Eyewitness memory for person, object and action information is mediated by interview environment and the presence or absence of rapport

Taylor, D.A. and Dando, C.J.

Eyewitness Memory for Person, Object and Action Information is mediated by Interview Environment and the presence or absence of Rapport.

Abstract

Purpose

Obtaining accurate and reliable information from witnesses and victims of crime is essential for guiding criminal investigations and for the successful prosecution of offenders and beyond. Here, we investigate the impact of prosocial rapport behaviours and retrieval environment on mock-eyewitness memory with an emphasis on the qualitative nature of information recalled in terms of persons, actions, objects and surroundings.

Design/Methodology

One hundred participants from the general population took part in mock witness research employing a 2 (Environment: face-to-face; virtual) X 2 (Rapport: present; absent) design. Participants individually viewed an event depicting a fight in a bar and were then interviewed 48 hours later according to condition.

Findings

Rapport and environment variously emerged as impactful, resulting in significant improvements in correct recall of persons, actions, and object information when rapport was present and in virtual environments during verbal communication between interviewer and interviewee who were visually represented as avatars. In many instances, the benefits of rapport and environment were additive, but not always. Surroundings information remained consistent across all conditions. Erroneous recall was extremely susceptible to inflation in face-to-face interviews when rapport was absent, in some instances increasing by over 100%. However, virtual interview spaces appeared to dilute the negative effects on retrieval when rapport was absent.

Limitations

We concur with others who have argued that comfortable witnesses are 'better' witnesses; however, understanding what 'comfort' looks and feels like remains a challenge.

Implications

Our results, alongside the findings of others, have implications for applied and social cognition and reveal avenues for future research centred on access to justice and professional interview training.

Keywords Eyewitness recall; Episodic Memory; Virtual Environment; Investigation Relevant Information; Investigative Interview; Avatar

Introduction

Obtaining accurate and reliable information from witnesses, victims, and survivors of crime (henceforth witnesses) is essential for guiding most criminal investigations from the very start and for the successful prosecution of offenders and beyond (Mendez & Areh, 2021). When giving their accounts, witnesses reconstruct and narrate personally experienced episodes or events situated in specific times and places, which is referred to as episodic memory (Tulving, 1993, 2002). Episodic memory includes details about what happened, when the event occurred, and where the incident took place. Episodic memory is a multifaceted, effortful, and complex cognition (Conway, 2009; Vrij et al., 2014) marked by a sensation of re-living or re-experiencing the event in question (Conway & Pleydell-Pearce, 2000; Tulving, 1993; 2002).

In the UK and elsewhere, eyewitness information is typically retrieved during an in-person investigative interview conducted by police officers or other professional investigators (Launay et al., 2022). There is a large body of applied memory research concerned with understanding how the retrieval process can be optimised to maximise the number of correct details recalled while minimising the reporting of incorrect or confabulated details (e.g., Dando, 2020; Greene et al., 2022; Mickes & Wixted, 2021; McNeil, 2024; Pezdeck et al., 2020). Accordingly, there now exists much guidance on witness-appropriate mnemonic retrieval techniques and on how to manage and structure an interview in an appropriate, witness-compatible manner using evidence-based techniques (e.g. Cognitive Interview, Fisher and Geiselman, 1992; PEACE model, Ministry of Justice, 2022 (both variously adapted for adults and children); NICHD, LaRooy et al., 2015 (generally used for children only).

An interview is inherently a social interaction, and so evidence-based practice encourages interviewers to consider how best to manage the social nature of the witness-interviewer interaction. A poorly managed social environment is seen as diverting finite cognitive resources away from the process of remembering (e.g., Gabbert et al., 2020; Fisher & Schreiber, 2017; Nahouli et al., 2020; Dando et al., 2023). Accordingly, evidence-based interviewing practice guidance highlights numerous prosocial behaviours towards reducing the real or perceived demand characteristics that can occur in

witness interviews where interviewees want to perform well (Tredoux et al., 2004) . This desire to 'help' investigators can result in witnesses offering information in response to questioning when they are not entirely confident about the answers or indeed may not actually know the answer (e.g., Caso et al., 2024; Holmberg 2004; Li, 2022; Mulayim et al., 2014). Rapport building and maintenance has been offered as one way of supporting social aspects of witness-interviewer interactions (see Gabbert et al., for a review), as such here we are concerned with a group of prosocial behaviours (described in Appendix A), collectively referred to as rapport.

The extant literature suggests that when an interviewer engages in rapport-building behaviours, witness memory is enhanced because a socially comfortable witness maybe more likely to cooperate since anxiety and stress may be lower, therefore memory can be optimised (e.g., Carol et al., 2021; Dando et al., 2023; Gabbert et al., 2020; Kieckhaefer et al., 2014; Nahouli et al., 2021; Wolfs et al., 2022). Rapport is generally ill-defined and not clearly operationalised for investigative interviewing purposes since interviews differ as a function of the to-be-remembered event, the retrieval environment, and individual witness characteristics. As such, rapport building is not prescriptive, but a dynamic process. However, there is consensus that rapport encompasses numerous widely accepted pro social, socially supportive verbal and non-verbal behaviours displayed by the interviewer at various stages of an interview in forensic settings (see Dando et al., 2023; Gabbert et al., 2020; Nahouli et al., 2021; Webster et al., 2021). Examples include empathic statements, interviewer self-disclosure, attentive listening, using names, eye contact, relaxed posture and facial expressions (Dando & Oxburgh, 2016; Gabbert et al., 2020; Nahouli et al., 2021).

Despite an apparent rapport superiority effect, that is typically improved recall and information gathering performance, (e.g., Carol et al., 2021; Dando et al., 2023; Gabbert et al., 2020; Kieckhaefer et al., 2014; Nahouli et al., 2021; Wolfs et al., 2022) the current understanding of facilitating a comfortable witness interview environment has largely emerged from face-to-face, in-person interactions since this approach maps onto traditional real-world practices and procedures (for a

review, Gabbert et al., 2020; Vallano & Schreiber Compo, 2011). However, technological advancements in communication over recent years have highlighted a need to consider the importance of rapport in interviewing contexts where the interviewee and interviewer do not necessarily engage in-person face-to-face (e.g., Taylor & Dando, 2018; Dando et al., 2023; Dando & Adam, 2024). In such instances, questions emerge centred on the appropriateness of rapport-building behaviours and whether they are expected, well received and as impactful as in-person face-to-face contexts. Previous research has indicated that rapport that does not 'feel' genuine is likely to be poorly received and so ineffective (e.g., Dando & Oxburgh, 2016; Webster et al., 2017). Thus, it is sensible to consider whether rapport is as important in remote interviews since physical distance may naturally alleviate some of the social demands inherent in in-person face-to-face interactions, and so memorial performance may be less sensitive to social context.

Some researchers have utilised remote computer-mediated applications such as Skype (Brown et al., 2021; Dickinson et al., 2021; Johnstone et al., 2024; Nash et al., 2014), where the interviewer and witness are face-to-face but situated in different physical locations and so communicate via a video platform. Others have removed the human interviewer, interviewing witnesses in different locations using a pre-programmed robot to gather witness information (Bethel et al., 2013; Kyriakidou, 2016). More recently, AI chatbots or contextual chatbots have been developed for gathering witness information (Dando & Adam, 2024; Minhas et al., 2022). Despite slight variations in remote interviewing paradigms, initial findings have indicated the relevance of rapport in terms of the presence of various rapport building behaviours (e.g., active listening, self-disclosure, and eye-contact) for improved witness memory, but rapport has not typically been experimentally controlled. Moreover, the environment within which the interview takes place has not necessarily been experimentally managed to allow comparisons across environments as a function of rapport to offer a nuanced understanding of the impact of context. Yet, given the nature of remote interviewing and because

rapport building is not prescriptive, it is sensible to consider whether rapport 'works' in the absence of a human interviewer.

Here we report further analysis of data emerging from our study of rapport in mock eyewitness interviews conducted in a Virtual Environment. The development and proliferation of increasingly sophisticated virtual environments (VE) offer a unique opportunity to examine the cognitive and social dynamics of witness interviews in settings that mimic real-world interactions where 'players' interact via avatars, thus affording greater control over environmental variables that can interfere with complex cognition (Meijer et al., 2021). Previously, we have reported that VE interviews have the potential to uphold the social dynamics essential for effective rapport building, thus supporting effortful remembering by allowing witnesses to devote more cognitive resources to invoke episodic retrieval mode (Dando et al., 2023; Taylor & Dando, 2018; Frith & Frith, 2012). The potential of VE interviewing is clear, albeit understanding is in the early stages (see Dando et al., 2023; Taylor & Dando, 2018; Georgiou & Dando, 2021) since while quantitative memorial performance appears improved, thus far little is known about the type of information recalled by mock eyewitnesses.

A review of the contemporary literature indicates researchers have typically focused on the quantitative nature of information recalled by mock witnesses, typically reporting correct, erroneous and confabulation metrics (e.g., Fisher et al., 2017; Gustafsson et al., 2024; Milne et al., 2019; Wixted et al., 2018). The qualitative nature of the information recalled is less scrutinised. Further, other than one or two studies, most have neither manipulated nor fully considered rapport in technology-mediated interactions (but see Dando et al., 2023; Muir et al., 2017). Some technology-mediated contexts do not naturally lend themselves to rapport building (e.g., conversational chatbots), but as we have previously reported (Taylor & Dando, 2018; Dando et al., 2023), avatar-to-avatar interviewing in virtual environments does allow for manipulation and management of some basic, well-operationalised rapport building behaviours. For example, previous research has reported benefits of prosocial behaviours such as active listening for improved memory in VE. The research reported here moves

towards filling a clear gap in understanding of the type of information that is recalled by eyewitnesses during interview by analysing the qualitative nature of mock eyewitness memory performance in face-to-face, in-person, and avatar-mediated interviews conducted in a virtual environment when rapport is present and absent. This type of analysis allows progression toward assessing the practical and investigative relevance for a criminal investigation or whether simply more peripheral information is recalled.

Understanding the quantity of correct (and incorrect) information remembered in various contexts establishes baselines for developing and refining best practices for interviewing witnesses, creating standardised methods that can be consistently applied in real-world settings (Memon et al., 2010; Wells & Loftus, 2003). However, focusing on the amount of correct information minimises individual differences and the impact of context, and does not take into account the range of information types of interest to investigators. Hence, there is value in examining the type of information provided by eyewitnesses (Wright & Skagerberg, 2007) because different types of information (e.g. descriptions of the perpetrator and details of the crime scene) have varying degrees of relevance to an investigation (Kebbell & Milne, 1998). Additionally, analysing the type of information remembered can offer theoretical and practical insights into how different types of memories (e.g. gist and verbatim) might be affected by variables, including retrieval context.

Investigation relevance is event-specific and more subjective in nature than the correctness of the information provided; one framework for categorising the information provided during forensic interviews has been suggested by Oxburgh and colleagues (2012). Focusing on the effect of question type on the information revealed by interviewees, investigation-relevant information was categorised into details that are recalled about a person, actions, objects and surroundings. It is this approach that has guided the analysis reported here, which is in line with a limited number of studies that have similarly analysed the type of information recalled by mock witnesses employing this method (Boon et al., 2020; Dando et al., 2011; National Institute of Justice, US, 1999).

In analysing mock-witness memorial performance, we are concerned with the qualitative nature of episodic recall in terms of the amount of investigation-relevant information provided by mock witnesses during face-to-face and virtual-reality-mediated interviews. Theoretical understanding and the existing empirical literature relevant to quality regarding the type of episodic recall (as opposed to quantity) as a function of interview environment and rapport do not support the formulation of hypotheses. Rather, we have developed two broad exploratory research questions. First, is there a qualitative difference in the type of information recalled as measured by person, object, action, and surroundings information reported by mock witnesses when rapport is present versus absent? Second, does context (VE versus in-person face-to-face) impact the quality of the information recalled as measured by person, object, action, and surroundings information? Quantitative differences have been previously reported that reveal significant rapport and context main effects and interactions for correct, erroneous and confabulated information recalled as a function of rapport and context (Dando et al., 2023; Taylor & Dando, 2018). However, the locus of these effects is unclear regarding information type losses and gains, so this research is novel in that it sheds new light on 'what', thus adding to the 'how much' applied memory literature.

Methods

Design

A mock witness 2 (Environment: face-to-face; virtual) X 2 (Rapport: present; absent) design was employed using five interviewers, as typically occurs in real-life cases where there are several witnesses. Interviewers were all experienced researchers in the domain of experimental investigative interviewing. Before conducting interviews for this research all underwent bespoke training (designed by the second author – see Procedure). Interviewers were counterbalanced across interview conditions. The mean number of interviews conducted by each interviewer was 20 (ranging from 11 to 32).

Participants individually watched a stimulus video, aware that they would later be interviewed about their recall of the events within the video. Participants were then randomly allocated to one of

the experimental interview conditions. Forty-eight hours later, participants were interviewed according to condition. The dependent variable was memory for the video, measured by the number of person, object, environment and action information items recalled and percentage accuracy (correct details as a function of overall details recalled). Ethical approval was obtained from the University of Westminster research ethics review committee.

Participants

An a-priori power analysis using G*Power 3.1 (Faul et al., 2007) indicated that a sample size of 100 mock witnesses would be adequate to detect large effects (assuming power = .80 and a = .05). Forty-four males and 56 females from the general population participated with a mean age of 25.8 years (SD = 7.5), ranging from 18 to 50 years. Given the applied nature of this experimental research, our inclusion criteria was that all participants had to be over 18 years of age with English as a first language. Participants were recruited by word of mouth, social media and advertisements in the vicinity of the University. There were no significant differences in mean age across conditions (rapport & environment), F = 1.46, p = .23. Participants were recruited through word of mouth, social media, and advertisements placed in the locality of the University. No incentives were offered for participating in the study.

Materials

Crime stimulus video. A pre-recorded video lasting 1 minute 40 seconds of a mock fight in a public bar was viewed individually by participants via a laptop computer (see https://youtu.be/4PumXJX1iZo). The video depicts a man buying drinks for a female friend while another female character walks over to chat about a coursework assignment. The second female character leaves, and the male and female then walk to the other side of the bar, where they sit down at a table. Their conversation is interrupted by two men, first talking and then shouting. One of the men pushes the other before punching him to the ground and repeatedly punching him. The male friend goes over and states he is unconscious. A woman who is sitting behind them calls an ambulance.

Interview Protocols. All interviews comprised two retrieval attempts in the same order. First, participants were asked to provide a free recall account of everything they could remember. This initial account was uninterrupted by the interviewer, who made bullet point notes regarding the topics recalled and the order in which they were recalled for use during the questioning phase that followed. In the questioning phase, each of the topics recalled in the preceding free recall phase was probed using either a Tell, Explain, or Describe (TED) question in turn. Hence, the number of TED questions asked during the questioning phase was predicated on free recall performance (see Caso et al., 2024; Dando et al., 2020: 2022; Vrij et al., 2014). For example, if the interviewee stated, 'there was a shop, and round the corner came a man and a woman' this would trigger three probing questions later in the interview based on the primary topics verbalised during the free recall, namely i) shop, ii) man, and iii) woman.

The free recall commenced with a pre-interview explanation phase, and the interview finished with a closure phase. Participants in the rapport condition experienced an additional rapport phase, with all rapport-building behaviours continuing throughout the interview. Participants in the no rapport conditions did not experience the rapport-building phase, and rapport behaviours were absent throughout the entire interview. Interview protocols are outlined in Appendix A (detailed protocols are available from the first author). The questioning phase commenced with a reminder of the four ground rules. Five experienced researchers conducted all the interviews verbatim, following the condition-appropriate protocols verbatim (see procedure).

Equipment. In the VE condition, the interviewer and participant were in different rooms within the same building and communicated using an Oculus Rift S virtual reality (VR) headset. The Oculus Rift creates a sense of complete immersion in a three-dimensional world (here, a bespoke interview environment) via 2560 X 1440 high-resolution OLED panels, one for each eye, which globally refresh at a rate of 90 Hz. An onboard Inertia Measurement Unit (IMU) positional camera allows transitional and rotational movement to be tracked with 6DoF. The headset tracks the movements of both head and

body, then translates them into VR with realistic precision. Verbal communication was via 3D positional audio built directly into the headset, which was digitally recorded for transcription and coding. A bespoke, virtual interview environment was developed for this research using Unreal Engine 4. The VE interview environment was purposely sparse and neutral, comprising a sofa, a table and chairs – one chair for the avatar interviewer, the other for the avatar participant (see Figs 1 and 2). Limited choice was offered to participants regarding the appearance of their avatar, likewise the interviewers, whereby they could appear as either male or female. Participants and all interviewers chose to match their avatar to their stated gender .

Figure 1 and Figure 2 go here

Procedure

Adult participants were recruited to take part in a mock eyewitness research investigating the use of virtual environments for investigating long-term memory performance. Once participants had consented, they accessed a one-time-only link, which allowed them to view the stimulus video. All participants were interviewed 48 hours later. Prior to the interview, participants were randomly assigned to one of the four interview conditions (face-to-face + rapport, face-to-face no rapport, VE + rapport, VE no rapport) and interviewed accordingly. Participants took part voluntarily and received no payment or other compensation for their time.

Prior to conducting interviews for this research all interviewers underwent bespoke (designed for this research by the second author) training comprising: i) a four-hour long classroom-based basic introduction to rapport, ii) 2 x 4-hour long instruction and practice sessions using the VE and VR headsets, iii) reading of theoretical and applied training materials produced for this research, iii) practice interviews (4 each) face-to-face and using the VR, which were digitally recorded to allow feedback and evaluation, and iv) instruction on reflective research practice. Once all researchers had

attended all training sessions and completed the required competencies, research interviews commenced.

Interview Coding

Interviews were digitally audio and video recorded, transcribed verbatim, anonymised, and coded for person, action, object and surroundings information, which was coded as either correct or erroneous (information relevant to the witnessed episode but described with error or reporting information that was not present in the film). Items recalled were only scored once (i.e., repetitions were not scored irrespective of the interview phase). Interviews were initially coded by one researcher blind to the conditions, using a coding grid. Five interviews from each condition (20 in total) were then randomly selected for recoding by a second independent coder also blind to the research questions but familiar with the scoring method. For example, if the interview states the man's jacket was brown, when in fact the man's jacket was black, 'brown' would be coded as one error, 'man' and 'jacket' would be each be coded as correct items.

Two-way mixed effects Intraclass Correlation Coefficient (ICC) analysis testing for absolute agreement between coders for the overall amount of correct, erroneous, and confabulated recall for each of the four information categories was conducted. Mean estimations with 95% CI reveal very good inter-rater reliability: Person correct, ICC = .894 (95% CI .836; .940); Person errors, ICC = 1.000 (95% CI 1.00; 1.00); Action correct, ICC = .944 (95% CI .889; .972); Action Errors, ICC = .865 (95% CI .498; .964); Object correct, ICC = 1.000 (95% CI 1.00; 1.00); Object errors, ICC = .937 (95% CI .867; .970); Surroundings correct, ICC = .920 (95% CI .876; .955); Surroundings Errors, ICC =.944 (95% CI .889; .972).

The same sample of 20 interviews was coded by a further two independent coders blind to study research questions for adherence to the interview protocol as a function of condition: that is, no rapport-building behaviours in the rapport absent (control) conditions and presence of rapport building behaviours in the rapport present conditions (see Table 1). A scoring sheet was used where each of the

behaviours were coded, ranging from 1 to 3 for each according to condition (e.g., 3 = fully implemented the open-ended self-disclosure behaviour, 2 = partially implemented the open-ended questions behaviour, 1 = did not implement) as a function of phase (e.g., see Nahouli et al., 2021). The rapport phase occurred only in the rapport-building condition, while the free recall and questioning phases were common to all conditions. In the rapport phase, six rapport behaviours were coded (see Table 1), and in the free recall and questioning phases, 4 rapport behaviours were coded. To score 1, the behaviour in question had to be absent. To score 2, the behaviour had to be present at least once but no more than twice. To score 3, the behaviour had to be present at least three times. Thus, each phase was awarded scores ranging from 6 to 18 for the rapport phase (in the rapport condition only), and ranging from 4 to 12 for each of the free recall and questioning phases.

Two-way mixed effects Intraclass Correlation Coefficient (ICC) analysis testing for absolute agreement between coders for the six rapport building behaviours expected to be present/absent in the rapport phase revealed good inter-rater reliability for each of the behaviours; open questions, ICC = .899 (95% CI, .593; .975), offering non-personal information, ICC = .862 (95% CI .443; .966), making eye contact, ICC = .862 (95% CI .443; .966), nodding, ICC = .865, (95% CI .498; .964), referring to interviewee by name, ICC = 1.00 (95% CI 1.00; 1.00) and thanking the interviewee, ICC = .757 (95% CI .096; .935). Good inter-rater reliability was also found for the four rapport building behaviours expected to be present/absent in the free recall phase: making eye contact, ICC = .938 (95% CI .843; .975), nodding, ICC = .883, (95% CI .705; .954), referring to interviewee by name, ICC = 1.00 (95% CI 1.00; 1.00) and thanking the interviewee, ICC = 1.00 (95% CI 1.00; 1.00); and questioning phase: making eye contact, ICC = .883 (95% CI .705; .954), nodding, ICC = .979, (95% CI .948; .992), referring to interviewee by name, ICC = 1.00 (95% CI 1.00; 1.00), and thanking the interviewee, ICC = 1.00 (95% CI 1.00; 1.00).

Rapport Manipulation Analysis

Means (SDs & 95% CIs) for rapport behaviours across phases common to both interview conditions (free recall, questioning) as a function of environment and interview condition are fully reported in Dando et al., 2023 and so are not fully repeated here. In brief, the rapport present main effect was non-significant for all four rapport-building behaviours across environments, in both the free recall, all Fs < 4.00, all ps > .059 and questioning phases, all Fs < 2.21, all ps > .144 revealing that all behaviours were similarly present across the two environments in both recall phases. Likewise, there were non-significant differences across environments in the rapport absent conditions for the four rapport-building behaviours in either the free recall, all Fs < 1.00, all ps > .322 or questioning phases, all Fs < .214, all ps > .646, and so all rapport behaviours were similarly absent.

Results

Person Information

Correct. There were significant main effects of Rapport, $F(1, 96) = 24.924$, $p < .001$, η_p^2 .21, and Environment, $F(1, 96) = 7.057$, $p = .009$, η_p^2 .07, for correct person information items recalled (see table 1 below for means, SDs and 95% CIs). Participants in the Plus Rapport conditions recalled more correct person items than those in the No Rapport conditions. Participants in the VE recalled more correct person items than those in the F2F environment. There was a significant Rapport X Environment interaction, $F(1, 96) = 11.179$, $p = .001$, η_p^2 .10. Participants in the VE Plus Rapport condition recalled more correct person items than those in the F2F Plus Rapport, $p < .001$. All other interactions were non-significant, $p = .247$.

Errors. There were significant main effects of Rapport, $F(1, 96) = 17.727$, $p < .001$, η_p^2 .16, and Environment, $F(1, 96) = 20.504$, $p < .001$, η_p^2 .18, for erroneous person information items (see table 1). Participants in the Plus Rapport condition recalled fewer person errors than those in the No Rapport condition and participants in the VE recalled fewer person errors than those in the F2F environment. There was a significant Rapport X Environment interaction, $F(1, 96) = 11.669$, $p = .001$,

η_p^2 .11. Participants in the F2F No Rapport recalled significantly more person errors than those in the VE No Rapport, $p < .001$. All other interactions were non-significant, $p = .434$.

Action Information

Correct. There was a significant main effect of Environment, F (1, 96) = 9.041, $p = .003$, η_p^2 .09, for correct action information items recalled. Participants in the VE recalled more correct action items than those in the F2F environment. The main effect of Rapport and the Rapport X Environment interaction were non-significant, with all Fs < 2.868 and all ps > .094 (see Table 1 for means, SDs and 95% CIs).

Errors. The main effects of Rapport and Environment were non-significant, with all Fs < 3.008 and all ps > .086 (see Table 1). There was a significant Rapport X Environment interaction, F (1, 96) = 10.480, $p = .002$, η_p^2 .10. Participants in the F2F No Rapport recalled significantly more action errors than those in the VE No Rapport, $p < .001$. All other interactions were non-significant, $p = .291$.

Table 1 Here

Object Information

Correct. The main effects of Rapport and Environment were non-significant for correct object information, all Fs < .506, all ps > .478. The Rapport X Environment interaction was also non-significant, $F = 2.025$, $p = .158$.

Errors. The main effect of Environment was significant for object errors, F (1, 96) = 6.816, $p = .010$, η_p^2 .097. Participants in the F2F condition recalled more object errors than those in the VE. The main effect of Rapport was non-significant, $F = 2.370$, $p = .127$. The Rapport X Environment interaction was significant, F (1, 96) = 10.759, $p = .001$, $\eta_p^2 = .10$. Participants in the F2F No Rapport

condition recalled significantly more object errors than those in the VE No Rapport condition, p = .001. All other interactions were non-significant (see Table 1 above for means, SDs and 95% CIs).

Surrounding Information

Correct. Main effects of Rapport and Environment were non-significant for correct surroundings information, all Fs < 3.099, all ps > .082. The Rapport X Environment interaction was also non-significant, F = .709, p =.402.

Errors. Main effects of Rapport and Environment were non-significant for erroneous surroundings information, all Fs < 4.449, all ps > .176. The Rapport X Environment interaction was also non-significant, F = 4.449, p =.038 (see Table 1 for means, SDs and 95% CIs).

Duration

Due to the additional rapport phase in the rapport present condition, there was a significant main effect of Rapport on interview duration (from start to finish including all interview phases), F (1, 96) = 108.22, p < .001, η_p^2 = .53. Globally, Rapport interviews were significantly longer (M = 44.96 mins, SD = 5.12, 95% CI [42.60, 46.17]) than the No Rapport interviews (M = 31.17 mins, SD = 9.15, 95% CI [29.38, 32.95]). The main effect of environment was non-significant, F(1, 96) = 1.13, p = . 291, as was the Environment X Rapport interaction, F(1, 96) = .02, p = . 887.

Discussion

Theoretical and evidence-based witness interviewing research has traditionally focused on the quantity of correct and erroneous information retrieved. However, there is a clear need to examine the qualitative aspects of episodic memory since recall performance for persons, objects, actions, and surroundings detail offers practical and applied memory performance insights. As far as we are aware, this research is the first to investigate types of information provided as a function of manipulating rapport and retrieval environment, thus shifting the focus from quantity of information to offering a

more nuanced understanding of applied memory performance. We found significant differences across our manipulations for recall of person, action, and object information, albeit that the pattern of results varied slightly across the three information types. Recall of surrounding information remained consistent across all manipulations. Person detail was proportionally the largest recall category across all conditions and manipulations.

Both rapport and environment variously emerged as impactful retrieval mediators. An average of 21% more correct person information was reported when basic rapport behaviours were present versus absent, and participants interviewed in the VE recalled 11% more correct person details. The VE + Rapport condition leveraged 23% more correct person information details than the F2F + Rapport condition without a concomitant increase in errors, suggesting the benefits of rapport and retrieval environment for improving recall may be additive in some instances. Participants in the F2F No Rapport condition reported 139% more erroneous person information items than participants in the VE No Rapport, indicating the applied importance of witness-appropriate retrieval environments for scaffolding effortful cognition, as has long been reported by others (e.g., Brown, 2003; Lane & Houston, 2019).

The relationship between rapport and environment is not well understood, but when rapport is absent, it may be that the retrieval environment becomes more important than previously considered. Here for example, correct action recall in the VE condition leveraged 20% more correct information than the F2F condition. Participants in the F2F No Rapport condition reported 120% more erroneous action information items than participants in the VE No Rapport, highlighting the importance of retrieval environment for effortful cognition when rapport is absent as was the case for person information. Similarly, environment emerged as relevant for recall of object information in the absence of rapport whereby more errors were reported in the F2F No Rapport condition than in the VE No Rapport with a 212% increase in the former condition.

Unpicking the locus of effect for this pattern of results is challenging, but two broad findings emerge that sit well with relevant previous research in this domain. First, interviews in traditional F2F contexts where rapport was absent were detrimental for invoking episodic retrieval mode and thus recall of an experienced event was variously compromised, typically reducing the quantity of information and inflating errors. (e.g., Dando et al., 2023; Kieckhaefer et al., 2014; Nahouli et al., 2021; 2023; Vallano & Schreiber Compo, 2015). Some form of prosocial behaviour in an F2F context is naturally expected. When prosocial rapport-building behaviours are absent, a social void occurs which 'feels' uncomfortable. In such situations, finite cognitive resources are believed to be diverted towards trying to understand and possibly improve the social interaction rather than trying to remember, the corollary being that episodic retrieval is less effortful (e.g., Dando et al., 2023; Nahouli et al., 2021).

Second, when basic prosocial rapport is absent, retrieval context or environment emerges as fundamental for improving recall. Here, the VE may have naturally alleviated some of the social demands inherent in F2F interviews since the witness and interviewer communicated via avatars. Avatar-to-Avatar communication may render rapport less important simply because prosocial behaviours may be less expected, thus mitigating demand characteristics associated with F2F witness interviews where performance is monitored, and witnesses are naturally striving to perform well and, in doing so, seek tacit encouragement/approval. VE affords a broader spectrum of communication than F2F, but as yet, the phenomena of non-verbal communication in VE and their influences on cognition and behaviours is under-researched despite the recent boom in commercial social VR applications (Maloney et al., 2020; McVeigh-Schlutz et al., 2018). More research is necessary to offer further insight into social and communication expectations and experiences in VE and the potential benefits of reducing environmental distractions and supporting complex cognition in criminal justice contexts.

The mean duration of rapport present interviews offers further insight. Rapport present interviews were 30% longer than when rapport was absent. But, as we have previously reported, the

number of probing questions asked immediately following the initial free recall account did not significantly differ as a function of rapport (see Dando et al., 2023). Accordingly, rapport apparently supports goal-directed cognitive effort because responses to questions were quantitively more detailed, and in such instances, context appears to become less important apart from correct person details, where the VE + Rapport interaction significantly increased correct remembering.

One possible interpretation for the dominance of person details as the largest proportion of recalled information emerges from social attention bias (see Heeren et al., 2015; Kovera & Evelo, 2021). The film stimulus used was of a mild, albeit in parts aggressive incident in a student bar, which is unlikely to have invoked emotional arousal, fear or stress related to the use of weapons or novel objects and actions. As such, social attention bias suggests a tendency to focus on people, often at the expense of inanimate objects or background details, supporting the Easterbrook hypothesis that individuals tend to remember more information central to an event rather than information peripheral to an event. One corollary is reducing memory for the stimuli's periphery and increasing memory for 'gist' event information, which may account for consistent surroundings results irrespective of our manipulations. The dominance of person details in mock witness research of this nature is common, and as such our findings sit alongside research conducted in traditional retrieval environments suggesting that performance has not been distorted as a function of our manipulations (e.g., Handler & Frühholz, 2021; Paulo et al., 2021; Sarwar et al., 2014; Sharps et al., 2009;).

The use of VE as interviewing spaces offers the potential for further integrating advanced technologies and cognitive theories of complex cognition, allowing a more nuanced examination of the impact of computer-mediated retrieval environments (Dando et al., 2023: 2024; Taylor & Dando, 2018). Our finding that the VE combined with rapport-building techniques apparently enhances memory performance in some instances suggests cognitive load and social dynamics theories might be adapted to include the impact of technological interfaces in terms of the unique affordances and constraints of virtual environments. This manuscript potentially contributes o the literature centred on

in terms of challenging the traditional view that physical presence is necessary for effective social interaction and memory retrieval.

Understanding of VE and basic rapport behaviours on complex cognition in eyewitness interview contexts is slow to emerge since, thus far, the empirical literature is scant. However, VE technology is both widely accessible and widely used. Traditional F2F interviews can be intimidating and anxiety inducing, and so VE might mitigate witness stress and anxiety by offering bespoke witness-centred choices. Moreover, virtual environments can be designed to accommodate various language and cultural challenges, further enhancing inclusivity and accessibility.

Having reported and then discussed our main findings, it is appropriate to highlight several of the more important limitations of this work. First, mock witness research of this nature does not mirror the real world. However, it does offer a way of controlling variables as a first step in understanding. Second, we have not controlled for any individual differences, and although our sample is adequate and recruited from the general population, much replication is needed. Finally, we did not measure immersion, so it is unclear how individual differences in immersive experience and use of VE headsets, for example, may have impacted our results. Future research should seek to mitigate these limitations and advance the encoding paradigm towards more ecologically valid empirical understanding.

Our findings and discussion offer several avenues for future research and applied practice. One area of interest could be the development of AI-driven virtual interviewers that can adapt their rapport-building strategies based on real-time analysis of witness responses. Further research might explore the long-term effects of VE interviews on memory retention and the potential for re-interviewing witnesses in virtual environments to gather additional details. Practical applications could also include the use of VE for training investigators in effective interviewing techniques and providing a safe and controlled setting for practice and feedback. Exploration of the ‘best’ VE design towards minimising

extraneous cognitive load and how different features of VE (e.g., level of immersion and types of avatars) might influence cognitive resource allocation and memory performance is also warranted.

The significant increase in correct person information in VE without an increase in errors is notable in terms of the practical implications for the administration of criminal justice practices and systems. Erroneous witness identification remains the single highest cause of wrongful convictions, so future research should aim to better understand the locus of effect concerning the use of VE and rapport in interview environments, potentially incorporating both person identification issues (line-ups in VE where crimes with a virtual environment have been witnessed) and suspect identification challenges (creating facial composites in VEs).

To conclude, we found clear qualitative and quantitative differences in the type of details recalled by witnesses across three information categories highlighting the benefits of VE interviews and basic rapport behaviours, albeit these benefits were not always additive. We concur with others who have argued that comfortable witnesses are 'better' witnesses (Vallano & Comp, 2011); however, understanding what 'comfort' looks and feels like and how this might be measured remains a challenge. Erroneous recall of persons, objects and action information was extremely susceptible to inflation in F2F contexts where rapport was absent. Conversely, correct recall of person information was significantly positively impacted by the retrieval environment and the presence of rapport behaviours. This pattern of results, alongside the findings of others, has implications for theories of applied and social cognition across environmental contexts and reveals avenues for future research, including widening access to justice and professional interview training.

References

Bethel, C. L., Eakin, D., Anreddy, S., Stuart, J. K., & Carruth, D. (2013, March). Eyewitnesses are misled by human but not robot interviewers. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 25-32). IEEE.

Boon, R., Milne, R., Rosloot, E., & Heinsbroek, J. (2020). Demonstrating detail in investigative interviews—An examination of the DeMo technique. *Applied Cognitive Psychology*, *34*(5), 1133-1142.

Brown, J. M. (2003). Eyewitness memory for arousing events: Putting things into context. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *17*(1), 93-106.

Carol, R. N., Kieckhaefer, J. M., Johnson, J., Peek, J., & Schreiber Compo, N. (2021). Being a good witness: The roles of benevolence and working memory capacity in rapport's effect on eyewitness memory. *Journal of Applied Social Psychology*, *51*(7), 730-745.

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*, *38*(3), e4197.

Dando, C., Taylor, D. A., 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*, *51*(2), 404-421.

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.

Fisher, R. P., & Schreiber, N. (2017). Interview protocols to improve eyewitness memory. In *The Handbook of Eyewitness Psychology: Volume I* (pp. 53-80). Psychology Press.

Frith, C. D., & Frith, U. (2012). Mechanisms of social cognition. *Annual review of psychology*, *63*, 287-313.

Gustafsson, P. U., Sikström, S., & Lindholm, T. (2024). The semantic structure of accuracy in eyewitness testimony. *Frontiers in Psychology*, *15*, 1211987.

Handler, A., & Frühholz, S. (2021). Eyewitness Memory for Person Identification: Predicting Mugbook Recognition Accuracy According to Person Description Abilities and Subjective Confidence of Witnesses. *Frontiers in Psychology*, *12*, 675956.

Heeren, A., Mogoaşe, C., Philippot, P., & McNally, R. J. (2015). Attention bias modification for social anxiety: A systematic review and meta-analysis. *Clinical psychology review*, *40*, 76-90.

Holmberg, U. (2004). *Police interviews with victims and suspects of violent and sexual crimes: interviewees' experiences and interview outcomes* (Doctoral dissertation, Dept. of Psychology, Univ.).

Joshua McVeigh-Schultz, Elena Márquez Segura, Nick Merrill, and Katherine Isbister. 2018. What's It Mean to Be Social in VR?: Mapping the Social VR Design Ecology. In Proceedings of the 2018 ACM Conference Companion Publication on Designing Interactive Systems. ACM, 289–294.

Kieckhaefer, J. M., Vallano, J. P., & Schreiber Compo, N. (2014). Examining the positive effects of rapport building: When and why does rapport building benefit adult eyewitness memory?. *Memory*, *22*(8), 1010-1023.

Kovera, M. B., & Evelo, A. J. (2021). Eyewitness identification in its social context. *Journal of Applied Research in Memory and Cognition*, *10*(3), 313-327. Lane, S. M., & Houston, K. A. (2019). Eyewitness memory. *Psychological science and the law*, 104-156.

La Rooy, D., & Dando, C. (2010). Witness interviewing. *Forensic psychology*, 195-209.

Launay, C., Brunel, M., & Bull, R. (2022). What do police interview eyewitnesses for? A review. *Policing: A Journal of Policy and Practice*, *16*(4), 602-614.

Li, L. (2022). How to tackle variations in elite interviews: Access, strategies, and power dynamics. *Qualitative research*, *22*(6), 846-861.

McNeil, S. (2024). Enhancing Justice and Efficiency: Advocating for the Adoption of Cognitive Interviewing Techniques Across Local and Federal Law Enforcement Agencies. *Journal of Police and Criminal Psychology*,1–6.

Mulayim, S., Lai, M., & Norma, C. (2014). *Police investigative interviews and interpreting: Context, challenges, and strategies*. CRC Press. Meijer, E., Hoogesteyn, K., Verigin, B., & Finnick, D. (2021). Rapport building: online vs in-person interviews. *Centre for Research and Evidence on Security Threats*.

Maras, K., Dando, C., Stephenson, H., Lambrechts, A., Anns, S., & Gaigg, S. (2020). The Witness-Aimed First Account (WAFA): A new technique for interviewing autistic witnesses and victims. *Autism*, *24*(6), 1449-1467.

Maloney, D., Freeman, G., & Wohn, D. Y. (2020). " Talking without a Voice" Understanding Non-verbal Communication in Social Virtual Reality. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW2), 1-25.

Mendez, J. E., & Areh, I. (Eds.). (2021). *Principles on effective interviewing for investigations and information gathering*. APT, Association pour la prévention de la torture.

Muir, K., Joinson, A., Cotterill, R., & Dewdney, N. (2017). Linguistic style accommodation shapes impression formation and rapport in computer-mediated communication. *Journal of Language and Social Psychology*, *36*(5), 525-548.

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.

Nahouli, Z., Mackenzie, J. M., Aresti, A., & Dando, C. (2023). Rapport building with offenders in probation supervision: The views of English probation practitioners. *Probation Journal*, *70*(2), 104-123.

National Institute of Justice (US). Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness evidence: A guide for law enforcement*. US Department of Justice, Office of Justice Programs, National Institute of Justice.

Nash, R. A., Houston, K. A., Ryan, K., & Woodger, N. (2014). Remembering remotely: would video-mediation impair witnesses' memory reports?. *Psychology, Crime & Law*, *20*(8), 756-768.

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.

Sarwar, F., Allwood, C. M., & Innes-Ker, Å. (2014). Effects of different types of forensic information on eyewitness' memory and confidence accuracy. *The European Journal of Psychology Applied to Legal Context*, *6*(1), 17-27.

Sharps, M. J., Janigian, J., Hess, A. B., & Hayward, B. (2009). Eyewitness memory in context: Toward a taxonomy of eyewitness error. *Journal of Police and Criminal Psychology*, *24*, 36-44.

Taylor, D. A. & Dando, C. J. (2018). Eyewitness Memory in Face-to-Face and Immersive Avatar-to-Avatar Contexts. *Frontiers in Psychology, 9,*507.

Tredoux, C. G., Meissner, C. A., Malpass, R. S., & Zimmerman, L. A. (2004). Eyewitness identification. *Encyclopedia of Applied Psychology*, *1*, 875-887.

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.

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.

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

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.

Wolfs, A. C., Sneyd, D., Vallano, J. P., Schreiber Compo, N., & Reinoso, L. (2022). The effects of building and maintaining rapport on cooperative mock eyewitness recall. *Journal of Investigative Psychology and Offender Profiling*, *19*(3), 151-166.

Appendix A.

Interview description

Phase	**Overview**
1.Explain	Explain the interview/research process prior to the commencement of the interview and offer the opportunity to ask questions.
2.Rapport *	Interviewer interacts with the participant using two behaviours - i) open-ended invitations to exchange information and ii) commencing with the interviewer offering some non-personal information about themself to begin this process; Interviewer displays two attentive physical behaviours - i) looking at interviewees/making eye contact when they were talking and ii) nodding when interviewees speak/answer questions; Interviewer displays two attentive verbal behaviours - i) referring to the interviewee by their first name and ii) thanking interviewees whenever they provided information/answered a question. * In the rapport present condition the two physical and two attentive verbal behaviours continued throughout the free recall and questioning phases.
3.No Rapport	The interviewer immediately moves from the Explain phase (phase 1) to the Free Recall phase. None of the above verbal or physical behaviours are used. This persists throughout the interview.
4.Free Recall	Commenced with an explanation of the four ground rules: 1. Report all/everything 2. Do not guess 3. Say if you do not know 4. Say if you do not understand Participants were then instructed to explain everything they could remember, uninterrupted by the interviewer. The interviewer made bullet point notes regarding the topics recalled and the order in which they were recalled for use during the questioning phase. Once interviewees had finished, they were asked if they wished to add anything else.
5.Questions	Commenced with a reminder of the four ground rules (above), following which participants were asked probing questions related to the main topics recalled in the free recall.
6.Close	Participants were thanked for their time and offered the opportunity to ask any further questions.