tell me about that in as much detail as you can.

TP006: I can remember it being quite heavy in the beginning of the video but I can't remember if it remained the same by the end of it

I: Thank you. I think I now have a good idea about what has happened. Just before we finish is there anything else you want to add or alter about what you have told me, anything at all?

TP006: No

	Free Recall	Cued Recall	Overall
Correct	23	12	35
Inaccurate	2	0	2
Confabulation	1	0	1

Appendix K

In-person Initial Account Interview Script

Hello, My name is XXXX and I am going to ask you some questions about the video you just watched, are you ready?

To start, I want you to tell me everything you can remember about the video, in as much detail as possible. Whatever you can remember, even if you can only remember partial information. I want to know everything.

- Let the participant talk until they have finished

Is there anything else you would like to add? Anything at all, even if it seems silly or unimportant. I want to know everything.

- if the participant provides more information pause, and ask again

How confident are you that this information is correct?

- answer

Was there anything obstructing your view?

- answer

Approximately how far away were you from what you saw?

- answer

When did this happen?

- answer

Where did this happen?

- answer

How many people did you see?

- number (X)

I am now going to ask you some questions about the people you have seen. In your mind label them person X (*person X and person X*). *(depending on answer above)*

- Please describe person (number) in as much detail as possible.
    - answer
- Have you seen this person before?
    - answer
        - who is this person? (if yes)

- Would you be able to recognise this person? (if no)

Can you describe their face?

- answer

What was this person wearing?

- answer

What was this person wearing on their top half?

- answer

What was this person wearing on their bottom half?

- answer

What was this person wearing on their feet?

- answer

Did they have anything with them?

- answer

Anything further you can tell me about this person? Anything at all, even if it seems silly or unimportant. I want to know everything.

- answer

How confident are you that this information is correct?

- answer

Thank you. I think I now have a good idea of what has happened. Just before we finish is there anything else you want to tell me about what you have witnessed, anything at all?

- Answer

Is there anything you want to change about what you have already told me, anything at all?

- Answer

Thank you that is the end of the interview

Appendix L

Ethics Application Decision Letter for Study 3 and Study 4

UNIVERSITY OF
FORWARD
THINKING
WESTMINSTER

Project title: ChatCharlie and the First Instance Interview.

Application ID: ETH2122-2235

Date: 27 Apr 2022

Dear Charlotte

I am writing to inform you that your significant amendments to protocol were considered by the Psychology Ethics Committee.

The proposal was approved.

Yours,

Samuel Evans

Psychology Ethics Committee

I am advised by the Committee to remind you of the following points:

Your responsibility to notify the Research Ethics Committee immediately of any information received by you, or of which you become aware, which would cast doubt upon, or alter, any information contained in the original application, or a later amendment, submitted to the Research Ethics Committee and/or which would raise questions about the safety and/or continued conduct of the research.

The need to comply with the Data Protection Act 2018 and General Data Protection Regulation (GDPR) 2018.

The need to comply, throughout the conduct of the study, with good research practice standards.

The need to refer proposed amendments to the protocol to the Research Ethics Committee for further review and to obtain Research Ethics Committee approval thereto prior to implementation (except only in cases of emergency when the welfare of the subject is paramount).

The desirability of including full details of the consent form in an appendix to your research, and of addressing specifically ethical issues in your methodological discussion.

The requirement to furnish the Research Ethics Committee with details of the conclusion and outcome of the project, and to inform the Research Ethics Committee should the research be discontinued. The Committee would prefer a concise summary of the conclusion and outcome of the project, which would fit no more than one side of A4 paper, please.

Appendix M

Participant Information Sheet ChatCharlie Condition for Study 3 and Study 4

Can technology be used to help support witness memory?

You are being invited to take part in a research project on the use of technology for collecting information from witnesses. This project is being conducted by Charlotte Adam, supervised by Prof. Coral Dando, at the University of Westminster.

Witness information is vital for the criminal justice system. Traditionally, eyewitness information is collected face-to-face through an interview with a police officer. However, in some circumstances where there are large numbers of eyewitnesses it can take weeks and weeks to collect all eyewitness information face-to-face. This delay can interrupt police investigations and can also result in witnesses forgetting some of the detail that they have seen or heard.

Computer mediated communication to collect information from witnesses may be one answer. The aim of this research is to investigate whether remote computer mediated communication is an effective and efficient method for collecting information from witnesses.

What is a chatbot?

A chatbot is a computer program designed to simulate conversation with human users, especially over the internet. Although it can be like with a person chatbots are not human and have a limited set of responses and questions they can ask. Examples of text based chatbots include the ‘help’ functions on banking websites or shopping websites before speaking to a real person. Examples of chatbots that can speak are Amazon Alexa or Apple Siri. Chatbots can have their own websites where you can connect with your computer and have a conversation. Other chatbots are

integrated in other websites, for instance, often companies will create chatbots that can interact with service users through their Facebook page.

The requirements for taking part in this study are:

- To take part in this study you must be 18-55 years old
- You must speak fluent English.
- You must have access to a computer/laptop/tablet with a keyboard.
- You must have access to a computer/laptop/tablet with the ability to video call
- You must be able to take part on two occasions one week apart
- You must have normal or corrected-to-normal vision.

Please note: To take part in this study must speak fluent English.

If this is not the case, then unfortunately you will not be able to take part at this time. To receive a £10 Amazon voucher, you much take part in both parts of this study.

What will I be asked to do?

This study requires you to participate twice.

Please make sure you have the time to take part in both elements.

If you would like to take part in this research, you will be asked to do the following:

Time 1:

1. You will receive an email with 3 links on the day you have agreed to take part.
2. When you are ready follow the first link and complete the consent form, watch a short one-minute film of a non-violent mock crime on your computer/laptop/tablet, and complete the short demographic survey.
3. After this, you will click on the second link. This will take you to a chatbot that will conduct a short interview about the video you just watched.

4. Immediately following this interview, you will then be asked to click on the third link and complete a short online survey on your experience of being interviewed.

Time 2:

1. One week later you will take part in an Interview with a researcher about the video you have previously watched. This will take place face-to-face over video call. You will be asked to explain what you can remember about the film. These interviews typically last between 30 to 40 minutes. The interviews will be audio recorded to allow us to understand how much you can remember about the film event.

2. Immediately following the interview, you will then be asked to complete a short online survey on your experience of being interviewed (duration 5 minutes approximately).

Please note:

This research is being conducted in accordance with the University of Westminster Code of Ethical Conduct, and the BPS Code of ethics. These documents are available online:

https://www.bps.org.uk/news-and-policy/bps-code-ethics-and-conduct

https://www.westminster.ac.uk/research/research-framework/research-ethics

- Throughout these processes your data will be labelled with your individual identification number which will be provided to you, you will not be personally identifiable.
- Your participation in this research is on an entirely voluntary basis, and you are able to withdraw without providing any reason, at any time up until the research

has been published, or submitted in any form of a report (e.g., conference presentation, dissertation, etc.)

- We will not store any personal identifying data. Should you wish to withdraw at any time (until publication) simply contact us with your personal identification number so that we can remove your contribution.
- We will not be able to give feedback, but we can provide all participants with a summary of the overall findings if requested.

If you would like to take part in this research, we will be very happy to provide further details, and to arrange a suitable date/time for you to participate.

Please contact Charlotte Adam:

c.adam1@westminster.ac.uk

Appendix N

Participant Information Sheet Control Condition Study 3 and Study 4

Can technology be used to help support witness memory?

You are being invited to take part in a research project on the use of technology for collecting information from witnesses. This project is being conducted by Charlotte Adam, supervised by Prof. Coral Dando, at the University of Westminster.

Witness information is vital for the criminal justice system. Traditionally, eyewitness information is collected face-to-face interview by a police officer. However, in some circumstances where there are large numbers of eyewitnesses it can take weeks and weeks to collect all eyewitness information face-to-face. This delay can interrupt police investigations and can also result in witnesses forgetting some of the detail that they have seen or heard.

Computer mediated communication to collect information from witnesses may be one answer. The aim of this research is to investigate whether remote computer mediated communication is an effective and efficient method for collecting information from witnesses.

The requirements for taking part in this study are:

- To take part in this study you must be 18-55 years old
- You must speak fluent English as a first language.
- You must have access to a computer/laptop/tablet with a keyboard.
- You must have access to a computer/laptop/tablet with the ability to video call
- You must be able to take part on two occasions one week apart
- You must live in the United Kingdom
- You must have normal or corrected-to-normal vision.

Please note: To take part in this study must speak fluent English.

If this is not the case, then unfortunately you will not be able to take part at this time. To receive a £10 Amazon voucher, you much take part in both parts of this study.

What will I be asked to do?

This study requires you to participate twice.

Please make sure you have the time to take part in both elements.

If you would like to take part in this research, you will be asked to do the following:

Time 1:

1. You will receive an email on the day you have agreed to take part.

2. When you are ready follow the link from the email and complete the consent form, watch a short one-minute film of a non-violent mock crime on your computer/laptop/tablet, and complete the short demographic survey.

Time 2:

1. One week later you will take part in an Interview with a researcher about the video you have previously watched. This will take place face-to-face over video call. You will be asked to explain what you can remember about the film. These interviews typically last between 30 to 40 minutes. The interviews will be audio recorded to allow us to understand how much you can remember about the film event.

2. Immediately following the interview, you will then be asked to complete a short online survey on your experience of being interviewed (duration 5 minutes approximately).

Please note:

This research is being conducted in accordance with the University of Westminster Code of Ethical Conduct, and the BPS Code of ethics. These documents are available online:

https://www.bps.org.uk/news-and-policy/bps-code-ethics-and-conduct

https://www.westminster.ac.uk/research/research-framework/research-ethics

- Throughout these processes your data will be labelled with your individual identification number which will be provided to you, you will not be personally identifiable.
- Your participation in this research is on an entirely voluntary basis, and you are able to withdraw without providing any reason, at any time up until the research has been published, or submitted in any form of a report (e.g., conference presentation, dissertation, etc.)
- We will not store any personal identifying data. Should you wish to withdraw at any time (until publication) simply contact us with your personal identification number so that we can remove your contribution.
- We will not be able to give feedback, but we can provide all participants with a summary of the overall findings if requested.

If you would like to take part in this research, we will be very happy to provide further details, and to arrange a suitable date/time for you to participate.

Please contact Charlotte Adam:

c.adam1@westminster.ac.uk

Appendix O

Debrief Sheet Chapters 4 and 5

Can a ‘quick’ initial account interview help improve witness memory?

You have taken part in a research project concerned with collecting information from witnesses quickly, and whether this helps improve memory later. This project is being conducted by Doctoral Researcher Charlotte Adam, supervised by Prof. Coral Dando, at the University of Westminster.

What was the purpose of this study?

Witness information is vital for the criminal justice system. Traditionally, eyewitness information is collected face-to-face interview by a police officer sometime after the event. However, in some circumstances, particularly where there are large numbers of eyewitnesses, for example, it can take weeks and sometimes months. This delay can interrupt police investigations and can result in witnesses forgetting some of the detail about what they have seen or heard.

Conducting a ‘first account’ quick interview may be one answer. The aim of this research is to investigate whether ‘first account’ quick interviews are helpful.

What did I do?

All participants in this study watched a short film and took part in an interviewed face-to-face a week later over video call. You completed a quick face-to-face, in person interview within one hour of watching the short video. Other participants answered questions using a secure online Chatbot, called ChatCharlie.

Please Note:

- Throughout these processes your data has been labelled with your individual identification number which has been provided to you, you will not be personally identifiable.

- I will not store any personal identifying data, so please keep this number safe.
- Your participation in this research is on an entirely voluntary basis, and you are able to withdraw without providing any reason, at any time up until the research has been published, or submitted in any form of a report (e.g., conference presentation, dissertation, etc.)
- We will not store any personal identifying data. Should you wish to withdraw at any time (until publication) simply contact us with your personal identification number so that we can remove your contribution.
- We will not be able to give feedback, but we can provide all participants with a summary of the overall findings if requested.

You will receive an email with your £10 Amazon voucher in the next few days. If you do not receive it in 7 days, please email the researcher (email below). After receiving your voucher, please credit your amazon account as soon as possible, after 3 months it will no longer be valid.

If you have any further questions, please email Charlotte Adam: c.adam1@westminster.ac.uk if you have any questions.

Appendix P

Chatbot Post-interview Survey Questions

Thank you for taking part in the Chatbot Interview, please answer the following questions about your experience. Please can you enter your unique anonymised participant number (If at any point you wish to remove your data from the full dataset - you will need to quote this number)

Q1. Overall, how easy did you find it to remember what happened in the video? (Very easy, Somewhat easy, Neither easy nor difficult, Somewhat difficult, Very difficult)

Q2. How easy was it to tell ChatCharlie when you could not remember? (Very easy, Somewhat easy, Neither easy nor difficult, Somewhat difficult, Very difficult)

Q3. How easy was it to tell ChatCharlie that you did not know the answer to a question? (Very easy, Somewhat easy, Neither easy nor difficult, Somewhat difficult, Very difficult)

Q3. Overall, how difficult did you find it to remember what happened in the video? (Very difficult, Somewhat difficult, Neither easy nor difficult, Somewhat easy, Very easy)

Q4. How difficult was it to tell ChatCharlie when you could not remember? (Very difficult, Somewhat difficult, Neither easy nor difficult, Somewhat easy, Very easy)

Q5. How difficult was it to tell ChatCharlie that you did not know the answer to a question? (Very difficult, Somewhat difficult, Neither easy nor difficult, Somewhat easy, Very easy)

Q6. How satisfied were you with ChatCharlie? (Very satisfied, Somewhat satisfied, Nether satisfied or dissatisfied, Somewhat dissatisfied, Very dissatisfied)

Q7. Would you recommend ChatCharlie for use in police investigations? (Definitely yes, Probably yes, Might or might not, Probably not, Definitely not)

Q8. How easy did you find it to communicate with ChatCharlie? (Very difficult, Somewhat difficult, Neither easy nor difficult, Somewhat easy, Very easy)

Q9. Do you think typing was a good way of conducting an interview? (Definitely yes, Probably yes, Might or might not, Probably not, Definitely not)

Q10. Do you think you would have provided more information if the chatbot had used speech (like Alexa) instead of texting? (Definitely yes, Probably yes, Might or might not, Probably not, Definitely not)

Q11. Do you think you would have provided more information if using a pen and paper instead of texting with a chatbot? (Definitely yes, Probably yes, Might or might not, Probably not, Definitely not)

Q12. Do you think you would have provided more information if the interview had been conducted by a person instead of a chatbot? (Definitely yes, Probably yes, Might or might not, Probably not, Definitely not)

Q13. Which format of messages did you prefer the most? ('Free text' or open-ended questions, Yes or No buttons, scaled 1-5 button questions, A mixture or 'free text' and buttons, Not sure)

Q14. I am confident that I remembered a lot of what I saw. (Not at all confident, Somewhat confident, Undecided, Quite confident, Completely confident)

Q15. I am confident that I did not make any errors. (Not at all confident, Somewhat confident, Undecided, Quite confident, Completely confident)

Q16. Have you witnessed a crime before? (Yes, No)

Q17. Have you ever been interviewed by a police officer as a witness or victim of a crime? (Yes, No)

Q18. Have you ever completed a paper self-administered questionnaire type interview after witnessing a crime? (Yes, No)

Thank you for completing this survey. If you have any questions, please email the researcher and they will be happy to help.

c.adam1@westminster.ac.uk

Appendix Q

Chatbot Post-Interview Survey Questions: Acceptability and Comparison

Thank you for taking part in the Skype Interview, please answer the following questions about your experience. Please can you enter your unique anonymised participant number (If at any point you wish to remove your data from the full dataset - you will need to quote this number)

Q1. Overall, how easy did you find it to remember what happened in the video today? (Very easy, Somewhat easy, Neither easy nor difficult, Somewhat difficult, Very difficult)

Q2. How easy was it to tell the interviewer when you could not remember? (Very easy, Somewhat easy, Neither easy nor difficult, Somewhat difficult, Very difficult)

Q3. How easy was it to tell the interviewer that you did not know the answer to a question? (Very easy, Somewhat easy, Neither easy nor difficult, Somewhat difficult, Very difficult)

Q4. Overall, how difficult did you find it to remember what happened in the video? (Very difficult, Somewhat difficult, Neither easy nor difficult, Somewhat easy, Very easy)

Q5. How difficult was it to tell the interviewer when you could not remember? (Very difficult, Somewhat difficult, Neither easy nor difficult, Somewhat easy, Very easy)

Q6. How difficult was it to tell the interviewer that you did not know the answer to a question? (Very difficult, Somewhat difficult, Neither easy nor difficult, Somewhat easy, Very easy)

Q7. Did you find the interviewer friendly? (Definitely yes, Probably yes, Maybe, Probably not, Definitely not)

Q8. Did you find the interviewer easy to communicate with? (Definitely yes, Probably yes, Maybe, Probably not, Definitely not)

Q9. I am confident that I remembered a lot of what I saw. (Not at all confident, Somewhat confident, Undecided, Quite confident, Completely confident)

Q10. I am confident I did not make any errors. (Not at all confident, Somewhat confident, Undecided, Quite confident, Completely confident)

Q11. Have you witnessed a crime before? (Yes, No)

Q12. Have you ever been interviewed by a police officer as a witness or victim of a crime? (Yes, No)

Q13. Have you ever completed a paper self-administered questionnaire type interview after witnessing a crime? (Yes, No)

Appendix R

Additional Chatbot Questions

Below are some questions about your experience of being interviewed by ChatCharlie and then, a week later, being interviewed by a person.

What is a chatbot?

A chatbot is a computer program designed to simulate conversation with human users, especially over the internet. Although it can be like with a person chatbots are not human and have a limited set of responses and questions they can ask. Examples of text based chatbots include the 'help' functions on banking websites or shopping websites before speaking to a real person. Examples of chatbots that can speak are Amazon Alexa or Apple Siri. Chatbots can have their own websites where you can connect with your computer and have a conversation. Other chatbots are integrated in other websites, for instance, often companies will create chatbots that can interact with service users through their Facebook page.

In this section, we would like to find out what you think about chatbots

Q1. Please indicate how much you agree or disagree with the statements below:

(The questions below in a matrix table)

I remembered more in the chatbot interview than the second interview

I remembered less in the chatbot interview than the second interview

I remembered more in the second interview than in the chatbot interview

I remembered less in the second interview

I remembered different things in the second interview

I remembered the same things in both interviews

Definitely agree, somewhat agree, neither agree nor disagree, somewhat disagree, definitely disagree)

Q2. Do you think the Chatbot interview helped you remember more about the video in the second interview? (Definitely yes, Probably yes, might or might, Probably not, Definitely not)

Q3. Do you think the ChatBot interview helped you remember the video more easily a week later? (Definitely yes, Probably yes, might or might, Probably not, Definitely not)

Q4. Do you think you would have remembered more information if you had not had the ChatBot interview? (Definitely yes, Probably yes, might or might, Probably not, Definitely not)

Q5. If you had not had the ChatBot interview how much do information do you think you would have remembered in the second interview?

- I would have remembered much more
- I would have remembered slightly more
- I would have remembered the same amount
- I would have remembered slightly less
- I would have remembered much less

Q6. How satisfied were you with ChatCharlie? (Very satisfied, Somewhat satisfied, Nether satisfied or dissatisfied, Somewhat dissatisfied, Very dissatisfied)

Q7. Would you recommend the ChatCharlie chatbot for use in police investigations? (Definitely yes, Probably yes, Not sure, Probably not, Definitely not)

Q8. How comfortable would you be texting with a chatbot using your phone to quickly provide some initial basic information about a crime to the police before being interviewed face-to-face, for example? (Very comfortable, Somewhat comfortable, Neither comfortable nor uncomfortable, Somewhat uncomfortable, Very uncomfortable)

Q9. How comfortable would you be typing with a chatbot using a computer to quickly provide some initial basic information about a crime event before being interviewed face-to-face, for example? (Very comfortable, Somewhat comfortable, Neither comfortable nor uncomfortable, Somewhat uncomfortable, Very uncomfortable)

Q10. How comfortable would you be speaking to a chatbot about a crime, but with no option to speak to a person later? (Very comfortable, Somewhat comfortable, Neither comfortable nor uncomfortable, Somewhat uncomfortable, Very uncomfortable)

Q11. How comfortable would you be disclosing information to a police chatbot about a real crime you had witnessed that used one of these platforms? (Very comfortable, Somewhat comfortable, Neither comfortable nor uncomfortable, Somewhat uncomfortable, Very uncomfortable)

(This is a matrix table)

- Facebook messenger
- WhatsApp
- Text message
- Email
- dedicated website (https://www.eviebot.com/en/)

Q12. How comfortable would you be disclosing the following information with a chatbot? (this is a matrix table) (Very comfortable, Somewhat comfortable, Neither comfortable nor uncomfortable, Somewhat uncomfortable, Very uncomfortable)

- Your name
- Your age
- Your email address

- Your telephone number

- Your current level of education

- Information about any physical disabilities

- Information about any learning difficulties

Q13. How comfortable would you be disclosing basic information about a real crime you had just witnessed to a chatbot, which would then quickly be accessible by police? (This is a matrix table) (Very comfortable, Somewhat comfortable, Neither comfortable nor uncomfortable, Somewhat uncomfortable, Very uncomfortable)

- A description of what happened

- Information about crime location

- Information about the people involved in the crime

- How confident you were in your memories of the crime

- If you were under the influence of alcohol or drugs when you witnessed the crime.