Alcohol Intoxication and Memory: The Reliability of Victims and Witnesses of Sexual Assault

By

Brittany Durnan Gibbs

Submitted to The University of Birmingham

for the degree of

Doctorate in Forensic Psychology Practice (ForenPsyD)

The University of Birmingham

September 2022

Abstract

The majority of crime in the United Kingdom takes place in establishments where people buy and consume alcohol. Owing to this, a large proportion of crime takes place against those who are under the influence of alcohol, as well as being witnessed by alcohol intoxicated bystanders. This thesis examines the relationships between alcohol intoxication and memory in the victims and witnesses of crime, and more specifically sexual assault. The aim of the thesis is to better understand the effects that alcohol intoxication has on the memory of victims and witnesses of crime by systematically exploring existing literature on the misinformation effect (Chapter Two). Next, a quantitative study examining the informativeness of uncertainty indicators in identifying accuracy in police interviews, and the impact of alcohol intoxication on accuracy in these interviews is presented (Chapter Three). The psychometric properties of the Alcohol Use Disorder Identification Test (AUDIT) are next discussed, following its use in the screening process of participants in Chapter Three (Chapter Four). Finally, the theoretical and practical implications of the overall findings of the study are presented (Chapter Five).

This thesis identifies areas for future research in order to better understand the effects of alcohol on victim and witness memory, as well as suggesting a novel method of understanding victim accuracy through uncertainty indicators in police interviews. The thesis goes on to identify potential aspects of the police interview process that may introduce error, and suggests ways in which this may be altered in the future to collect full and accurate accounts from victims of sexual assault. The requirement for education of those within the Criminal Justice System regarding the effects of alcohol on memory are also discussed with the aim of improving the perceived credibility of intoxicated victims and witnesses.

Acknowledgements

Firstly, I would like to thank my supervisor, Dr Melissa Colloff, whose immense support throughout my thesis has been incredible. I truly admire her kindness and encouragement, as well as her knowledge, and dedication to the fine details. Also, to Dr Heather Flowe for her expertise throughout Chapter Three and providing the data for analysis. Finally, I would like to thank the wider psychology department who have supported me through not only my academic work, but personally through every step of the way.

To my dad, whose unwavering confidence that I would achieve great things pushed me to apply for the Doctorate programme in the first place. While you never got the chance to see the things I went on to achieve, it is the part of you that lives on in me forever that takes such pride in my work, and finds the fun and laughter in life at every opportunity. I hope I will continue to make you proud and be the sparkle in your life.

To Tyler, I will forever be grateful for the support you have provided me. I could never have completed this thesis without your practical and emotional support and there will never be the right words to fully express my gratitude. You, Bobbie, and Katy will always be my family.

I must also thank my dearest friends, old and new, for the grounding and comic relief they have provided on days where I could not see the forest for the trees. For constantly reminding me that it would all be worth it, and the never-ending encouragement.

Ollie, thank you for holding my hand, loving me, and reminding me how capable I am in every moment that I have needed you to, throughout the writing of this thesis. I look forward to supporting you in your adventures forever and hope that I can do half as good of a job as you have for me.

And of course, to Pippin. Forever my best friend.

Publications/Contributions

During my postgraduate study within the School of Psychology, University of Birmingham, the following article was submitted for publication. The article uses the coding that I completed while conducting my own research in Chapter Three. For the paper, the co-authors advised on study design, data analysis and editing of the paper.

Ingham, M.P., **Gibbs, B.,** Colloff, M.F., Stevens, L., Rockowitz, S., Butt, M., Hayre, R., & Flowe, H.D. (under review). Harnessing meta-memory processes in police interviews as behavioural indicators of accurate recall. *Frontiers*.

Table of Contents

List of Tables and Figures

Chapter One

Research and data have demonstrated a strong link between alcohol and crime, with England and Wales statistics reporting more alcohol related offences than none alcohol related offences in 2017 (Office for National Statistics, 2018; ONS). Monds et al. (2021) conducted a survey of individuals who had witnessed a crime involving intoxication and reported that the most frequently implicated substance was alcohol (19.4%), followed by amphetamines, ecstasy, cocaine, and sedatives, with cannabis reported as being involved least often (1.7%). Crime occurs most commonly in places where alcohol is sold and consumed (Leaonard et al., 2002). The Crime Survey for England and Wales combined data sets for 2012/2013 and 2013/2014 and reported that 93% of alcohol related violent offences took place in or around a pub or club, 70% took place in a public space, and 51% of incidents occurred on the street (ONS, 2015). As well as violent incidents occurring in these settings, research has also found that bars and clubs often serve as antecedent settings to sexual aggression, including rape, attempted rape, stalking, sexual assault, and sexual harassment (Anderson et al., 2007; Fox & Sobol 2000; Graham & Wells, 2001). The Scottish Crime and Justice Agency (2021) report that in 44% of violent offences, offenders were believed to be under the influence of alcohol, and survey results showed 25% of victims consumed alcohol prior to the offence. This picture appears to be consistent globally, with 52.9% of police offers in the US stating that it is common to encounter intoxicated witnesses, and 20.2% reporting that this is very common, with the remaining officers stating it was either unusual (17.6%) or very unusual (5.9%; Evans et al., 2009). Further, despite the prevalence of intoxicated victims and witnesses being interviewed by the police (44% in the UK; Crossland et al., 2018), police officers report that they lack standardised guidelines for interviewing these witnesses and victims and that they are unclear about the procedures and their effectiveness (Crossland et al., 2018; Evans et al., 2009).

Eyewitness testimony can be essential evidence when solving crimes, and inaccurate or unreliable testimony can mislead professionals in the Criminal Justice System, or potentially lead to miscarriages of justice within the Criminal Justice System (Monds et al., 2021). When presented with eyewitness testimony, an important consideration for professionals in the system (such as solicitors) and jurors is the credibility of the witness or victim. Perceptions about this credibility may have implications not only for the approach taken to investigating the offence, but also eventual court proceedings, with intoxication affecting verdicts through perceptions of impairment (Evans & Schreiber Compo, 2010). When other evidence of guilt is lacking, the reduced credibility of a victim or witness can impact whether a suspect is prosecuted, and how they are prosecuted (Flowe, 2011). Lindsay (1994) additionally considered the effects of the severity of intoxication. They presented jurors with an eyewitness who had consumed one, eight, or twenty-four beers in four hours, and in different scenarios, the eyewitness was either confident or not confident. It was found that the verdicts of the mock jurors were not influenced by alcohol consumption by the confident eyewitness, but that guilty verdicts decreased as alcohol consumption increased by the not confident juror. Monds et al. (2021) report that the majority of survey respondents believed that alcohol had a 'large negative effect' on memory, with a higher proportion of individuals reporting that alcohol intoxicated witnesses were less credible than sober witnesses. It was also found that participants with experience in serving alcohol (such as bar staff) were more likely to report that they believed alcohol intoxicated witnesses were just as credible as sober witnesses. This perspective is in line with a review by Jores et al. (2019) that found alcohol intoxication reduced the quantity, but not accuracy, of information recalled by those that are alcohol intoxicated. Participants in the survey (Monds et al., 2021) may have had more experience in dealing with alcohol intoxicated individuals, leading to a potentially more accurate perception of their recall abilities. This demonstrates that

perceptions of witness and victim testimonies may differ dependent on personal experiences. It also indicates the importance of educating both professionals and jurors on the actual effects of alcohol intoxication on memory.

Erroneous eyewitness testimony has been found to be a leading cause of wrongful convictions by The Innocence Project (2020), further lending support to the importance of obtaining accurate testimony from witnesses and victims. This may be especially pertinent for allegations of rape where there are often conflicting versions of events in the absence of other strong, corroborating evidence (Lundrigan et al., 2019). Between 2006 and 2009 in England and Wales, 47% of offenders charged with rape who pleaded not guilty and went to trial were not convicted of rape, and these figures have further dropped to 41% in 2012, and 36% in 2017 (ONS, 2018). These falling conviction rates for sexual crimes are thought to be affected by a number of issues. Community views, stigma, and the acceptance of rape myths (attitudes and beliefs that suggest victims are at fault for being raped; Brownmiller, 2005) may reduce the rate of initial reporting (Hill & Marshall, 2018), and the decisions made by the police and legal professionals about which cases will proceed through the Criminal Justice System, based on the likelihood of conviction, can lower rates of prosecution (Pattavina et al., 2016). Unanimous verdicts must be reached by jurors in the United Kingdom, and these must be reached beyond reasonable doubt, however, misconceptions about alcohol and memory in relation to sexual offences may introduce doubt (Hamilton & Tidmarsh, 2020). This suggests a need to better understand how jurors view intoxicated victims and witnesses.

The Crown Prosecution Service (CPS) guidelines state that individuals can provide evidence in criminal proceedings if they are able to understand the questions put to them, and others can understand their responses (CPS, 2018). Credibility when referring to eyewitnesses relates to the individual's trustworthiness (how honest they are perceived to be)

and their personal expertise (a perception of the eyewitness's knowledge of the event, their accuracy, and their ability to give evidence; McGuire, 1985). Reliability is defined as being consistently good in quality or performance (Waite, 2012). Evans et al. (2019) aimed to explore mock jurors' perceptions of both witnesses and victims of crime at different levels of intoxication. They found that the mock jurors perceived the intoxicated witness and victims to be less credible and that this perception of credibility did not vary between mildly and severely intoxicated witnesses and victims. Similar findings were seen in a study by Schuller and Stewart (2000), who found that when victims were alcohol intoxicated, police officers blamed the victims more and perceived the victims as less credible. This effect is also seen elsewhere in the Criminal Justice System: 90% of psychology and law experts (from both the American Psychology-Law Society, and the European Association of Psychology and Law) agree that alcohol impairs recall (Kassin et al., 2001), and judges believe that alcohol intoxication reduces reliability (Houston et al., 2013).

Theories of alcohol induced memory impairment

A growing body of research suggests that alcohol may reduce the completeness of eyewitness memory, but not its accuracy, at low to moderate blood alcohol concentrations (Altman et al., 2018; Colloff & Flowe, 2016; Crossland et al., 2018; Flowe et al., 2016, 2017; Jores et al., 2019; Monds et al., 2019; Van Oorsouw & Merckelbach, 2012). There are a number of theories that examine why alcohol may affect the memory of those that ingest it. Alcohol Myopia Theory (Steele & Josephs, 1990) states that alcohol limits the number of external and internal cues that an individual can both perceive and process. It goes on to state that immediate and salient experiences, therefore, have a greater influence on decision-making and behaviour. These immediate and salient experiences are referred to as central details, and those outside of this are referred to as peripheral details. The theory suggests that alcohol intoxicated individuals will have better memory for central details, such as the

actions of a key figure in a scenario, than peripheral details, such as the colour of the front door of the house where the event took place. This theory is supported by Schreiber Compo et al. (2011), who found that intoxicated participants recalled fewer accurate peripheral details than sober participants, however, there was no significant difference between the number of central details recalled across groups. Alcohol Myopia Theory is, however, not supported by other studies. Van Oorsouw and Mercklebach (2012), for example, found that moderately and severely intoxicated participants recalled fewer central details. In addition, only severely intoxicated participants demonstrated reduced recall of peripheral details. Further, Flowe et al. (2016) did not find a significant difference in the recall of central or peripheral details between alcohol intoxicated or sober groups, however, they found that in general, all participants recalled central memory more accurately, regardless of experimental group. The mixed results when considering Alcohol Myopia Theory raise questions regarding the use of the theory in explaining the effects of alcohol on the memory of witnesses and victims.

Hypervigilance theory (Testa et al., 2006) states that an individual's beliefs and expectations about the impact that alcohol may have on their memory can change their behaviours and cognitions. This theory states that this effect exists without the physiological impact of alcohol itself in order to compensate for a perceived deterioration in cognitive functioning (Fillmore & Blackburn, 2002; Testa et al., 2006). Testa et al. (2006) found that women who believed they had consumed alcohol showed more cautious behaviours in a sexual assault scenario than women who did not believe they had consumed alcohol. The women that believed they had consumed alcohol, but had not, displayed lower intentions to engage in sexual approach behaviours and higher intentions to resist sexual advances from a male. It was suggested that this may be due to the women who expected they had drank alcohol recognising their own vulnerability and adapting their behaviour accordingly.

Schreiber Compo et al. (2011) lent further support for the theory, finding that placebo participants expressed more uncertainty in their answers than both sober and intoxicated participants, suggesting that they may have been regulating their memory more stringently following concerns that alcohol had impaired their memory. Gawrylowicz et al. (2019) additionally used a reverse placebo group in which participants expected to be in the sober group but were given alcohol. This group performed worse than the alcohol, placebo, and control groups during cued recall, although no significant differences were seen in free recall. That is, the reverse placebo group gave more incorrect responses and made more errors (providing inaccurate information) during the cued recall portion of a memory test than any other group. Together, this research suggests that alcohol may affect cognition, in turn weakening memory, but that knowingly consuming alcohol may allow for individuals to compensate for this. Flowe et al.'s (2019) findings demonstrated that women who expected to receive alcohol provided less complete accounts during recall than those who expected to receive tonic, regardless of beverage group. This suggests that they may have been compensating for memory impairment from alcohol by omitting any information they were not certain of. These women also showed better discriminability between correct and incorrect details on a recognition test compared to those that expected tonic water, however, their recall performance did not differ. These findings suggest that individuals who know they are consuming alcohol will compensate for the physiological effects of alcohol on memory by paying closer attention to tasks, and omitting information of which they are not certain.

Thesis aim

It is clear the importance of understanding the effect of alcohol on eyewitness testimony due to the high rates of alcohol intoxicated witnesses and victims encountered by the police (Crossland et al., 2018). The effects of alcohol on the memory of witnesses and

victims, however, are unclear, as there are mixed findings from studies in the literature to date. The literature suggests that there are conflicting perceptions of alcohol intoxicated victims and witnesses held by jurors and professionals in the Criminal Justice System. The mechanisms behind the effect of alcohol on memory reporting are also unclear. The aim of this thesis is to better understand the impact of alcohol intoxication on the memory of witnesses and victims of sexual assault. This will be achieved by reviewing the existing literature and methodologies for exploring this topic, exploring the accuracy of information given by those in a police-style interview related to a mock sexual assault, and by assessing whether metamemory indicators can provide insight into the accuracy of these witnesses and victims. The ultimate goal of work in this area is to provide clear advice to the police, jurors, and to other parts of the Criminal Justice System on how best to collect and interpret memory evidence from intoxicated witnesses and victims of crime.

It is noted that throughout this thesis, witnesses and victims are often conflated. This is a reflection of the current literature surrounding the topic, which has not clearly differentiated between the terms. While current studies have aimed to explore the impact of alcohol on memory in a forensic setting, such as witnessing or being victims to a crime, it must be noted that there is, in the real world, a clear distinction between these two groups. Victimisation is, in general, highly likely to be perceived as more traumatic than witnessing a crime, and research differs in its opinions of whether the impact of this trauma affects memory in a negative way. Some research suggests that stressful and traumatic memories can lead to impairments in memory, as these memories can be challenging to retrieve as coherent verbal narratives (Herman, 1992). However, more recent research has suggested that trauma may not impair, or may even enhance, stressful memories (Christianson, 2014). Rivera-Velez et al. (2014) noted that victims of sexual abuse did not demonstrate memory deficits, where the contents of the memory were trauma related, however no studies have compared the

impact of trauma on the memory of witnesses compared with victims. As much research suggests no deficits in trauma-related memories, and no distinction has been made between the two groups in research, the terms are used interchangeably throughout this thesis.

Chapter summaries

To achieve the thesis aim, Chapter Two takes a systematic literature review approach to exploring existing knowledge related to alcohol and misinformation. Misinformation is misleading information that individuals are exposed to, following an event (see Chapter Two for a full explanation of the concept). The review explores research that has examined the effects of alcohol on the acceptance of misinformation in those who are witness or victims of crime. The review presents mixed findings, with some studies finding that alcohol intoxication increases the acceptance of misinformation in witnesses or victims, and other studies finding no detrimental effects of alcohol intoxication. The dose dependent effects of alcohol were reviewed where possible. The review also discusses factors that might influence misinformation acceptance in sober versus alcohol intoxicated witnesses, such as delayed testing. Chapter Three uses quantitative techniques to explore the effects of alcohol intoxication on the memory of participants that have experienced a hypothetical rape scenario. The study considered metamemory, which is the knowledge one has about their own memory, as well as the process of self-monitoring memory (Pannu & Kaszniak, 2005). Specifically, it explores whether participants' use of uncertainty indicators can be used to identify accurate or inaccurate memory details reported throughout a Cognitive Interview. Chapter Four reviews the psychometric and scientific properties (e.g., reliability, validity) of the Alcohol Use Disorders Identification Test (AUDIT). The AUDIT is used not only in healthcare settings to screen for alcohol use disorders, but is particularly relevant, and frequently used, as a tool for researchers to screen participants in alcohol intoxication studies and was used when screening the participants in Chapter Three. Finally, Chapter Five

summarises findings from the previous chapters and draws conclusions from across the thesis, discussing practical implications and recommendations for future research and Criminal Justice System Policy.

Chapter Two

A Literature Review Following a Systematic Approach: The Effects of Alcohol Intoxication on the Suggestibility of Eyewitnesses

Abstract

Alcohol is often involved in criminal offences, and these offences often take place in establishments where witnesses are also likely to have consumed alcohol. Research related to alcohol and memory demonstrates that in basic memory studies, alcohol may reduce the amount of information recalled (memory completeness) but not the accuracy of this information (Jores et al., 2019). Basic memory research utilises word lists and other recognition tests to explore participants' memory. Although the evidence base related to alcohol and eyewitness memory is growing, there are still important aspects of this context to be explored. The effects of alcohol intoxication on the reporting of misinformation are still relatively under researched when considering the importance of minimising misinformation reported by witnesses in the Criminal Justice System. This chapter aimed to systematically review the literature relating to the effects of alcohol intoxication on eyewitness suggestibility. Five electronic databases were searched for relevant literature, initially yielding 363 results. Eleven papers met the inclusion criteria for this review. The quality assessments of the eleven identified papers suggested that, despite the limited amount of research, the papers included were of high quality. Comparison of the results across these studies showed mixed findings, possibly due to the vastly different methodologies employed across the research. The effects of delayed testing, differing intoxication levels, and the timing of intoxication in relation to the introduction of misinformation are discussed, along with the strengths and limitations of current methodologies. Recommendations for future research are made, and the practical implications of this research are discussed.

Introduction

Previous research has shown that alcohol use is closely linked to offending behaviour (Fast et al., 1999). According to the Ministry of Justice (2015), 52% of adults who frequently visit pubs or bars have witnessed crime, compared with 25% of adults who do not visit these establishments regularly. This would suggest that the numbers of intoxicated witnesses encountered by the police are high. Although this is a common occurrence, it is noted that there is little guidance for the police regarding the interviewing of intoxicated witnesses or victims (Crossland et al., 2018). In England, formal interviews cannot be conducted with witnesses when they are intoxicated, however, some officers collect initial statements before the witness is sober (Crossland et al., 2018). As well as having an impact on the evidence gained to progress an investigation, the intoxication of a witness can also impact the beliefs of those in the court process about the accuracy of the information provided. This is important as witnesses often provide central leads in cases (Fisher & Schreiber, 2007). Benton et al. (2006) stated that jurors in particular could have limitations in their knowledge of factors affecting the accuracy of eyewitness testimony. Jurors may perceive intoxicated witnesses to have been more cognitively impaired than sober witnesses, affecting how they assess the quality of their accounts (Evans & Schreiber Compo, 2010). This has also been seen in a survey of psychology and law experts (from both the American Psychology-Law Society, and the European Association of Psychology and Law) by Kassin et al. (2001). This survey found that 90% of respondents agreed that alcohol intoxication would impair the witnesses' performance when giving a statement, and 65% stated they would be willing to testify about the negative effects of alcohol. Despite the commonly held beliefs about the negative impact of alcohol on memory, research findings in this area are not always simple or consistent.

Research shows that intoxicated eyewitnesses may be less able to attend to multiple cues at the time of encoding and therefore may be less able to draw inferences from incoming information, or relate it to knowledge they already have (Marinkovic et al., 2004). It has been shown that alcohol intoxication at encoding decreases participants ability to discriminate previously seen items, from not previously seen items in recognition tasks (e.g., Ray & Bates, 2006). Alcohol intoxication at encoding may impact the hippocampus' ability to manage new input, meanwhile other recently formed memories (prior to consuming alcohol) are protected from retroactive interference (memory interference by any post-learning material) that they would otherwise encounter from the encoding of new information (Wixted, 2005).Others agree that disrupted hippocampus function impacts the formation of autobiographical memories and, have therefore also concluded that alcohol intoxication during encoding harms memory performance more that alcohol intoxication during retrieval (White, 2003; Mintzer, 2007). This effect of alcohol on encoding has been reliably demonstrated in basic memory research using word lists and other recognition tests. However, the methods used in basic memory research sometimes make it difficult to apply the findings of this research to forensic contexts. Unlike in basic research, in forensic contexts there is likely to be increased emotional arousal arising from witnessing a crime and witnesses' memories are tested differently, as they are asked to freely recall information during a police interview. Real-world intoxicated witnesses' memory is also likely to be more complex, containing information of both high and low salience. These limitations make it challenging to apply findings from basic memory research to real-world scenarios.

Because of the differences between basic research and forensic contexts, researchers have studied the relationship between alcohol and eyewitness memory, particularly trying to replicate the effects in real-world eyewitness contexts (Soraci et al., 2007). Applied memory literature exploring the effects of alcohol on memory accuracy and overall completeness

yield mixed results (Hagsand et al., 2013). Some studies have found that alcohol can decrease the amount of information recalled but that it does not affect the accuracy of the information that is given when compared with sober participants (Flowe et al., 2019). Other studies, however, have found no decrease in the completeness of information provided by intoxicated eyewitnesses (Schreiber Compo et al., 2011). Some studies examining the accuracy of participants' memories, regardless of completeness, have found reduced levels of accuracy in intoxicated participants compared to sober participants (Bartlett et al., 2022; Read et al., 1992; Van Oorsouw & Mercklebach, 2012). In contrast, a meta-analysis conducted by Jores et al. (2019) found that overall, alcohol significantly decreased the number of items of correct information provided by intoxicated individuals, however, did not increase the amount of incorrect information reported, meaning that while less information is reported, it is not less accurate. The Jores et al. results suggest that alcohol intoxication does affect the completeness of memory but not the accuracy. These results suggest that alcohol intoxication does affect the completeness of memory but not the accuracy. It is important to note, however, that many studies exploring the effect on alcohol of eyewitness memory occur in the laboratory where researchers are ethically limited in the amount of alcohol they can give to participants. Jores et al. (2019) state that the level of intoxication experienced by individuals can impact on memory, suggesting that any effects are, at least in part, dose dependent. Hagemann et al. (2013) state that many witnesses have a blood alcohol concentration (BAC) of over 0.18% however, participants in laboratory studies have BAC's considerably less than this, making it difficult at times to generalise laboratory findings to real-world instances of witnessing a crime.

The findings of laboratory studies of alcohol and memory can be explained through theories such as Alcohol Myopia Theory (AMT). AMT's (Steele & Joseph, 1990) suggestion that alcohol causes an individual's attention to be narrowed to salient cues, while reducing

their ability to remember peripheral information may be clear in a forensic context. In this context, it is likely to be details about the offender that are salient or central. This has been demonstrated by research that has shown that intoxicated witnesses are just as accurate as sober witnesses when identifying the perpetrator of a crime from a photographic line-up (Kneller & Harvey, 2016).

Alcohol hypervigilance may also be helpful in understanding the relationship between alcohol and eyewitness memory in a forensic context. The theory suggests that people who are aware or expect that they have consumed alcohol are aware of its negative impact on memory and expect to experience this, therefore being conservative in their reporting, or not reporting information of which they are unsure (Schreiber Compo et al., 2011). This expected impairment may impact an individual's performance to a greater extent than any actual alcohol effects (Evans et al., 2017) and may be seen in the reporting of fewer pieces of information in situations such as police interviews.

Despite the existing research exploring the relationship between alcohol and eyewitness memory, there is little research exploring the relationship between alcohol intoxication and witness suggestibility. Suggestibility in eyewitnesses refers to the vulnerability in memory which can lead to reporting of erroneous details (misinformation) when questioned about the scenario that was witnessed, either through suggestive questioning or from exposure to misleading post-event information, that individuals may accept or reject (Ridley & Gudjonsson, 2013). Misinformation is defined as misleading information presented following an event. Exposure to this can create a misinformation effect in which individuals are less accurate in their reporting of information because they include this misleading post-event information in their memory reports (Lindsay, 2008). Methods in research exist to produce suggestion-induced misinformation, such as the misinformation paradigm. In the misinformation paradigm, participants are exposed to an event and are then

provided with false information about the event. Later, it is determined whether or not these false memories are reported during memory tests (Loftus, 2005). The majority of studies examining the misinformation effect have found that the effect does exist, with the magnitude of the effect depending on the characteristics of the study and the memory test used (Blank & Launay, 2014). For example, research suggests that misinformation is more likely to be reported when the misinformation is repeated (Foster et al., 2012), or when the misinformation is presented through a narrative rather than through questions (LaPaglia & Chan, 2013)

How might witnesses' suggestibility to misinformation differ in witnesses who have versus have not been alcohol intoxicated? On the one hand, the negative effects that alcohol can have on the encoding of memories and on the completeness of a witness's memory mean it is possible that witnesses who were intoxicated compared to sober at encoding would be less able to detect discrepancies between stored and suggested information (Van Oorsouw et al., 2015). Nash and Takarangi (2011) also suggest that participants may then look to less credible sources of information to fill in gaps in their memory caused by cognitive impairments at encoding. These factors may in turn, lead to intoxicated witnesses being more likely to report misinformation compared to their sober counterparts. Put another way, because people seem to be more suggestible when their memory for the original event is weak, previously intoxicated compared to sober witnesses may be more suggestible because alcohol has disrupted encoding conditions for the original event (Zaragoza et al., 2007). It is however of note that alcohol intoxication can disrupt the consolidation of memory traces (Soraci et al., 2007) and may reduce the depth of memory traces (Birnbaum et al., 1978), leading to impaired completeness of memory. This suggests that alcohol may impair memory through a number of mechanisms.

On the other hand, it is possible that alcohol intoxication will reduce suggestibility to misinformation. According to retrograde facilitation (Wixted, 2004; 2005), if alcohol is administered after encoding but before exposure to misinformation, the individual will be less likely to report misinformation at test. This is due to alcohol disrupting the formation and encoding of new (false) memories after intoxication. The misinformation would, therefore, not disrupt the recently formed memory of the event and, in turn, decrease retrograde interference, improving memory accuracy. This is supported by Gawrylowicz et al. (2017), who demonstrated the positive effects of alcohol on reduced reporting of misinformation by an intoxicated group and a reverse placebo group where participants did not expect but were given alcohol. Clearly, the timing of alcohol consumption in relation to witnessing an event is important.

The current study

The results of existing research exploring alcohol and eyewitness suggestibility are mixed possibly due to differing methodologies employed by researchers in each study. This review aims to present a clearer picture of the research thus far. Suggestibility will be considered along with other variables that may impact the effects of alcohol intoxication such as the timing of alcohol administration and intoxication levels. It is important to understand alcohol's effects on suggestibility to misinformation as reporting misinformation when being interviewed by the police or in court may seriously impact the accuracy of potential prosecutions. Also, misinformation may be introduced in a number of ways after an individual has witnessed a crime, such as discussing what was witnessed with others (Eisen et al., 2017). This may be particularly important when considering that establishments where alcohol is consumed are often visited by pairs or groups of friends.

A preliminary search of PsychArticles, PsychInfo, Web of Science, Scopus, and Pubmed was undertaken on the 12th of February 2020 and no systematic reviews exploring

the link between alcohol intoxication and suggestibility were found. A later review by Kloft et al. was published in 2021, however, the review has been critiqued by Flowe and Schreiber Compo (2021) stating that there were concerns regarding the external validity of many of the studies referenced in the review, as well as concerns that some results are highlighted over others. While some of the same studies are included within this review, the external validity of these studies is commented on and caution is exercised in drawing conclusions that relate to police and practice, owing to this. The review was also narrative, rather than systematic, meaning only weak conclusions can be drawn (Pae, 2015). As such, the current systematic review was necessary. The main aim of this systematic review was to explore the relationship between suggestibility and alcohol, and any factors that may moderate this relationship if found.

Method

Systematic search

The systematic search for this literature review began with a scoping exercise to determine whether or not this review was viable by determining the current extent of the literature regarding suggestibility and alcohol intoxication. It was determined that there was literature in this area but that the research was limited, suggesting that a review was warranted to determine gaps in the literature to inform future research.

A search of five bibliographic databases was conducted as all were deemed to contain research related to psychology, other social sciences, and medicine which was deemed relevant to alcohol research (Booth et al., 2016). The databases used were PsychArticles, PsychInfo, Web of Science, Scopus, and Pubmed. No further papers were found through additional methods of searching, such as the use of Google Scholar or through the reference

lists of existing papers. Papers found were also discussed with an expert in the field who agreed that, in their opinion, no key papers were missing.

After preliminary searching to ensure these databases were relevant, search terms that would yield appropriate research were developed by mapping key terms from relevant papers to subject headings. Key terms can be found in Table 1. The terms were then applied to each database using Boolean operators (Appendix 1).

Table 1

Search Terms

Alcohol	**Witness**	**Suggestibility**
Alcohol	Eyewitness*	Suggestibility
Alcohol myopia	Legal testimon*	Misinformation effect
Alcohol intoxicat*	Legal evidence	False memor*
Drunk*	Witness testimon*	
Blood alcohol concentration	Evidence	
	Cross examin*	

After preliminary searching, a population, intervention, comparator, and outcome (PICO) framework (Richardson et al., 1995) was developed to ensure that only literature relevant to the aims of this review were included (Table 2).

Table 2

Inclusion and exclusion criteria

PICO	Inclusion Criteria	Exclusion Criteria	Rationale
Population	Any gender / sex Any nationality Any ethnicity Age range 18+	Below legal drinking age in the country in which the study was conducted	Any group above the legal drinking age in each country was considered likely to witness a crime while potentially under the influence of alcohol, or consuming alcohol after the event. The only exclusion criterion was that participants should be over the legal drinking age as studies would be required to give alcohol to participants or measure BAC in public houses where those consuming alcohol should be doing so legally.
Intervention/ Exposure	Participants consume alcohol or believe	Any drug aside from alcohol used	The effects of alcohol were the focus of this review. Studies in a preliminary data search have

	they are consuming alcohol Confederates (actors working for the researchers that participants believe are fellow participants) in study consume or are believed to be consuming alcohol Misinformation paradigm	Studies that did not present misinformation to participants	showed different effects of other drugs such as cannabis or stimulants that could not be compared to alcohol's effects
Comparator	Sober participants who believe they are sober Or Sober participants who believe they are intoxicated		There were no exclusion criteria for comparator due to the limited research in the area and the greatly varying methodologies used by researchers

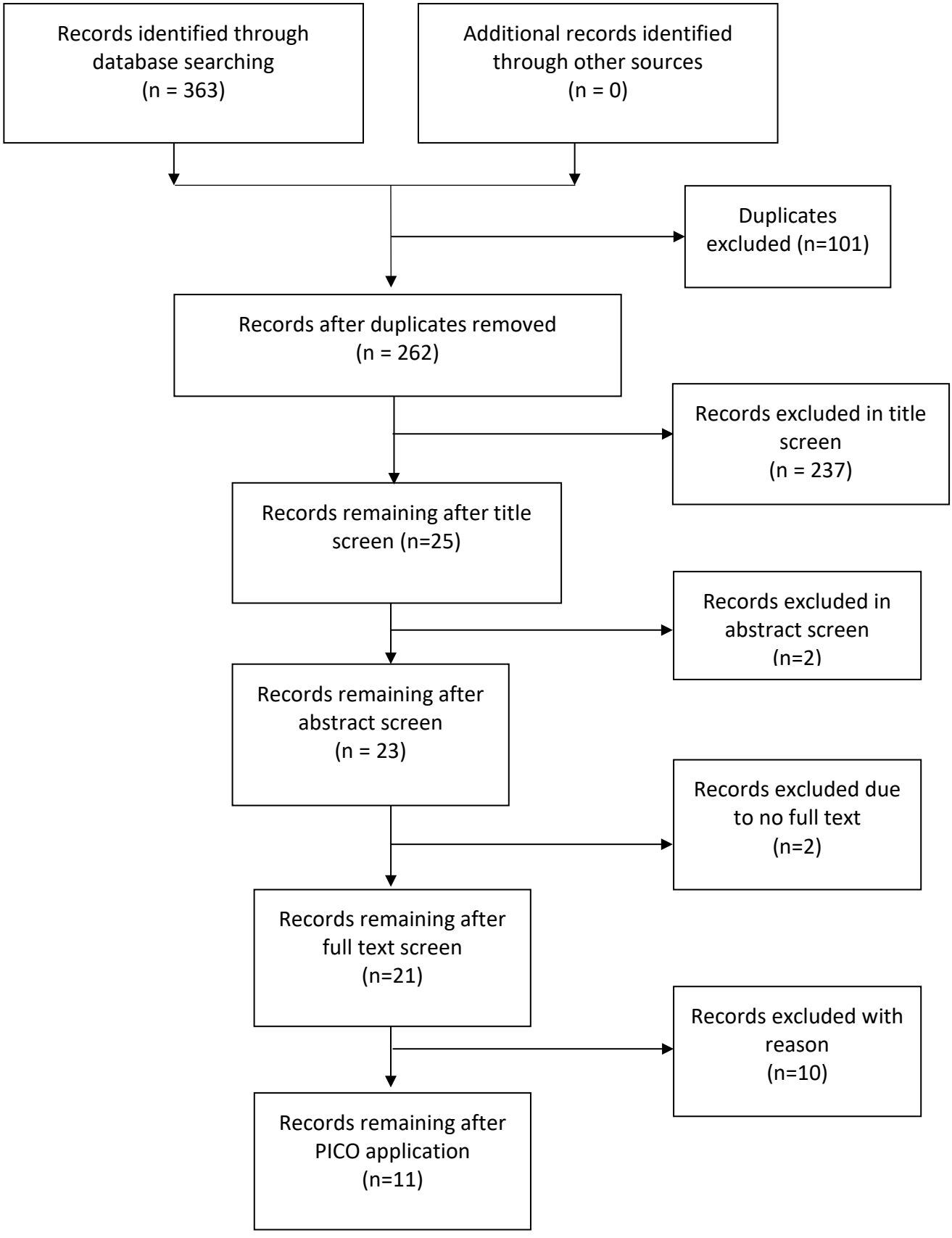
	Or Participants as own control i.e. sober testing on a different occasion		
Outcome	Misinformation reported by the participant Amount of misinformation detected at test		Outcome was defined as the amount of misinformation reported by participants at test to measure suggestibility
Study Design	Either laboratory studies or field studies		Both laboratory and field studies were included in this review, again due to differing methodology across research. Laboratory studies were most often randomised control trials.
Other Factors	Language of publication: English	Any other language	Due to time constraints, only studies that were published in English were used due to the

	Peer reviewed studies	lengthy process of translation. Studies were included, regardless of whether or not they were peer reviewed. Peer reviewed, none peer reviewed, and status unknown studies were included in order to reduce sampling bias, and due to the limited amount of research in this area.

Searches were conducted on the 18th of July 2022. Once databases had been searched and papers had been found, all duplicates were removed and papers were first sifted through by title to ensure that studies met the inclusion criteria. Two hundred and thirty-seven papers were removed due to irrelevance in title, leaving twenty-five papers. Next, papers that did not meet the PICO inclusion/exclusion framework according to their abstract were excluded. Two papers were excluded on this basis. Full texts were sourced for the remaining twenty-three papers from the University of Birmingham online library or journal websites. The remaining papers were read in full and were evaluated using the PICO framework and a further ten studies were excluded (appendix 2). No additional relevant papers were found in the references of others. Figure 1 shows the number of papers at each stage of this method and how many were removed on each basis. Full text for two papers (Ashok et al., 2016; Reich, 2003) could not be found and attempts were made to contact the researchers via Research Gate however no reply was received (appendix 3).

Figure 1

Flow chart of excluded studies



Quality assessment

The remaining eleven papers (containing twelve studies) were quality assessed. The quality assessment tools were based on checklists obtained from the Critical Appraisal Skills Programme (CASP, 2018). Two different checklists were used, because the papers being reviewed were either laboratory or field studies. This made the CASP (2018) frameworks for randomised control trials (RCT) and case cohort studies most appropriate when assessing laboratory and field studies, respectively. Both checklists were adapted to be relevant to the studies being examined (please find adapted checklists in appendix 4). The CASP (2018) case cohort study checklist states that it is designed for use with longitudinal studies. None of the included studies in this review were longitudinal studies, therefore questions 6a and 6b, relating to follow-ups, were removed. The question "What are the results of this study?" was removed from the field study checklist, as scoring for this question was likely to introduce bias through giving higher or lower scores depending on the support found for the hypothesis. In the RCT checklist, two questions were added to allow comparison between the RCT and case cohort checklists. These questions were 'have the authors considered all important confounding variables?' and 'are the results in line with other evidence?'. It was deemed appropriate to use two separate checklists to quality assess the studies, as the combining of questions into one checklist caused issues with scoring. It was possible that studies may have been scored as a lower quality for not meeting requirements of the checklist due to study type. For example, the question "Was the cohort recruited in an acceptable way?" would have differing thresholds in both laboratory and field studies. In the field, only opportunistic sampling of approaching individuals would have been accessible and this cannot be compared to the recruitment process of laboratory studies.

As CASP (20108 state that the checklists are an educational tool, no scoring system is suggested. Therefore, a scoring system was devised whereby a "yes" response scored 2,

where data was not reported, or the answer was unclear, this scored 1, a “no” response scored 0. These scores were then converted into a percentage based on the total possible score for each checklist (26 for RCT’s and 20 for field studies). When scoring effect size, a small to medium effect (i.e., >0.2<0.5) scored 1 and any effect size greater than medium (i.e., >0.5) using Cohen’s (2013) cut-off scores, scored 2. Where effect size was not reported, a score of 0 was given. In general, partial support of the hypothesis scored 1, while full support scored 2. A rating of 0 was not used for this question as finding no support for the hypothesis was not deemed to represent a poor quality study. The question “How precise were the results?” was judged based on the reporting of confidence intervals, when 95% confidence intervals were reported a score of 2 was given, no confidence interval reporting scored 0.

There was no quality assessment cut-off score that would have been used to exclude papers. No papers were excluded based on the quality assessment due to the limited amount of research in the field and a lack of clear methodological approach to this research. Quality assessment was, however, considered when interpreting the results of the studies. All studies had a quality score between 65 and 96%. All quality assessments can be found in appendix 5.

Data Extraction

A form for data extraction was created and collected data regarding the study’s aims, the population studied, methodology, BAC of participants that received alcohol, the format of MI effect testing, and the study outcome. The extracted data are displayed in Table 3.

Table 3

Data Extraction

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
Title: Absolute memory distortions: alcohol placebos influence the misinformation effect Author: Assefi, Garry Year: 2003 Location: New Zealand	To investigate whether subjects who were falsely told they were consuming alcohol would be more susceptible to misleading post event information than their counterparts who were told they were not consuming alcohol	Sample size: 117 Age: Unknown – undergraduate students Gender: Unknown	Laboratory or field study? Laboratory Control group? Yes Timing of alcohol consumption? N/A How much alcohol was given/BAC? N/A Intoxicated at encoding? No Intoxicated at misinformation? No Intoxicated at test? No	ANOVA	Yes - no alcohol given, but told alcohol group were significantly more likely to be misled	Alcohol expectancy increased suggestibility	No alcohol given, only studied expectancy	65%

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
			Format of memory test? 19 item forced choice test containing misinformation and correct information, with confidence ratings How many times tested? Once Times between tests? N/A					
Title: The effects of alcohol and co-witness information on memory reports: a field study	To examine the tendency of sober and intoxicated mock-witnesses to incorporate correct and	Sample size: 67 Age: 18-65 M= 33.4 Gender: 26 females	Laboratory or field study? Field Control group? No Timing of alcohol consumption?	Correlations with bootstrapping of 1000 samples	No	Increased intoxication levels were related to reduced overall accuracy and less complete	Only a brief delay between encoding and recall was used which is not consistent	95%

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
Author: Bartlett, Albery, Frings, & Gawrylowicz Year: 2022 Location: UK	misleading post-event information from a sober co-witness	36 males (5 participants did not disclose their gender)	Before being approached by researchers How much alcohol was given/BAC? N/A Mean BAC = 0.05% Rang = 0.01 – 0.19% Intoxicated at encoding? Yes Intoxicated at misinformation? Yes Timing of misinformation? Immediately after encoding Intoxicated at test? Yes			accounts, but did not make participants more likely to report misinformation.	with the longer delays in real-world scenarios of police interviews following a crime	

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility ?	Relationship between alcohol and suggestibility ?	Limitations	Quality Assessment
			Format of memory test? Free recall portion and 12 cued recall questions related to the video watched How many times tested? Once Times between tests? N/A					
Title: The intoxicated co-witness: effects of alcohol and dyadic discussion on memory conformity	To explore whether participants that discussed an event in pairs would incorporate more misinformation in their	Sample size: 122 Age: M= 24.10 Gender: 106 females 16 males	Laboratory or field study? Laboratory Control group? Yes Timing of alcohol consumption? Before encoding	Log-linear analysis including beverage, discussion, and misinformation reported	Yes	The dyad condition did increase misinformation compared to the individual condition. Intoxicated dyads were no more likely to	A low to moderate dose of alcohol was used as opposed to higher BAC levels in other	96%

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
and event recall Author: Bartlett, Gawrylowicz, Frings, & Albery Year: 2021 Location: UK	reports than individuals, and whether intoxicated dyads would report more misinformation that sober dyads.		How much alcohol was given/BAC? 0.6g/kg, BAC= 0.01 - 0.1% , M=0.06% Intoxicated at encoding? Yes Intoxicated at misinformation? Yes Timing of misinformation? Immediately after encoding in discussion with co-witness Intoxicated at test? Yes	Chi-square analysis		report misinformation than sober dyads and intoxicated participants were no more likely to report misinformation that sober participants	laboratory studies	

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
			Format of memory test? Questionnaire based on the video watched How many times tested? Once Times between tests? N/A					
Title: The impact of alcohol intoxication on witness suggestibility immediately and after a delay Author: Evans, Schreiber Compo,	To examine the effect of intoxication and expectancy, both at encoding and retrieval, on suggestibility. Participants were tested immediately (while in the same	Sample size: 210 Age: 21-47 M=24 Gender: 56.66% female, 43.34% male	Laboratory or field study? Laboratory Control group? Yes – sober control, placebo, and intoxicated groups Timing of alcohol consumption? Prior to encoding	Between participant ANOVA's	Yes	Witnesses who were intoxicated at encoding and retrieval were more likely to display MI effect, this relationship was driven by delayed recall participants. No significant effect of	Suggestibility questionnaire used had not been previously used therefore validity and reliability had not been evaluated	88%

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility ?	Relationship between alcohol and suggestibility ?	Limitations	Quality Assessment
Carol, Nichols-Lopez, Holness, & Furton Year: 2018 Location: USA	intoxication state as encoding), or one week later (either intoxicated, placebo, or sober).		How much alcohol was given/BAC? M BrAC= 0.08 Intoxicated at encoding? Yes Intoxicated at misinformation? Yes for immediate group, no for delayed testing Intoxicated at test? Yes for immediate testing condition, no for delayed testing Format of memory test? Open-ended free recall, nonleading			intoxication on overall acceptance of misinformation was found		

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
			cued questions, 16 forced choice questions inducing misinformation How many times tested? Once Times between tests? One week					
Title: An experimental examination of the effects of alcohol consumption and exposure to misleading post-event information on remembering a hypothetical rape scenario	To explore the impact of alcohol intoxication on the number of accurate and inaccurate details recalled by participants, whether alcohol intoxicated	Sample size: 80 Age: 18-31 M=20.36 Gender: Female	Laboratory or field study? Laboratory Control group? Yes Timing of alcohol consumption? First phase of study How much alcohol was	ANCOVA	No	Participants in the delayed condition were more likely to report MI however this was not found to be higher for intoxicated participants compared with sober participants	Participants were only interviewed on one occasion, whereas when the police interviews are used in a real-world context, victims are asked to	96%

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
	participants		given/BAC?				provide this	
Author:	would reports		0.04-0.09%				evidence on	
Flowe,	more		M=0.06%				multiple	
Humphries,	misleading						occasions	
Takarangi,	information,		Intoxicated at					
Zelek,	and whether		encoding? Yes					
Karoglu,	there would							
Gabbert, &	be a weaker		Intoxicated at					
Hope	relationship		misinformation?					
	between		No					
Year: 2019	confidence							
	and accuracy							
Location: UK			Intoxicated at					
			test? No					
			Format of					
			memory test?					
			Cognitive					
			interview or self-administered					
			interview and					
			multiple-choice					
			recognition test					
			How many times					
			tested? Once					

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
			Times between tests? 1 week for delayed condition. Misinformation was introduced at this time (i.e. 1 week after reading the scenario)					
Title: Alcohol-induced retrograde facilitation renders witnesses of crime less suggestible to misinformation Author: Gawrylowicz,	To examine the effects of alcohol intoxication after witnessing an event (encoding) but before being introduced to misleading information. This was	Sample size: 83 Age: 18-58 M=27.38 Gender: 60 females, 23 males	Laboratory or field study? Laboratory Control group? Yes Expected no alcohol and received no alcohol (control group), expected no alcohol and received alcohol (reverse placebo),	Mixed ANOVA	Yes	Participants in the intoxicated condition reported significantly fewer misinformation items on the memory test	Only one day between intoxication and test, unclear if this effect deteriorates over time	96%

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
Ridley, Albery, Barnoth, & Young Year: 2017 Location: UK	examined at test 24 hours later, once participants were sober again		and expected alcohol and received alcohol (alcohol group) Timing of alcohol consumption? After watching crime video How much alcohol was given/BAC? 0.065% Intoxicated at encoding? No Intoxicated at misinformation? Yes Intoxicated at test? No					

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility ?	Relationship between alcohol and suggestibility ?	Limitations	Quality Assessment
			Format of memory test? Cued recall memory test How many times tested? Once Times between tests? N/A					
Title: No evidence that low levels of intoxication at both encoding and retrieval impact scores on the Gudjonsson Suggestibility Scale Author: Mindthoff,	To examine whether intoxicated participants would recall the fewest details in free recall, whether they would make the most confabulation eorrors, and whether they would be	Sample size: 303 N=55 alcohol N= 57 placebo N= 53 control Age: 21-49 M= 24 Gender: 57.5% female 42.5% male	Laboratory or field study? Laboratory Control group? Yes and placebo group Timing of alcohol consumption? Before taking GSS 2 How much alcohol was	One-way ANOVAS Posthoc LSD pairwise comparisons used when significant differences were found to identify differences across the	No	Intoxicated participants recalled fewer correct details than did placebo and control participants but did not make more confabulation errors. No effects of intoxication on	Only moderate levels of intoxication were achieved which is not representative of the populations usually encountered by police. Lacked high external	85%

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
Evans, Schreiber Compo, Polanco, & Hagsand Year: 2021 Location: USA	more suggestible when presented with misleading information. Placebo participants suggestibility compared to sober participants suggestibility was also explored		given/BAC? Peak BAC of 0.06-0.08% Intoxicated at encoding? Yes (alcohol group) Intoxicated at misinformation? Yes (alcohol group) Intoxicated at test? Yes (alcohol group) Format of memory test? Gudjonsson Suggestibility Scale (GSS2) How many times tested?	three conditions		suggestibility measures emerged.	validity due to use of GSS and the lack of emotional subject which may have reduced encoding.	

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
			Completed the GSS2 once Times between tests? Delayed portion of test took place 40 minutes after the first portion of the test					
Title: Intoxicated eyewitnesses: Better than their reputation? Author: Schreiber Compo, Evans, Carol, Villalba, Ham, Garcia, & Rose	To explore possible differences in overall event recall and to explore differences in misinformation effects among intoxication levels	Sample size: 93 Age: M=24 Gender: 63% female, 37% male	Laboratory or field study? Laboratory Control group? Yes – sober participants Timing of alcohol consumption? 40 minutes prior to encoding	Mixed model ANOVA	No	Misinformation effect did not vary with intoxication level	Different methodology compared to other studies	73%

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
Year: 2011 Location: USA			How much alcohol was given/BAC? M= 0.07g/210 L BrAC Intoxicated at encoding? No Intoxicated at misinformation? Yes Intoxicated at test? Yes Format of memory test? Interviews (Open-ended narrative, open-ended cued, mixed)					

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
			How many times tested? Once Times between tests? N/A					
Title: Alcohol intoxication impairs eyewitness memory and increases suggestibility: Study one Author: Van Oorsouw, Broers, Sauerland Year: 2019 Location: the Netherlands	To investigate whether immediate repeated testing of previously intoxicated participants affected the accuracy and completeness of memory reports	Sample size: 86 Age: 18-44 M=22.9 Gender: 40 women, 46 men	Laboratory or field study? Field Control group? Yes – sober subjects found in the bar setting Timing of alcohol consumption? Before approached How much alcohol was given/BAC? Range of BAC between 0.00% and 0.16% M=0.06%	Generalised Estimating equations	Yes	Alcohol intoxication increased the acceptance of misinformation with a positive relationship between BAC and tendency to adopt misinformation. This was higher in the immediate, intoxicated test compared to the repeated sober test	Potential ceiling effects due to a limited number of questions. Delays between first and second memory varied between participants due to responding by email.	90%

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
			Intoxicated at encoding? Yes Intoxicated at misinformation? Yes Intoxicated at test? Yes Format of memory test? Photo line-up and memory test of cued, and open questions How many times tested? 44 participants responded to a follow up email to repeat the					

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
			memory test. All other participants were only tested once Times between tests? Within one week					
Title: Alcohol intoxication impairs eyewitness memory and increases suggestibility: Study two Author: Van Oorsouw, Broers, Sauerland Year: 2019	To examine the relationship between intoxication at the time of testing and to inform about the best time frame for interviewing a previously intoxicated witness	Sample size: 189 Age: 17-54 M= 26.48 Gender: 92 women, 97 men	Laboratory or field study? Field Control group? Yes – sober subjects found in the bar settings Timing of alcohol consumption? Before approached How much alcohol was	Simple slopes analysis	Partially	Higher BAC increased the tendency to adopt misinformation in the delayed only testing condition, but not the immediate testing condition	Could not disentangle the negative effects of delay and the positive effects of sobering up.	90%

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
Location: the Netherlands			given/BAC? Range= .00%-.20%, M= 0.058% in intoxicated participants Intoxicated at encoding? Yes Intoxicated at misinformation? Yes Intoxicated at test? Some tested immediately when intoxicated and after a delay, others tested only after a delay when sober Format of memory test?					

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
			Cued and open questions How many times tested? Some twice, some once Times between tests? Approximately 1 week					
Title: Alcohol impairs memory and increases suggestibility for a mock crime: a field study Author: Van Oorsouw, Merckelbach, & Smeets	To explore whether alcohol would impair memory of a mock crime event, whether alcohol intoxication would increase suggestibility, and whether	Sample size: 67 Age: 18-36 M= 22.2 Gender: 13 women, 57 men	Laboratory or field study? Field Control group? Yes Timing of alcohol consumption? Prior to approach How much alcohol was given/BAC?	ANOVA	Yes	Severely intoxicated participants displayed a stronger tendency to go along with misleading questions (yield)	Cannot control for how much alcohol was consumed after they left, possible confounding variable	90%

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
Year: 2015 Location: the Netherlands	intoxicated participants would remember more details when interviewed immediately, or when sober in a repeated interview		0.00-0.26% M=0.09% Intoxicated at encoding? Yes Intoxicated at misinformation? Yes Intoxicated at test? Yes Format of memory test? Free recall, cued recall, misleading questions How many times tested? Twice Times between tests? 3-5 days					

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility ?	Relationship between alcohol and suggestibility ?	Limitations	Quality Assessment
Title: Trusshtme, I know what I sshaw: The acceptance of misinformation from an apparently unreliable co-witness Author: Zajac, Dickson, Munn Year: 2016 Location: New Zealand	To examine whether the consumption of alcohol by a co-witness would increase the likelihood of accepting misinformation	Sample size: 100 Age: 18-35 M=20.6 Gender:63 female, 37 males	Laboratory or field study? Laboratory Control group? Yes, half of the participants were randomly assigned to a sober co-witness group Timing of alcohol consumption? N/A How much alcohol was given/BAC? N/A Intoxicated at encoding? No Intoxicated at misinformation? No	Poisson log linear regression analysis	Yes	Overall, there was no significant difference in the reporting of misinformation between the sober and intoxicated eyewitness groups. An effect was found when participants responses were discrepant with the misinformation. When a discrepancy was detected, participants were more likely to	Did not directly test participants with alcohol however are important when considering participants that may be intoxicated and looking to others for “accurate” information	81%

Key information	Aims	Population	Method	Statistical analysis used	Effect on suggestibility?	Relationship between alcohol and suggestibility?	Limitations	Quality Assessment
			Intoxicated at test? No Format of memory test? Cued recall questions and a photo ID line-up How many times tested? Once Times between tests? N/A			accept misinformation from a sober co-witness.		

Descriptive overview of results

Methodologies/Population

Of the eleven papers in this review, four were field studies (one paper contained two separate field studies; (Bartlett et al., 2022; Van Oorsouw et al., 2015; Van Oorsouw et al., 2019) and eight were laboratory studies (Assefi & Garry, 2003; Bartlett et al., 2021; Evans et al., 2018; Flowe et al., 2019; Mindthoff et al., 2021; Schreiber Compo et al., 2011; Zajac et al., 2016). The total number of participants studied across all research was 1,025. Of these participants, 26.9% were recruited in the field in public houses, one study did not specify where participants were recruited (Assefi & Garry, 2003), and the rest were recruited through university research credit schemes or universities in general (Bartlett et al., 2021; Evans et al., 2018; Flowe et al., 2019; Mindthoff et al., 2021; Schreiber Compo et al., 2011; Zajac et al., 2016) with participants being staff and students at the university. The ages of participants across the studies ranged from 17-65 years old. It is of note that despite the PICO inclusion criteria of participants being of the legal age to consume alcohol, one study did include data from a participant who was 17 years of age in a country where the legal drinking age is 18 years of (Van Oorsouw et al., 2019). There is no information explaining why this participant was included in the study to begin with, however, it is assumed that due to this being a field study, researchers would have assumed that to be in the establishment and drinking alcohol, the participant would have been of legal drinking age. The researchers note that by the second session of data collection, this participant had turned 18 years old, therefore, the data were still included in the study. Differences in gender of the participants across the studies could not be accurately calculated as one study (Assefi & Garry, 2003) did not report participant gender, and other papers reported gender splits in percentages with no decimal places resulting in inaccurate totals (i.e., summed less or more than 100%). Four of the studies were conducted in the UK (Bartlett et al., 2021; Bartlett et al., 2022; Flowe et al.,

2019; Gawrylowicz et al., 2017), two in the USA (Evans et al., 2018; Schreiber Compo et al., 2011), two in New Zealand (Assefi & Garry, 2003; Zajac et al., 2016), and three in the Netherlands (Van Oorsouw et al., 2015; Van Oorsouw et al., 2019).

In the studies, the misinformation was introduced to participants through overhearing misleading information or discussion with a confederate (an actor working for the researchers who participants believed was also a participant; Bartlett et al., 2021; Bartlett et al., 2022; Schreiber Compo et al., 2011; Zajac et al., 2016), reading misleading narratives (Assefi & Garry, 2003; Flowe et al., 2019; Gawrylowicz et al., 2017), suggestive questions (Van Oorsouw et al., 2015; Van Oorsouw et al., 2019), or misleading answers being provided at test, in this case, being provided with a test sheet that already had incorrect answers circled (Evans et al., 2018).

Overall findings relating to alcohol and misinformation

Below are the findings for each study related simply to the effect of alcohol intoxication on the reporting of misinformation at test. Table 4 categorises the studies and the effects of alcohol on misinformation reporting.

Gawrylowicz et al. (2017) explored the potential protective effects of alcohol when participants became intoxicated after encoding but before receiving misinformation. They found that the control group (who did not receive any alcohol) were more likely than the alcohol (p = .001) and reverse placebo groups (who did receive alcohol, but were told they had not; p = .002) to report misinformation. A main effect of condition was found in a mixed ANOVA ($F(2, 79) = 3.51$, $p = .035$, $\eta 2 = .088$). This suggests that the physiological effect of alcohol has a protective effect on witness suggestibility, when the participant is intoxicated at the time of misinformation presentation, for example if a witness or victim drank alcohol as a means of coping after an event.

Van Oorsouw et al. (2015) studied the suggestibility of participants in the field, some of whom were highly intoxicated. They found that severely intoxicated participants (at encoding and the introduction of misinformation) were the most likely to go along with misleading questions both at immediate and follow up tests when sober ($t(39)=2.94$, $p=.01$). This lends support for the suggestion that more misinformation may be reported, when participants are at higher levels of intoxication than can be achieved ethically in a laboratory.

The first field study conducted by Van Oorsouw et al. (2019) found a main effect of BAC on misinformation acceptance (i.e., the number of misled items reported, $b = 0.87$), demonstrating that an increase in BAC during encoding increased the likelihood of reporting misinformation. This alcohol effect was increased upon immediate test, when participants were still intoxicated. In study two, simple slopes analyses showed "marginally significant" correlations between BAC and the tendency to adopt misinformation ($r = .15$; small effect size) and significant correlations between BAC and memory completeness ($r = -.40$; small-medium effect size). It is, however, of note that when comparing acceptance of misinformation between the immediate and delayed testing only groups, an increase in BAC only increased the tendency to adopt misinformation in the delayed only testing group. Overall, these findings lend partial support for increased BAC at encoding leading to increased reporting of misinformation.

Evans et al. (2018) tested all of their participants only once, however, some participants were tested immediately after the misinformation was given and some were only tested after a delay allowing them to sober up. Three intoxication groups were used, intoxicated, sober, and placebo. In the delayed testing group, participants were randomly assigned to one group during encoding, and one group during retrieval (i.e. the state they were in at encoding was not necessarily the state they were in at retrieval). Evans et al. (2018)

found increased suggestibility in participants that were intoxicated at encoding and retrieval, however, this effect was mainly seen in those who were in the delayed testing condition. Those who were interviewed immediately (the intoxicated group) were less likely to agree with inaccurate suggested responses (d = 2.03; large effect size) than those who were interviewed sober, after a delay. This study did not disentangle the effects of delayed testing and the misinformation effect as it was unclear whether the poor performance of the delayed testing group that were intoxicated at encoding and sober at retrieval was due to the time before testing, or the change of intoxication state. However, these findings suggest that misinformation may be provided by previously intoxicated witnesses due to less complete memories after a delay in which they sobered up.

Schreiber Compo et al. (2011) did not find any significant differences in misinformation reporting across their intoxication groups. Specifically, there was no difference across the three intoxication levels in the percentage of misinformation that each group provided ($F(2, 75) = .43$, *ns*). There was also no interaction between believability ratings (whether or not participants believed they were intoxicated at the .08 BAC level) and intoxication level for the percentage of misinformation given ($F(2, 76) = 1.15$, *ns*). This demonstrates that whether or not participants believed the mock crime did not account for the lack of difference observed between the alcohol and no alcohol groups. There was also no significant difference found in the percentage of correct control (i.e., non-misled) items reported across the intoxication levels ($F(2, 90) = .13$, *ns*), which indicates that intoxicated participants did not show noticeably impaired memory compared to sober witnesses or witnesses who believe they are intoxicated.

Flowe et al. (2019) reported similar findings to Evans et al. (2018). They did not find a main effect of beverage, but found that there was a significant main effect of delayed

testing ($F(1, 64 = 7.38, p = .01)$ with more misinformation being reported after a longer delay in testing (Pearson's r = 0.33). They, therefore, concluded that alcohol was not associated with recalling more misinformation in this context, but that it was the effects of delayed testing that caused an increase in misinformation reporting.

Mindthoff et al. (2021) explored the effect of alcohol intoxication on participant suggestibility with the use of a validated suggestibility measure (Gudjonsson Suggestibility Scale 2) when participants' intoxication state was the same at encoding and retrieval. While intoxicated participants recalled the fewest number of correct details at immediate and delayed free recall, none of the suggestibility measures differed significantly across intoxication groups. It was suggested by Mindthoff et al. (2021) that the intoxicated individuals reported less correct information due to omitting information that they were not certain about. Moreover, it is possible that the lack of difference in susceptibility to misinformation across the alcohol groups may have been due to the low-moderate intoxication levels employed in this study.

Bartlett et al. (2022) explored the effects of co-witnesses and alcohol intoxication on witness memory during a field study. They found that acute alcohol intoxication did negatively impact eyewitness memory performance, with increased alcohol intoxication being associated with decreased completeness of accounts, and poorer accuracy. However, they found that participants with higher BAC's were not more suggestible and did not report more misinformation than participants with lower BACs, however, they were less confident in their responses. Bartlett et al. (2022) suggest that the lack of significant difference in the reporting of misinformation suggests that alcohol intoxication does not diminish participants' source-monitoring ability. The intoxicated participants were not less able to determine the source of their memories, and whether they came from another individual.

Assefi and Garry (2003) found that participants who believed they had consumed alcohol were significantly more likely to report misinformation than those who were told they had not consumed alcohol, even though neither group had received alcohol. Specifically, there was a significant interaction between drink condition and post-event information with a small to medium effect size ($f = 0.38$). They found no significant difference in performance on control items across the drink expectancy conditions. Together, this indicates that the suggestion that participants had consumed alcohol was enough to make them more susceptible to misinformation, but not more likely to provide incorrect answers regarding information on which they had not been misinformed, not alcohol itself.

Bartlett et al. (2021) also explored the effects of both alcohol intoxication and co-witness discussion on eyewitness memory and suggestibility during a laboratory study. They found that participants who took part in discussions with a co-witness were more likely to report misinformation, but that this finding was not influenced by alcohol intoxication. Intoxicated individuals were as susceptible to reporting misinformation as sober participants.

Zajac et al. (2016) used apparent co-witness intoxication to study participants' willingness to accept misinformation from reliable or unreliable sources. This was achieved through the participant witnessing the confederate drinking what appeared to be alcoholic beverages. Although the number of pieces of misinformation reported was not related to participant ratings of the co-witnesses' ability to complete the tasks, results showed that participants were more likely to accept misinformation from a sober co-witness than a drunk co-witness. This alcohol effect was, however, only present when the witness's pre-misinformation response was clearly discrepant with the misinformation. To explain, the co-witness in the study asked the participants about a detail in the video before introducing the misinformation. If the participant answered 'I don't know', they were no more likely to

accept information from a sober compared to an intoxicated co-witness. If the participant gave an answer, however, they were presumably able to detect a discrepancy between their answer, and that of the co-witness, and were more likely to report misinformation from the sober co-witness than the intoxicated co-witness, suggesting a source credibility effect. They found that more reports of misinformation were incorporated into accounts of participants that had interacted with a sober co-witness (64%) than a co-witness who appeared intoxicated (32%) when misinformation was discrepant with pre-misinformation response. When this discrepancy was not clear, (i.e., an "I don't know" answer was given), co-witness condition did not appear to play a role in reporting misinformation (45% in the sober co-witness condition; 43% in the intoxicated co-witness condition).

Table 4

Categorised findings of studies

Key information	Field or laboratory study?	Effects on suggestibility	Relationship between alcohol and suggestibility	Quality assessment
Negative effects of alcohol on suggestibility				
Title: The impact of alcohol intoxication on witness suggestibility immediately and after a delay Author: Evans, Schreiber Compo, Carol, Nichols-Lopez, Holness, & Furton Year: 2018 Location: USA	Laboratory	Yes	Increase in misinformation reporting for participants in the alcohol group, driven by delayed testing. No significant effect of intoxication on overall acceptance of misinformation was found	88%
Title: Alcohol intoxication impairs eyewitness	Field	Yes	A positive correlation between BAC and misinformation reporting was found.	90%

memory and increases suggestibility: Study one Author: Van Oorsouw, Broers, Sauerland Year: 2019 Location: the Netherlands			This was higher at immediate, intoxicated test, compared to the repeated, sober test.	
Title: Alcohol intoxication impairs eyewitness memory and increases suggestibility: Study two Author: Van Oorsouw, Broers, Sauerland Year: 2019 Location: the Netherlands	Field	Yes	Higher BAC increased the tendency to adopt misinformation in the delayed only testing condition, but not the immediate testing condition	90%
Title: Alcohol impairs memory and increases suggestibility for a mock crime: a field study Author: Van Oorsouw, Merckelbach, & Smeets Year: 2015 Location: the Netherlands	Field	Yes	Severely intoxicated participants displayed a stronger tendency to go along with misleading questions (yield)	90%

No effect of alcohol on suggestibility

Title: The effects of alcohol and co-witness information on memory reports: a field study Author: Bartlett, Albery, Frings, & Gawrylowicz Year: 2022 Location: UK		No	Increased intoxication levels were related to reduced overall accuracy and less complete accounts, but did not make participants more likely to report misinformation.	95%
Title: The intoxicated co-witness: effects of alcohol and dyadic discussion on memory conformity and event recall Author: Bartlett, Gawrylowicz, Frings, & Albery Year: 2021 Location: UK	Laboratory	No	The dyad condition did increase misinformation compared to the individual condition. Intoxicated dyads were no more likely to report misinformation than sober dyads and intoxicated participants were no more likely to report misinformation that sober participants	96%
Title: An experimental examination of the effects of alcohol consumption and exposure to misleading post-event information on remembering a hypothetical rape scenario Author: Flowe, Humphries, Takarangi, Zelek, Karoglu, Gabbert, & Hope Year: 2019 Location: UK	Laboratory	No	Participants in the delayed condition were more likely to report MI however this was not found to be higher for intoxicated participants compared with sober participants	96%
Title: No evidence that low levels of intoxication at both encoding and retrieval impact	Laboratory	No	Intoxicated participants recalled fewer correct details than did placebo and control participants	85%

scores on the Gudjonsson Suggestibility Scale Author: Mindthoff, Evans, Schreiber Compo, Polanco, & Hagsand Year: 2021 Location: USA			but did not make more confabulation errors. No effects of intoxication on suggestibility measures emerged.	
Title: Intoxicated eyewitnesses: Better than their reputation? Author: Schreiber Compo, Evans, Carol, Villalba, Ham, Garcia, & Rose Year: 2011 Location: USA	Laboratory	No	Misinformation effect did not vary with intoxication level	73%
Positive effect of alcohol on suggestibility				
Title: Alcohol-induced retrograde facilitation renders witnesses of crime less suggestible to misinformation Author: Gawrylowicz, Ridley, Albery, Barnoth, & Young Year: 2017 Location: UK	Laboratory	Yes	Participants in the intoxicated condition reported significantly fewer misinformation items on the memory test	96%
Social effects of alcohol on suggestibility				
Title: Absolute memory distortions: alcohol placebos influence the misinformation effect	Laboratory	Yes - no alcohol given, but told alcohol group were significantly more likely to be misled	Alcohol expectancy increased suggestibility	65%

Author: Assefi, Garry Year: 2003 Location: New Zealand				
Title: Trussht me, I know what I sshaw: The acceptance of misinformation from an apparently unreliable co-witness Author: Zajac, Dickson, Munn Year: 2016 Location: New Zealand	Laboratory	Yes	Overall, there was no significant difference in the reporting of misinformation between the sober and intoxicated eyewitness groups. An effect was however found when participant's responses were clearly discrepant with the misinformation). When a discrepancy was detected, participants were more likely to accept misinformation from a sober co-witness.	81%

Discussion

The aim of the current systematic review was to examine research that explores the relationship between alcohol intoxication and the suggestibility of eyewitnesses of crime. The studies that were deemed relevant after PICO criteria had been applied had a wide range of methodologies and results were also varied. A total of eleven papers were included with one paper including two separate, but related, studies. Three of the papers included studied suggestibility in the field in pubs and bars (Bartlett et al., 2022; Van Oorsouw et al., 2015;

Van Oorsouw et al., 2019), while the remaining eight studies were laboratory studies (Assefi & Garry, 2003; Bartlett et al., 2021; Evans et al., 2018; Flowe et al., 2019; Gawrylowicz et al., 2017; Mindthoff et al., 2021; Schreiber Compo et al., 2011; Zajac et al., 2016). Only two studies did not include a condition in which participants were intoxicated (Zajac et al., 2016; Assefi & Garry, 2003).

Ten of the twelve studies investigated whether the physiological effects of alcohol increased the reporting of misinformation (Bartlett et al., 2021; Bartlett et al., 2022; Evans et al., 2018; Flowe et al., 2019; Gawrylowicz et al., 2017; Mindthoff et al., 2021; Schreiber Compo et al., 2011; Van Oorsouw et al., 2015; Van Oorsouw et al., 2019) and four found that it did (three of which were field studies; Evans et al., 2018; Van Oorsouw et al., 2015; Van Oorsouw et al., 2019 Study 1-2). Van Oorsouw et al. (2019) concluded that intoxication during encoding and the introduction of misinformation did make individuals more prone to reporting misinformation. They concluded that this effect was mediated by memory completeness, supporting the idea of discrepancy detection theory. The theory states that it is more difficult for individuals to detect discrepancies between their memory and suggested details when the memory of the event is poor (Schooler & Loftus, 1986). This is further supported by the second study in this paper (Van Oorsouw et al., 2019) which noted that although intoxication increased misinformation reporting, this was more likely in the delayed only testing condition. This suggests that the delay, coupled with alcohol intoxication, resulted in less complete memory, causing more misinformation to be reported.

Further support for this suggestion comes from Evans et al. (2018), who also found increased suggestibility in participants that were intoxicated at encoding, but also found this difference in suggestibility to be more apparent when testing was delayed by one week. They state that these findings have important implications for police guidance as the increase in

suggestibility after a delay suggests that officers should attempt to obtain initial statements containing less information when intoxicated individuals are questioned, as soon as possible to the time of the event. Other findings, however, did not seem to support the recommendation that intoxicated witnesses should be interviewed as soon as possible to limit misinformation effects. Van Oowsouw et al. (2015) found that severely intoxicated individuals (intoxicated at encoding, misinformation, and test) were more likely to endorse suggested events both at immediate testing when still intoxicated and at delayed testing when sober. They also found that severely intoxicated individuals performed better on cued recall memory tests when they had sobered up compared to when they were tested while intoxicated (Van Oorsouw et al., 2015).

Four of the studies reviewed did not find a negative effect of alcohol on the acceptance of misinformation, including Flowe et al. (2019). Flowe et al. (2019) found that participants in the delayed testing condition were more likely to report misinformation at test, however, this was not found to be more likely for participants that had been intoxicated. This suggests that it is the decay of memory over time that contributed to the reporting of misinformation in this study, rather than the direct effects of alcohol intoxication. This finding is in line with the others in this review (Bartlett et al., 2022; Mindthoff et al., 2021; Schreiber Compo et al., 2011) and lends support for questioning witnesses soon after the event, even if intoxicated, rather than delaying their questioning.

Bartlett et al. (2022), Mindthoff et al. (2021), and Schreiber Compo et al. (2011) further support Flowe et al.'s (2019) finding regarding misinformation and alcohol intoxication. Bartlett et al. (2022) found no difference in the amount of misinformation reported by sober and intoxicated participants, as did Mindthoff et al. (2021). Results from Schreiber Compo et al.'s (2011) study found that there was no significant difference in the

amount of misinformation reported by participants across the three intoxication groups (alcohol intoxicated, placebo, and control), and further, did not find any memory impairment in intoxicated participants. Bartlett et al. (2021) found no effect of alcohol intoxication on the reporting of misinformation. Participants were asked to recall either alone or with a co-witness. The acceptance of misinformation by participants did not differ by alcohol intoxication, however, there was an increase in the reporting of misinformation in those that took part in discussion with a co-witness. These findings suggest that alcohol may not negatively affect the likelihood of witnesses and victims to report misinformation, and that other factors should be considered when reviewing the research, as discussed in themes.

Uniquely, Gawrylowicz et al. (2017) explored the effect of consuming alcohol following witnessing an event, but before the introduction of misleading information. Findings suggest that alcohol consumption at this stage has a protective effect on participants' memory, with control participants (who did not receive alcohol) being more likely to report misinformation. While this study was the only to consider the protective effects of alcohol, it lends support for the notion that alcohol intoxication is not necessarily detrimental to memory.

Two of the three studies that investigated the social aspects of alcohol, found effects. Assefi and Garry (2003) also found negative impacts of alcohol on misinformation reporting, however, this study did not give participants any alcohol therefore those that reported more misinformation at test were only under the impression that they were intoxicated. This lends support to the suggestion that some aspects of memory can be affected by expected psychological effects as well as physiological factors. This suggests that memory is not only a cognitive function but that expectancy also plays a role in alcohol and memory research (Neisser & Hyman, 2000). It is possible that those who thought they were intoxicated were

less confident in their own memories and, therefore, looked to another source. Zajac et al. (2016) also suggested source credibility effects to explain their findings that witnesses were more likely to accept misinformation from a sober co-witness than an intoxicated co-witness in some conditions. This does also suggest a social component of accepting misinformation, perhaps due to the widespread belief that alcohol impairs memory. However, it is of note that this cannot directly support Assefi and Garry's (2003) findings as the participants in Zajac et al.'s study did not expect to receive any alcohol themselves. It is also of note that Assefi and Garry's (2003) scored lowest across all studies in quality assessment, being given a score of 65%, which calls into question the weight that these results can be given in the systematic review.

Further examining the social effects of alcohol intoxication on the acceptance of misinformation, Zajac et al. (2016) found that although participants in both the sober and intoxicated eyewitness conditions were similarly likely to report misinformation, when witnesses did not have a pre-misinformation response (i.e., an "I don't know" response to a question") they were no more likely to accept misinformation from a sober or intoxicated co-witness. Participants who did provide a pre-misinformation response that was discrepant with the misinformation were more likely to accept misinformation from a sober co-witness than an intoxicated co-witness This lends support for social variables such as source credibility, whereby the credibility of a source may determine whether a witness detects a discrepancy between their account, and that of a co-witness, and how they resolve this discrepancy, for example, accepting or rejecting the discrepant information from the co-witness (Lampinen & Smith, 1995). In this case, a participant may have viewed the sober witness as credible and accepted the discrepant information they provided. The opposite may have been true with intoxicated co-witnesses due to the individual's strong belief held by many that alcohol will impair their memory (Assefi & Garry, 2003). This has practical implications due to the co-

witness discussions that often occur after a crime has been witnessed (Paterson & Kemp, 2006). When it is considered that witnesses of crime have often consumed alcohol, these findings become even more relevant when considering their acceptance of misinformation. This demonstrates the fallibility of eyewitness information when eyewitness memory has been contaminated (Zajac et al., 2016).

Themes

Across the twelve studies, various themes were introduced and discussed, as well as several factors that may interact with the effect of alcohol on suggestibility, such as the timing of misinformation introduction. The themes of alcohol consumption, and delayed testing are discussed next.

Level of intoxication

There is some evidence to suggest that the deleterious effects of alcohol on suggestibility are due to level of intoxication. The four field studies included in this review were able to test participants who had higher BAC's than those who were tested in the laboratory, as laboratory studies were not able to dose higher than a moderate level of intoxication to gain ethical approval. Hagemann et al. (2013) state that many real-life witnesses have a BAC of over 0.18%, therefore, field studies may arguably be a more appropriate method of researching in this area. Out of the four field studies included in the review, two (Van Oorsouw et al., 2015; Van Oorsouw et al., 2019) found that these levels of intoxication (i.e., over 0.06% BAC on average) did increase the levels of misinformation reported by participants. This lends some support to the idea that witnesses typically encountered by police forces would be more prone to suggestibility. The final field study, by Bartlett et al. (2021), examined participants with BAC's ranging from 0.01 to 0.19%. They

found that participants with higher BAC's were no more likely to report misinformation than those with a lower BAC. They did, however, find that as BAC increased, the completeness and accuracy of information recalled decreased. This suggests that while participants with higher BAC's may provide less information and information that is generally less accurate, they are no more likely to provide misinformation that had been presented to them in their reports. Together, this suggests that the effects of alcohol dose on the acceptance of misinformation are unclear. The average dose reported by Van Oorsouw et al. (2019) was .06% which is comparable dose to laboratory studies that have not found an effect of alcohol intoxication on the reporting of misinformation (Flowe et al., 2019; Mindthoff et al., 2021; Schreiber Compo et al., 2011) or even fewer instances of misinformation reporting (Gawrylowicz et al., 2017). It is also of note that despite the high intoxication levels seen in Van Oorsouw et al.'s (2015) study, whereby severe intoxication was related to an increase in suggestibility, it is important to note that cause and effect cannot be determined in field studies due to other uncontrolled variables. For example, it is possible that these effects may be due to other underlying characteristics of those that drink heavily in these settings. Further support comes from Mindthoff et al. (2021), who did not find significant differences in suggestibility between beverage groups and hypothesised that this was in part due to low intoxication levels. The control of other factors that may influence suggestibility which is only possible in the laboratory.

When considering the quality of the studies, both conducted by Van Oorsouw et al. (2019) and the study conducted by Van Oorsouw et al. (2015) are high quality papers, with all having a quality assessment score 90%. This is, however, also true of Bartlett et al. (2021), scoring 95% on quality assessments and finding no increase of misinformation reporting at higher levels of intoxication. Overall, the quality of the four field studies are comparable and of high quality, despite their mixed findings, therefore it is likely that confounding variables

that could not be controlled in a field setting contributed to the mixed findings. While confounding variables such as noise in the places where the study took place, as well as varying emotional salience of the material (for example, in Van Oorsouw et al.'s 2019 study, the event to be remembered was a conversation which is likely to have lower emotional salience than witnessing a crime) were acknowledged and discussed by researchers, the impact on outcomes cannot be determined.

The studies conducted in laboratory settings by Flowe et al. (2019), Mindthoff et al. (2021), Schreiber Compo et al. (2011), that found no effects of alcohol intoxication ranged in quality assessment scored, scoring 96%, 85%, and 73% respectively. While the quality assessment of these laboratory studies was lower, it is also of note that confounding variables can be controlled better in a laboratory setting. Positive effects of alcohol intoxication were also found by Gawrylowicz et al. (2017; quality assessment score of 96%). Therefore, these findings overall remain somewhat unclear.

Timing of alcohol consumption

Across the studies, the timing of alcohol consumption varied in relation to encoding, the introduction of misinformation, and test. In eight of the studies, participants were intoxicated at the time of encoding (Bartlett et al., 2021; Bartlett et al., 2022; Evans et al., 2018; Flowe et al., 2019; Mindthoff et al., 2021; Van Oorsouw et al., 2015; Van Oorsouw et al., 2019). In nine of the studies, participants were intoxicated when misinformation was introduced (Bartlett et al., 2021; Bartlett et al., 2022; Evans et al., 2018; Gawrylowicz et al., 2017; Mindthoff et al., 2021; Schreiber Compo et al., 2011; Van Oorsouw et al., 2015; Van Oorsouw et al., 2019). However, it is of note that only some of the participants in Evans et al.'s (2018) study were intoxicated when misinformation was introduced as this was part of the testing phase which was delayed for some participants, allowing them to sober up before-

hand. Eight studies had at least some participants intoxicated at test (Bartlett et al., 2021; Bartlett et al., 2022; Evans et al., 2018; Mindthofff et al., 2021; Schreiber Compo et al., 2011; Van Oorsouw et al., 2015; Van Oorsouw et al., 2019). Not all participants were intoxicated at test in Study 2 of Van Oorsouw, et al.'s (2019) study or in Evans et al. (2018) due to being in placebo or control groups. Assefi and Garry (2003) and Zajac et al. (2016) did not provide any alcohol to participants, therefore, none were intoxicated at any point throughout the study in order to test social components of suggestibility. In Gawrylowicz et al. (2017) study, participants in the alcohol group were only intoxicated when misinformation was introduced, not at encoding or retrieval.

Of the seven studies where participants were intoxicated at encoding, four found that suggestibility increased with alcohol consumption. Van Oorsouw et al. (2019; study one) found that alcohol intoxication did significantly increase suggestibility in participants and found a significant positive correlation between BAC and the tendency to adopt misinformation with a medium effect size (r = .27) In the second study conducted in the paper, it was found that a higher BAC increased the acceptance of misinformation but only in the delayed testing group, and not in the immediate testing group (Van Oorsouw et al., 2019). Evans et al. (2018) also found increased suggestibility with alcohol intoxication, finding a significant main effect of intoxication on the amount of misinformation specifically at encoding ($\eta p^2 = 0.13$; large effect size) with no main effect of intoxication condition at retrieval. Van Oorsouw et al. (2015) found that it was only severely intoxicated participants that displayed a stronger tendency to report misinformation than other intoxication levels or sober participants. However, Flowe et al. (2019) tested participants that were intoxicated at encoding and did not find an increase in suggestibility with intoxication. They found no main effect of intoxication on increased misinformation intrusions with a small effect size ($\eta p^2 =$

0.032). This was also found by Bartlett et al. (2021) and Bartlett et al. (2022) who found no increase in suggestibility in participants that were intoxicated at encoding.

The three studies that found no significant effect of alcohol intoxication at encoding on the reporting of misinformation were slightly higher in quality than those who found a negative impact. Bartlett et al. (2021) had a quality assessment score of 96%, Bartlett et al. (2022) scored 95%, and Flowe et al. (2019) scored 96%. Scores for those that found a negative impact of alcohol intoxication (Evans et al., 2018; Van Oorsouw et al., 2019; Van Oorsouw et al., 2015), had quality assessment scores ranging between 88% and 90%. While these differences in quality assessment score are small, three of the four studies that found a negative impact were field studies, while only one of the three studies that found no negative impact were. This may suggest that more weight should be placed on higher quality studies, and those with less confounding variables (laboratory studies as opposed to field studies). Other variables such as noise and distractions in the field setting may have had a greater negative impact on the reporting of misinformation, than alcohol itself.

In studies where participants were intoxicated during the presentation of misinformation, but not at encoding, the following effects were found. Schreiber Compo et al. (2011) found that across all levels of intoxication, participants reported higher proportions of misinformation on questions they had been misinformed about, compared to control items. They did not find that participants that were intoxicated at the time misinformation was given, were more likely to report this misinformation than sober groups ($F(2, 90)=1.36$, *ns*). Gawrylowicz et al. (2017) did not find any negative effects of intoxication at misinformation either, however, unlike Schreiber Compo et al. (2011) they found positive effects. Their results demonstrated that the control group provided significantly more misinformation at test than the alcohol and reverse placebo group (who expected not to get alcohol but did).

Participants were more likely to provide a misled response to misinformation items than control items with a main effect of condition (η^2 = .088; a medium effect size). Together, these two studies suggest a protective effect of alcohol when alcohol is introduced only at the misinformation stage.

When considering this impact of alcohol intoxication on the reporting of misinformation when participants were intoxicated at the introduction of misinformation, results from the two studies above are mixed. It should however be considered that the study by Schreiber Compo et al. (2011) was given a quality assessment score of 73%, while the study by Gawrylowicz et al. (2017) was given a quality assessment score of 96%. This suggests that more weighting should be placed on the positive effects of alcohol intoxication at misinformation introduction, found by Gawrylowicz et al. (2017) due to the studies overall higher quality.

Gawrylowicz et al. (2017) dosed participants with alcohol after encoding but before they were presented with misinformation. They found that participants in the intoxicated group were less likely to report misinformation than those in the control or reserve placebo group. This supports the theory of retrograde facilitation (Wixted, 2004; 2005). This demonstrates the importance of timing in relation to determining the accuracy of later questioning of eyewitnesses. This demonstrates that intoxicated eyewitnesses encountered by the police are not always likely to be more prone to suggestion, and may be less prone to suggestion, depending on when they consumed alcohol.

Two studies in the review did not give participants any alcohol. The first (Assefi & Garry, 2003) told some participants they had consumed alcohol when they had not. There was no significant difference between the performance on control items for those who had

believed they had consumed alcohol, and those who knew they had not. But there was a significant drink condition and misinformation interaction (small to medium effect size *f*= 0.38), indicating that those in the told alcohol condition were more likely to report misinformation at test. It is however of note that this study was the lowest quality paper included in the review, with a quality assessment score of 65%, meaning that caution should be used in the extrapolation of these findings to other scenarios. This is particularly challenging, as there were no studies with similar methodologies within the review.

Zajac et al. (2016) also did not give participants alcohol but half of the participants observed a co-witness drinking what they believed to be alcohol, and half witnessed them drink lemonade. They found that accuracy for misled items was lower than for control items across all conditions, however, they found no significant effect of co-witness alcohol condition on the amount of misinformation reported by participants (Wald v^2 (*df* = 1) = 1.44, *p* = .23). However when a clear discrepancy could be detected between the participants pre-misinformation response and the response given by the co-witness, participants were more likely to accept misinformation from a sober co-witness (x^2 (1, *N* = 20) = 3.33, Fisher's exact *z*, *w* = .41). This suggests that participants were no more likely to be misled by a sober compared to an intoxicated co-witness when there was no discrepancy between their response and the co-witnesses, but the opposite was true when a discrepancy existed. This study received a quality assessment score of 91%, far higher than that of the study conducted by Assefi and Garry (2003), the only other study that did not provide participants with alcohol. The same challenges, however, do remain in that no direct comparisons to other research included in the review can be made, due to the unique methodology.

Five studies also explored the effects of testing participants when intoxicated. Van Oorsouw et al. (2019) explored an interaction between time of testing and BAC and found

that participants who were intoxicated at the time of testing were no more likely to adopt misinformation (t=0.05, p=.958). Mindthoff et al. (2021) also used same state methods (where participants intoxicated at encoding were intoxicated at test, and sober participants at encoding were sober during test) when testing participants. Participants were given alcohol then, when intoxicated, misinformation was introduced during testing that occurred immediately after intoxication. Misinformation acceptance was measured using yield scores where a score of 1 was given if a participant yielded to the misleading question by the researcher. They found no significant effects of intoxication on yield scores the first time that participants were tested, or the second time they were tested (still while intoxicated; yield 1: ηp^2 = 0.02; yield 2: ηp^2 =0.03). This finding is consistent with other studies that have found no effect of alcohol intoxication on the acceptance of misinformation (Bartlett et al., 2022; Flowe et al., 2019; Schreiber Compo et al., 2011). It is, however, of note that overall correct responses given at test were very low across all groups in the Mindthoff et al. (2021) study. It is hypothesised that due to this, participants were less likely to be able to detect discrepancies when misinformation was introduced as they were already unclear of an answer. If they had given an answer (rather than responding that they did not know), they may have been more likely to detect discrepancies with misinformation. The scenario used in the study was of a married couple witnessing a neighbour lose their bike and, therefore, did not have the same emotional impact that scenarios in other studies were likely to have which may have reduced the amount of information encoded. Other studies used scenarios such as a first-person hypothetical rape (Flowe et al., 2019) or witnessing a staged crime first hand (Schreiber Compo et al., 2011). This makes these findings potentially less generalisable to the legal context as well as being less comparable to the other studies discussed in this review.

The quality assessment scores of all studies that explored the testing of participants while intoxicated (Bartlett et al., 2022; Flowe et al., 2019; Mindthoff et al., 2021; Schreiber Compo et al., 2011; Van Oorsouw et al., 2019) were all similar in their quality assessment scores. The lowest quality assessment score received by one of these studies was that by Schreiber Compo et al. (2011), receiving a score of 73%, however, the findings of this study were comparable to those of the remaining five, which all scored >85%. This suggests that these findings are reliable, as they are consistent across studies of high quality.

Delayed testing

Five of the twelve studies in this review used delayed testing. Van Oorsouw et al. (2019) used immediate testing only and delayed testing only conditions in their second field study. Their results showed that a higher BAC at encoding and misinformation lead to participants providing more misinformation at test in the delayed testing group when they were sober (t=3.60, p=.001) but not in the immediate testing group when they were still intoxicated (t=0.05, p=.958). This is further supported by Van Oorsouw et al. (2015) who found that severely intoxicated participants were the most likely to go along with misleading questions both at immediate and follow up tests when sober (t(39)=2.94, p=.01).

Both of the studies (Van Oorsouw et al., 2015; study two of Van Oorsouw et al., 2019) received a quality assessment score of 90%, suggesting that considerable weight can be placed on their findings, however it is of note that the delayed testing by Van Oorsouw et al. (2015) and Van Oorsouw et al. (2019) were conducted by telephone and email, respectively. Van Oorsouw et al. (2015) note that the portion of testing that was completed over the phone may have enhanced performance compared to the testing that was conducted in person, as participants may have been experiencing less anxiety during the call due to being in their own surroundings. On the other hand, however, it could however be argued in both studies that

performance on these delayed tests may have been negatively impacted by the participants' surroundings. It is possible that participants may have been distracted, particularly when filling in the email task. It was also noted in Van Oorsouw et al. (2015) that the follow up period varied by a number of days between participants, and the authors state that this may have affected performance. Both Van Oorsouw et al. (2019) and Van Oorsouw et al. (2015) found performance was worse in the delayed testing conditions, with more misinformation being reported.

It is important to note the confound that in intoxicated groups of participants, immediate testing is while intoxicated and delayed testing occurs when sober. This was noted by Van Oorsouw et al. (2019), who state that this makes is challenging to disentangle the negative effects of a delay with the effects of participants sobering up. While this can be overcome by using state dependent methods, this requires a laboratory study.

Evans et al. (2018) also employed a delayed testing condition in their study and found there was a significant effect of time of test, however, this was seen to be more apparent for participants that were in the same state at both encoding and retrieval. Participants who were tested immediately while still intoxicated were less likely to agree with misinformation than those who were tested after a delay, allowing them to sober up before testing (d = 2.03; large effect size). The quality assessment of this study was 88%. While this is lower than that of Van Oorsouw et al. (2015) and Van Oorsouw et al. (2019), the difference in quality assessment score is marginal, suggesting that the increase in misinformation reporting after delayed testing is a reliable finding, however the driving factor behind this increase in misinformation reporting remains unclear, as same-sate testing was not carried out in either study by Van Oorsouw et al. (2015) or Van Oorsouw et al. (2019) Note that both Flowe et al. (2019) and Gawrylowicz et al. (2017) tested all of their

participants after a delay, to allow participants to sober up before testing. Therefore, in those two studies there was no comparison group of immediate testing and so the effect of this delayed testing in these studies is not clear.

Strengths and limitations of the current review

Attempts were made to ensure that the search terms used in this review were as inclusive and comprehensive as possible by noting the key terms of papers that were deemed relevant in the preliminary searching and keywords mentioned throughout these papers. Database search headings and a thesaurus were also used to ensure search terms were as inclusive as possible. No further papers were found through additional methods of searching, such as the use of Google Scholar or through the reference lists of existing papers. The search yielded a reasonable number of initial results considering the size of this subject area and all key papers that had been identified by academics in the field were seen in the results, demonstrating the validity of the key terms used.

Every effort was made to ensure that data extraction and quality assessment were robust, however, due to time constraints these processes were carried out by only one individual. This could potentially have more impact on the quality assessment checklist scores due to the subjective nature of some questions included, however, the process used by the researcher to answer these questions was consistent across all papers, therefore, mitigating this as much as is possible. When creating the data extraction form, questions were added when deemed necessary and previous studies re-evaluated to ensure that all relevant information was captured within this form. Overall, it is believed that the search was conducted in the most thorough way possible, strengthening this review.

It is also of note that some studies included in this review had slightly differing focuses. For example, the study by Zajac et al. (2016) was included in the review. While its focus differed from other studies which explored the direct impact of alcohol on the individual who was to be tested, this study studied misinformation when a confederate was thought to be consuming alcohol. This was deemed relevant for this review as suggestibility and misinformation were still being studied. This research may also be highly useful in situations where groups of friends have witnessed a crime and later discuss this before giving statements. This study demonstrates the potential impact of witnesses having discussed information with friends that may or may not have been intoxicated at the time of witnessing the event. This is relevant when the concept of memory completeness is considered with individuals wishing to fill these gaps in their memory.

Moreover, the methodologies varied greatly across all available research which presents difficulties in making direct comparisons. There were a wide number approaches taken to measuring memory recall across studies. Moreover, alcohol was introduced at different points across studies (e.g., before or after encoding, and some studies using repeat and/or delayed testing. These complexities made it difficult to disentangle the results and make comparisons or draw overall conclusions from the wider material. For example, Evans et al. (2018) noted that the suggestibility questionnaire used in their study had not been used in previous research and therefore the validity and reliability of it as a measure had not been evaluated. There is currently no best practice for measuring the suggestibility of participants in studies such as these. The same can be said for the introduction of misinformation. For example, in Gawrylowicz et al. (2017) study, misinformation was provided in the form of scrambled sentences of a narrative containing misleading information that participants were required to order and read out loud. In other studies, such as Schreiber Compo et al. (2011) and Van Oorsouw et al. (2019), misinformation was introduced through participants

overhearing information. The amount of attention given to the misinformation in both cases may be different which could in turn account for differing amounts of misinformation being reported.

It is also of note that none of the studies included in the review compared any sex differences in this research. Previous research has found that women can be affected differently by alcohol even when dose is determined by their height and weight due to their higher body fat percentage and lower water weight compared to men (Mumenthaler, Taylor, O'Hara, & Yesavage, 1999). This suggests that women may be more significantly intoxicated or experience more deleterious effects of alcohol when witnessing a crime compared to men in the same setting. Research has also suggested that women are more likely to experience alcohol induced blackouts (Rose & Grant, 2010). It, therefore, may be beneficial for future research to compare misinformation effects across sexes.

Conclusion

Overall, the findings of the studies included in this review demonstrate that eyewitness memory *may* be negatively affected by alcohol intoxication when exposed to misinformation. There are, however, discrepancies between studies that appear to be impacted by the delay of testing, BAC levels, and timing of alcohol consumption. There were also studies that did not find a negative effect of alcohol on the reporting of misinformation (Evans et al., 2018; Flowe et al., 2019; Gawrylowicz, 2017; Mindthoff et al., 2021; Schreiber Compo et al., 2011).

The effects of BAC levels remain somewhat unclear, however the findings of this review suggests that higher quality papers, and predominantly those conducted in a laboratory, found no effects of alcohol intoxication at encoding. Alcohol intoxication at the

introduction of misinformation appeared to reduce the amount of misinformation reported by participants, when considering the highest quality of two papers that explored this affect. Consistent findings from high quality papers demonstrated no significant effect of alcohol intoxication at test, and, once again, findings consistently demonstrate negative effects of delayed testing, however, the driving factors of this are unclear.

It is suggested that future research should aim to disentangle these effects further to give police clearer guidance regarding the interviewing and initial statement taking from intoxicated witnesses. The research reviewed here supports claims that eyewitness testimony may not always be entirely reliable. Future research specifically examining the effects of intoxication on suggestibility will enable the Criminal Justice System to more appropriately manage the statement taking and interviewing of witnesses, allowing for stronger leads in cases being investigated and more reliable evidence in the courts.

Chapter Three

Alcohol and remembering rape: Examining the effects of acute alcohol intoxication during rape on metamemory processes during police interviews

Police interviews are designed to obtain evidence. In the case of interviewing victims or witnesses, this is with the view to obtaining evidence that can be used in an investigation to identify and/or prosecute the perpetrator. One case in which obtaining clear evidence may be especially pertinent is in cases of sexual offending. The Ministry of Justice (2013) estimates that 473,000 adults in England and Wales are victims of sexual offences each year. However, it is suggested that the actual figure of victimisation may be higher due to victim concerns about not being believed, retribution by the offender, and feelings of shame and embarrassment (Taylor & Gassner, 2010). Of the sexual offences that are reported, only 12.5% of these cases result in a conviction (Daly & Bouhours, 2010). Similar statistics have been reported in more recent years, with Rape Crisis England and Wales (2022) reporting the highest ever number of rapes were reported in a 12-month period, ending in September 2022, with 70,633 reported, but only 2,616 charges brought. The Office for National Statistics (2023) reported a 22% increase in sexual offences from the year ending March 2020, in September 2022. Victims of sexual offences that have reported their experience to the police have expressed that they did not feel as though the police had believed them and that they found the legal process to be stressful and, in some cases, further traumatising (Lees, 1993; McMillan & Thomas, 2009). Particularly in rape cases, victims are frequently under the influence of alcohol at the time of the offence (Testa, 2002) which can call into question the accuracy of their memory of the offence (Kassin et al., 2001). Although formal interviews cannot take place while witnesses or victims are intoxicated, some police officers do take initial statements when they first have contact with the individual, which may be while they

are still intoxicated (Crossland, Kneller, & Wilcock 2018). Although psychological research has demonstrated that eyewitness memory is more reliable than it was once thought, the view that it is unreliable is still held by the Criminal Justice System (Wixted et al., 2018). This is particularly true of intoxicated witnesses, who are widely considered to be unreliable by jurors (Evans & Schreiber Compo, 2010). This chapter examines the reliability of memory recall from intoxicated witnesses across a simulated police interview by measuring metamemory.

Metamemory

Metamemory is the knowledge that one has about their own memory capabilities, the strategies they have that can aid their memory, as well as the process involved in self-monitoring memory (Pannu & Kaszniak, 2005). Metamemory plays an important role in human learning throughout development and therefore has been the subject of much research. Metamemory research provides an understanding of how individuals use memory throughout their daily life as it provides essential strategies for encoding and retrieval of memories. When learning new information, metamemory helps people to determine whether or not they have studied the material sufficiently to recall this at a later date (Metcalfe, 2000). Metamemory also helps people to determine how likely their memory recall is to be accurate (Kelemen, 2000), for example, “how sure am I that I turned off the oven?”.

Eyewitness metamemory in forensic settings (metamemory in the context of witnessing or being the victim of a crime) is a rapidly expanding field of research (e.g., Wixted & Wells, 2017). One way of measuring metamemory is to collect confidence judgements and examine how confidence relates to accuracy. If confidence increases with accuracy, then one is said to have good metamemory ability, as they are able to monitor when their memories are and are not likely to be accurate. It has previously been argued by researchers that there is a poor correlation between confidence and accuracy, suggesting

metamemory cues are not a useful predictor of accuracy (Cutler & Penrod, 1989; Leippe, 1980). Smith et al. (1989), for example, stated that confidence was not a predictor of accuracy in witnesses or their statements. It has since been proposed, however, by various researchers that confidence ratings *are* a reliable predictor of the accuracy of an individual's recognition memory and that this is the case even for weak memories, such as from older adults or alcohol intoxicated witnesses (Colloff et al., 2017; Flowe et al., 2017). The findings of these more recent studies may be due to advances in measuring confidence and accuracy (see Wixted & Wells, 2017), as well as new approaches to statistical inferences (Saraiva, 2019).

According to Signal Detection Theory (SDT), confidence is a useful proxy for memory strength (Wixted & Mickes, 2010). SDT uses a 2 X 2 matrix in human performance tasks that require participants to discriminate between stimulus presence, versus stimulus absence (Macmillan, 2002). When applied to eyewitness research, Lee and Penrod (2019) used SDT to examine the accuracy of eyewitnesses when presented with a lineup task in which the perpetrator was sometimes present and sometimes absent (i.e., a recognition memory task). They found that participants who reported higher confidence were more able to discriminate between guilty suspects and fillers, while those with low confidence showed poor discriminability between guilty suspects and fillers. Confidence has also been shown to be predictive of accuracy in recall when the confidence judgments are collected after free-recall. Roberts and Higham (2002) examined this relationship by asking participants to watch a videotape of a crime and then freely recall a narrative of the scenario. Participants were then asked follow up questions about the information they gave and asked to rate their confidence in the information they had given on a 7-point scale, from low to high. In this study, high accuracy was associated with high confidence. Odinot et al. (2009) further provided support for this relationship in forensic settings. They interviewed eyewitnesses of

an armed robbery in a supermarket 3 months after the event and, again, asked them to rate their confidence in the information provided on a 7-point scale. It was found that accuracy raised from 61% to 85% as confidence raised from low to high (See Wixted et al., 2018 for an updated review of these studies). This provides further support for a meaningful confidence-accuracy relationship in forensic settings, which Wixted et al. (2018) state that the field of psychology has been slow to appreciate.

Research also suggests that people use metamemory to choose whether or not to volunteer their memories based on how strong they feel that memory is. Strategic memory process theories suggest that when in a free recall situation, individuals will volunteer the memories when they feel certain that they are correct in their recollection of them (Goldsmith et al., 2005). The strongest memories—the ones they are most confident, and therefore most accurate in—will be recalled first, and then, over time throughout the free recall, weaker memories will be recalled. The information that is recalled first tends to be better learnt and remembered in greater detail, with higher certainty (Mickes et al., 2013).

Moreover, evidence suggests that the general meaning of information is forgotten at a slower rate than more precise and detailed information (Reyna & Kiernan, 1994). Memory performance differs dependent on the coarseness (level of detail) at which it is measured. This is supported by Neisser (1998), who explained the superiority of free recall as opposed to recognition testing as allowing for participants to recall information at a level of generality where they are not mistaken (coarse-grained information). In contrast, recognition tasks require participants to make fine discriminations. For example, witnesses may choose to provide more coarse-grained information after a delay, therefore maintaining stable accuracy over time, but providing less detailed answers over time. Coarse-grained information is that which is imprecise or broad, such as stating that an offender is driving a dark coloured car, whereas fine-grained information is precise, such as that an offender is driving a dark blue

coloured car (Brewer et al., 2018). Indeed, memory research has found that participants demonstrate stable recall accuracy over lengths of time as great as 6 years (Hudson & Fivush, 1991). In the legal system, it may be more important that the information provided by witnesses remains accurate but less detailed, instead of witnesses being encouraged to provide more detailed but potentially inaccurate information. This however, is often not what is expected of those giving evidence in the Criminal Justice System. Bell and Loftus (1989) suggest that witnesses that provide accounts rich in detail, are more likely to be perceived by a jury as reliable and trustworthy. It is also of note that it is common practice for witnesses and victims to be quizzed about gaps in their recollection during cross-examination in court, and these gaps being presented as evidence of unreliability (Ellison & Munro, 2017). This suggests that those under cross-examination may be more likely to fill those gaps with information of which they are less certain.

Based on the above research, it is proposed that victims that are interviewed by the police will strive to volunteer their strongest memories and will withhold weak memories where they are less confident of their accuracy (Evans & Fisher, 2011; Wixted et al., 2018). It is, however, noted that during police interviews, Achieving Best Evidence (ABE) guidance encourages efforts to obtain as much detail as possible during the interview. This may be inadvertently reducing the quality of the evidence given by the victim by encouraging them to recall weak and potentially inaccurate memories. It is likely that the information recalled in the question phase of the interview will be weak in comparison to the memories recalled in the free recall phase due to stronger memories being volunteered first by victims.

Alcohol intoxication and metamemory

As previously discussed, victims of sexual offences may be intoxicated. Research has examined the effects of alcohol on memory accuracy and completeness. Accuracy is the amount of correct versus incorrect information reported, and completeness refers to the

amount of information recalled in total. Completeness of memory is undermined by the omission of details (Smeets et al., 2004). Research examining the effects of alcohol on metamemory, memory accuracy, and memory completeness are mixed (Hagsand et al., 2013). It is known that alcohol can impair an individual's ability to attend to multiple cues at the time of encoding, therefore weakening memories (Marinkovic et al., 2004). However, it is possible that previously alcohol intoxicated witnesses are still reliable, because they are aware of their poorer memory compared to sober witnesses and are able to report only memories of which they are certain (Flowe et al., 2016). A meta-analysis by Jores et al. (2019) suggests that alcohol can significantly decrease the amount of correct information that is recalled by witnesses (i.e., completeness) but that it does not increase the reporting of false information. Put another way, the completeness of the information given decreases, however the accuracy of the information does not. This suggests that intoxicated witnesses may be able to monitor their memories and only report details of which they are sure are correct. Flowe et al. (2017) also found that the confidence increased with accuracy in alcohol intoxicated witnesses in a line-up identification task. Again, this suggests that witnesses that are intoxicated at encoding are able to monitor their likely memory accuracy.

What has not yet been examined, is metamemory in the context of police interviews with victims that were alcohol intoxicated at the time of the offence. This is important as jurors are less likely to believe the report of a witness that appears uncertain (Brewer & Burke, 2002). Uncertainty in a witness may be seen in the form of uncertainty indicators, such as the use of word fillers (e.g. "like"), linguistic hedges (e.g. "I think", "I guess"), or long delays in speech when recalling memories (Krahmer & Swerts, 2005). It is hypothesised that alcohol intoxicated witnesses will be more likely to express uncertainty indicators, particularly in the question phase of an interview, however they will be equally likely to

provide accurate information overall (Flowe et al., 2016). This is because they will be aware that they have possibly been able to encode less information, due to an inability to attend multiple cues, and therefore express uncertainty. This is hypothesised to be particularly the case in the question phase of the interview when witnesses are being encouraged to report memories of which they are less certain. An evidence base that suggests uncertainty cues are indicative of inaccuracy may aid police officers or jurors in decision making. For example, a lack of uncertainty indicators may indicate more confidence in the information being given. In the case of police decision making, details given with high confidence, or a lack of uncertainty, in testimony should then be prioritised for further investigation and in the collection of corroborating evidence. If, for example, a witness is sure that an offender was not wearing gloves and touched their jacket, but aren't sure if they touched their trousers, it may allow the police to prioritise sending the jacket for forensic testing, over the trousers. It is however of note that this evidence base should be utilised to prioritise the collection of evidence, not to dismiss the evidence given my intoxicated victims or witnesses. While uncertainty indicators may highlight areas of lower confidence, and therefore, lower accuracy, it is important to note that they are likely not indicative of a completely inaccurate detail and care should be taken to avoid disadvantaging intoxicated witnesses and victims. This prioritisation may allow for better allocation of police resources in a system whereby these resources are limited. Research is also required to assess how weaker memories may be expressed during police interviews and how weak memories can be identified (i.e., through metamemory indicators) to achieve best evidence.

The Current Research

The current research will use a unique data set from previous research in which participants experienced a simulated rape scenario in a written line-by-line format, and narrated by a female voice, known as the participant choice paradigm (Flowe et al., 2007;

Flowe et al., 2019). In this research, participants were either alcohol intoxicated or sober at encoding. Half of the sober participants were told they had received alcohol, and half the alcohol group were told they had received tonic water in order to induce alcohol expectancy effects. The scenario depicted a female (the participant) interacting with a man who began as flirtatious, with the scenario gradually introducing sexual activity. The interactive scenario allowed the participants to control how much consensual interaction they had with the man in the scenario and they could opt out of the scenario and "call it a night" at any point. When participants opted out of the scenario, a legally definable act of rape was described to them. The participant choice paradigm used to present the scenario allowed participants to become more actively involved in the scenario due to its high levels of psychological realism (Mook, 1983). Participants that completed the study expressed that they had felt distressed by the scenario and that they had traumatic stress symptoms, as measured by the Impact of Events Scale (Palmer et al., 2013; please see ethical issues outlined below). After one week, participants returned to be interviewed, as they would be during a police interview, using Ministry of Justice Guidance (2011). The interview consisted of a free recall phase, and a question phase. In the free recall phase, participants were asked to provide as much information as possible related to the scenario. In the question phase, participants were asked to recall specific details about the perpetrator and the location in which the offence took place. After these phases, participants then took a recognition test to systematically measure their memories about the scenario as a whole. Along with each answer given in the recognition test, participants also rated their confidence in this answer.

Flowe et al. (2019) found that participants were well-calibrated in terms of their confidence ratings being predictive of their accuracy on the recognition test. In the free recall data, Flowe et al. (2019) measured only the number of correct and incorrect details recalled by the participants (31.04 correct details and 4.19 incorrect details on average) to measure

memory output and accuracy and found no variation in this across intoxication conditions. However, the length of the transcripts derived from the participant interviews were between 643.68 and 1186.55 words in length, suggesting there is far more rich data to analyse, especially considering the recent advancements regarding witness metamemory.

The aim of the current study is to use new indices to examine metamemory processes over the course of a police interview to test the efficacy of current police interview practices in allowing victims of crime to provide the most accurate and full memory evidence. Metamemory indicators are analysed against the accuracy of the information provided to examine if metamemory cues are informative of accuracy. This study also aims to assess whether victims are more accurate during the free recall phase or question phase of the interview and how this is affected by alcohol intoxication at the time of encoding.

It is hypothesised that participants will recall information with less uncertainty when they are accurate (hypothesis 1). This would be demonstrated by less uncertainty indicators associated with accurate information. It is also hypothesised that participants will express uncertainty less often in the free recall phase of the interview compared to the question phase (hypothesis 2). Finally, it is hypothesised that the group that received alcohol at encoding will recall memories with greater uncertainty, particularly in the question phase of the interview (hypothesis 3), compared to those who received tonic water at encoding, demonstrated by using uncertainty indicators more often.

Finally, it was explored whether intoxicated witnesses and sober witnesses that do not express uncertainty (i.e. are confident) would be equally likely to be accurate.

Design and Method

The current research used secondary data analysis to examine a unique data set previously collected by Flowe et al. (2019). The original study and its method of data collection can be found below, followed by the method of secondary analysis.

Participants

The participants that took part in the original study were a sample of 80 women recruited from The University of Leicester, where advertisements were circulated around the campus. Participants were informed that the study was examining sexual and dating behaviours of women. Further information was received upon contacting the researcher to express interest in taking part. Each of the women in the study were between 18 and 31 years of age (M = 20.36, SD = 2.41 years) and they were each paid £6 per hour for their participation.

Procedure

Participants completed an initial prescreening online which included the Alcohol Use Disorders Identification Test (AUDIT) to detect harmful alcohol consumption (Babor et al., 2001; see Chapter 4) and a general health questionnaire developed by the researchers, as well as a list of prescription medications they were taking. This was included to account for any potential liver or heart issues, or psychiatric disorders. Participants were only invited to take part in the study if they scored less than 10 on the AUDIT and had no existing health issues or medications that interact with alcohol. Secondary screening also occurred in the laboratory, where answers to the online screening were verified, photographic identification was checked to ensure participants were over 18 years of age, and a pregnancy test was administered which was required to be negative to continue. It was also confirmed that participants had not eaten in the 4 hours prior to the study, and had not consumed alcohol for 24 hours prior.

Participant height and weight measurements were recorded to inform later alcohol dosage. The Alcohawk Slim Digital Alcohol Breath Tester was used to take a blood alcohol concentration (BAC) reading by converting deep-lung air alcohol into an estimated BAC measurement.

A 2 beverage (tonic water or alcohol) x 2 expectancy (told alcohol or told tonic) x 4 information type (consistent, neutral, misled, and control) x 4 perpetrator x 4 scenario (his house, her house, a bar, a party) mixed design was used. Three different images of men were mixed with four descriptions of their hometown, hobbies, occupation, etc. to create 4 different perpetrators that were shown to participants (while two looked the same, they had different hobbies and occupations). These perpetrator descriptions were balanced between four different scenarios of being at his house, her house, a bar, and a party. Data were collapsed over the perpetrator and scenario conditions for the analysis due to the small number of participants in each condition in this reanalysis. No significant differences in memory recall were noted across scenario man or scenario version in the original Flowe et al. (2019) study. The only within-subjects factor was information type. All participants were randomly assigned to conditions.

An initial breathalyser reading of 0.00% BAC was confirmed and participants were either given 3 alcoholic drinks of vodka (37.5% proof) and tonic in a 1:5 ratio or plain tonic water. The necessary dose of alcohol for each participant was based on their height and weight as per Curtin and Fairchild (2003). This level of intoxication was used due to reported attention-allocation disruptions at this level (Harvey et al., 2013). All cups were rimmed with vodka and contained vodka-soaked limes. Participants did not see their drinks being prepared and were asked to consume them at the rate of 1 cup every 5 minutes. Half the participants were told they would receive alcohol, half were told they would receive tonic water and cups were labelled accordingly to induce expectancy.

Thirty minutes after commencing drinking, participants were breathalysed and took part in the interactive scenario. At this point, the mean BAC in the tonic water group was 0.00% and was 0.06% in the alcohol group (range: 0.04-0.09%, SD = 0.02).

Sixteen versions of the scenario were presented via the participant choice paradigm (Flowe et al., 2007). This encouraged active involvement from each participant and allowed her to determine how much interaction she had with the man, including consensual sexual contact. There was a total of 25 scenario stages and at the end of each, the participant was given the opportunity to opt out or continue to interact with the man. The scenario was presented as written text on a computer as well as being narrated by a female voice. The first stage presented background information about the man alongside a photograph of him. As the scenario progresses, the man is flirtatious and then sexual activity occurs. The participant could choose to consent or opt out. If she opted out, a legally definable act of rape was then described. If she remained in the scenario until the end, consensual sexual activity took place.

Regardless of experimental condition, all participants were breathalysed at 30-minute intervals throughout the study. Participants in the tonic water condition also stayed in the laboratory for 2 hours after beverage consumption so they could not infer their beverage group. Those in the alcohol group were not permitted to leave until their BAC was below 0.02% and were advised not to drive or operate heavy machinery for the remainder of the day.

Participants were interviewed 7 days after they completed the scenario. They first read a written postevent narrative about the man's background information. They were then told they would be interviewed as the study was looking to examine police interview techniques, and that the excerpt they had just read was from another participant's interview to demonstrate the next portion of the study. The narrative contained 6 items that were consistent with the original information, 6 items that were neutral (had no specific details), and 6 misleading items about the original information. Participants were then randomly assigned to be interviewed by the Self-Administered Interview (SAI) or the Cognitive Interview. The SAI is a written interview that consists of seven sections that give instructions

designed to facilitate memory recall (Hope et al., 2011). The CI is an in-person interview protocol commonly used by police in the United Kingdom that uses various mnemonic techniques, such as recalling the series of events backwards to support memory recall (Dando et al., 2009). In both interviews, participants were encouraged not to make guesses but to give the most complete and accurate description they were able to. First, participants were asked to give as much information as they remember about the event. In the next phase of the interview nonleading prompts were used to encourage further recall regarding perpetrator appearance, location, etcetera. Finally, there was a question phase which asked the participant to recall information they may not have already considered. The average number of words written in the free recall phase of the SAI was 643 (SD = 502.41 words) and 1,186 words in the question phase (SD = 573.93). The average length of interview for the Cognitive Interview was 16.78 minutes (SD = 4.58 minutes). All Cognitive Interviews were audio recorded and transcribed for coding purposes.

After the interviews were administered, participants completed a recognition test. The test was multiple choice and was designed to assess participants' memory for consistent, neutral, and misled items (Flowe et al., 2016). There were 30 questions about 18 critical items, and participants had the option to answer "I don't know" to each question. For each answer, participants also included a confidence rating from 0-100% (with 0 being "not at all confident", and 100 being "completely confident"). After the recognition test, participants were asked whether they believed their encounter with the male was rape. The response was given via an 11-point likert scale with 1 being "definitely no" and 11 being "definitely yes".

Once participants had completed all above stages of the study, they were informed that the aim of the study was to investigate whether the degree of intoxication influenced women's interactions in the encounter, and their recall of the encounter afterwards. The study took 6 hours in total.

In the original study, all transcripts of the Cognitive Interviews and SAI's were coded using a coding system from previous research (Holliday, 2003; Wright & Holliday, 2007) and details given were coded as either incorrect or correct. Misinformation items that were recalled were coded separately as misinformation intrusions. Accuracy was calculated by dividing the number of correct details by the sum of the total incorrect and correct details recalled.

Secondary Data Analysis

The secondary data analysis focused on participants in the Cognitive Interview condition (n = 35). This is because metamemory indicators could not be identified within Self-Administered Interviews, as individuals are highly unlikely to use non-word fillers or similar indicators in a written statement. Transcripts of Cognitive Interviews were collapsed over all scenario types as metamemory indicators could be identified throughout, and there was no theoretical reason to suggest that scenario type would influence the findings. First, audio and transcripts were compared to ensure that transcripts were accurate. As Flowe et al. (2019) did not code the data for uncertainty indicators (e.g., pauses, filler words), it was possible that these may have been missed in parts of the original transcribing process in the original study, therefore pauses and other uncertainty indicators were added where missing. Any audio files that did not have transcripts associated with them were transcribed. Next, the data were coded.

The transcripts were coded for accuracy in relation to the scenario that participants were shown. Indications of uncertainty, which would be used to demonstrate metamemory processes, were also coded. Unlike in the original study, misinformation was not coded in this secondary analysis (for accuracy or metamemory indicators) as misinformation was reported so infrequently by participants and with no significant differences between conditions in the original study. Therefore, the secondary data analysis focused on correct

and incorrect information in relation to each scenario. Details of the metamemory indicator coding are outlined, next.

Metamemory indicators

The coding scheme was developed in advance of reviewing the data in depth. This was developed based on research by Lindholm et al. (2018) which suggested that delays, word fillers, non-word fillers, and hedges were predictive of eyewitness (in)accuracy. While it is not possible to definitively determine whether certain words such as "like" are being used to mark uncertainty, or as part of habitual speech patterns, research has given support for its common use as a filler word indicating uncertainty (Laserna et al., 2014; Seals & Coppock, 2022).

Additionally, it is of note that in this study delays between the end of a question and the beginning of a participant's answer were not recorded as delays. This decision was made as only audio recordings of the interviews were available. It was not possible to determine whether there were other factors that may influence a delay before the beginning of an answer, such as the interviewer looking at notes as if they may say more, but not doing so. Interviewers were also noted to have added further questioning after a delay, it was not clear whether this was in order to prompt the participant, or whether it was clear that more was to be added to the question. There were also some interruptions noted in the recording, so it could not be ruled out that participants were distracted by something else, or for example, taking a drink at this time. To remove this ambiguity, delays were only included throughout answers given.

Each metamemory indicator was coded separately. These measures and the method of coding for each are listed below:

Short delays. A short delay was defined as an uncertainty indicator demonstrated by a pause of between 3 and 5 seconds. Pauses were only counted when they occurred during

recall, for example, when the statement in response to a question has begun. Any pause between hearing the question and beginning to respond was not counted.

Long delays. A long delay was defined as an uncertainty indicator demonstrated by a pause of 6 seconds or more. Pauses were only counted when they occurred during recall, for example when the statement in response to a question has begun. Any pause between hearing the question and beginning to respond was not counted.

Word fillers. Word fillers were defined as words that did not add meaning to the participant's response when used during recall effort. These included "you know", "well", and "like". Even when more than one word was used (such as the phrase "you know") this was counted as one word filler.

Nonword fillers. Nonword fillers were defined as interjections or sounds that were made during recall effort. These included "hmm" and "umm". Repetitions of these sounds were counted separately and all accompanying pauses were coded separately as either a long or short delay.

Word hedges. Word hedges were defined as words or phrases that suggested low confidence or a lack of willing to commit, as well as diminishing the value of an assertion. These included phrases such as "I'm not sure but", "kind of", "sort of", etc. Each phrase was coded as one word hedge.

In addition, we coded measures of certainty from each transcript as follows:

Verbal Certainty. Operationally, we defined certainty as the point at which information was given in the absence of any uncertainty indicators.

Ethical Issues

Full ethical approval was granted from the University of Leicester for the original study after considerations were made about the sensitive nature of the fictional scenario in the study, and the administration of alcohol to participants. Details of the current study and previous ethical approval from the University of Leicester were shared with the University of Birmingham, and the Ethics Chair confirmed that they were happy to accept this ethical approval in lieu of further review at the University of Birmingham. The data analysed for the current project are minimally sensitive as the interviews do not give any personal information about the participants themselves. The researcher only had access to the audio recordings and transcripts of the interviews, all of which were anonymised. The data were shared and stored in accordance with GDPR guidelines.

Results

In the free recall phase, the uncertainty indicator with the highest average use across participants was non-word fillers, used on average 4.34 times, and the least frequently used were long delays which were not used by any participants. In the question phase, the uncertainty indicator with the highest average use across participants was non-word fillers, used an average of 2.04 times. Short delays were used the least frequently by participants, used an average of 0.06 times. In the free recall phase, participants used an uncertainty indicator on average 9.34 times, regardless of information accuracy. In the question phase, participants used an uncertainty indicator on average 7.72 times, regardless of information accuracy. Therefore, the following analysis collapsed data across all uncertainty indicator types due to the infrequency of some of these indicators in transcripts.

Two mixed analysis of variance (ANOVA's) were used to analyse data. First, to test hypotheses one and three, data were analysed using a 2 beverage (alcohol, tonic) x 2 phase (free recall, question phase) x 2 indicator (indicator used, no indicator) mixed ANOVA. The

dependent variable was the proportion of correct details given by a participant. Absence and presence of an uncertainty indicator was included as a factor in the ANOVA. Therefore, in order to run the ANOVA, it was necessary to calculate what proportion of the correct details were given along with an uncertainty indicator, and what proportion were given without an uncertainty indicator. The proportion of correct details given with an uncertainty indicator was calculated by dividing the number of correct details given with an indicator, by the sum of the number of correct details given with an indicator and the number of incorrect details given with an indicator. The proportion of correct details given without an indicator was calculated by dividing the number of correct details given without an indicator, by the sum of the number of correct details given without an indicator and the number of incorrect details given without an indicator. These are referred to as 'proportion correct' from this point on. The data met assumptions required for an ANOVA in level of measurement, random sampling, independence of observations, and normal distribution. When performing Levene's test for equality of variance, the test was violated for one group (question phase, using an indicator). The mixed ANOVA was still deemed to be an appropriate analysis, however, because ANOVA is robust to violations, providing the sample size of each group is similar in size (Pallant, 2016) which was true in this case.

The analysis found a main effect of indicator, $F(1,30) = 4.493$, $p=.042$, $\eta_p^2 = .130$. The proportion of correct details was higher when a detail was given without an indicator ($M = .865$, $SE = .021$, 95% CI [.822, .908]) compared to with an indicator ($M = .816$, $SE = .021$, 95% CI [.773, .858]). This suggests that participants recalled information with less uncertainty when they were accurate (supporting hypothesis 1). A main effect of phase was also found, $F(1,30) = 18.059$, $p<.001$, $\eta_p^2= .376$. The proportion of correct details was higher in the free recall phase ($M =.908$, $SE = .015$, 95% CI [.878, .938]), compared to the question

phase (*M* =.772, *SE* = .030, 95% CI [.711, .834]), suggesting that participants were more accurate in the free recall phase of the interview.

There was no main effect of beverage, $F(1,30) = 0.195$, p=.662, $\eta_p^2 = .006$, suggesting that alcohol intoxication (*M* = .833, *SE* = .026, 95% CI [.779, .886]) compared to being sober (*M* = .848, *SE* = .023, 95% CI [.801, .895]) did not affect the proportion of correct details given by participants. There was also no interaction between beverage and phase, $F(1,30) = 1.187$, p=.285, η_p^2= .038. This suggests that alcohol intoxication did not affect the proportion of correct details given in the free recall phase (alcohol intoxicated, *M* = .918, *SE* = .022, 95% CI [.872, .963]; sober, *M* = .898, *SE* = .020, 95% CI [.858, .939]), compared to the question phase of the interview (alcohol intoxicated, *M* =.747, *SE* = .045, 95% CI [.565, .839]; sober, *M* =.798, *SE* =.040, 95% CI [.717, .878]).

There was no interaction between indicator and beverage, $F(1, 30) = 2.588$, $p = .118$, $\eta_p^2 = .079$. This suggests that the use of uncertainty indicators were equally informative in predicting of accuracy information reported in both the alcohol intoxicated group (with an indicator, *M* = .789, *SE* = .031, 95% CI [.725, .853]; without an indicator, *M* = .876, *SE* = .032, 95% CI [.811, .941]) and the sober group (with an indicator, *M* = .842, *SE* = .028, 95% CI [.786, .899]; without an indicator, *M* = .854, *SE* = .028, 95% CI [.797, .911]). Put another way, when participants did not use an uncertainty indicator (i.e. they were confident) they were equally likely to be accurate in both the sober and alcohol intoxicated group.

There was no interaction between indicator and phase, $F(1, 30) = 1.527$, $p = .226$, $\eta_p^2 = .048$. This suggests that the informativeness of uncertainty indicators in predicting accuracy was not affected by phase. There was also no interaction between indicator, beverage, and phase, $F(1, 30) = 1.037$, $p = .317$, $\eta_p^2 = .033$. This suggests that the informativeness of indicators is not affected by the phase of the interview, regardless of whether participants are alcohol intoxicated or sober.

Second, to examine hypotheses two and three, a 2 beverage (alcohol, tonic) x 2 phase (free recall, question phase), mixed ANOVA was conducted. This time, the dependent variable was calculated by dividing the number of items recalled with an indicator, by the total number of items recalled. Or, put another way, the proportion of details recalled with an indicator, regardless of whether or not the detail was correct. The data met assumptions required for an ANOVA in level of measurement, random sampling, independence of observations, and normal distribution. When performing Levene's test for equality of variance, the test was violated for one group (question phase, using an indicator). Again, this was however accepted due to analysis of variance being robust to violations, providing the sample size of each group is similar in size (Pallant, 2016) which was true in this case.

A main effect of phase was found, $F(1, 30)$, 16.042, p= <.001, η_p^2 = .348. More uncertainty indicators were used in the question phase of the interview (M = .547, SE = .037, 95% CI [.472, .622]) compared to the free recall phase (M = .384, SE = .027, 95% CI [.328, .440]). This suggests participants were more uncertain in the question phase than the free recall phase (supporting hypothesis 2). There was no main effect of beverage, $F(1, 30)$ = 0.001, p= .980, η_p^2 = .000, suggesting that participants were no more likely to use uncertainty indicators in the alcohol intoxicated group (M = .466, SE = .033, 95% CI [.398, .534]) compared to the sober group (M = .465, SE = .038, 95% CI [.388, .542]). There was also no interaction between beverage and phase, $F(1, 30)$ = 1.760, p=.198 η_p^2 = .055, suggesting that alcohol intoxication did not effect of the likelihood of participants using uncertainty indicators in the free recall phase (alcohol intoxicated, M = .410, SE = .041, 95% CI [.327, .494]; sober, M = .357, SE = .036, 95% CI [.284, .431]) compared to the question phase (alcohol intoxicated, M = .519, SE = .055, 95% CI [.407, .632]; sober, M = .575, SE = .049, 95% CI [.475, .674], not supporting hypothesis 3).

Discussion

This study investigated the use of metamemory indicators as a predictor of memory accuracy in mock-victims of sexual assault, both when sober and alcohol intoxicated at encoding. It was hypothesised that participants would recall information with less uncertainty when they were accurate (hypothesis 1). It was also hypothesised that participants would express uncertainty less often in the free recall phase of the interview compared to the question phase (hypothesis 2). It was hypothesised that the group that received alcohol at encoding would recall memories with greater uncertainty, particularly in the question phase of the interview, compared to the tonic water group (hypothesis 3). Finally, it was explored whether intoxicated witnesses and sober witnesses that do not express uncertainty (i.e. are confident) would be equally likely to be accurate.

Support was found for hypothesis one and hypothesis two, but did not find support for hypothesis three. When exploring whether intoxicated witnesses and sober witnesses that do not express uncertainty are equally likely to be accurate, participants were equally likely to be accurate in both the sober and alcohol intoxicated group when not expressing uncertainty. These findings will now be discussed in relation to the wider literature and their implications in practice.

In finding support for hypothesis one, it was seen that a higher proportion of correct details provided by participants were not accompanied by an uncertainty indicator. The metamemory indicators used by participants were indicators of uncertainty (Teigan & Brun, 2003), therefore a lack of these indicators was likely to suggest certainty, or higher confidence in a detail. The relationship between confidence and accuracy has been demonstrated in previous research. For example, lineup identifications made with higher confidence are often found to be higher in accuracy than those made with low confidence (e.g., Flowe et al., 2017).

In their review, Wixted and Wells (2017) found that confidence was highly informative of accuracy. They suggest that identifications in a lineup made with low accuracy were less than 70% accurate, while those made with high confidence were approximately 97% accurate. The same relationship between confidence and accuracy has also been found in freely recalled narratives (Roberts & Higham, 2002). The current study also found uncertainty indicators were informative of accuracy consistent across interview phase, meaning that uncertainty indicators suggested lower confidence and lower accuracy in details in both the free recall and question phase of the interview. This suggests that individuals use uncertainty indicators in both free recall and cued recall tasks to indicate their confidence in the information being given, in turn indicating their accuracy. This finding was consistent with Flowe et al. (2019), who reported an increase in accuracy as confidence increased.

In finding support for hypothesis two, it was seen that participants were more likely to express uncertainty in the question phase of the interview than the free recall phase. Pansky et al. (2005) reported similar results, suggesting that individuals will report only what they feel they remember during free recall. This is supported by studies that demonstrated increased accuracy during free recall, when compared to a 'forced recall' (or question phase) portion (Koriat & Goldsmith, 1994; Koriat et al., 2001). Mickes et al. (2013) suggest that information recalled first is better learnt and remembered in more detail than information recalled later on. Indeed, consistent with this, accuracy was also higher in the free recall phase than the question phase in the current study. In a police interview-type structure where participants' desire is to report accurate information, it is likely that more accurate information would then be given first in the free recall phase, rather than the question phase. Building on findings from hypothesis one, if participants are more likely to use uncertainty indicators when their accuracy is lower, and if accuracy is lower in the question phase, participants express more uncertainty in the question phase.

Findings from hypotheses one and two can be supported by Goldsmith et al. (2002). Goldsmith et al.'s (2002) strategic memory regulation framework suggests that individuals can decide whether to maximise the accuracy or the informativeness of the information they are providing through the use of coarse- and fine-grained information. Coarse-grained information is that which gives minimal detail (for example, the perpetrator was wearing dark bottoms) and fine-grained answers provide more specific information (for example, the perpetrator was wearing navy bottoms with a white logo). Goldsmith et al. (2002) state that course-grained answers are more likely to be accurate than fine-grained answers although provide less information, therefore coarse-grained answers maximise accuracy over informativeness. The framework states that witnesses attempt to retrieve fine-grained answers, and assess their accuracy by using their confidence in this. They then compare this to a criterion value which defines when a memory signal is strong enough to report. If the witness's confidence exceeds this value, the fine-grained answer is reported, if not, attempts are made to retrieve a coarse-grained answer. The process of assessing confidence against criterion value then occurs again to decide whether the coarse-grained answer should be reported or withheld. Ackerman and Goldsmith (2008) report that witnesses prefer to report fine-grained answers than coarse-grained answers, however this tendency can be mitigated through factors such as penalties for incorrect answers (McCallum et al., 2016). Butt et al. (2020) found that witnesses were significantly less confident in their fine- compared to coarse-grained answers. They also reported that participants were more likely to volunteer fine-grained information to the police, maximising informativeness.

This framework suggests that in the free recall phase of the interview, participants would be more likely to volunteer fine-grained information, in which they felt more confident. It is also possible that the criterion value for assessing whether to report this information would be raised in a police-style interview whereby there are significant consequences for reporting

inaccurate information. This would be consistent with the current study findings that participants were more certain of the information provided in the free recall phase, where they would rather report fine-grained information and have a higher criterion value to do so. In the question phase, participants may be asked about details in which they are less confident, therefore assessing confidence against the criterion value and potentially reporting coarse-grained information of which they are less certain, leading to an increase in uncertainty indicators. If participants are reporting information in which they are more confident, this would suggest they are more accurate, leading to a higher proportion of details recalled correctly without an uncertainty indicator.

Hypothesis three stated that participants in the alcohol group would express more uncertainty than the sober group, particularly in the question phase of the interview. This study did not find support for this hypothesis. Alcohol intoxication did not affect the likelihood of participants expressing uncertainty, and these uncertainty cues were found to be equally informative of accuracy across the alcohol intoxicated and sober groups. Alcohol intoxication did not make participants more uncertain overall, or more uncertain particularly in the question phase. This may make sense, however, as the study also found that alcohol did not affect the proportion of correct details given by participants in the free recall phase, or the question phase. That is, memory accuracy was not found to be significantly different across the sober and intoxicated groups.

Research on the effect of alcohol on eyewitness memory accuracy is mixed. For example, Bartlett et al. (2022), Evans et al. (2018), Van Oorsouw et al. (2018), and Van Oorsouw et al. (2019) found negative effects of alcohol on eyewitness memory accuracy (see also Chapter 2 for further discussion of mixed findings of alcohol on eyewitness memory accuracy). However, other researchers have found no significant differences between alcohol intoxicated, and sober participants (Schreiber Compo, 2011; Flowe et al., 2019). Flowe et al.

(2019) found no significant difference in memory accuracy between sober and intoxicated groups, as well as an increase in accuracy with confidence, regardless of beverage group.

It is possible that the mixed findings across the literature are dependent on the dose of alcohol given to participants. Altman et al. (2018) found that alcohol decreased accuracy of recall in participants, however, this effect was found in participants with a higher BAC than would be typically found in the laboratory, with BAC's between .00 and .29% (M = .08%). This suggests that there may be, at least in part, dose dependant effects of alcohol on eyewitness memory. Further support comes from Van Oorsouw and Merckelbach (2012), who conducted a field study in which they found that severely intoxicated participants were less accurate than sober participants, and that this effect was more prominent in the question phase of recall, than the free recall phase. This effect was, however, found in the moderately intoxicated group, with similar levels of intoxication to the current study (M=0.06%). No effects of alcohol on memory accuracy were found at mild to moderate levels of intoxication by Schreiber Compo et al. (2012). Findings from Flowe et al. (2016) support the findings of the current study, suggesting that alcohol intoxicated participants may give fewer details, but that the accuracy of these details were not impacted in comparison to a sober control group. Although we did not measure the completeness of memory in this secondary data analysis, the results showed no statistically significant difference in accuracy across the alcohol and tonic groups.

Practical Implications

The findings of this study have important implications for witnesses or victims that are alcohol intoxicated at the time of a criminal event. The lack of alcohol intoxication effects found in this research suggest that alcohol intoxicated witnesses can be as accurate as sober witnesses in both free recall and questioning. This finding is contradictory to the beliefs of many in the Criminal Justice System (Crossland et al., 2021) and therefore may be beneficial

in educating both professionals and jurors in cases where victims of witnesses of crime have been intoxicated with alcohol.

The current research, and the research of others (e.g., Goldsmith et al., 2005; Mickes et al., 2013) demonstrate that victims often volunteer their strongest memories first in free recall, withholding weaker memories of which they are less confident and accurate, for the question phase. The ABE—which is considered to be the gold standard interviewing technique in England and Wales—strives to obtain as much detail as possible from victims in the question phase (Ministry of Justice, 2011) however, this detailed questioning may encourage victims to provide weaker and potentially less accurate memories, therefore reducing the quality of memory evidence they are providing. This reduced quality of memory evidence could result in erroneous convictions or victims being discredited by legal teams in court in relation to inaccurate evidence given in the question phase. Inaccurate information from the question phase of interviews may be highlighted by the defence to demonstrate the unreliability of a witness, disregarding the potentially greater accuracy of information given in the free recall phase. Findings of this study may therefore provide an evidence base for legal teams to consider the inaccuracy brought about by questioning, rather than the unreliability of a witness.

Garett's (2009) analysis of 250 DNA exoneration cases showed that in cases where individuals were wrongly convicted of sexual offences (67% of the cases in this study were rape), the convictions were often brought about through reliance on weak memory evidence. It was, however, recognised by Wixted et al. (2018) that victims in these cases who made errors in identification had not expressed high confidence in their identification. Similarly, in the current study, it was found that uncertainty indicators appear to be a useful measure of a witness or victims confidence in the information they are providing during free recall or questioning throughout the cognitive interview. This shows a clear example of how metamemory processes could be harnessed to increase the informativeness of evidence given to the police and aid

police officers in determining which information to rely upon based on witness metamemory cues. That is, if memory reports are weak, the police might require further corroborating evidence before charging an individual, or, conversely, if some memory for details are strong, the police could rely on this information to further their investigation. Moreover, metamemory processes could be harnessed to improve other legal decision-making, such as in juror instructions to inform jurors that confidence at time of initial memory report is informative of a witness's memory accuracy. Douglass et al. (2010) suggest that jurors find the confidence of witnesses and victims to be persuasive. As a lack of uncertainty indicators (and therefore confidence in a detail) appears to be associated with the accuracy of the detail given, this research may be beneficial in further increasing juror confidence in initial memory reports from witnesses or victims.

Findings from this and future research could allow knowledge to be disseminated to police forces, allowing for more accurate victim statements. Suggestions may be made not to push for information in the question phase of the interview, and to use metamemory indicators to estimate witness or victim certainty. Findings may also increase the likelihood of police officers giving more weight to information provided by witnesses or victims that were alcohol intoxicated at the time of the offence. Ultimately, reliance on accurate memory evidence could strengthen prosecutors' cases in those that are notoriously challenging to prosecute, such as rape and sexual assault, where witness statements are often used to discredit the witness (Daly & Bouhours, 2010).

Limitations and future research

Due to ethical issues, the BAC's of participants in the original Flowe et al.'s (2019) study was lower than often encountered in individuals drinking alcohol in pubs and bars, where they may witness crime. Specifically, participants in the original study had a mean BAC of .06%, compared with a mean BAC of .08% in Altman et al.'s (2018) field study, where bar

patrons had self-intoxicated. There is currently mixed support regarding the dose dependent effects of alcohol intoxication on eyewitness memory (Jores et al., 2019). Therefore, future research should consider an adaptation of the study in a field setting in order to achieve higher BAC's in participants and assess the metamemory processes during a subsequent interview.

Future research may also want to consider the informativeness of different metamemory indicators of uncertainty. In this study, long delays, short delays, word fillers, non-word fillers, and word hedges were all collapsed into indicators of uncertainty due to the limited frequency of some indicators used by participants. Long delays, for example, were not used by participants in the free recall phase. Examining the informativeness of each uncertainty indicator, however, may provide even more beneficial information to the police and the Criminal Justice System when considering the confidence and accuracy of witnesses and victims.

Future research may also benefit from increasing the ecological validity of the study. While participants reported being emotionally affected by the rape scenario presented to them through the participant choice paradigm (Takarangi et al., 2013), this experience was, of course, not emotionally comparable to experiencing rape. The use of a video, or other method, to present the simulated rape scenario may increase the psychological realism of the study, as well as potentially impacting the recall of participants. This may strengthen the application of the findings to real-world scenarios.

In the development of the coding scheme used in the current study, there was difficulty in coding delays between the end of a question, and the beginning of a participants answer, due to only having audio recordings of the interviews (e.g., participants may have been drinking, looking elsewhere). This could potentially lead to important data not being captured. It is recommended that future research use video recordings to capture this data and eliminate ambiguity in different reasons for delays before participants generate a response.

Finally, the current study collapsed data across expectancy conditions. This was due to the small number of participants in the secondary data analysis, and the uneven participant numbers this created in expectancy conditions. The original study by Flowe et al. (2019) found that women who were told they had consumed alcohol, and had consumed alcohol, recalled fewer incorrect details whereas those who consumed tonic water but were told they consumed alcohol, reported more incorrect details. Flowe et al. (2019) did, however, not find these results to be statistically significant. They also did not find any effects of expectancy on misinformation reporting. While the original study found no statistically significant expectancy effects, it is possible that expectancy effects may have impacted on the findings from the secondary data analysis. Future research should consider exploring expectancy effects when considering the use of metamemory indicators in police interviews.

Conclusion

Overall, when examining the effects of alcohol intoxication on metamemory processes during police interviews, no effects of intoxication on memory accuracy were found. Furthermore, alcohol intoxicated participants were no more uncertain in their recall than sober participants. Effects of the interview phase were found, suggesting that participants, regardless of alcohol intoxication, were more accurate in the free recall portion of the interview and that uncertainty indicators were informative of accuracy throughout the interview. Findings suggest that alcohol intoxicated victims and witnesses are as accurate as their sober counterparts at police interview. The findings also suggest that the structure of police interviews, asking questions about details not recalled in free recall, may introduce inaccuracy into the information provided by witnesses or victims. The informativeness of metamemory indicators however may be helpful in identifying less accurate information given by individuals throughout the police interview process.

Chapter Four

A Psychometric Critique of the Alcohol Use Disorders Identification Test

The word 'psychometric' translates to measurement of the mind (Sireci et al., 2005). Hammond (2006) states that psychometric tests are often used in psychological research as they are designed to test a broad range of psychopathologies, symptomologies, personality traits, and more. They are used as measurement tools to access psychological variables to be studied in research. Part of what makes them preferred in this setting is their ease of use when collecting and scoring data. This makes them simple to use over large populations to allow for data collection from a larger number of participants, or greater screening ability before studies. DeVon et al. (2007) state that the foundation of research is the use of measurement tools that are psychometrically sound, and that the reliability and validity of these tools must be high in order to ensure the integrity of a study. This chapter will consider the use of the Alcohol Use Disorders Identification Test (AUDIT) in research screening processes, considering its reliability and validity in practice.

The Alcohol Use Disorders Identification Test (AUDIT) was developed by the World Health Organisation (WHO) in a six-country collaborative study and has been used worldwide since 1989 (Babor et al., 2001; WHO, 2001). It was designed to examine an individual's relationship with alcohol across three different domains: alcohol intake, potential dependence on alcohol, and experience of alcohol-related harm (WHO, 2001). Alcohol has been known to be directly linked to morbidity and mortality, including depression, anxiety, cancer, violence, and suicide (Saunders et al., 1993a; Saunders et al., 1993b) making it important that healthcare settings are able to identify difficult relationships with alcohol and provide care as necessary. The AUDIT was initially developed for use in primary care settings, however, has since been validated for use in other settings, including in the community and other healthcare settings such as psychiatric inpatient hospitals (Lima et al., 2005).

The AUDIT is a ten-item questionnaire developed from an original bank of 150 questions that were studied. It asks individuals about the amount and frequency of their alcohol intake and is scored by summing the value associated with each response option on a scale of 0 to 4, allowing for possible scores between 0 and 40 with a generally accepted cut-off score of 8 (Nadkarni et al., 2019). The WHO state that this cut-off score provides good sensitivity to alcohol use disorders and suggests the following classification of scores: hazardous drinking (8-15), harmful drinking (16-19), and dependent drinking (>20) (Babor et al., 2001). Although a score of 8 is the generally accepted cut-off there are many questions in the literature regarding this. Some studies have found better specificity at a cut-off score of 10 (Babor et al., 2001) however in some populations lower cut-off scores such as 7 have been recommended (Reinert & Allen, 2007).

The AUDIT manual (WHO, 2001) states that as well as being used to identify those suffering with diagnosable alcohol use disorders (AUD) and signposting them to treatment, it can also be used as a tool for education. Those with elevated scores but that do not reach a clinical cut-off can be signposted to educational resources and given support. The measure can also be used to identify whether healthcare providers need to be concerned about how an individual's drinking may interact with any medication they currently take. The measure was designed to be self-administered however can also be used in an oral interview. There are benefits and limitations to both of these approaches, however self-administration is usually a quicker way to administer the test, while an interview may be beneficial for those that struggle to read or write.

A systematic review conducted by Nakarni et al. in 2019 reviewed the use of differing cut-off scores in different populations and cultures. The study by Nakarni et al. (2019) will be discussed in relation to the effect these differing scores may have on the reliability and validity

of the AUDIT. It will then be considered how differing cut-off scores may impact the use of the AUDIT in research, such as in the original study from which the data were generated for Chapter 3 (Flowe et al., 2019) and the impact of this on alcohol research more generally.

Reliability

Reliability is concerned with the ability of an instrument to measure consistently under consistent testing conditions (Kline, 1986). There are different types of reliability, including test-retest reliability, internal consistency, and interrater reliability. As the AUDIT was designed to be self-administered, there are no studies evaluating interrater reliability. Interrater reliability is the extent to which raters consistently can distinguish between items on a measurement scale, with the general trend of ratings being considered important, as opposed to absolute values assigned (Gisev et al., 2013). For example, if three raters rating the communication skills of a pharmacist on a scale on 1-10, and all three have the same ranked order, the interrater reliability would be considered high (Gisev et al., 2013). Test-retest reliability and internal consistency will now be discussed in turn.

Test-retest reliability is concerned with the stability of test results over time, however, questions have been raised about the appropriateness of this method with some tests. Repeated testing may sensitise the respondent to the questions asked, therefore, changing results (Joppe, 2000). It is also possible that the construct being measured will change over time (Bannigan & Watson, 2009). In the case of the AUDIT, those that produce high scores may begin receiving interventions (with their consent), therefore, lower scores may be expected after the re-test period. This may give the illusion of lower reliability of scores, however, this may not be the case.

When using test-retest scores to determine the reliability of the AUDIT, results indicated high reliability (r=.86; Sinclair et al., 1992). The acceptable value for test-retest values is considered to be greater than .7 (Guilford, 1956). It is, however, of note that there is

no standard time interval between the implementation of tests, therefore, it can be difficult to determine the amount of time required to wait so as to reduce confounding variables such as participants remembering items from the test. If participants remember their answers from the first test and repeat these, this inflates test re-test reliability scores. A systematic review conducted by de Meneses-Gaya et al. (2009) identified three studies that had all used a re-test period of one month. The first study by Selin (2003) found a test-retest correlation of between 0.6 and 0.8 which is considered between good and excellent. The only item that showed a much lower correlation was item nine (asking when the respondent or those around them have been injured due to their drinking) which had a correlation of only .29. Overall, Selin (2003) saw a test-retest correlation of .84, clearly within the acceptable range. The second study (Dybeck et al., 2006) also found the lowest correlation on item nine of the test (.39) suggesting potential issues with this question. It may be that incidents of this nature occurred during the re-test period but further investigation may be required to establish hypotheses for this low correlation. The study also found that using the WHO suggested cut-off score of 8, 87.5% of participants that were screened as having AUD and 98.9% of those that tested negative, were classified the same in the re-test. These values were similar when using a cut-off score of 5 (as cut-off scores often vary across studies). Three studies have evaluated the test-retest reliability of the AUDIT using the same cut-off scores of 8 or above. Lennings (1999) and Maisto et al. (2000) used a re-test period of 2 weeks and found correlation coefficients of .92 and .64, respectively. Daeppen et al. (2000) used a re-test period of 6 weeks and found a correlation coefficient of .81. These correlation coefficients are within acceptable range and demonstrate stability of the AUDIT test results over time.

Internal consistency is the degree of homogeneity of items in the test to ensure they are measuring the same construct. Internal consistency can be measured using Cronbach's alpha, which is the most widely used statistic to show internal consistency by indicating how well the

items on a tool fit together conceptually (DeVon et al., 2007). It is the only reliability index that can be performed when a test has only been administered once, eliminating issues such as those seen in test-retest methods (Ferketich, 1990). Classical test theory (Spearman, 1907) was developed in order to improve the reliability of tests. The theory states that each respondent on a test has a true score and an observed score. The observed score (what is seen at the end of the test) is the true score but with error added to this (Traub, 1997). The reliability of test scores increases as the proportion of error variance is lowered, therefore reliability can be determined using Cronbach's alpha which provides a lower bound for reliability.

Opinions regarding the ideal Cronbach's alpha value differ, with a value of .70 being seen to be acceptable for new scales (DeVellis, 2003), however, other authors have recommended much higher values, stating a preferred Cronbach's alpha of .90 or in some cases .95 (Nunnally & Berntein, 1994; Polit & Beck, 2004). A lower alpha value may be due to factors such as a low number of questions, poor inter-relatedness between items, or heterogeneous constructs (Tavakol & Dennick, 2011). This can be determined through correlating test items with the total test score and removing items with a low correlation. If, however, alpha is very high, this may indicate that some questions on the test are not necessary as they are asking the same question and therefore can also be removed (Streiner, 2003). Reinert and Allen (2002) conducted a review of studies that reported psychometric properties of the AUDIT. They reported that of the studies reviewed that reported a Cronbach's alpha value, a median value of well within .80 was found, and this is stated to be an acceptable level. One of the initial validation studies carried out on the ADUIT for use in primary healthcare settings (Valladolid et al., 1998) found an alpha value of .86. With an acceptable alpha value of greater than .70 (Kline, 1999; DeVellis, 2003), this demonstrates that internal reliability of the AUDIT is within an acceptable range.

In sum, the Cronbach's alpha values found for the AUDIT across meta-analyses and the test-retest scores found in systematic reviews are all within acceptable ranges suggesting that the AUDIT can be considered to be a reliable measure. Questions remain regarding the low test-retest correlation seen on item nine but overall, no concerns are raised.

Validity

Joppe (2000) defines validity as a measure of whether or not a psychometric test is measuring the construct it was designed to measure and whether the inferences made from the results are appropriate, meaningful, and useful.

Content validity determines the extent to which a measure is representing and testing all measures of a construct and therefore how well it relates to the theory surrounding the construct. The AUDIT was developed using items from a variety of self-report, laboratory, and clinical procedures that have been used across a variety of countries to detect AUD (WHO, 2001). After this process was complete, another multinational study (Saunders et al., 1993a) was conducted in order to determine the best features from each of these approaches at distinguishing low risk drinkers from those with harmful alcohol consumption. As items were developed, their correlation was checked against daily alcohol intake in a group of participants, some of whom had been diagnosed as alcohol dependent. Items were also selected based on having good face validity, clinical relevance, and to ensure they covered the three relevant domains (alcohol use, alcohol dependence, and adverse consequences of drinking). Finally, it was ensured that these items were applicable across genders and could be generalised across cultures to the best possible extent (due to differences in drinking cultures across countries).

Once the items were developed, they were then tested for sensitivity (the percentage of positive cases correctly identified) and specificity (the percentage of negative cases that were correctly identified). Various cut-off scores were also used when computing specificity and

sensitivity in order to find the most appropriate cut-off score which is judged by the WHO to be 8. When using a cut-off score of 8, Saunders et al. (1993b) found the specificities (percentages of correctly identified negative cases) averaged in the 0.80's, making this a valid cut-off value. As development took place through a large multinational study of measures that were already able to appropriately identify harmful drinking, the test can be considered to have good content validity. The test is also consistent with the ICD-10 definitions of alcohol dependence and harmful alcohol use (Babor et al., 1994; WHO, 1993). While the test was not developed based on specific theory, it has been developed in line with the ICD-10 which is used by clinicians to diagnose AUD and was itself developed based on a clearly defined construct. Further, there are systematic reviews such as de Meneses-Gaya (2009) that support the use of the AUDIT to appropriately measure the construct of AUD's. It can therefore be considered to have high content validity.

Criterion validity is a measure of how well a psychometric can predict an outcome, in this example it would be a measure of how well the AUDIT can predict AUD (Vogt & Johnson, 2011). Evidence for criterion validity can be seen in the test remaining sensitive throughout differences in the populations being tested. The AUDIT has been developed and used multinationally and has been tested in a number of distinct populations. In primary care settings, where the AUDIT was originally designed to be used, the AUDIT performs well with large areas under the receiver operating curve (ROC) being observed (Cherpitel, 1995, 1997; Volk et al., 1997). The area under the curve is a measure of how well the measure can distinguish between two groups, in this case those with AUD's and those without. This demonstrates that a large proportion of the times the test is used, it is correctly discriminating between positive hits and false alarms, therefore it is correctly identifying those with AUD's. Bischof et al. (2005) used ROC analysis and found an area under the curve of .98, suggesting excellent discriminatory validity of the AUDIT to tell the difference between those who did

and did not have AUD. The area under the curve when testing the AUDIT has also seen to be greater than that of other AUD identification psychometrics such as the CAGE (.70) or CRAFFT (.79; Cook et al., 2005), suggesting that it is overall a more valid method of identifying AUD's (Reinert & Allen, 2002).

Systematic reviews (Nadkarni et al., 2019) have supported the use of the AUDIT in a large number of populations. Differences, however, do start to appear when looking at the AUDIT's use across gender and culture. Nadkarni et al.'s (2019) cross-cultural systematic review identified fifty-four studies that did not use the WHO recommended cut-off scores and found that the majority of these studies did not measure the psychometric properties of the AUDIT when the WHO-recommended cut-off score or the modified cut-off was used. Twelve of these studies were validation studies that used a range of cut-off scores to detect different levels of drinking. The scores used to detect hazardous drinking ranged from >3 to <5, harmful drinking ranged from >5 to <16 and dependent drinking ranged from >7 to <24. An additional four studies also recommended different cut-off scores for males and females. The review also noted that of the four studies identified that used standardised terminology and psychometric data, no cut-off scores clearly outperformed others. Dependent drinking was measured using the widest range of cut-off scores from >7 to <24, and all but the score of 7 provided high psychometric properties with sensitivity ranging from 81 to 100% and specificity from 28.6 to 94.1% (Santis et al., 2009). This suggests that further exploration of the appropriate cut-off scores across gender and culture are required in order for consistency in the cut-off scores used, and to ensure the validity of the AUDIT.

While the AUDIT was designed to be effective across cultures and groups, there is limited recent research that reviews the validity of the psychometric, with very small numbers of studies testing the reliability and validity of non-English versions of the test (de Meneses-

Gaya et al., 2009). Non-English versions of the test include Nigerian dialects, Hindi, Spanish, Chinese, German, French, and Vietnamese (de Meneses-Gaya et al., 2009). The AUDIT has been found by some studies to require a different cut-off score dependent on gender (Reinert & Allen, 2007) while others have found this not to be the case (Cook et al., 2005) leading to questions about the validity of cut-off scores used across genders. When using the AUDIT across different age groups, cut-off scores appear to change. Knight et al. (2003) found that at a cut-off score of 8, the AUDIT when used with teenagers (14 to 18 years old) was almost insensitive, but yielded better results at a cut-off score of 2. In elderly patients, the AUDIT demonstrated an area under the curve of .96 and showed sensitivity of .67 and specificity of .96 but did not use one cut-off score (scores of 7 or 8 were used; Philpot et al., 2003). In patients experiencing first episode psychosis, a cut-off score of 10 was preferable, yielding a sensitivity value of .85, specificity value of .91, and area under the curve of .86 (Cassidy et al., 2008). While the AUDIT has been tested in a number of varying samples and has produced satisfactory results demonstrating its reliability and validity as a measure, further research is required to establish different cut-off scores across groups. Currently, determining the overall validity of the measure is somewhat difficult as this relies heavily on the population being tested. Practitioners may have difficulty in determining the validity of the measure in their population without validation studies in these specific groups, meaning they may also be using inappropriate cut-off scores.

Construct validity is the degree to which a measure is appropriately measuring the construct it claims to measure (Vogt & Johnson, 2011). The AUDIT manual describes the measure as having three distinct factors based on conceptual domains: hazardous alcohol use, dependence symptoms, and adverse consequences of drinking. Maisto et al. (2000) used results from primary care settings as the AUDIT's standardised sample and concluded that although the three-factor structure can be supported by data, there may be a better model fit when using

a two-factor structure of dependence/consequences and alcohol consumption. Other studies have continued to find support for the two-factor model (Karno et al., 2000; O'Hare & Sherrer, 1999), finding that items 1-3 load onto a single factor of alcohol consumption. While this does suggest that the measure is appropriately measuring the constructs it claims to, the more specific domains may not be being as accurately measured as once thought. Although a three-factor model can fit, the support for a two-factor model suggests that only two domains are being appropriately measured during administration.

Use of AUDIT in research

The AUDIT has been used in many psychological memory studies to screen participants prior to them potentially being given alcoholic beverages in the study (e.g., Flowe et al., 2019; Miller et al., 2018; Sharma, Albery, & Cook, C, 2001). For example, the data that were reanalysed in Chapter Three were collected by Flowe et al. (2019), who used the AUDIT to identify suitable participants. As with many laboratory studies carried out in the United Kingdom, the participant sample was mainly university students, from the University where the research was being carried out due to the ease of accessing these participants. As the items on the AUDIT are dependent on cultural attitudes towards drinking behaviours, it is necessary to consider the cultural norms of university students in the United Kingdom in regards to alcohol consumption in interpreting AUDIT scores. A study by Livingstone et al., (2011) found that health related behaviours such as alcohol consumption are related to social identity and norms. They state that heavy alcohol consumption is an important part of social identity in university students in the United Kingdom. They found that heavy drinking was endorsed by in-group members and that this positive attitude towards drinking led to a greater intention to drink heavily in social situations. It was seen that high identifiers were more likely to exclude non-drinkers. As the prevalence of heavy drinking in UK university students can be inferred to be high, this raises questions about the validity of a

cut-off score of 8 in University participants presenting for screening to take part in studies. As the AUDIT was originally developed to be used in primary health settings where people are often presenting due to AUD's, it must be considered how cultural norms affect the scores seen in these students who are not presenting with ill health related to drinking behaviours. There are clear differences between the cut-off scores required for different groups when using the AUDIT therefore it is difficult to generalise its validity across groups

DeSimone (2005) states that drinkers are likely to be less risk averse, although it is not clear whether individuals who drink more frequently or more heavily are less risk averse. In Flowe et al.'s (2019) study, participants took part in a hypothetical rape scenario using the participant choice paradigm. In the scenario, the participant was interacting with an unknown man in a social setting and chose whether to continue the interaction until the end, in which case they have consensual sexual intercourse, or withdraw prior to this, in which case they are informed a legally definable act of rape is committed against her. If drinkers who are above the cut-off score of 8 on the AUDIT are less risk averse, it may be that they would behave differently in the scenario that was given in the study conducted by Flowe et al. (2019), such as spending more time with the man and perhaps letting him into her home. If they were to continue further into the participant choice scenario, this would mean they would be required to accurately remember more information than a participant that withdrew from the scenario sooner. This raises questions about the generalisability of the results found by Flowe et al. (2019) to target samples such as university students that may be drinking heavily. Conversely, it is seen that although some cut-off scores for the AUDIT are higher or lower than 8, this cut-off score generally provides adequate specificity and sensitivity, suggesting that the AUDIT is a useful measure when screening research participants in order to comply with ethical procedures and greatly minimise the risk of providing alcohol to those with AUD's. Overall, it seems that future research should consider the use of the AUDIT in research screening

processes and the populations often used in these such as university students in order to provide validated cut-off scores and better guidance for researchers.

Conclusion

The AUDIT as a measure is easy to administer because it can either be used as a self-report psychometric or be administrated by an oral interview that is completely structured. This helps to standardise the test as no administrator skill is required to administer or score the measure. The AUDIT is also considered objective as the items are easy to answer and the scores it provides enable clear communication between professionals and researchers alike.

Together, the research evidence suggests that the AUDIT is seen to be a reliable measure with high internal reliability, with a Cronbach's alpha value that is on average >.80 (Reinert & Allen, 2002). It also demonstrates good test-retest correlation values, making it overall an acceptably reliable measure. The AUDIT has been developed in a manner that allows for high content validity by drawing items from a number of multinational procedures and measures, and matching the final items to the ICD-10 definitions of AUD's. The AUDIT also demonstrates good sensitivity and specificity, with high area under the ROC values when used in its standardised setting (primary healthcare) when using the WHO recommended cut-off score of 8. It is, however, of note that the AUDIT is used frequently in a large range of settings and across different ages, genders, and cultures, and debate surrounds the cut-off scores necessary to yield good specificity and sensitivity in these populations. It is because of this that the validity of the AUDIT in some settings is not known but its use likely continues in practice. Caution is needed when using the AUDIT in different settings where it has not been validated and future research should seek to further examine appropriate cut-off scores. Research should also examine the AUDIT's use in screening processes for alcohol research, as no literature currently exists examining this topic specifically.

Chapter Five

Discussion

This chapter provides an overview of the findings of previous chapters. Also discussed are the overall findings of this thesis, and how these may be applied to future theory, policy, and practice.

Overall Aims

The overall aim of this thesis was to better understand the impact of alcohol intoxication on the memory of witnesses and victims of sexual assault. The thesis sought to address this aim through the systematic review of current literature and methodologies for exploring this topic, exploring the accuracy of information given by those in a police-style interview related to a mock sexual assault, and assessing the informativeness of metamemory indicators to provide insight into the accuracy of these witnesses and victims. This work may, in the future, be used to help develop clear, consistent, and evidence-based guidance for the Criminal Justice System when collecting and interpreting memory evidence from intoxicated witnesses and victims.

Summary of Findings

Chapter One introduced the requirement for research examining the relationship between alcohol and memory by demonstrating the strong link between alcohol and crime. It was noted that crime often occurs in settings where victims and witnesses have consumed alcohol (Leaonard et al., 2002). The chapter also discussed the negative beliefs of those in the Criminal Justice System in relation to the effects of alcohol on memory (Houston et al., 2013; Kassin et al., 2001; Schuller & Stewart, 2000). This, coupled with the low rates of prosecution of sexual assault, demonstrate the importance of research that explores the accuracy of memory in alcohol intoxicated witnesses and victims of sexual assault.

Chapter Two used a systematic approach to reviewing the literature relating to alcohol intoxication and misinformation. The review of twelve studies exploring the impact of alcohol intoxication on eyewitness memory found mixed results. It was concluded that while alcohol may have a negative effect on the acceptance of misinformation in eyewitness testimony, issues of blood alcohol concentration (BAC; i.e., the dose of alcohol received), timing of alcohol consumption, and delays in testing may have all contributed to the acceptance of misinformation. Of the twelve studies, five found negative physiological effects of alcohol intoxication on misinformation acceptance (Evans et al., 2018; Van Oorsouw et al., 2015; Van Oorsouw et al., 2019), one found a protective physiological effect of alcohol (Gawrylowicz et al., 2017), four found no physiological effects of alcohol (Bartlett et al., 2022; Flowe et al., 2019; Mindthoff et al., 2021; Schreiber Compo et al., 2011), and three found social effects of alcohol intoxication (Assefi & Garry, 2003; Bartlett et al., 2021; Zajac et al., 2016). The review noted that studies with a higher quality assessment score, and more often conducted in a laboratory setting, found no effect of alcohol intoxication at encoding on misinformation acceptance. When comparing the two studies examining alcohol's effect when introduced at the same time as misinformation, the paper of significantly higher quality found positive effects, with participants who consumed alcohol reporting less misinformation. Explorations of the effects of alcohol intoxication at test, consistently found no significant increase in the amount of misinformation reported, and all studies were of good quality. Findings of no significant effect of delayed testing were also found, however the reasons for this appear unclear.

The undertaking of future research to further disentangle the effects of dose, timing of alcohol consumption, and testing delays was suggested in order to appropriately manage the collection of statements from intoxicated witnesses and victims in the Criminal Justice System in order to obtain the most accurate information.

Chapter Three presented a study examining the effects of acute alcohol intoxication during a mock- rape on metamemory processes during police interviews. The study examined the use of new indicators to examine metamemory processes over the course of a police-style interview to determine the informativeness of the indices in indicating memory accuracy, as well as the efficacy of the police interview in allowing witnesses and victims to provide full and accurate memory evidence. The study also examined the effects of alcohol intoxication at encoding on the accuracy of participants in the free recall and question phases of the interview. The study was a secondary data analysis of Flowe et al.'s (2019) study, in which participants taking part in a hypothetical rape scenario were either alcohol intoxicated or sober, and were interviewed using the Cognitive Interview one week later. The study coded for metamemory indicators and examined the memory accuracy of participants across interview phase and beverage group.

No effects of alcohol intoxication on memory accuracy were found in the study. Furthermore, when examining metamemory indicators, alcohol intoxicated participants were no more uncertain in the information they provided than sober participants. This suggests that alcohol intoxicated victims and witnesses are as accurate and confident as their sober counterparts during police interviews. Effects of interview phase were seen, with participants being more accurate in the free recall phase of the interview than in the question phase. Uncertainty indicators were found to be informative of memory accuracy throughout the interview, with more uncertainty indicators expressed in the question phase than in the free recall phase. Memory accuracy was also poorer in the question phase than in the free recall phase. These findings call into question the potential efficacy of police interview techniques in obtaining full and accurate evidence, as the additional probing for information in the question phase of the interview is likely to introduce uncertainty and inaccuracy, into witness and victim accounts.

Chapter Four examined the use of the Alcohol Use Disorders Identification Test (AUDIT) test in assessing alcohol use disorders (AUD's). The measure is easy to use and is standardised due to its ability to be used as a self-report measure or as a structured interview. It is seen to be a reliable measure, with high internal reliability (Reinert & Allen, 2002) and good test-retest correlation values. It is also seen to have high content validity, drawing on items from a variety of self-report, laboratory, and clinical procedures that have been used multinationally to detect AUD (WHO, 2001). The AUDIT's sensitivity and specificity are also seen to be good. While the review of the literature found good reliability, validity, specificity, and sensitivity of the AUDIT, this was primarily seen using the World Health Organisation (WHO) cut-off score of 8, and in settings for which the AUDIT was designed, such as primary healthcare settings. The chapter discussed the frequent use of the AUDIT outside of these settings, as well as across different cultures, ages, and genders. The cultural norms surrounding alcohol consumption and, therefore, debate surrounding appropriate cut-off scores, were considered. The chapter suggests that future research should focus on the cut-off scores required to achieve good sensitivity and specificity across populations, as well as the validity of the scale's scores using these cut-offs. The chapter finally discusses the AUDIT's frequent use in screening prior to alcohol and memory research, such as in Flowe et al. (2019). Currently, no studies examine the AUDIT's use in research screening processes.

Research and Theoretical Implications

Research into the effects of acute alcohol intoxication on the memory of eyewitnesses has grown rapidly in recent years, with research generally suggesting that alcohol intoxication negatively effects the completeness of memory, but not its accuracy (Jores et al., 2019). Chapter Two, however, demonstrates that the effects of alcohol on suggestibility are less clear. It is clear that more research is needed in order to disentangle the effects of alcohol, delayed testing, and the timing of alcohol consumption. It is, however, of note that

the memory of alcohol intoxicated witnesses may be more reliable than once thought. A number of studies reviewed in Chapter Two support strategic memory reporting theories (Bartlett et al., 2021; Flowe et al., 2019; Gawrylowicz et al., 2017; Mindthoff et al., 2021). Strategic memory theories suggest that while an impact of alcohol on memory may exist, victims and witnesses are aware of this potential cognitive impairment and may omit information they are uncertain of. Individuals may volunteer memories of which they are more certain, which is often information that is learnt and remembered in greater detail, and is often volunteered first (Goldsmith et al., 2005; Mickes et al., 2013). Support for this theory in Chapter two came from studies that found that while there was no significant difference in the accuracy of alcohol intoxicated participants, their accounts were less complete (Bartlett et al., 2021; Flowe et al., 2019; Gawrylowicz et al., 2017; Mindthoff et al., 2021).

Further support comes from the study presented in Chapter Three, in which intoxicated participants were not significantly less accurate than sober participants, and that participants were more accurate in the free recall phase of the interview than the question phase, where they were able to volunteer their strongest memories. This is also in line with Goldsmith et al.'s (2002) strategic memory reporting framework, in which it is suggested that individuals can maximise the accuracy or the informativeness of their accounts through using either course- (minimal detail) or fine-grained (specific detail) information. Coarse-grained information is more likely to be accurate but is less informative in the details provided in a setting such as a police interview. For example, a coarse-grained piece of information may state that the offender was wearing dark clothing, whereas a fine-grained details would be that they were wearing a navy top and black bottoms. The framework suggests that individuals attempt to recall fine-grained information but when they are uncertain of the information, they may resort to recalling coarse-grained information, or no information at all if their confidence is particularly low. The findings in Chapter Three suggest that participants

may have used strategic memory when presenting information through the use of uncertainty indicators. More uncertainty indicators were seen in the question phase of the interview, suggesting that participants were less certain of the information they were providing in the question phase, compared to the free recall phase. This was further supported by reduced accuracy in the question phase of the interview, demonstrating a confidence-accuracy relationship.

It is, however, important to note that some studies do not support strategic memory reporting theories. For example, Assefi and Garry (2003), a study reviewed in Chapter Two, found that participants who believed they had consumed alcohol reported more misinformation than those who believed they were sober. This suggests that participants were not adjusting the information they reported based on the belief that alcohol would impair their memory, and not withholding memories they were not certain of. Nevertheless, overall, this thesis suggests that strategic memory reporting theories are important to consider in researching the impact of alcohol intoxication on eyewitness memory.

Another theory-Alcohol Myopia Theory (AMT; Steele & Joseph, 1990)-has been put forward as an explanation for the impact of alcohol intoxication on memory. AMT suggests that alcohol narrows the attention of the individual to central rather than peripheral cues, therefore an individual's memory of central details, such as details of an offender's appearance, would be better than those of peripheral details, such as the colour of the sofa as the crime scene. While recall of central details was not specifically reviewed throughout this thesis, AMT was not supported by the studies in Chapter Two that explored this (Flowe et al., 2019; Van Oorsouw et al., 2019). Flowe et al. (2019) did not find any decrease in the number of details reported by intoxicated participants however no distinction between central and peripheral details were made. This could have been achieved through the classification of central and peripheral details within the scenario, and coding these in transcripts. Future

research may wish to explore this. Van Oorsouw et al. (2019) found a reduction in central details recalled by intoxicated participants, and results for the recall of peripheral details were mixed. Memories for both central and peripheral details decreased in intoxicated participants when comparing immediate and delayed testing. In the immediate testing condition, participants did remember central details better, however intoxication reduced the amount of both central and peripheral details. Overall, these findings suggest that AMT may not always be a useful theoretical approach when researching the effect of alcohol on the memory of eyewitnesses and victims.

Throughout Chapter Two, studies also considered the impact of delays in testing on witness accuracy, as well as the impact of intoxication at the time of questioning. Van Oorsouw et al. (2019) found that immediate testing was preferable to delayed only testing, with delayed only testing reducing accuracy and increasing susceptibility to misinformation. This however is not supported by Van Oorsouw et al. (2015), who found better performance by sober and intoxicated participants on free recall after a sobering delay. Evans et al. (2019) also found an effect of delay in intoxicated witnesses, hypothesising that both delay and intoxication had weakened memory, leading to participants being unable to detect discrepancies, and therefore including misinformation. Mindthoff et al. (2021) found no detrimental effects of short delays on intoxicated participants' accuracy, but noted that this effect was only explored with mild to moderately intoxicated participants. Research regarding the suggestion that highly intoxicated witnesses may be more likely to report misinformation (Van Oorsouw et al., 2019), taking into account blood alcohol concentration, is required. Overall, the effects of delayed testing a mixed, and further research is needed to examine this effect. In Chapter Three, a one-week delay between encoding and recall was used. Future research may examine the informativeness of metamemory indicator changes for those in different intoxication states at encoding, after varying delays.

Studies presented in Chapter Two of this thesis also discussed the social effects of alcohol intoxication on memory. Bartlett et al. (2021) found that alcohol did not increase susceptibility to misinformation but that discussion with a co-witness did when compared to participants who recalled alone. Zajac et al. (2016) also examined the influence of co-witness discussion. While overall findings suggested that the intoxication state of the co-witness did not impact susceptibility to misinformation, there was an effect seen when participant responses were clearly discrepant with that of the co-witness. Discrepancy detection theory (Tousignant et al., 1986) states that when an individual's memory is weak and they cannot detect discrepancies between original memories and misinformation, this may increase the likelihood of filling in the blanks with misinformation. When participants in Zajac et al.'s (2016) study did detect a discrepancy, they then used source credibility to resolve this discrepancy. In this situation, participants were more likely to accept misinformation from sober co-witnesses than those who appeared intoxicated. These findings suggest that as well as further need for exploration of the physiological effects of alcohol, more research is needed to understand the social impacts of alcohol intoxication, and the discussions that take place between co-witnesses following an event.

There is an additional need for research to be conducted to understand the impact of trauma on the memory of alcohol intoxicated individuals. As previously discussed, the literature at this time often conflates witnesses and victims throughout research. Studies that involve participants in a mock crime as though they are victims, such as Flowe et al. (2019), do report participants feeling traumatised by the study materials, however this is not comparable to the trauma that would be experienced by a real-world victim of these crimes. Existing research presents conflicting results. Spiegel (1997) reported that traumatic memories are resistant to retrieval, unlike memories of neutral events, however a review by Laney and Loftus (2005) states that traumatic memories are not unique and simply deteriorate

over time. The review (Laney & Loftus, 2005) also suggests that high emotional arousal may lead to better memory for an event, as opposed to an event that does not produce any emotional arousal. Studies have demonstrated that individuals who are highly traumatised by experiences of child sexual abuse have better memory for the details of this abuse, than individuals who report feeling less traumatised by child sexual abuse (Alexander et al., 2005). Similarly, when specifically considering sexual trauma, Peace et al. (2008) found that memories of sexual trauma contained the most detail, compared with a memory of non-sexual trauma, or a positive memory. At this time, no research considers the impact of trauma on alcohol intoxicated witnesses of crime, compared with alcohol intoxicated victims of crime. This research would be beneficial in further considering the impact of trauma on the memory of those giving evidence in the Criminal Justice System.

Additionally, the need to research the psychometric properties of the AUDIT for use in different populations and within research are presented (Livingstone et al., 2011). The use of appropriate and validated cut-off scores has important implications to maintain the ethical standards of research in which participants are given alcohol. Incorrect cut-off scores may lead to those with AUD's being provided with alcohol, however, cut-off scores that are lower than necessary may present researchers with a group of participants that are less risk averse than others that may score higher on the AUDIT without having an AUD (DeSimone, 2005).

Practical Implications

The findings of this thesis in chapters Two and Three indicate that individuals use strategic memory reporting strategies, suggesting that intoxicated witnesses and victims of crime are no less accurate than those who are sober. The findings showed that while accounts were no less accurate, they were less complete and individuals may have appeared less confident (Bartlett et al., 20121; Flowe et al., 2019; Van Oorsouw et al., 2015). This less complete account and decreased confidence may reinforce the beliefs held by police, jurors,

and other professionals in the Criminal Justice System that intoxicated witnesses and victims are less accurate than those who are sober (Kassin et al., 2001).

When considering police interviews, if the findings of the study presented in Chapter Three can be replicated, this may be important education for police officers. If police officers were able to use the informativeness of uncertainty indicators to consider the accuracy of elements of the information given by witnesses and victims, this may allow for more accurate leads to be followed in investigations. Conversely, a lack of uncertainty indicators may then increase the credibility of the victim or witness in the eyes of the police. It may be particularly informative for police officers if this study could be replicated, but using a much larger sample size which would allow researchers to avoid collapsing over uncertainty indicator type. This may allow researchers to draw important conclusions about the informativeness of particular types of uncertainty indicator.

The findings presented may also be used to educate the wider Criminal Justice System, including jurors. If jurors have an understanding that intoxicated witnesses may be more accurate than once thought, and that less complete accounts are not a sign of inaccuracy, this may help better guide the verdicts of guilt made.

Due to the mixed findings of research into the effects of delayed testing of intoxicated witnesses, caution should be used in the decision to interview intoxicated witnesses, taking into account their level of intoxication, following suggestions that highly intoxicated (>.10% breath alcohol concentrations) may decrease accuracy. While Flowe et al. (2019) found that early questioning was advisable for ensuring accuracy, they state that this may not be the case when leading questions are used.

It has been demonstrated throughout Chapter Two that witnesses and victims of crime can be susceptible to misinformation acceptance. It is because of this that police officers and other professionals in the Criminal Justice System should refrain from using leading

questions that may introduce misinformation to the individual. When considering the increased acceptance of misinformation by witnesses that discuss the event with co-witnesses, the suggestion is made that co-witness discussions should be prevented wherever possible to maintain witness accuracy (Bartlett et al., 2022). Witnesses of crimes often discuss what they have seen with other witnesses (Paterson & Kemp, 2006) and this is often noted to occur before the police respond to the incident (Paterson & Kemp, 2005). Despite this, it may still be beneficial for police officers to separate witnesses upon arrival and instruct them not to discuss what they have witnessed, in order to minimise the negative effects of co-witness discussion. Paterson et al. (2011) suggest that as co-witness discussion is challenging to entirely eliminate, research should focus on improving interview techniques that elicit accurate memories and minimise the reporting of misinformation.

The findings presented in Chapter Three raise further questions regarding the police interview process. Victims were less accurate and expressed more uncertainty during the question phase of the interview. This suggests that caution should be used in questioning victims and witnesses as questioning may elicit information of which the individual is less certain and therefore is less likely to be accurate. Use of the cognitive interview has been seen to increase accuracy and reduce the acceptance of misinformation (Memon et al., 2010)

While it is challenging to provide specific recommendations for policy throughout the Criminal Justice System based on these mixed findings, it is clear that intoxicated eyewitnesses and victims are not as inaccurate as thought by professionals in the system. It is recommended that education for professionals in the system is provided to dispel widely held beliefs about the inaccuracy of intoxicated witnesses and victims. Caution is also urged in the interviewing of victims and witnesses so as not to obtain inaccurate information.

Conclusion

To conclude, the thesis aim of better understanding the impact alcohol on eyewitness memory was achieved through a systematic review of existing literature that suggested that eyewitnesses' acceptance of misinformation *may* in some circumstances be increased by alcohol intoxication, but that the effects of other factors must be disentangled before conclusive assertions are reached (Chapter Two). The accuracy of eyewitnesses and victims was not seen to differ significantly between intoxication states, and uncertainty indicators throughout police interviews were found to be informative of memory accuracy (Chapter Three). The use of the AUDIT in alcohol research was examined and suggestions of further research into its psychometric properties in this area are recommended (Chapter Five). The mixed findings of research presented in this thesis make it challenging to provide specific guidance to aid the police in establishing policies, such as for the time of interviewing intoxicated witnesses. It does, however, clearly demonstrate the requirement for further education to be provided to all professionals in the Criminal Justice System, as well as jurors. If these individuals become more aware of the accuracy of intoxicated victims and witnesses, it is hoped that this may increase the chances of prosecution in the trial of sexual assault cases where victims are often regarded as inaccurate due to their consumption of alcohol.

References

Alexander, K. W., Quas, J. A., Goodman, G. S., Ghetti, S., Edelstein, R. S., Redlich, A. D., ... & Jones, D. P. (2005). Traumatic impact predicts long-term memory for documented child sexual abuse. *Psychological science, 16*(1), 33-40. https://doi.org/10.1111/j.0956-7976.2005.00777.x

Altman, C. M., Schreiber Compo, N., 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

Anderson, Z., Hughes, K., & Bellis, M. A. (2007). Exploration of young people's experience and perceptions of violence in Liverpool's nightlife. *Liverpool: Centre for Public Health, Liverpool John Moores University*.

Assefi, S. L., & Garry, M. (2003). Absolut (R) memory distortions: Alcohol placebos influence the misinformation effect. *Psychological Science, 14*(1), 77-80. https://doi.org/10.1111/1467-9280.01422

Babor, F, Campbell, R., Room, R., & Saunders, J. (1994). *Lexicon of Alcohol and Drug Terms*. Geneva, Switzerland: World Health Organization.

Babor, T. F., de la Fuente, J. R., Saunders, J., & Grant, M. (2001). The Alcohol Use Disorders Identification Test: Guidelines for use in. *Primary care*.

Babor, T. F., Higgins-Biddle, J. C., Saunders, J. B., & Monteiro, M. G. (2001). The alcohol use disorders identification test (AUDIT). *Guidelines for use in primary care*, *2*, 1-17.

Bannigan, K., & Watson, R. (2009). Reliability and validity in a nutshell. *Journal of clinical nursing*, *18*(23), 3237-3243. https://doi.org/10.1111/j.1365-2702.2009.02939.x

Bartlett, G., Albery, I. P., Frings, D., & Gawrylowicz, J. (2022). The effects of alcohol and co-witness information on memory reports: a field study. *Psychopharmacology*, 1-9. https://doi.org/ 10.1007/s00213-022-06179-5

Bartlett, G., Gawrylowicz, J., Frings, D., & Albery, I. P. (2021). The intoxicated co-witness: effects of alcohol and dyadic discussion on memory conformity and event recall. *Psychopharmacology*, *238*(6), 1485-1493. https://doi.org/ 10.1007/s00213-021-05776-0

Bell, B. E., & Loftus, E. F. (1989). Trivial persuasion in the courtroom: The power of (a few) minor details. *Journal of personality and social psychology, 56*(5), 669. https://doi.org/10.1037/0022-3514.56.5.669

Benton, T. R., McDonnell, S., Ross, D. F., Thomas, W. N. III, & Bradshaw, E. (2007). Has eyewitness research penetrated the American legal system? A synthesis of case history, juror knowledge, and expert testimony. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *The handbook of eyewitness psychology, Vol. 2. Memory for people* (p. 453–500). Lawrence Erlbaum Associates Publishers.

Bischof, G., Reinhardt, S., Grothues, J., Dybek, I., Meyer, C., Hapke, U., John, U., & Rumpf, H. J. (2005). Effects of item sequence on the performance of the AUDIT in general practices. *Drug and Alcohol dependence*, *79*(3), 373-377. https://doi.org/10.1016/j.drugalcdep.2005.03.002

Birnbaum, I. M., Parker, E. S., Hartley, J. E., & Noble, E. P. (1978). Alcohol and memory: Retrieval processes. *Journal of Verbal Learning and Verbal Behavior, 17,* 325–335. doi:10.1016/S0022- 5371(78)90210-4

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/10.1016/j.jarmac.2014.03.005

Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic approaches to a successful literature review.* Sage.

Brewer, N., & Burke, A. (2002). Effects of testimonial inconsistencies and eyewitness confidence on mock-juror judgments. *Law and Human Behavior, 26(3), 353-364.* https://doi.org/10.1023/A:1015380522722

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

Brownmiller, S. (2005). Against Our Will: Men, Women and Rape (1975). In R. K. Bergen, J. L. Edleson, & C. M. Renzetti, *Violence against women: Classic papers* (pp. 5–8). Pearson Education New Zealand.

Butt, M. M., Colloff, M. F., Magner, E., & Flowe, H. D. (2022). Eyewitness memory in the news can affect the strategic regulation of memory reporting. *Memory, 30*(6), 763-774. https://doi.org/10.1080/09658211.2020.1846750

Cassidy, C. M., Schmitz, N., & Malla, A. (2008). Validation of the alcohol use disorders identification test and the drug abuse screening test in first episode psychosis. *The Canadian Journal of Psychiatry*, *53*(1), 26-33. https://doi.org/10.1177/070674370805300105

Cherpitel, C. J. (1995). Screening for alcohol problems in the emergency room: a rapid alcohol problems screen. *Drug and alcohol dependence*, *40*(2), 133-137. https://doi.org/10.1016/0376-8716(95)01199-4

Cherpitel, C. J. (1997). Comparison of screening instruments for alcohol problems between black and white emergency room patients from two regions of the

country. *Alcoholism: Clinical and Experimental Research*, *21*(8), 1391-1397. https://doi.org/ 10.1111/j.1530-0277.1997.tb04467.x

Christianson, S.-A. (2014). *The Handbook of Emotion and Memory: Research and Theory*. New York, NY: Psychology Press https://doi.org/10.4324/9781315807454

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Colloff, M. F., & Flowe, H. D. (2016). The effects of acute alcohol intoxication on the cognitive mechanisms underlying false facial recognition. *Psychopharmacology*, *233*(11), 2139-2149. https://doi.org/10.1007/s00213-016-4263-4

Colloff, M. F., Wade, K. A., Wixted, J. T., & Maylor, E. A. (2017). A signal-detection analysis of eyewitness identification across the adult lifespan. *Psychology and Aging*, *32*(3), 243. https://doi.org/10.1037/pag0000168

Cook, R. L., Chung, T., Kelly, T. M., & Clark, D. B. (2005). Alcohol screening in young persons attending a sexually transmitted disease clinic: Comparison of AUDIT, CRAFFT, and CAGE instruments. *Journal of general internal medicine*, *20*(1), 1-6. https://doi.org/ 10.1111/j.1525-1497.2005.40052.x

Critical Appraisal Skills Programme. (2018). CASP cohort study checklist. https://casp-uk.net/casp-tools-checklists/

Crossland, D., Kneller, W., & Wilcock, R. (2016). Intoxicated witnesses: Testing the validity of the alcohol myopia theory. *Applied Cognitive Psychology*, *30*, 270–281. https://doi.org/10.1002/acp.3209

Crossland, D., Kneller, W., & Wilcock, R. (2018). Intoxicated eyewitnesses: prevalence and procedures according to England's police officers. *Psychology, Crime & Law*, *24*(10), 979-997. https://doi.org/10.1080/1068316X.2018.1474216

Crossland, D., Kneller, W., & Wilcock, R. (2021). Mock Juror Perceptions of Intoxicated Eyewitness Credibility. Journal of Police and Criminal Psychology, 1-12. https://doi.org/10.1007/s11896-021-09430-5

Crown Prosecution Service (2018). The Code for Crown Prosecutors https://www.cps.gov.uk/sites/default/files/documents/publications/Code-for-Crown-Prosecutors-October-2018.pdf

Curtin, J. J., & Fairchild, B. A. (2003). Alcohol and cognitive control: Implications for regulation of behavior during response conflict. *Journal of abnormal psychology*, *112*(3), 424. https://doi.org/10.1037/0021-843X.112.3.424

Cutler, B. L., & Penrod, S. D. (1989). Forensically relevant moderators of the relation between eyewitness identification accuracy and confidence. *Journal of Applied Psychology*, *74*(4), 650. https://doi.org/10.1037/0021-9010.74.4.650

Daeppen, J. B., Yersin, B., Landry, U., Pécoud, A., & Decrey, H. (2000). Reliability and validity of the Alcohol Use Disorders Identification Test (AUDIT) imbedded within a general health risk screening questionnaire: results of a survey in 332 primary care patients. *Alcoholism: Clinical and Experimental Research*, *24*(5), 659-665. https://doi.org/ 10.1111/j.1530-0277.2000.tb02037.x

Daly, K., & Bouhours, B. (2010). Rape and attrition in the legal process: A comparative analysis of five countries. *Crime and justice*, *39*(1), 565-650. https://doi.org/ 10.1086/653101

Dando, C., Wilcock, R., & Milne, R. (2009). The cognitive interview: The efficacy of a modified mental reinstatement of context procedure for frontline police investigators. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *23*(1), 138-147. https://doi.org/10.1002/acp.1451

de Meneses-Gaya, C., Zuardi, A. W., Loureiro, S. R., & Crippa, J. A. S. (2009). Alcohol Use Disorders Identification Test (AUDIT): An updated systematic review of psychometric properties. *Psychology & Neuroscience*, *2*(1), 83. https://doi.org/10.1590/S1983-32882009000100012

DeSimone, J. (2005). Drinking and academic performance in high school. *Applied Economics*, *42*(12), 1481-1497. https://doi.org/ 10.1080/00036840701721554

DeVellis, R. (2003). Scale development: theory and applications *Lomdon: SAGE, 2003*

DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., Savoy, S. M., & Kostas-Polston, E. (2007). A psychometric toolbox for testing validity and reliability. *Journal of Nursing scholarship*, *39*(2), 155-164. https://doi.org/ 10.1111/j.1547-5069.2007.00161.x

Douglass, A. B., Neuschatz, J. S., Imrich, J., & Wilkinson, M. (2010). Does post-identification feedback affect evaluations of eyewitness testimony and identification procedures?. *Law and Human Behavior*, *34*(4), 282-294. https://doi.org/10.1007/s10979-009-9189-5

Dybek, I., Bischof, G., Grothues, J., Reinhardt, S., Meyer, C., Hapke, U., John, U., Broocks, A., Hohagen, F., & Rumpf, H. J. (2006). The reliability and validity of the Alcohol Use Disorders Identification Test (AUDIT) in a German general practice population sample. *Journal of studies on alcohol*, *67*(3), 473-481. https://doi.org/10.15288/jsa.2006.67.473

Eisen, M. L., Gabbert, F., Ying, R., & Williams, J. (2017). “I Think He Had A Tattoo On His Neck”: How Co-Witness Discussions About A Perpetrator's Description Can Affect Eyewitness Identification Decisions. *Journal of Applied Research in Memory and Cognition*, *6*(3), 274-282. https://doi.org/0.1016/j.jarmac.2017.01.009

Ellison, L., & Munro, V. E. (2017). Taking trauma seriously: Critical reflections on the criminal justice process. *The International Journal of Evidence & Proof*, *21*(3), 183-208. https://doi.org/10.1177/1365712716655168

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

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

Evans, J. R., Schreiber Compo, N., & Russano, M. B. (2009). Intoxicated witnesses and suspects: Procedures and prevalence according to law enforcement. *Psychology, Public Policy, and Law*, *15*(3), 194. https://doi.org/10.1037/a0016837

Evans, J. R., Schreiber Compo, N., Carol, R. N., Nichols-Lopez, K., Holness, H., & Furton, K. G. (2018). The impact of alcohol intoxication on witness suggestibility immediately and after a delay. *Applied Cognitive Psychology.* https://doi.org/10.1002/acp.3502

Evans, J. R., Schreiber Compo, N., Carol, R. N., Schwartz, B. L., Holness, H., Rose, S., & Furton, K. G. (2017). Alcohol intoxication and metamemory: Little evidence that moderate intoxication impairs metacognitive monitoring processes. *Applied Cognitive Psychology*, *31*(6), 573-585. https://doi.org/10.1002/acp.3373

Fast, D. K., Conry, J., & Loock, C. A. (1999). Identifying fetal alcohol syndrome among youth in the Criminal Justice System. *Journal of Developmental and Behavioral Pediatrics*, *20*(5), 370-372. https://doi.org/10.1097/00004703-199910000-00012

Ferketich, S. (1990). Focus on psychometrics: Internal consistency estimates of reliability. *Research in nursing & health*, *13*(6), 437-440.

Fillmore, M. T., & Blackburn, J. (2002). Compensating for alcohol-induced impairment: alcohol expectancies and behavioral disinhibition. *Journal of studies on alcohol*, *63*(2), 237-246. https://doi.org/10.15288/jsa.2002.63.237

Fisher, R. P., & Schreiber, N. (2007). Interviewing protocols to improve eyewitness memory. In M. Toglia, R. Lindsay, D. Ross, & J. Reed (Eds.), *The handbook of eyewitness psychology: Vol. 1. Memory for events* (pp. 53–80). Mahwah, NJ: Erlbaum Associates.

Flowe, H. D., & Compo, N. S. (2021). The lack of robust evidence for the effects of alcohol on false memory. *Neuroscience & Biobehavioral Reviews*, *127*, 332-333. https://doi.org/ 0.1016/j.neubiorev.2021.04.029

Flowe, H. D., Colloff, M. F., Karoğlu, N., Zelek, K., Ryder, H., Humphries, J. E., & Takarangi, M. K. (2017). The effects of alcohol intoxication on accuracy and the confidence–accuracy relationship in photographic simultaneous line-ups. *Applied Cognitive Psychology*, *31*(4), 379-391. https://doi.org/10.1002/acp.3332

Flowe, H. D., Ebbesen, E. B., & Putcha-Bhagavatula, A. (2007). Rape shield laws and sexual behavior evidence: effects of consent level and women's sexual history on rape allegations. *Law and human behavior*, *31*(2), 159. https://doi.org/10.1007/s10979-006-9050-z

Flowe, H. D., Humphries, J. E., Takarangi, M. K., Zelek, K., Karoğlu, N., Gabbert, F., & Hope, L. (2019). An experimental examination of the effects of alcohol consumption and exposure to misleading postevent information on remembering a hypothetical rape scenario. *Applied cognitive psychology*, *33*(3), 393-413. https://doi.org/ 10.1002/acp.3531

Flowe, H. D., Stewart, J., Sleath, E. R., & Palmer, F. T. (2011). Public house patrons' engagement in hypothetical sexual assault: A test of Alcohol Myopia Theory in a field setting. *Aggressive Behavior*, *37*(6), 547-558. https://doi.org/10.1002/ab.20410

Flowe, H. D., Takarangi, M. K., Humphries, J. E., & Wright, D. S. (2016). Alcohol and remembering a hypothetical sexual assault: Can people who were under the influence of alcohol during the event provide accurate testimony?. *Memory*, *24*(8), 1042-1061. https://doi.org/10.1080/09658211.2015.1064536

Foster, J. L., Huthwaite, T., Yesberg, J. A., Garry, M., & Loftus, E. F. (2012). Repetition, not number of sources, increases both susceptibility to misinformation and confidence in the accuracy of eyewitnesses. *Acta psychologica*, *139*(2), 320-326. https://doi.org/10.1016/j.actpsy.2011.12.004

Fox, J. G., & Sobol, J. J. (2000). Drinking patterns, social interaction, and barroom behavior: A routine activities approach. *Deviant Behavior*, *21*(5), 429-450. https://doi.org/10.1080/01639620050085834

Garrett, B. L. (2009). DNA and Due Process. *Fordham L. Rev.*, *78*, 2919.

Gawrylowicz, J., Ridley, A. M., Albery, I. P., Barnoth, E., & Young, J. (2017). Alcohol-induced retrograde facilitation renders witnesses of crime less suggestible to misinformation. *Psychopharmacology, 234*(8), 1267-1275. http://dx.doi.org/10.1007/s00213-017-4564-2

Gawrylowicz, J., Scoboria, A., Teodorini, R., & Albery, I. P. (2019). Intoxicated eyewitnesses: the effect of a fully balanced placebo design on event memory and metacognitive control. *Applied Cognitive Psychology*, *33*(3), 344-357. https://doi.org/10.1002/acp.3504

Gisev, N., Bell, J. S., & Chen, T. F. (2013). Interrater agreement and interrater reliability: key concepts, approaches, and applications. Research in Social and Administrative Pharmacy, 9(3), 330-338. https://doi.org/10.1016/j.sapharm.2012.04.004

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

Graham, K., & Wells, S. (2001). Aggression among young adults in the social context of the bar. *Addiction Research & Theory*, *9*(3), 193-219. https://doi.org/10.3109/16066350109141750

Guilford, J. (1956). Fundamental Statistics in Psychology and Education, *(3rd ed.). McGraw-Hill.*

Hagsand, A., Hjelmsäter, E. R. A., Granhag, P. A., Fahlke, C., & Söderpalm-Gordh, A. (2013). Bottled memories: On how alcohol affects eyewitness recall. *Scandinavian Journal of Psychology*, *54*(3), 188-195. https://doi.org/10.1111/sjop.12035

Hamilton, G., & Tidmarsh, P. (2020). Misconceptions of sexual crimes against adult victims: Barriers to justice. *Trends and Issues in Crime and Criminal Justice [electronic resource]*, (611), 1-18. https://doi.org/10.52922/ti04824

Hammond, S. (2006). Using psychometric tests. *Research methods in psychology*, *3*, 182-209.

Harvey, A. J., Kneller, W., & Campbell, A. C. (2013). The elusive effects of alcohol intoxication on visual attention and eyewitness memory. *Applied Cognitive Psychology*, *27*(5), 617-624. https://doi.org/10.1002/acp.2940

Herman, J. L. (1992). *Trauma and Recovery.* New York: NY: Basic Books

Hill, S., & Marshall, T. C. (2018). Beliefs about sexual assault in India and Britain are explained by attitudes toward women and hostile sexism. *Sex Roles*, *79*, 421-430. https://doi.org/10.1007/s11199-017-0880-6

Holliday, R. E. (2003). The effect of a prior cognitive interview on children's acceptance of misinformation. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *17*(4), 443-457. https://doi.org/10.1002/acp.879

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

Houston, K. A., Hope, L., Memon, A., & Don Read, J. (2013). Expert testimony on eyewitness evidence: In search of common sense. *Behavioral Sciences & the Law*, *31*(5), 637-651. https://doi.org/10.1002/bsl.2080

Hudson, J. A., & Fivush, R. (1991). As time goes by: Sixth graders remember a kindergarten experience. *Applied Cognitive Psychology*, *5*(4), 347-360. https://doi.org/10.1002/acp.2350050405

Huntley, M. S. Jr. (1974). Effects of alcohol, uncertainty and novelty upon response selection. *Psychopharmacologia, 39*(3), 259–266.

Joppe, G. (2000). Testing reliability and validity of research instruments. *Journal of American Academy of Business Cambridge*, *4*(1/2), 49-54.

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

Karno, M., Granholm, E., & Lin, A. (2000). Factor structure of the Alcohol Use Disorders Identification Test (AUDIT) in a mental health clinic sample. *Journal of studies on alcohol*, *61*(5), 751-758.

Kassin, S. M., Ellsworth, P. C., & Smith, V. L. (1989). The" general acceptance" of psychological research on eyewitness testimony: A survey of the experts. *American Psychologist*, *44*(8), 1089. https://doi.org/10.1037/0003-066X.44.8.1089

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

Kelemen, W. L. (2000). Metamemory cues and monitoring accuracy: Judging what you know and what you will know. *Journal of Educational Psychology*, *92*(4), 800. https://doi.org/10.1037/0022-0663.92.4.800

Kline, P. (1986). A Handbook of Test Construction: Introduction to Psychometric Design. New York: Methuen.

Kline, T. J. (1999). The team player inventory: Reliability and validity of a measure of predisposition toward organizational team-working environments. *Journal for specialists in Group Work*, *24*(1), 102-112. https://doi.org/ 10.1080/01933929908411422

Kloft, L., Monds, L. A., Blokland, A., Ramaekers, J. G., & Otgaar, H. (2021). Hazy memories in the courtroom: a review of alcohol and other drug effects on false memory and suggestibility. *Neuroscience & Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2021.02.012

Kneller, W., & Harvey, A. J. (2016). Lineup identification accuracy: The effects of alcohol, target presence, confidence ratings, and response time. *The European Journal of*

Psychology Applied to Legal Context, *8*(1), 11-18. https://doi.org/10.1016/j.ejpal.2015.09.001

Knight, J. R., Sherritt, L., Harris, S. K., Gates, E. C., & Chang, G. (2003). Validity of brief alcohol screening tests among adolescents: a comparison of the AUDIT, POSIT, CAGE, and CRAFFT. *Alcoholism: Clinical and experimental research*, *27*(1), 67-73. https://doi.org/10.1111/j.1530-0277.2003.tb02723.x

Koriat, A., & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General*, *123*(3), 297. https://doi.org/10.1037/0096-3445.123.3.297

Koriat, A., Goldsmith, M., Schneider, W., & Nakash-Dura, M. (2001). The credibility of children's testimony: Can children control the accuracy of their memory reports?. *Journal of Experimental Child Psychology*, *79*(4), 405-437. https://doi.org/10.1006/jecp.2000.2612

Krahmer, E., & Swerts, M. (2005). How children and adults produce and perceive uncertainty in audiovisual speech. *Language and speech*, *48*(1), 29-53. https://doi.org/10.1177/00238309050480010201

Lampinen, J. M., & Smith, V. L. (1995). The incredible (and sometimes incredulous) child witness: Child eyewitnesses' sensitivity to source credibility cues. Journal of Applied Psychology, 80, 621–627. https://doi.org/10.1037/0021-9010.80.5.621

Laney, C., & Loftus, E. F. (2005). Traumatic memories are not necessarily accurate memories. *The Canadian Journal of Psychiatry, 50*(13), 823-828. https://doi.org/10.1177/070674370505001303

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & Van Knippenberg, A. D. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and emotion*, *24*(8), 1377-1388. https://doi.org/10.1080/02699930903485076

LaPaglia, J. A., & Chan, J. C. (2013). Testing increases suggestibility for narrative-based misinformation but reduces suggestibility for question-based misinformation. *Behavioral sciences & the law*, *31*(5), 593-606. https://doi.org/10.1002/bsl.2090

Laserna, C. M., Seih, Y. T., & Pennebaker, J. W. (2014). Um... who like says you know: Filler word use as a function of age, gender, and personality. Journal of Language and Social Psychology, 33(3), 328-338. https://doi.org/10.1177/0261927X1452699

Lee, J., & Penrod, S. D. (2019). New signal detection theory-based framework for eyewitness performance in lineups. *Law and human behaviour*, *43*(5), 436. https://doi.org/10.1037/lhb0000343

Lees, S. (1993). Judicial rape. In *Women's studies international forum*. Vol. 16, No. 1, pp. 11-36. Pergamon. https://doi.org/10.1016/0277-5395(93)90077-M

Leippe, M. R. (1980). Effects of integrative memorial and cognitive processes on the correspondence of eyewitness accuracy and confidence. *Law and Human behaviour*, *4*(4), 261-274. https://doi.org/10.1007/BF01040618

Lennings, C. J. (1999). An evaluation of the Leeds dependence questionnaire. *Journal of Child & Adolescent Substance Abuse*, *8*(3), 73-87. https://doi.org/10.1300/J029v08n03_05

Leonard, K. E., Quigley, B. M., & Collins, R. L. (2002). Physical aggression in the lives of young adults: Prevalence, location, and severity among college and community samples. *Journal of interpersonal violence*, *17*(5), 533-550. https://doi.org/10.1177/0886260502017005004

Lima, C. T., Freire, A. C. C., Silva, A. P. B., Teixeira, R. M., Farrell, M., & Prince, M. (2005). Concurrent and construct validity of the AUDIT in an urban Brazilian sample. *Alcohol and Alcoholism*, *40*(6), 584-589. https://doi.org/10.1093/alcalc/agh202

Lindholm, T., Jönsson, F. U., & Liuzza, M. T. (2018). Retrieval effort cues predict eyewitness accuracy. *Journal of Experimental Psychology: Applied*, *24*(4), 534. https://doi.org/10.1037/XAP0000175

Lindsay, R.C.L. 1994. "Expectations of eyewitness performance: Jurors verdicts do not follow their beliefs". In *Adult Eyewitness Testimony*, Edited by: Ross, D., Read, J. and Toglia, M. 362–384. Cambridge: Cambridge University Press.

Lindsay, D. S. (2008). Source monitoring and eyewitness memory. In B. L. Cutler (Ed.), *Encyclopedia of psychology and law* (ed., Vol. 1) (pp. 749–750). Thousand Oaks, CA: SAGE Publications, Inc. https://doi.org/10.4135/9781412959537.n291

Livingstone, A. G., Young, H., & Manstead, A. S. (2011). "We Drink, Therefore We Are" The role of group identification and norms in sustaining and challenging heavy drinking "Culture". *Group Processes & Intergroup Relations*, *14*(5), 637-649. https://doi.org/10.1177/1368430210392399

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & memory*, *12*(4), 361-366. https://doi.org/10.1101/lm.94705

Lundrigan, S., Dhami, M. K., & Agudelo, K. (2019). Factors predicting conviction in stranger rape cases. *Frontiers in psychology*, 526. https://doi.org/10.3389/fpsyg.2019.00526

Maisto, S. A., Carey, M. P., Carey, K. B., Gordon, C. M., & Gleason, J. R. (2000). Use of the AUDIT and the DAST-10 to identify alcohol and drug use disorders among adults with a severe and persistent mental illness. *Psychological assessment*, *12*(2), 186. https://doi.org/10.1037/1040-3590.12.2.186

Marinkovic, K., Halgren, E., & Maltzman, I. (2004). Effects of alcohol on verbal processing: An event-related potential study. *Alcoholism, Clinical and Experimental Research, 28*, 415–423. https://doi.org/10.1097/01.ALC.0000117828.88597.80.

McGuire, W. J. (1985). Attitudes and attitude change. *The handbook of social psychology*, 233-346.

McMillan, L., & Thomas, M. (2009). Police interviews of rape victims: Tensions and contradictions. *Rape: Challenging contemporary thinking*, 255-280.

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

Metcalfe, J. (2000). Metamemory: Theory and data. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 197–211). Oxford University Press.

Mickes, L., Seale-Carlisle, T. M., & Wixted, J. T. (2013). Rethinking familiarity: Remember/Know judgments in free recall. *Journal of memory and language*, *68*(4), 333-349. https://doi.org/10.1016/j.jml.2013.01.001

Miller, M. B., Merrill, J. E., DiBello, A. M., & Carey, K. B. (2018). Distinctions in alcohol-induced memory impairment: A mixed methods study of en bloc versus fragmentary blackouts. *Alcoholism: Clinical and Experimental Research*, *42*(10), 2000-2010. https://doi.org/10.1111/acer.13850

Mindthoff, A., Evans, J. R., Compo, N. S., Polanco, K., & Hagsand, A. V. (2021). No evidence that low levels of intoxication at both encoding and retrieval impact scores

on the Gudjonsson Suggestibility Scale. *Psychopharmacology*, 1-12. https://doi.org/10.1007/s00213-021-05797-9

Ministry of Justice (2013). An Overview of Sexual Offending in England and Wales retrieved 17th June 2020 from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/214970/sexual-offending-overview-jan-2013.pdf

Ministry of Justice (2011). Achieving Best Evidence in Criminal Proceedings. Retrieved 17th June 2020 from https://www.cps.gov.uk/sites/default/files/documents/legal_guidance/best_evidence_in_criminal_proceedings.pdf

Ministry of Justice, (2015). Witnessing Crime – Findings from the Crime Survey for England and Wales 2013/14. Retrieved 21st April 2020 from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/449584/witnessing-crime-findings-from-crime-survey.pdf

Mintzer, M. Z. (2007). The acute effects of alcohol on memory: A review of laboratory studies in healthy adults. *International Journal on Disability and Human Development*, 6, 397–403. https://doi.org/10.1515/IJDHD.2007.6.4.397

Monds, L. A., Cullen, H. J., Kloft, L., van Golde, C., Harrison, A. W., & Flowe, H. (2021). Memory and credibility perceptions of alcohol and other drug intoxicated witnesses and victims of crime. *Psychology, Crime & Law*, *28*(8), 820-840. https://doi.org/10.1080/1068316X.2021.1962871

Monds, L. A., Kloft, L., Sauer, J. D., Honan, C. A., & Palmer, M. A. (2019). No evidence that alcohol intoxication impairs judgments of learning in face recognition. *Applied Cognitive Psychology*, *33*(3), 325-333. https://doi.org/10.1002/acp.3534

Mook, D. G. (1983). In defense of external invalidity. *American psychologist*, *38*(4), 379. https://doi.org/10.1037/0003-066X.38.4.379

Mumenthaler, M. S., Taylor, J. L., O'Hara, R., & Yesavage, J. A. (1999). Gender differences in moderate drinking effects. *Alcohol Research & Health*, *23*(1), 55.

Nadkarni, A., Garber, A., Costa, S., Wood, S., Kumar, S., MacKinnon, N., Ibrahim, M., Velleman, R., Bhatia, U., Godwin, F., Rane, A., & Weobong, B. (2019). Auditing the AUDIT: A systematic review of cut-off scores for the Alcohol Use Disorders Identification Test (AUDIT) in low-and middle-income countries. *Drug and alcohol dependence*, *202*, 123-133. https://doi.org/10.1016/j.drugalcdep.2019.04.031

Nash, R. A., & Takarangi, M. (2011). Reconstructing alcohol-induced memory blackouts. *Memory*, 19, 566–573. https://doi.org/10.1080/09658211.2011.590508

Neisser, U. (1998). Stories, Selves, and Schemata: A Review of Ecological. *Theories of memory*, *2*, 171.

Neisser, U., & Hyman, I.E., Jr. (Eds.). (2000). *Memory observed* (2nd ed.). New York: Worth.

Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory 3rd edn. *New York [etc].: McGraw-Hill*.

O'Hare, T., & Sherrer, M. V. (1999). Validating the alcohol use disorder identification test with college first-offenders. *Journal of Substance Abuse Treatment*, *17*(1-2), 113-119. https://doi.org/ 10.1016/S0740-5472(98)00063-4

Office for National Statistics. (2018, December 3). Data on alcohol related incidents, years ending March 2011 to March 2017, Crime Survey for England and Wales. https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/adhocs/009372dataonalcoholrelatedincidentsyearsendingmarch2011tomarch2017crimesurveyforenglandandwales

Office of National Statistics. (2016, February 11). Crime and justice: Focus on violent crime and sexual offences, year ending March 2015. https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/compendium/focusonviolentcrimeandsexualoffences/yearendingmarch2015

Office of National Statistics. (2023, January 23). Crime in England and Wales: year ending September 2022. https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/crimeinenglandandwales/yearendingseptember2022#domestic-abuse-and-sexual-offences

Pae, C. U. (2015). Why systematic review rather than narrative review?. *Psychiatry investigation*, *12*(3), 417. https://doi.org/ 10.4306/pi.2015.12.3.417

Pallant, J. (2016). Spss survival manual (6 uppl.). *Maidenhead: McGraw-Hill Education*.

Palmer, F. T., Flowe, H. D., Takarangi, M. K., & 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. https://doi.org/10.1037/lhb0000010

Pannu, J. K., & Kaszniak, A. W. (2005). Metamemory experiments in neurological populations: A review. *Neuropsychology review*, *15*(3), 105-130. https://doi.org/10.1007/s11065-005-7091-6

Pansky, A., Koriat, A., & Goldsmith, M. (2005). Eyewitness Recall and Testimony. In N. Brewer & K. D. Williams (Eds.), *Psychology and law: An empirical perspective* (pp. 93–150). The Guilford Press.

Paterson, H., & Kemp, R. (2006). Co-witnesses talk: A survey of eyewitness discussion. *Psychology, Crime and Law*, 12, 181–191. https://doi.org/10.1080/10683160512331316334

Paterson, H. M., & Kemp, R. I. (2005). Co-witness discussion: A survey of police officers' attitudes, knowledge, and behaviour. *Psychiatry, Psychology and Law, 12*(2), 424-434. https://doi.org/10.1375/pplt.12.2.424

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

Pattavina, A., Morabito, M., & Williams, L. M. (2016). Examining connections between the police and prosecution in sexual assault case processing: Does the use of exceptional clearance facilitate a downstream orientation?. *Victims & Offenders*, *11*(2), 315-334. https://doi.org/10.1080/15564886.2015.1046622

Peace, K. A., Porter, S., & Brinke, L. T. (2008). Are memories for sexually traumatic events "special"? A within-subjects investigation of trauma and memory in a clinical sample. *Memory, 16*(1), 10-21. https://doi.org/10.1080/09658210701363583

Philpot, M., Pearson, N., Petratou, V., Dayanandan, R., Silverman, M., & Marshall, J. (2003). Screening for problem drinking in older people referred to a mental health service: a comparison of CAGE and AUDIT. *Aging & Mental Health*, *7*(3), 171-175. https://doi.org/ 10.1080/1360786031000101120

Polit, D. F., & Beck, C. T. (2004). *Nursing research: Principles and methods*. Lippincott Williams & Wilkins.

Poole, D. A., & White, L. T. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology*, *27*(6), 975. ttps://doi.org/10.1037/0012-1649.27.6.975

Rape Crisis England and Wales. (2023). Statistics about rape, sexual assault and sexual abuse. https://rapecrisis.org.uk/get-informed/statistics-sexual-

violence/#:~:text=Only%201%20in%20100%20rapes,Let%20alone%20a%20convicti on.

Ray, S., & Bates, M. E. (2006). Acute alcohol effects on repetition priming and word recognition memory with equivalent memory cues. *Brain and Cognition, 60*(2), 118-127. https://doi.org/10.1016/j.bandc.2005.07.009

Reinert, D. F., & Allen, J. P. (2002). The alcohol use disorders identification test (AUDIT): a review of recent research. *Alcoholism: Clinical and Experimental Research*, *26*(2), 272-279. https://doi.org/10.1111/j.1530-0277.2002.tb02534.x

Reinert, D. F., & Allen, J. P. (2007). The alcohol use disorders identification test: an update of research findings. *Alcoholism: Clinical and Experimental Research*, *31*(2), 185-199. https://doi.org/10.1111/j.1530-0277.2006.00295.x

Reyna, V. F., & Kiernan, B. (1994). Development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval. *Developmental Psychology*, *30*(2), 178. https://doi.org/10.1037/0012-1649.30.2.178

Richardson, W. S., Wilson, M. C., Nishikawa, J., & Hayward, R. S. A. (1995). The well-built clinical question: A key to evidence-based decisions. *ACP Journal Club, 123,* A12-13.

Ridley AM, Gudjonsson GH (2013) Suggestibility and individual differences: psychosocial and memory measures. In: Ridley AM, Gabbert F, La Rooy DJ (eds) *Suggestibility in legal contexts: psychological research and forensic implications*. Wiley-Blackwell, Chichester, pp 85–106.

Rivera-Vélez, G. M., González-Viruet, M., Martínez-Taboas, A., & Pérez-Mojica, D. (2014). Post-traumatic stress disorder, dissociation, and neuropsychological performance in

Latina victims of childhood sexual abuse. *Journal of Child Sexual Abuse, 23*(1), 55-73. https://doi.org/10.1080/10538712.2014.864746

Roberts, W. T., & Higham, P. A. (2002). Selecting accurate statements from the cognitive interview using confidence ratings. *Journal of Experimental Psychology: Applied*, *8*(1), 33. https://doi.org/10.1037/1076-898X.8.1.33

Rose, M. E., & Grant, J. E. (2010). Alcohol-induced blackout: phenomenology, biological basis, and gender differences. *Journal of addiction medicine*, *4*(2), 61-73. https://doi.org/10.1097/ADM.0b013e3181e1299d

Rubio Valladolid, G., Bermejo Vicedo, J., Caballero Sánchez-Serrano, M. C., & Santo-Domingo Carrasco, J. (1998). Validación de la prueba para la identificación de trastornos por uso de alcohol (AUDIT) en atención primaria. *Revista Clínica Española*, *198*(1), 11-14.

Santis, R., Garmendia, M. L., Acuña, G., Alvarado, M. E., & Arteaga, O. (2009). The Alcohol Use Disorders Identification Test (AUDIT) as a screening instrument for adolescents. *Drug and Alcohol Dependence*, *103*(3), 155-158. https://doi.org/ 10.1016/j.drugalcdep.2009.01.017

Saraiva, B. R. (2019). The role of metamemory in eyewitness testimony performance. *Portsmouth Research Portal*

Sarmah, H. K., & Hazarika, B. B. (2012). Determination of Reliability and Validity measures of a questionnaire. *Indian Journal of Education and information management, 1*(11), 508-517.

Saunders, J. B., Aasland, O. G., Amundsen, A., & Grant, M. (1993a). Alcohol consumption and related problems among primary health care patients: WHO collaborative project on early detection of persons with harmful alcohol consumption—I. *Addiction*, *88*(3), 349-362. https://doi.org/ 10.1111/j.1360-0443.1993.tb00822.x

Saunders, J. B., Aasland, O. G., Babor, T. F., De la Fuente, J. R., & Grant, M. (1993b). Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction*, *88*(6), 791-804. https://doi.org/ 10.1111/j.1360-0443.1993.tb02093.x

Schneider, W., Dumais, S. T., & Shiffrin, R. M. (1984). Automatic and control processing and attention. In R. Parasuraman, R. Davies, & J. Beatty (Eds.), Varieties of attention (pp. 1–27). New York: Academic Press.

Schooler, J. W., & Loftus, E. F. (1986). Individual differences and experimentation: Complementary approaches to interrogative suggestibility. *Social Behaviour, 1*, 105–112.

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. 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

Schreiber Compo, N., Evans, J. R., Carol, R. N., Villalba, D., Ham, L. S., Garcia, T., & Rose, S. (2012). Intoxicated Eyewitnesses: Better than Their Reputation? *Law & Human Behavior, 36*(2), 77-86. https://doi.org/ 10.1037/h0093951

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

Schuller, R. A., & Stewart, A. (2000). Police responses to sexual assault complaints: The role of perpetrator/complainant intoxication. *Law and human behavior*, *24*(5), 535-551. https://doi.org/10.1023/A:1005519028528

Scottish Crime and Justice Agency (2021). Scottish Crime and Justice Survey 2019/20: main findings. https://www.gov.scot/publications/scottish-crime-justice-survey-2019-20-main-findings/

Seals, D. R., & Coppock, M. E. (2022). We, um, have, like, a problem: excessive use of fillers in scientific speech. Advances in Physiology Education, 46(4), 615-620. https://doi.org/10.1152/advan.00110.2022

Selin, K. H. (2003). Test-retest reliability of the alcohol use disorder identification test in a general population sample. *Alcoholism: Clinical and experimental research*, *27*(9), 1428-1435. https://doi.org/ 10.1097/01.alc.0000085633.23230.4a

Sharma, D., Albery, I. P., & Cook, C. (2001). Selective attentional bias to alcohol related stimuli in problem drinkers and non-problem drinkers. *Addiction*, *96*(2), 285-295. https://doi.org/10.1046/j.1360-0443.2001.96228512.x

Sinclair, M., McRee, B., & Babor, T. F. (1992). Evaluation of the Reliability of AUDIT. *University of Connecticut School of Medicine, Alcohol Research Center.*

Sireci, S. G., Wainer, H., & Braun, H. (2005). Psychometrics, overview. *Encyclopedia of biostatistics, 6*. https://doi.org/10.1002/0470011815.b2a10075

Smeets, T., Candel, I., & Merckelbach, H. (2004). Accuracy, completeness, and consistency of emotional memories. *The American journal of psychology*, 595-609. https://doi.org/10.2307/4148994

Smith, V. L., Kassin, S. M., & Ellsworth, P. C. (1989). Eyewitness accuracy and confidence: Within- versus between-subjects correlations. *The Journal of Applied Psychology*, *74*, 356–359. https://doi.org/10.1037/0021-9010.74.2.356

Soraci, S. A., Carlin, M. T., Read, J. D., Wakeford, Y., Pogoda, T., Cavanaugh, S., & Shin, L. (2007). Psychological impairment, eyewitness testimony, and false memories: Individual differences. In M. P. Toglia, J. D. Read, D. F. Ross, & R. C. L. Lindsay (Eds.), *Handbook of eyewitness psychology: Vol. II: Memory for events* (pp. 261–297). Mahwah, NJ: Erlbaum Associates.

Spearman, C. (1907). Demonstration of formulae for true measurement of correlation. *The American Journal of Psychology*, 161-169. https://doi.org/10.2307/1412408

Spiegel, D. (1997). Trauma, dissociation, and memory. In R. Yehuda & A. C. McFarlane (Eds.), *Psychobiology of posttraumatic stress disorder* (pp. 225–237). New York Academy of Sciences.

Steele, C. M., & Josephs, R. A. (1990). Alcohol myopia: Its prized and dangerous effects. *American psychologist*, *45*(8), 921. https://doi.org/10.1037/0003-066X.45.8.921

Streiner, D. L. (2003). Starting at the beginning: an introduction to coefficient alpha and internal consistency. *Journal of personality assessment*, *80*(1), 99-103. https://doi.org/10.1207/S15327752JPA8001_18

Taylor, S. C., & Gassner, L. (2010). Stemming the flow: challenges for policing adult sexual assault with regard to attrition rates and under-reporting of sexual offences. *Police Practice and Research: An International Journal*, *11*(3), 240-255. https://doi.org/10.1080/15614260902830153

Teigen, K. H., & Brun, W. (2003). Verbal probabilities: a question of frame?. *Journal of Behavioral Decision Making*, *16*(1), 53-72. https://doi.org/10.1002/bdm.432

Testa, M. (2002). The impact of men's alcohol consumption on perpetration of sexual aggression. *Clinical psychology review*, *22*(8), 1239-1263. https://doi.org/10.1016/S0272-7358(02)00204-0

Testa, M., Fillmore, M. T., Norris, J., Abbey, A., Curtin, J. J., Leonard, K. E., ... & Hayman, L. W. (2006). Understanding alcohol expectancy effects: Revisiting the placebo condition. *Alcoholism: Clinical and experimental research*, *30*(2), 339-348. https://doi.org/10.1111/j.1530-0277.2006.00039.x

The Innocence Project (2020). 200 Exonerated, too many wrongfully convicted. Retrieved 5th September 2022 from https://www.innocenceproject.org/wp-content/uploads/2016/10/ip200.pdf

Tousignant, J. P., Hall, D., & Loftus, E. F. (1986). Discrepancy detection and vulnerability to misleading postevent information. *Memory & Cognition*, *14*(4), 329-338. https://doi.org/10.3758/BF03202511

Traub, R. E. (1997). Classical test theory in historical perspective. *Educational Measurement*, *16*, 8-13. https://doi.org/ 10.1111/j.1745-3992.1997.tb00603.x

Van Oorsouw, K., & Merckelbach, H. (2012). The effects of alcohol on crime-related memories: A field study. *Applied Cognitive Psychology*, *26*(1), 82-90. https://doi.org/10.1002/acp.1799

Van Oorsouw, K., Broers, N. J., & Sauerland, M. (2019). Alcohol intoxication impairs eyewitness memory and increases suggestibility: Two field studies. *Applied Cognitive Psychology, 33*(3), 439-455. https://doi.org/ 10.1002/acp.3561

Van Oorsouw, K., Merckelbach, H., & Smeets, T. (2015). Alcohol intoxication impairs memory and increases suggestibility for a mock crime: A field study. *Applied Cognitive Psychology, 29*(4), 493-501. https://doi.org/10.1002/acp.3129

Vogt, W. P., & Johnson, B. (2011). *Dictionary of statistics & methodology: A nontechnical guide for the social sciences*. Sage.

Volk, R. J., Steinbauer, J. R., Cantor, S. B., & Holzer III, C. E. (1997). The Alcohol Use Disorders Identification Test (AUDIT) as a screen for at-risk drinking in primary care

patients of different racial/ethnic backgrounds. *Addiction*, *92*(2), 197-206. https://doi.org/10.1111/j.1360-0443.1997.tb03652.x

Waite, M. (Ed.). (2012). *Paperback Oxford English dictionary*. Oxford University Press.

Weiss, K. G. (2010). Too ashamed to report: Deconstructing the shame of sexual victimization. *Feminist Criminology*, *5*(3), 286-310. https://doi.org/10.1177/1557085110376343

White, A. M. (2003). What Happened? Alcohol, memory blackouts, and the brain. *Alcohol, Research and Health, 27*, 186–196.

Wilson, B. M., Seale-Carlisle, T. M., & Mickes, L. (2018). The effects of verbal descriptions on performance in lineups and showups. *Journal of Experimental Psychology: General*, *147*(1), 113. https://doi.org/10.1037/xge0000354

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. (2005). A theory about why we forget what we once knew. *Current Directions in Psychological Science*, *14*(1), 6-9. https://doi.org/10.1111/j.0963-7214.2005.00324.x

Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychological Review*, *117*(4), 1025–1054. https://doi.org/10.1037/a0020874

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

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

World Health Organization. (1993). *The ICD-10 classification of mental and behavioural disorders: diagnostic criteria for research* (Vol. 2). World Health Organization.

World Health Organization. (2001). *AUDIT: The alcohol use disorders identification test: Guidelines for use in primary health care* (No. WHO/MSD/MSB/01.6 a). World Health Organization.

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

Zajac, R., Dickson, J., Munn, R., & O'Neill, S. (2016). Trussht me, I know what I sshaw: The acceptance of misinformation from an apparently unreliable co-witness. *Legal and Criminological Psychology, 21*(1), 127-140. https://doi.org/10.1111/lcrp.12032

Zaragoza, M. S., Belli, R. F., & Payment, K. E. (2007). Misinformation effects and the suggestibility of eyewitness memory. In M. Garry, & H. Hayne (Eds.), *Do justice and let the sky fall* (pp. 35–64). New Jersey: Mahwah.

Appendices

Appendix 1

Search terms by database

Psych Articles

Alcohol OR "Alcohol myopia" OR "Alcohol intocicat*" OR Drunk* OR "Blood alcohol concentration"	AND	Eyewitness* OR Legal testimon* OR "Legal evidence" OR "Witness testimon*" OR Evidence OR "Cross examin*"	AND	Suggestibility OR "Misinformation effect" OR "False memor*"

PsychInfo

Alcohol OR "Alcohol myopia" OR "Alcohol intocicat" OR Drunk* OR "Blood alcohol concentration"	AND	Eyewitness* OR Legal testimon* OR "Legal evidence" OR "Witness testimon*" OR Evidence OR "Cross examin*"	AND	Suggestibility OR "Misinformation effect" OR "False memor*"

Pubmed

Alcohol OR "Alcohol myopia" OR "Alcohol intocicat*" OR Drunk* OR "Blood alcohol concentration"	AND	Eyewitness* OR Legal testimon* OR "Legal evidence" OR "Witness testimon*" OR Evidence OR "Cross examin*"	AND	Suggestibility OR "Misinformation effect" OR "False memor*"

Scopus

Alcohol OR "Alcohol myopia" OR "Alcohol intocicat*" OR Drunk* OR "Blood alcohol concentration"	AND	Eyewitness* OR Legal testimon* OR "Legal evidence" OR "Witness testimon*" OR Evidence OR "Cross examin*"	AND	Suggestibility OR "Misinformation effect" OR "False memor*"

Web of Science

Alcohol OR "Alcohol myopia" OR "Alcohol intocicat*" OR Drunk* OR "Blood alcohol concentration"	AND	Eyewitness* OR Legal testimon* OR "Legal evidence" OR "Witness testimon*" OR Evidence OR "Cross examin*"	AND	Suggestibility OR "Misinformation effect" OR "False memor*"

Appendix 2

Studies excluded by full text

Study	Reason excluded
Reich, R. R., Goldman, M. S., & Noll, J. A. (2004). Using the false memory paradigm to test two key elements of alcohol expectancy theory. *Experimental and Clinical Psychopharmacology*, *12*(2), 102.	Testing method was not applicable to eyewitness memory
Desmarais, S. L., & Read, J. D. (2011). After 30 years, what do we know about what jurors know? A meta-analytic review of lay knowledge regarding eyewitness factors. *Law and Human Behavior*, *35*(3), 200-210.	No mention of suggestibility
Curran, H. V., & Hildebrandt, M. (1999). Dissociative effects of alcohol on recollective experience. *Consciousness and Cognition*, *8*(4), 497-509.	No mention of suggestibility
Steffens, M. C., & Mecklenbräuker, S. (2007). False Memories. *Zeitschrift für Psychologie/Journal of Psychology*, *215*(1), 12-24.	No reference to alcohol
Mintzer, M. Z., & Griffiths, R. R. (2001). Alcohol and false recognition: a dose-effect study. *Psychopharmacology*, *159*(1), 51-57.	No mention of suggestibility
Yuille, J. C., & Tollestrup, P. A. (1990). Some effects of alcohol on eyewitness memory. *Journal of Applied Psychology*, *75*(3), 268.	Suggestibility not studied, only false ID's
Evans, J. R., Schreiber Compo, N., & Russano, M. B. (2009). Intoxicated witnesses and suspects: Procedures and prevalence according to law enforcement. *Psychology, Public Policy, and Law*, *15*(3), 194.	Suggestibility not studied
Flowe, H. D., & Compo, N. S. (2021). The lack of robust evidence for the effects of alcohol on false memory. *Neuroscience & Biobehavioral Reviews*.	Relevant subject but was a commentary piece on another paper

Kloft, L., Monds, L. A., Blokland, A., Ramaekers, J. G., & Otgaar, H. (2021). Hazy memories in the courtroom: a review of alcohol and other drug effects on false memory and suggestibility. *Neuroscience & Biobehavioral Reviews*.	A review of relevant literature but therefore secondary data and not included in the data extraction process

Appendix 3

Studies where full text was unavailable:

Ashok, J., Nair, M., & Friedman, R. (2016). Drug-facilitated sexual assaults. In *Sexual offending: Predisposing antecedents, assessments and management* (pp. 67-77). New York, NY: Springer Science + Business Media; US.

Reich, R. R. (2003). Alcohol expectancies and implicit memory: Using a false memory approach to examine the implicit nature of alcohol expectancies in different contexts. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 64*(1-B), 429.

The following email was sent via Research Gate:
Good afternoon, I was hoping to access this full text as I am currently writing a systematic literature review looking at the effects of alcohol on suggestibility at the University of Birmingham. Thank you - Brittany Gibbs

Appendix 4

Quality assessment checklists

Randomised Control Trial Checklist

Criteria	Yes	No	Other (CD, NR, NA)*
1. Did the trial address a clearly focussed issue?			
2. Was the assignment of participants to treatments randomised?			
3. Were all of the participants who entered the trial properly accounted for at conclusion?			
4. Were study personnel blind to treatment?			
5. Were the groups similar at the start of the trial?			
6. Aside from the experimental intervention, were the groups treated equally?			
7. How large was the treatment effect?			
8. How precise were the results?			
9. Can the results be applied to other populations?			
10. Were all important outcomes considered?			
11. Are the benefits worth the harms and costs?			
12. Have the authors identified all important confounding factors?			
13. Are results in line with other evidence?			

Field study checklist

Criteria	**Yes**	**No**	**Other (CD, NR, NA)***
1. Did the study address a clearly focussed issue?			
2. Was the cohort recruited in an acceptable way?			
3. Was the exposure accurately measured to minimise bias?			
4. Was the outcome accurately measured to minimised bias?			
5. Have the authors identified all important confounding factors?			
6. Have they taken account of the confounding factors in the design and/or analysis?			
7. How precise are the results?			
8. Do you believe the results?			
9. Can the results be applied to the local population?			
10. Do the results of this study fit with other available evidence?			

Appendix 5

Completed quality assessments and scores

Absolute memory distortions: alcohol placebos influence the misinformation effect

Criteria	**Yes**	**No**	**Other (CD, NR, NA)***
1. Did the trial address a clearly focussed issue?	X		
2. Was the assignment of participants to treatments randomised?	X		
3. Were all of the participants who entered the trial properly accounted for at conclusion?	X		
4. Were study personnel blind to treatment?		X	
5. Were the groups similar at the start of the trial?			Unknown
6. Aside from the experimental intervention, were the groups treated equally?	X		
7. How large was the treatment effect?		0.38	
8. How precise were the results?			Not reported
9. Can the results be applied to other populations?		x	
10. Were all important outcomes considered?	x		
11. Are the benefits worth the harms and costs?			N/A
12. Have the authors identified all important confounding factors?	X		

13. Are results in line with other evidence?	X		

Score= 17, 65%

The intoxicated co-witness: effects of alcohol and dyadic discussion on memory conformity and event recall

Criteria	**Yes**	**No**	**Other (CD, NR, NA)***
1. Did the trial address a clearly focussed issue?	X		
2. Was the assignment of participants to treatments randomised?	X		
3. Were all of the participants who entered the trial properly accounted for at conclusion?	X		
4. Were study personnel blind to treatment?		X	
5. Were the groups similar at the start of the trial?	X		
6. Aside from the experimental intervention, were the groups treated equally?	X		
7. How large was the treatment effect?	X		
8. How precise were the results?	X		
9. Can the results be applied to other populations?	X		
10. Were all important outcomes considered?	X		
11. Are the benefits worth the harms and costs?	X		
12. Have the authors identified all	X		

important confounding factors?			
13. Are results in line with other evidence?	X		

Score = 25, 96%

The impact of alcohol intoxication on witness suggestibility immediately and after a delay

Criteria	**Yes**	**No**	**Other (CD, NR, NA)***
1. Did the trial address a clearly focussed issue?	X		
2. Was the assignment of participants to treatments randomised?	X		
3. Were all of the participants who entered the trial properly accounted for at conclusion?	X		
4. Were study personnel blind to treatment?		X	
5. Were the groups similar at the start of the trial?	X		
6. Aside from the experimental intervention, were the groups treated equally?	X		
7. How large was the treatment effect?	.13		
8. How precise were the results?			Not reported
9. Can the results be applied to other populations?	X		
10. Were all important outcomes considered?	X		
11. Are the benefits worth the harms and costs?	X		
12. Have the authors identified all	X		

important confounding factors?			
13. Are results in line with other evidence?	X		

Score = 23, 88%

An experimental examination of the effects of alcohol consumption and exposure to misleading postevent information on remembering a hypothetical rape

Criteria	**Yes**	**No**	**Other (CD, NR, NA)***
1. Did the trial address a clearly focussed issue?	X		
2. Was the assignment of participants to treatments randomised?	X		
3. Were all of the participants who entered the trial properly accounted for at conclusion?	X		
4. Were study personnel blind to treatment?		X	
5. Were the groups similar at the start of the trial?	X		
6. Aside from the experimental intervention, were the groups treated equally?	X		
7. How large was the treatment effect?	.10		
8. How precise were the results?	X		
9. Can the results be applied to other populations?	X		
10. Were all important outcomes considered?	X		
11. Are the benefits worth the harms and costs?	X		
12. Have the authors identified all	X		

important confounding factors?			
13. Are results in line with other evidence?	X		

Score = 25, 96%

Alcohol-induced retrograde facilitation renders witnesses of crime less suggestible to misinformation

Criteria	**Yes**	**No**	**Other (CD, NR, NA)***
1. Did the trial address a clearly focussed issue?	X		
2. Was the assignment of participants to treatments randomised?	X		
3. Were all of the participants who entered the trial properly accounted for at conclusion?	X		
4. Were study personnel blind to treatment?		X	
5. Were the groups similar at the start of the trial?	X		
6. Aside from the experimental intervention, were the groups treated equally?	X		
7. How large was the treatment effect?	.088		
8. How precise were the results?	X		
9. Can the results be applied to other populations?	X		
10. Were all important outcomes considered?	X		
11. Are the benefits worth the harms and costs?	X		
12. Have the authors identified all	X		

important confounding factors?			
13. Are results in line with other evidence?	X		

Score = 25, 96%

No evidence that low levels of intoxication at both encoding and retrieval impact scores on the Gudjonsson Suggestibility Scale

Criteria	**Yes**	**No**	**Other (CD, NR, NA)***
1. Did the trial address a clearly focussed issue?	X		
2. Was the assignment of participants to treatments randomised?	X		
3. Were all of the participants who entered the trial properly accounted for at conclusion?	X		
4. Were study personnel blind to treatment?		X	
5. Were the groups similar at the start of the trial?	X		
6. Aside from the experimental intervention, were the groups treated equally?	X		
7. How large was the treatment effect?			N/A – not significant
8. How precise were the results?	X		
9. Can the results be applied to other populations?	X		
10. Were all important outcomes considered?	X		
11. Are the benefits worth the harms and costs?	X		

12. Have the authors identified all important confounding factors?	X		
13. Are results in line with other evidence?		X	

Score = 22, 85%

Intoxicated eyewitnesses: Better than their reputation?

Criteria	**Yes**	**No**	**Other (CD, NR, NA)***
1. Did the trial address a clearly focussed issue?	X		
2. Was the assignment of participants to treatments randomised?	x		
3. Were all of the participants who entered the trial properly accounted for at conclusion?			Unknown
4. Were study personnel blind to treatment?	X		
5. Were the groups similar at the start of the trial?	X		
6. Aside from the experimental intervention, were the groups treated equally?	X		
7. How large was the treatment effect?			N/A – results not significant
8. How precise were the results?	X		
9. Can the results be applied to other populations?			N/A
10. Were all important outcomes considered?	X		
11. Are the benefits worth the harms and costs?	X		

12. Have the authors identified all important confounding factors?	X		
13. Are results in line with other evidence?		X	

Score = 19, 73%

Trussht me, I know what I sshaw: The acceptance of misinformation from an apparently unreliable co-witness

Criteria	**Yes**	**No**	**Other (CD, NR, NA)***
1. Did the trial address a clearly focussed issue?	X		
2. Was the assignment of participants to treatments randomised?	X		
3. Were all of the participants who entered the trial properly accounted for at conclusion?	X		
4. Were study personnel blind to treatment?		X	
5. Were the groups similar at the start of the trial?	X		
6. Aside from the experimental intervention, were the groups treated equally?	X		
7. How large was the treatment effect?			No effect size stated
8. How precise were the results?			Not reported
9. Can the results be applied to other populations?	X		
10. Were all important outcomes considered?	X		
11. Are the benefits worth the harms and costs?	X		

12. Have the authors identified all important confounding factors?	X		
13. Are results in line with other evidence?	X		

Score = 21, 81%

Alcohol intoxication impairs eyewitness memory and increases suggestibility: Study one

Criteria	**Yes**	**No**	**Other (CD, NR, NA)***
1. Did the study address a clearly focussed issue?	x		
2. Was the cohort recruited in an acceptable way?	x		
3. Was the exposure accurately measured to minimise bias?	x		
4. Was the outcome accurately measured to minimised bias?	x		
5. Have the authors identified all important confounding factors?	x		
6. Have they taken account of the confounding factors in the design and/or analysis?	x		
7. How precise are the results?			Not reported
8. Do you believe the results?	X		
9. Can the results be applied to the local population?	x		
10 Do the results of this study fit with other available evidence?	x		

Score = 18, 90%

Alcohol intoxication impairs eyewitness memory and increases suggestibility: Study two

Criteria	**Yes**	**No**	**Other (CD, NR, NA)***

1. Did the study address a clearly focussed issue?	x		
2. Was the cohort recruited in an acceptable way?	x		
3. Was the exposure accurately measured to minimise bias?	x		
4. Was the outcome accurately measured to minimised bias?	x		
5. Have the authors identified all important confounding factors?	x		
6. Have they taken account of the confounding factors in the design and/or analysis?	x		
7. Do you believe the results?	X		
8. How precise are the results?			Not reported
9. Can the results be applied to the local population?	x		
10. Do the results of this study fit with other available evidence?	x		

Score = 18, 90%

Alcohol intoxication impairs memory and increases suggestibility for a mock crime: a field study

Criteria	**Yes**	**No**	**Other (CD, NR, NA)***
1. Did the study address a clearly focussed issue?	x		
2. Was the cohort recruited in an acceptable way?	x		
3. Was the exposure accurately measured to minimise bias?	x		

4. Was the outcome accurately measured to minimised bias?	x		
5. Have the authors identified all important confounding factors?	x		
6. Have they taken account of the confounding factors in the design and/or analysis?	x		
7. How precise are the results?			Not reported
8. Do you believe the results?	X		
9. Can the results be applied to the local population?	x		
10. Do the results of this study fit with other available evidence?	x		

Score=18, 90%

The effects of alcohol and co-witness information on memory reports: a field study

Criteria	**Yes**	**No**	**Other (CD, NR, NA)***
1. Did the study address a clearly focussed issue?	X		
2. Was the cohort recruited in an acceptable way?	X		
3. Was the exposure accurately measured to minimise bias?	X		
4. Was the outcome accurately measured to minimised bias?	X		
5. Have the authors identified all important confounding factors?	X		
6. Have they taken account of the confounding factors in the design and/or analysis?	X		

7. How precise are the results?	X		
8. Do you believe the results?	X		
9. Can the results be applied to the local population?	X		
10. Do the results of this study fit with other available evidence?		X	

Score = 19, 95%