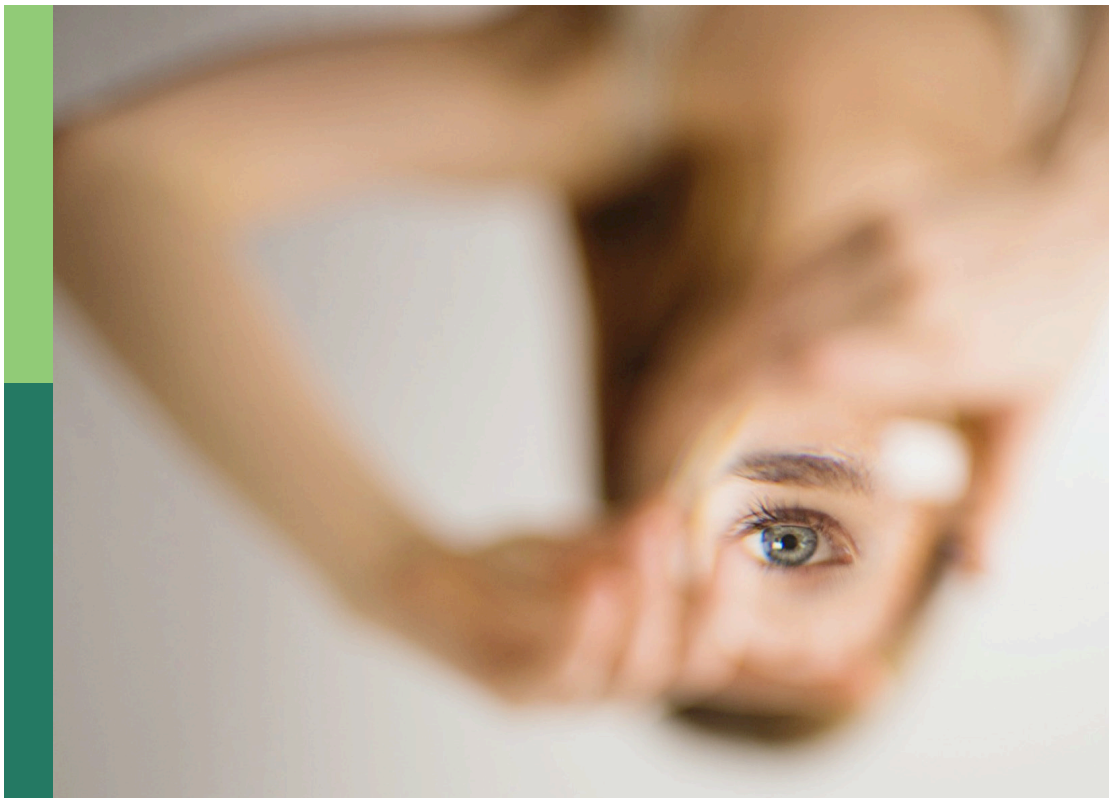
frontiers | Research Topics

The impact of internal and external influences on memory and their relevance to legal decisions

Edited by
Fabiana Battista, Antonietta Curci, Ivan Mangiulli and Henry Otgaar

Published in
Frontiers in Psychology



ISSN 1664-8714
ISBN 978-2-8325-4996-4
DOI 10.3389/978-2-8325-4996-4

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

The impact of internal and external influences on memory and their relevance to legal decisions

Topic editors

Fabiana Battista — University of Bari Aldo Moro, Italy
Antonietta Curci — University of Bari Aldo Moro, Italy
Ivan Mangiulli — University of Bari Aldo Moro, Italy
Henry Otgaar — Maastricht University, Netherlands

Citation

Battista, F., Curci, A., Mangiulli, I., Otgaar, H., eds. (2024). *The impact of internal and external influences on memory and their relevance to legal decisions*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-4996-4

Table of contents

frontiers | Frontiers in Psychology

TYPE Editorial
PUBLISHED 29 May 2024
DOI 10.3389/fpsyg.2024.1408797

OPEN ACCESS

EDITED AND REVIEWED BY
Jamal K. Mansour,
University of Lethbridge, Canada

*CORRESPONDENCE
Fabiana Battista
✉ fabiana.battista@uniba.it

RECEIVED 28 March 2024
ACCEPTED 20 May 2024
PUBLISHED 29 May 2024

CITATION
Battista F, Mangiulli I, Otgaar H and Curci A (2024) Editorial: The impact of internal and external influences on memory and their relevance to legal decisions.
Front. Psychol. 15:1408797.
doi: 10.3389/fpsyg.2024.1408797

Editorial: The impact of internal and external influences on memory and their relevance to legal decisions

Fabiana Battista [1]*, Ivan Mangiulli [1,2], Henry Otgaar [2,3] and Antonietta Curci [1]

[1]Department of Education, Psychology, Communication, University of Bari Aldo Moro, Bari, Italy, [2]Faculty of Law and Criminology, Leuven Institute of Criminology, Katholieke Universiteit (KU) Leuven, Leuven, Belgium, [3]Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Netherlands

KEYWORDS

internal influences, external influences, forgetting, amnesia, memory distortions, false memories

Editorial on the Research Topic
The impact of internal and external influences on memory and their relevance to legal decisions

Introduction

Cases of wrongful convictions based on unreliable testimonies, as shown by data from Innocence Projects of different countries (i.e., projects aiming to assure a fair process to people wrongfully convicted), show the deleterious effects of inaccurate memories in the legal context. Both external and internal influences can make memories inaccurate. For example, an abundance of research has shown that exposure to misleading and suggestive information can undermine memory for the original event even resulting into formation of false memories (for a review, Pickrell et al., 2016). Similarly, studies on deception have also found undermining effects in terms of both forgetting and false memories for the event (for a review, Battista and Otgaar, 2022). In the current Research Topic, we provide a unique assemblage of empirical and theoretical papers on these different influences on memory and the impact of memory studies in the courtroom. Specifically, in this Research Topic, papers on emotions and memory, traumatic memories, memory conformity, the misinformation effect, lying and memory, and developmental trends in false memories are presented.

Articles on external influences

Specifically, Marr et al. wrote a critical view on how acute stress can influence the retrieval of events. Their article is a reply to Pezdek and Reisberg's (2022) manuscript on whether or not stress can improve memory retrieval. The authors reviewed the literature on acute stress and memory, concluding that evidence on the link between them is mixed thereby arguing that the relationship between stress and memory depends on several moderators.

Davis et al. also wrote a review on the role of stress on memory but accounting for how stress affects attention which, in turn, can impact the encoding and retrieval of emotional events. In their review, they challenged common claims regarding tunnel memory, the effects of attention narrowing under stress, and the accuracy of memory for emotionally intense events, discussing potential biases in the acceptance of these claims by legal professionals. Finally, they reflected on the role of emotion in the interpretation and memory of sexual consent as well as the potential of trauma-informed interviewing strategies to induce memory distortions.

Dodier et al. provided a new framework to understand the phenomenon of recovered memories. They proposed that recovered memories can be seen as a form of involuntary autobiographical memories whose retrieval is triggered by internal (e.g., age and internal cues) or external cues (e.g., suggestion in therapy, suggestion during interview, and contextual cue). Using this framework, they proposed a new way to evaluate the validity of recovered memories in legal context and provided guidelines for practitioners to correctly apply this novel approach.

Kękuś et al. demonstrated that the classical memory conformity effect occurs also in online situations (i.e., MORI-v) and the effect is comparable to the effects obtained from in-person studies. They also showed the role of individual traits, such as susceptibility to social influence, need for closure, and self-esteem, on the memory conformity effect.

In two studies, Cullen et al. examined the influence of different types of misinformation (i.e., pro-prosecution, pro-defense, or contradictory) on juror decision-making and memory. Specifically, they tested the effects of congruent misinformation on jurors' evaluation of the credibility and verdict for a fictitious trial record of an alleged sexual assault as well as for the recall of the case.

Shah and Knott presented an experiment aiming to test the influence of retention interval and arousal for negative events on the exposure to gist or verbatim misleading information. They demonstrated that the misinformation effect is strong enough to persist over time for negative highly arousing event. According to the authors, these results further suggest the urge to avoid suggestive interviews, especially when arousing events are at stake.

O'Donnell et al. explored two possible aspects affecting the effects of misinformation on memory: The misinformation repetition and the source of misinformation. In two experiments, they readapted Foster et al.'s (2012) procedure and consistently detected in both studies that repetition did affect people's proneness to report misinformation in their recall for the original event, while source of misinformation did not.

Deering et al. further tested the misinformation effect by combining the misinformation procedure with a procedure used in line-up identification studies. They demonstrated that the viewing angle congruency between the perpetrator seen in the encoding phase and the one seen in the misinformation phase did not affect the identification accuracy. The authors concluded that the congruency between encoded and misleading information does not determine either an increase or decrease in the misinformation effect.

Jones et al. extended research on the identification of a culprit in line-up situations by taking into account the phenotypic bias (e.g., tendency to associate people with more Afrocentric -as opposed to Eurocentric- features with criminality). In their study, phenotypic bias did not undermine the correct recognition of the culprit when the culprit had more Afrocentric, rather than Eurocentric, features. Instead, participants were more able to identify the culprit when the phenotype was incongruent between the culprit and the line-up fillers, suggesting that practitioners (i.e., police) need to keep in mind the importance of matching facial phenotype between suspects and fillers when they arrange line-ups.

Articles on internal influences

Dianiska and Meissner investigated the effect of lying on memory accuracy and consistency. In addition, they examined whether the type of interview (i.e., Structure Interview, Reverse Order Interview) influences these two memory outcomes. Overall, lying made people's recall of the original event less accurate along with making people less consistent across interviews. Moreover, interviewing people with a Reverse Order technique reduced inconsistencies, in terms of omissions.

Articles on internal and external influences

Rosendaul et al. summarized studies on the two lines of research of the normative developmental position and the reverse developmental position. By reviewing internal (e.g., source misattributions, inferential reasoning) and external (e.g., valence, suggestion) influences that affect people's proneness to false memories, the authors argued there are no conclusive findings on how age determines memory accuracy, as such both children and adults can be reliable sources of information during legal proceedings.

Relatedly, Otgaar et al. provided an overview of studies on the role of suggesting non-occurrence and non-experience (i.e., external influences) and on the effect of deception (i.e., internal influence) on forgetting and false memories. According to this research, the authors map the outcomes associated to both influences underlining that, although differently, both types of influences can lead to similar mnemonic effects. Cognitive dissonance is put forward as the mechanism behind and operating both at an interpersonal or intrapersonal level.

Conclusions

The articles collected in this Research Topic show the wide researchers' interest in both internal and external influences, with a higher inclination to study external influences than internal ones. This could be due to several reasons. For one thing, while it is easier to explore external influences compared to internal ones, there is a clear necessity for fresh insights in these areas. Hence, there is a call to replicate previous findings in different contexts or through methodological adaptations to uncover novel insights. In addition, the few studies on internal influences and memory invite future investigation in this regard.

To conclude, the studies in this Research Topic provide legal practitioners with practical information on how to avoid

detrimental effects on memory and legal decisions. We firmly think that the current Research Topic can inspire future studies, contributing to disseminate knowledge among legal professionals.

Author contributions

FB: Conceptualization, Writing – original draft, Writing – review & editing. IM: Writing – review & editing. HO: Writing – review & editing. AC: Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Battista, F., and Otgaar, H. (2022). Research on the effects of lying on memory: a scientometric analysis and a call for new studies. *Front. Psychol.* 13:837265. doi: 10.3389/fpsyg.2022.837265

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

Pickrell, J. E., McDonald, D. L., Bernstein, D. M., and Loftus, E. F. (2016). "Misinformation effect," in *Cognitive Illusions*. New York, NY: Psychology Press, 406–423.

frontiers | Frontiers in Psychology

TYPE Opinion
PUBLISHED 17 April 2023
DOI 10.3389/fpsyg.2023.1078021

OPEN ACCESS

EDITED BY
Bernhard Pastötter,
University of Trier, Germany

REVIEWED BY
Charlie Frowd,
University of Central Lancashire,
United Kingdom
Marcus Möschl,
Technical University Dresden, Germany

*CORRESPONDENCE
Carey Marr
✉ carey.marr@unsw.edu.au

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Forensic and Legal Psychology, a section of the journal Frontiers in Psychology

RECEIVED 23 October 2022
ACCEPTED 29 March 2023
PUBLISHED 17 April 2023

CITATION
Marr C, Otgaar H, Quaedflieg CWEM, Sauerland M and Hope L (2023) Correcting myths about stress and memory: a commentary on Pezdek and Reisberg, 2022. *Front. Psychol.* 14:1078021. doi: 10.3389/fpsyg.2023.1078021

Correcting myths about stress and memory: a commentary on Pezdek and Reisberg, 2022

Carey Marr[1]*, Henry Otgaar[2,3†], Conny W. E. M. Quaedflieg[4†], Melanie Sauerland[2†] and Lorraine Hope[5†]

[1]School of Clinical Medicine, Discipline of Psychiatry and Mental Health, University of New South Wales, Sydney, NSW, Australia, [2]Department of Clinical Psychological Science, Maastricht University, Maastricht, Netherlands, [3]Leuven Institute of Criminology, Catholic University of Leuven, Leuven, Belgium, [4]Department of Neuropsychology and Psychopharmacology, Maastricht University, Maastricht, Netherlands, [5]Department of Psychology, University of Portsmouth, Portsmouth, United Kingdom

KEYWORDS

acute stress, eyewitness memory, psychological myths, emotional memory, forensic settings

Introduction

The question of how acute stress might affect memory has applied value because witnesses, victims, and perpetrators often report experiencing stress or associated emotions (e.g., fear) during a crime. They might also experience acute stress when they are interviewed by the police. It is therefore important that legal professionals and memory scientists, particularly those acting as expert witnesses, can rely on evidence-based knowledge concerning the acute effects of stress on memory.[1] Pezdek and Reisberg (2022) recently published an article aimed at debunking six psychological myths about evidence in the legal system. In their article, they argued that the idea that high stress improves the accuracy of memory is a myth (Myth #2). We take issue with this assertion on the basis that such a conclusion is not empirically warranted and does not accurately reflect the current state of research. In this commentary, we lend some critical nuance regarding the complex stress-memory relationship in eyewitness contexts.

Discussion

In their article, Pezdek and Reisberg noted that they "focused on myths of which the contrary evidence seems particularly clear" (p. 144) and that the evidence they provided showed that "these widely held beliefs are (at least) without basis and, in many cases, flatly false" (p. 143). Although research on stress and memory has been ongoing over several decades, evidence or consensus on this topic is not as clear-cut as suggested. In a recent survey of 73 memory experts, Marr et al. (2021a) showed that 95% of eyewitness experts and 81% of fundamental memory experts generally agreed that "*Very high levels of stress impair the accuracy of eyewitness testimony*." However, in their study, only 61% of these

1 In this commentary—like Pezdek and Reisberg—we focus on *acute* rather than *chronic* stress, where a stressor is long-lasting and continuous. Though not discussed further in this commentary, future studies should consider the unique effects of chronic stress on memory performance (Wolf, 2008; Finsterwald and Alberini, 2014) and the interaction between acute and chronic stress, which may be relevant to legal settings when considering ongoing or repeated events such as family violence crimes.

experts deemed the statement reliable enough to present in court (see also Kassin et al., 2001). Importantly, the opinions of eyewitness memory experts and fundamental memory experts diverged widely regarding stress effects during encoding. While 78% of fundamental memory experts agreed that "*Experiencing stress during an event (i.e., at encoding) enhances memory for that event*," only 32% of eyewitness experts did, highlighting the lack of consensus even amongst memory experts.

Pezdek and Reisberg acknowledged the complexity of the stress-memory relationship by referring to a meta-analysis (Shields et al., 2017) that suggests that encoding stress may enhance memory for stressor-relevant information when there is no or little delay between encoding and the stressor. However, Pezdek and Reisberg concluded that these conditions for encoding stress improving memory were limited to a "narrow focus" (p. 145) and implied that situations where stress impaired memory were more common. In reality, though, eyewitnesses frequently experience stress and encoding simultaneously, and the type of to-be-remembered information is often directly related to the stressor in a crime situation. These factors are in line with the moderating conditions for memory enhancements within the meta-analytic findings. Both of these factors also align with neurobiological theories and findings of many acute stress studies in the fundamental memory field suggesting memory enhancements (e.g., Joëls et al., 2006; Marr et al., 2021b, for a review).

To provide evidence against Myth #2, Pezdek and Reisberg cited findings from the eyewitness memory field suggesting that encoding stress impairs memory. However, this past work suffers from serious methodological limitations (Sauerland et al., 2016; Marr et al., 2021b). Many eyewitness studies conduct the memory retrieval test within minutes after the stressor/encoding phase (e.g., Brigham et al., 1983; Stanny and Johnson, 2000; Davis et al., 2019; Pezdek et al., 2020; Price et al., 2022). Because stress has an opposite effect on memory encoding (i.e., enhancing) and retrieval (i.e., impairing), this lack of sufficient retention interval obstructs any conclusions about the effects of encoding stress effects on memory.[2] Additionally, the majority of eyewitness studies (e.g., Davis et al., 2019; most studies in Deffenbacher et al., 2004; Pezdek et al., 2020) have relied on self-reports of stress rather than more objective, physiological measures, such as blood pressure or cortisol. Self-reported measures are valuable for application to real life, where physiological, objective measures are often unobtainable. However, for experimental lab studies, this measurement issue raises the question of whether the effect of encoding *stress* on memory was actually captured—or merely an effect of *arousal* (or a number of other cognitive phenomena). Researchers should strive to ensure that stress is properly induced and verified by using objective measures wherever possible, alongside self-report measures (cf. Shields et al., 2017; Marr et al., 2021b). If physiological measures of stress cannot be included, researchers should be cautious in using the term "stress" with respect to its effects on memory without noting this limitation. This care in terminology is particularly important for eyewitness studies, which often involve complex scenarios that likely produce many other effects, including the impact of arousal, divided attention, perceptual phenomena, or cognitive load. More studies examining links between self-reported levels of stress and physiological states of stress would be helpful for improving the construct validity of self-report measures, and in turn, will improve application to reality (e.g., Weber et al., 2022).

Pezdek and Reisberg (2022) also discredited the (ecological) validity of stress induced by the Trier Social Stress Test (TSST; see footnote 2 on p. 146). Dozens of studies collecting physiological measures alongside self-reports and recent meta-analyses have confirmed the validity of the TSST for inducing a full stress response (e.g., Goodman et al., 2017; Seddon et al., 2020; Gu et al., 2022). In contrast, it is currently unclear how "stressors" used in many eyewitness studies score on these dimensions (e.g., emotional pictures, violent videos, false fire alarms, Joëls et al., 2006; Marr et al., 2021b). Given that to-be-remembered materials that are directly related to the stressor elicit stronger effects (Shields et al., 2017), this should motivate eyewitness memory and stress researchers to collaborate in designing studies that combine the best of both fields to study the effects of encoding stress on memory (cf. Marr et al., 2021a). However, the fact that stress elicited in the TSST is not directly related to the to-be-remembered material does not justify discarding all findings that derive from its use—or effectively throwing the baby out with the bath water.

Conclusion and implications

We conclude that the empirical research base to date does not allow for any strong conclusions about the effect of encoding stress on memory. Rather, whether acute stress impairs, enhances, or does not reliably affect memory performance is dependent on many moderators, most of which still need to be more thoroughly investigated in future research (Marr et al., 2021b). Eyewitness reports from those who have been through a stressful experience should not be immediately accepted *or* discounted without examining the surrounding context and keeping the findings from *both* the eyewitness and fundamental memory fields in mind.

Future research on this topic will provide a clearer understanding of the factors that critically contribute to the relationship between stress and memory and the direction of that relationship. In the meantime, however, it is important to acknowledge the existing shades of gray when discussing stress effects on memory, particularly in applied legal settings. That being said, certain sub-topics relevant to the stress-memory relationship in eyewitness settings do show greater expert consensus than others (e.g., those related to stress severity and detail type; Marr et al., 2021a). Additionally, strong expert consensus exists regarding the inaccuracy of certain widespread layperson beliefs, including ideas that police officers are less influenced by acute stress or that stressful experiences can cause memory repression. These incorrect beliefs can and should be countered where relevant by expert witnesses in court.

2 Note that this type of research is important in its own right—if the research question aims to examine how acute stress may affect immediate memory performance (e.g., Krix et al., 2016). However, if researchers specifically aim to examine effects of encoding stress, a retention interval of at least 24 h is needed to properly separate the encoding and retrieval memory stages due to the lengthy timeline of a physiological stress response (Joëls and Baram, 2009).

Author contributions

CM drafted the commentary. All other authors provided critical feedback and contributed to the development and finalization of the commentary.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Brigham, J. C., Maass, A., Martinez, D., and Wittenberger, G. (1983). The effect of arousal on facial recognition. *Basic Appl. Soc. Psychol.* 4, 279–293. doi: 10.1207/s15324834basp0403_6

Davis, S. D., Peterson, D. J., Wissman, K. T., and Slater, W. A. (2019). Physiological stress and face recognition: Differential effects of stress on accuracy and the confidence-accuracy relationship. *J. Appl. Res. Memory Cognit.* 8, 367–375. doi: 10.1016/j.jarmac.2019.05.006

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., and McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law Hum. Behav.* 28, 687–706. doi: 10.1007/s10979-004-0565-x

Finsterwald, C., and Alberini, C. M. (2014). Stress and glucocorticoid receptor-dependent mechanisms in long-term memory: from adaptive responses to psychopathologies. *Neurobiol. Learn. Memory* 112, 17–29. doi: 10.1016/j.nlm.2013.09.017

Goodman, W. K., Janson, J., and Wolf, J. M. (2017). Meta-analytical assessment of the effects of protocol variations on cortisol responses to the Trier Social Stress Test. *Psychoneuroendocrinology* 80, 26–35. doi: 10.1016/j.psyneuen.2017.02.030

Gu, H., Ma, X., Zhao, J., and Liu, C. (2022). A meta-analysis of salivary cortisol responses in the Trier Social Stress Test to evaluate the effects of speech topics, sex, and sample size. *Compreh. Psychoneuroendocrinol.* 16, 100125. doi: 10.1016/j.cpnec.2022.100125

Joëls, M., and Baram, T. Z. (2009). The neuro-symphony of stress. *Nat. Rev. Neurosci.* 10, 459–466. doi: 10.1038/nrn2632

Joëls, M., Pu, Z., Wiegert, O., Oitzl, M. S., and Krugers, H. J. (2006). Learning under stress: how does it work? *Trends Cognit. Sci.* 10, 152–158. doi: 10.1016/j.tics.2006.02.002

Kassin, S. M., Tubb, V. A., Hosch, H. M., and Memon, A. (2001). On the "general acceptance" of eyewitness testimony research: a new survey of the experts. *Am. Psychol.* 56, 405–416. doi: 10.1037/0003-066X.56.5.405

Krix, A. C., Sauerland, M., Raymaekers, L. H. C., Memon, A., Quaedflieg, C. W. E. M., and Smeets, T. (2016). Eyewitness evidence obtained with the self-administered interview is unaffected by stress. *Appl. Cogn. Psychol.* 30, 103–112. doi: 10.1002/acp.3173

Marr, C., Otgaar, H., Sauerland, M., Quaedflieg, C. W. E. M., and Hope, L. (2021a). The effects of stress on eyewitness memory: a survey of memory experts and laypeople. *Memory Cognit.* 49, 401–421. doi: 10.3758/s13421-020-01115-4

Marr, C., Sauerland, M., Otgaar, H., Quaedflieg, C., and Hope, L. (2021b). The effects of acute stress on eyewitness memory: an integrative review for eyewitness researchers. *Memory* 29, 1091–1100. doi: 10.1080/09658211.2021.1955935

Pezdek, K., Abed, E., and Cormia, A. (2020). Elevated stress impairs the accuracy of eyewitness memory but not the confidence-accuracy relationship. *J. Exp. Psychol. Appl.* 158–169. doi: 10.1037/xap0000316

Pezdek, K., and Reisberg, D. (2022). Psychological myths about evidence in the legal system: how should researchers respond? *J. Appl. Res. Memory Cognit.* 11, 143. doi: 10.1037/mac0000037

Price, H. L., Tottenham, L. S., Hatin, B., Fitzgerald, R. J., and Rubínov,á, E. (2022). Effects of stress on eyewitness identification in the laboratory. *Appl. Cognit. Psychol.* 36, 191–202. doi: 10.1002/acp.3910

Sauerland, M., Raymaekers, L. H. C., Otgaar, H., Memon, A., Waltjen, T. T., Nivo, M., et al. (2016). Stress, stress-induced cortisol responses, and eyewitness identification performance. *Behav. Sci. Law* 34, 580–594. doi: 10.1002/bsl.2249

Seddon, J. A., Rodriguez, V. J., Provencher, Y., Raftery-Helmer, J., Hersh, J., Labelle, P. R., et al. (2020). Meta-analysis of the effectiveness of the trier social stress test in eliciting physiological stress responses in children and adolescents. *Psychoneuroendocrinology* 116, 104582. doi: 10.1016/j.psyneuen.2020.104582

Shields, G. S., Sazma, M. A., McCullough, A. M., and Yonelinas, A. P. (2017). The effects of acute stress on episodic memory: a meta-analysis and integrative review. *Psychol. Bull.* 143, 636–675. doi: 10.1037/bul0000100

Stanny, C. J., and Johnson, T. C. (2000). Effects of stress induced by a simulated shooting on recall by police and citizen witnesses. *Am. J. Psychol.* 113, 359–386. doi: 10.2307/1423364

Weber, J., Angerer, P., and Apolinário-Hagen, J. (2022). Physiological reactions to acute stressors and subjective stress during daily life: a systematic review on ecological momentary assessment (EMA) studies. *PLoS ONE* 17, e0271996. doi: 10.1371/journal.pone.0271996

Wolf, O. T. (2008). The influence of stress hormones on emotional memory: Relevance for psychopathology. *Acta Psychol.* 127, 513–531. doi: 10.1016/j.actpsy.2007.08.002

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 26 June 2023
DOI 10.3389/fpsyg.2023.1184055

OPEN ACCESS

EDITED BY
Ivan Mangiulli,
University of Bari Aldo Moro, Italy

REVIEWED BY
Richard S. John,
University of Southern California, United States
Kate Houston,
Texas A&M International University,
United States

*CORRESPONDENCE
Christian A. Meissner
✉ cameissn@iastate.edu

RECEIVED 10 March 2023
ACCEPTED 22 May 2023
PUBLISHED 26 June 2023

CITATION
Dianiska RE and Meissner CA (2023) The effect of credibility assessment techniques on consistency and subsequent memory for the truth.
Front. Psychol. 14:1184055.
doi: 10.3389/fpsyg.2023.1184055

The effect of credibility assessment techniques on consistency and subsequent memory for the truth

Rachel E. Dianiska[1] and Christian A. Meissner[2]*

[1]Department of Psychological Science, University of California, Irvine, Irvine, CA, United States, [2]Department of Psychology, Iowa State University, Ames, IA, United States

Repeated interviews are common during an investigation, and perceived consistency between multiple statements is associated with an interviewee's credibility. Furthermore, research has shown that the act of lying can affect a person's memory for what truthfully occurred. The current study assessed the influence of lying on memory during initial and repeated interviews, as well as how an interviewer's approach might affect between-statement consistency for true and false statements. Participants performed a scavenger hunt at two sets of buildings on a university campus and then were either dismissed or interviewed (with a Reverse Order instruction or a Structured Interview) about their activities. Participants chose one set to tell the truth about and then created a lie about activities in another area of campus that had not been visited. One week later, all participants provided a second free recall statement about their activities during the scavenger hunt, and then a final truthful description of both areas that were visited during the scavenger hunt. Truthfully rehearsed experiences were associated with more accurate recall of information learned during the scavenger hunt as well as more consistent and more detailed statements. The Structured Interview led to initially more detailed statements, but more inconsistencies in the form of omissions.

KEYWORDS

lying, memory accuracy, consistency, credibility assessment, evidence-based interviewing

1. Introduction

Investigators routinely conduct repeated interviews with the same suspect during an investigation (Kassin et al., 2007). In this context, alterations to a statement, regardless of the intention behind them, are often used to question the reliability of a witness's statement (Brewer et al., 1999) and to identify a subject who may be providing a false statement (Granhag and Strömwall, 2001). Repeated interviews therein pose a quandary for both innocent and guilty suspects.

Innocent suspects may choose to be strategically forthcoming and cooperative with an investigation when telling the truth (Hartwig et al., 2007; Granhag and Hartwig, 2008). When interviewed again at a later time, inconsistencies might appear if the individual were to offer new information that was not provided in a previous statement (i.e., a *reminiscent* detail), or if they failed to recall information that was provided previously (i.e., a *forgotten* detail). Though reminiscence and forgetting reflect natural underlying cognitive processes that can arise as a

result of repeated retrieval (Ballard, 1913), such inconsistencies can lead an investigator to question a truth-teller's credibility.

Guilty suspects, on the other hand, will need to remember a previously provided false statement in order to maintain consistency across interviews. False statements that a guilty suspect can provide include false descriptions (or fabrications), and false denials of events (e.g., simple denials or simulated amnesia; Otgaar and Baker, 2018). In the former, a suspect may describe an event or an experience differently than how it actually occurred or describe an event that never occurred. Lying, in this instance, requires the suspect to confabulate details to create a plausible account. Alternatively, the suspect may lie by falsely denying that an event occurred, despite the fact that the event did take place. Psychological research has shown that the type of false statement that is provided can carry implications for one's ability to remember that lie, and that the act of lying can change a person's memory for the truth (Otgaar and Baker, 2018; Dianiska et al., 2019; Battista et al., 2020; Dianiska and Meissner, 2022). For example, denying or simulating amnesia can lead to more errors of omission, while lying by describing can lead to more errors of commission. As such, the manner in which a guilty suspect provides a false statement could influence not only their ability to appear credible (i.e., consistent) on subsequent interviews, but also their memory for what truthfully occurred.

False descriptions and false denials differ in the extent to which effortful, constructive mental processing is required. As a result, these two types of lies tend to differentially affect both accurate memory and false memory (Vieira and Lane, 2013; Battista et al., 2020). Lies that are told via false description are more likely to be correctly remembered due to the constructive processes involved in generating the descriptions (Riesthuis et al., 2020; Battista et al., 2021; Dianiska and Meissner, 2022). Providing a brief false denial, on the other hand, requires less effort to produce and is therefore more easily forgotten (Otgaar and Baker, 2018; Dianiska et al., 2019). In addition to denials being less effortful, poor memory for denials may also be due to an inhibitory mechanism (Anderson and Neely, 1996; Debey et al., 2015). However, memory for false denials can improve when the denials are repeated (Vieira and Lane, 2013; Dianiska and Meissner, 2022). Due to the constructive processes involved in lying by describing, false descriptions can paradoxically be more likely to be *misremembered* as the truth should the act of generating a description as a lie (rather than as a truth) be forgotten (Polage, 2012; Vieira and Lane, 2013). This process is likely a result of source misattribution, where one mistakes the origin of that description (Johnson et al., 1993). In this case, the content of the lie is retained, but the reason for its generation (e.g., to tell a lie) is not.

1.1. Memory, consistency, and perceived credibility

Relying on consistency as an indicator of truthfulness can negatively affect innocent suspects who seek to be cooperative with an interviewer. Truth-tellers' statements will be grounded in their memory for an event, and the reconstructive nature of memory increases the likelihood of errors (Bartlett, 1932). Should an innocent person provide an initially mistaken alibi statement due to faulty memory and come to realize their error, any attempt to correct their statement by providing contradictory information might lead to suspicion as a result of that inconsistency (Crozier et al., 2017). As such, unwarranted mistrust of an inconsistent (but innocent) alibi provider could potentially redirect the course of an investigation away from pursuing a different suspect. Investigators must consider not only the presence and type of an inconsistency in a statement, but also the role of memory recall inherent to producing that statement. Although some statement-enhancing questioning techniques strategically support and capitalize on an interviewee's memory, the impact of such tactics have not yet been fully assessed with respect to possible misattributions of deception and guilt due to inconsistencies across statements.

Regardless of guilt, the interaction between lying and memory has implications for a suspect's experience with the criminal justice system. For instance, whether guilty suspects are able to maintain their false narratives over time could have significant downstream consequences that lead to their conviction. On a subsequent interview, the ability to remember (and repeat; Granhag and Strömwall, 1999) what was said in an initial interview is extremely important given the common perception that inconsistency is associated with deception (Vredeveldt et al., 2014).

1.2. Consistency across repeated interviews for truthful and deceptive accounts

Though inconsistencies are often treated by laypeople and professionals as indicators of deception, research suggests that it is the *type* of inconsistency that is a more important indicator of deception, rather than inconsistency itself (Fisher et al., 2013; Vredeveldt et al., 2014). Across repeated interviews, engaging in varied retrieval can contribute to the reminiscence of details not previously reported. Gilbert and Fisher (2006) examined the effects of varied retrieval across a repeated interview context on inconsistencies in the form of contradictions, reminiscences, and omissions. Varying the retrieval cues between two event recall opportunities increased the amount of reminiscent information reported and decreased the number of items that were omitted on the second event recall. The amount of consistent and contradictory items that were recalled were similar. Gilbert and Fisher also found that the accuracy of inconsistent-reminiscent and inconsistent-omitted details was fairly high (87 and 93%, respectively). Consistent details, however, were still associated with the highest accuracy (95%). Few contradictory details were reported overall, but when they were reported, they were associated with low accuracy (49%).

For guilty suspects, there are different types of (in)consistency that can induce suspicion, including the perceived consistency within a suspect's statement and across multiple statements. Inconsistencies can also arise between statements elicited from multiple suspects, or between a suspect's statement and the available evidence. Interviewers can use strategic questioning approaches to encourage the production of some inconsistencies to facilitate credibility assessment. Consistency across statements has been suggested to be indicative of liars who have rehearsed their statement (Vrij et al., 2009; Masip et al., 2016), liars who underestimate the extent to which forgetting occurs (i.e., stability bias; Harvey et al., 2017a,b), and/or liars who deliberately repeat the same statement given previously to avoid being exposed (Granhag and Strömwall, 1999). However, manipulating the way in which a suspect provides a statement can prevent a liar from using a "repeat" strategy to appear consistent.

Liars are likely to be *inconsistent* when faced with varied retrieval, such as when they must report an event differently between multiple interviews. For example, (Leins et al., 2012) asked liars and truth-tellers to describe their activities in an initial interview either verbally, by providing an initial free recall and then answering specific questions from an interview, or pictorially, by producing a sketch drawing of the task room and the location of as many items as possible. After a 10-min delay, participants provided the interviewer with an additional statement about their activities in the same or different reporting method. Truth-tellers were more consistent than liars (when only items that were *contradictory* were compared to items that were consistent); however, liars were even less consistent when the retrieval method differed between interviews.

1.3. Evidence-based interviewing techniques and consistency

Researchers and practitioners have advocated for the use of evidence-based interviewing techniques to increase cooperation and disclosure of information in investigative interviewing (Vrij et al., 2014; Vrij and Granhag, 2014; Meissner et al., 2017; Brandon et al., 2018). Such interviewing tactics have been assessed as both tools to improve the quality of an interviewee's memory report as well as to magnify differences in verbal content between liars and truth-tellers that aid lie detection (Vrij, 2015), particularly given that the most successful training protocols for lie detection and credibility assessment focus on such verbal content (see Hauch et al., 2016). Examples include eye closure instructions (Perfect et al., 2008), mental context reinstatement (Smith and Vela, 2001), recalling an event in reverse temporal order (Vrij et al., 2008), and asking subjects to sketch along with their statement (Deeb et al., 2022).

The primary goal of these techniques is to increase the amount of information obtained from an interview without a commensurate decrease in accuracy. Techniques that encourage a speaker to elaborate, however, can sometimes lead to the provision of information that may not be true (or information that they might be unsure of; Koriat and Goldsmith, 1996) due to an interviewee reporting incorrect information (i.e., errors in describing a witnessed detail) or confabulating novel details (i.e., errors in describing unwitnessed details). Should an interviewee report such erroneous information on a subsequent interview (or amend a prior statement to correct an error), an interviewer could note a difference between the two statements and infer deception on the part of the subject. However, an error that persists could become incorporated into the subject's memory for what truthfully occurred (e.g., self-generated misinformation; Pickel, 2004), irreparably affecting their credibility if the information is revealed to be inaccurate.

One tactic that has been evaluated as a credibility assessment tool is a *reverse-order recall* instruction (Vrij et al., 2008; Evans et al., 2013). After an interviewee has provided an initial free narrative, they are asked to recall the event once more in reverse chronological order. Recalling an event from multiple retrieval perspectives, in particular one that is counter to an initial schema-guided retrieval attempt (Geiselman and Callot, 1990), can allow for a previously inaccessible memory trace to be accessed and therein increase the amount of information reported. Asking for an event description in reverse-order increases cognitive load more so for liars than truth-tellers, thereby magnifying discernible verbal and nonverbal behaviors between the two (Vrij et al., 2008; Evans et al., 2013). However, when compared with a request for an open-ended narrative, recalling an event in reverse-order can sometimes increase confabulations and decrease overall statement accuracy (Dando et al., 2011). Errors that are produced as a result of a reverse-order instruction could persist across repeated interviews, leading to further consequences for interviewees with respect to perceived inconsistency (Fisher et al., 2013).

Interviewers' use of reverse-order recall can induce inconsistencies in both liars and truth-tellers (Gilbert and Fisher, 2006; Hudson et al., 2019). Hudson and colleagues examined consistency between two statements provided in close succession to each other. When a reverse-order recall instruction was administered, both liars and truth-tellers provided more omissions and fewer repetitions. Overall, truth-tellers provided more details across the two interviews, and specifically more reminiscent details during a second interview than did liars. Liars, in contrast, made significantly more omissions when a reverse-order recall instruction was administered during an interview, compared to when a chronological order recall instruction was administered.

1.4. Present study

The purpose of the present study was to examine the potential for evidence-based interviewing tactics to foster the generation of inconsistencies across multiple interviews, as well as the potential detrimental influence of providing a false statement on memory for the truth. We used a behavioral paradigm in which participants completed a series of complex tasks prior to being interviewed about them (see Figure 1). Participants experienced two distinct events (scavenger hunt tasks in two areas of a university campus) and then had the opportunity to choose which event they would rehearse truthfully. They then created a false description about activities about a third event, a set of building that had not been visited. Participants were then randomly assigned to be initially interviewed using one of two forensic interview protocols (a Structured Interview or a Reverse-Order recall instruction), or to not complete an initial interview. Seven days later, all participants returned for a second session, during which time they were interviewed about their activities the week prior. In the second interview, participants were provided an open-ended prompt to freely recall each of the two events that they experienced in the first session, describing each event truthfully (or deceptively). Lastly, participants provided a final truthful description of the lied-about event, as well as a final account of the truthfully rehearsed event.

We examined how lying on a prior interview affects one's memory for what truthfully occurred, and how interviewing techniques might affect the consistency of information reported across repeated interviews. It was hypothesized that fewer correct details would be recalled about a lied-about event during the final all-truth statement, compared to events that were rehearsed truthfully. We further expected that truthfully-described events would be associated with an overall greater amount of detail, and associated with greater consistency and/or inconsistencies in the form of reminiscences. With respect to the Reverse Order instruction,

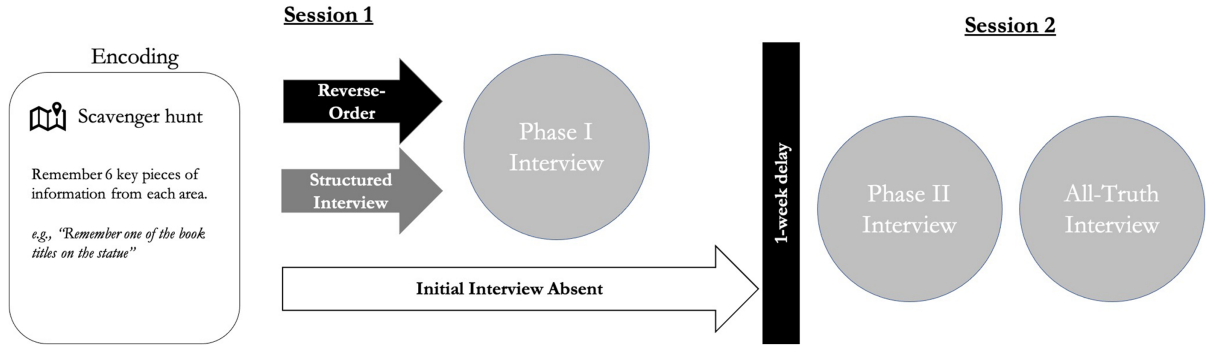


FIGURE 1
Experiment procedure.

we predicted that this interviewing technique would lead to more overall total statement detail, as well as more inconsistencies in the form of omissions (i.e., details present in an initial statement, but not repeated during the second).

2. Methods

2.1. Pre-registration

This study was pre-registered on OSF (https://doi.org/10.17605/OSF.IO/AJ296) to include sample size, methods, primary hypotheses, and planned analyses.

2.2. Participants

A total of 112 participants (56 female) were recruited from a large Midwestern state university, and 105 completed the full experiment ($n = 7$ dropped out between Session 1 and Session 2). Data from six additional participants were excluded for not complying with interview instructions for either Phase I or Phase II interviews. Thus, the final sample included slightly uneven cells for Interview Absent ($n = 29$), Reverse Order ($n = 32$), and Structured Interview ($n = 38$) conditions. Participants age varied between 18 and 28 years of age ($M = 19.38$, $SD = 1.39$).

Due to University closure in the Spring of 2020 in response to the COVID-19 virus, data collection ended prematurely. The target sample of 144 research participants ($n = 48$ per group) specified in the pre-registration would provide sufficient power to detect a relatively small within-between interaction effect size ($f = 0.15$) with power of 0.90 (Faul et al., 2009). This effect size was chosen based on prior work demonstrating differences in consistency for liars and truth-tellers across repeated interviews (e.g., $f = 0.31$ in Leins et al., 2012) and robust increases in total detail following strategic interviewing techniques (e.g., $f = 0.20$ when comparing chronological recall and reverse order recall in Hudson et al., 2019). To appropriately power an interaction between Veracity and Interview Technique, a more conservative effect size was used ($f = 0.15$) than has been observed in prior work. Data analyzed and presented here represent those collected prior to the university closure in March of 2020. Because of ongoing institutional changes with respect to research procedures and the permanent closure of areas of the university campus included in the current study, the remaining participants needed to fulfill the proposed target sample were unable to be collected. Had the power analysis been conducted to be less conservative (0.80), the current sample size would have been sufficient to detect the anticipated effect size.

2.3. Design

A 3 (*Initial Interview Technique*: Absent, Reverse-Order, Structured Interview) × 2 (*Veracity*: Lie, Truth) × 2 (*Interview Time*: Phase I, Phase II) mixed design was used. Initial Interview Technique was manipulated between-participants, while Veracity and Interview Time were manipulated within-participants.

2.4. Procedure

Participants completed two sessions conducted 1 week apart (see Figure 1). The first session comprised the Encoding Phase and the Phase I Interview (for initially-interviewed conditions). Participants visited four buildings (two pairs of buildings total) on the university campus and completed a scavenger hunt for information within each building. After completing the scavenger hunt, some participants were interviewed about their activities (Reverse Order and Structured Interview conditions) and some were dismissed from the session (Absent condition). Before being interviewed, participants were instructed that they would truthfully tell the interviewer about one pair of buildings of their choice; they would not discuss the other pair of buildings they visited, and instead were instructed to lie about a specific set of buildings that were not visited during the experiment.

2.4.1. Encoding phase

Upon arrival to the session, participants received instructions and provided informed consent to complete the experiment. Before beginning the Encoding Phase, participants completed a brief survey assessing their familiarity with six buildings on the University campus on scale from 1 (I have never been there/Not familiar) to 7 (I know the ins and outs of the building/Extremely familiar). During the Encoding Phase, participants completed what they believed to be a study assessing people's memory for previously performed activities.

Participants received instructions that they would be going to different buildings on the university campus and performing a scavenger hunt at each one. Participants then navigated to two "areas" of campus (i.e., two buildings near each other) and completed a series of brief tasks at each one. Throughout the course of the Encoding Phase, participants were tasked to remember six key pieces of information that they learned in each area. Three versions of the scavenger hunt were created, such that each pair of buildings was equally presented to participants as the first area or second area to which they navigated. All tasks and instructions for the scavenger hunt can be found on the OSF repository.

When participants arrived back to the lab, those in the Interview-Absent condition were dismissed and asked to return 1 week later to complete Phase II. Those in the Interview-Present conditions (Reverse-Order, Structured Interview) received instructions for the initial interview phase. Participants were told that they would be interviewed about their actions after leaving the lab. For the interview, they were asked to tell the truth about one area of campus (meaning one "pair" of buildings) and lie about another area of campus. The participants could choose to tell the truth about either the first pair of buildings or the second pair of buildings that they visited but were instructed that they must lie about another, pre-specified set of buildings.

For the lied-about event, participants were instructed that they would create a detailed, believable cover story. Participants were provided with a worksheet with minimal information about the buildings they were tasked with lying about (taken from the public access building information available on the University's Facilities Planning and Management website; see OSF) and given 5–6 min to write down details that could be provided in their narratives. To motivate participants to lie well during the task, participants were told that their interviews would be evaluated by other people after the session has concluded, and the person who was judged to be most believable will win a $25 reward. After preparing for the interview, the experimenter confirmed that the participant understood the instructions for the interview task and then left the room to notify the interviewer.

2.4.2. Phase I interview

The participants interviewed in Phase I were randomly assigned to be interviewed with a Reverse Order Instruction or a Structured Interview. Interviews always began by asking for an initial open-ended narrative of participants' activities at the first area of campus, and then an open-ended narrative for the second area of campus.

In the *Reverse Order* condition, the interviewer followed up the initial request by asking the participant interviewee to recall their activities in the two areas once again in reverse chronological order, beginning from the last temporal detail that they provide for each area. In the *Structured Interview* condition, the interviewer followed up the initial request by asking three probing questions about details the participant had mentioned for each area of campus.

After the conclusion of the interview, participants completed a brief post-interview questionnaire. In addition to demographic information, participants reflected on how well they remembered the tasks that they had completed, what strategy they used to select which event to truthfully describe, how motivated they were to be perceived as truthful, if they did anything in particular to convince the interviewer that they were telling the truth, how comfortable they are with lying in everyday life, as well as global perceptions of the interviewer.

2.4.3. Phase II interview

One week later, all participants (Interview-Absent, Reverse-Order, Structured Interview) returned to the lab to complete the Phase II interview. At the beginning of the session, the experimenter informed all participants that they would be interviewed (for the first time, for Interview-Absent participants; or again, for Reverse-Order and Structured Interview participants) about their activities during the first session of the experiment. Participants were asked to provide a free recall narrative of the two areas of campus that they visited the week prior. At this time, Interview-Absent conditions were given the same lie-truth instructions and cover story preparation time as participants who were interviewed in Phase I. All other participants (Reverse-Order and Structured Interview participants) were instructed to continue to respond truthfully or deceptively for each area of campus as they had in Phase I. During the Phase II interview, the interviewer requested only an open-ended narrative from participants recalling as much information as possible about their activities in both areas of campus.

2.4.4. Final all-truth interview

After describing the two areas of campus truthfully and deceptively, the interviewer informed participants that it was known they were told to lie about their activities in the previous session. Therefore, the participant's final experimental task was to describe both events as they had *actually* occurred. In addition to providing a third and final statement about their truthfully rehearsed event, participants were told to cease responding deceptively (about their chosen, lied-about event) and to describe their activities in the second, visited area truthfully and in as much detail as possible. This final truthful recall allowed us to assess the influence of having previously lied about an event on subsequent recall of the experienced details.

At the conclusion of the Phase II interview, participants completed a similar post-experiment questionnaire as in the earlier session. These questions reflected overall memory for the tasks in Phase I, strategy use, motivation, comfort with lying in everyday life, and perceptions of the interviewer and the interview experience. Further, they were asked to what extent they expected to be interviewed again, as well as the extent to which they expected the second half of the interview (i.e., the Final All-Truth interview) and how difficult was it to truthfully recall their activities from the first session. Participants who completed an initial interview were also asked the extent to which they attempted to repeat everything they had said previously about their activities during Phase I (i.e., to be consistent) and to what extent they attempted to provide new information about the first and second areas they visited. For the Interview-Absent participants, this questionnaire contained the same questions as the post-Phase I interview questionnaire. Finally, participants completed a cued-recall test for the details that they were tasked to remember during Phase I. Before being debriefed, participants were asked whether they had rehearsed their story or discussed the experiment with anyone since completing Phase I.

2.4.5. Coding of interview statements

Video recorded interviews for each phase (Phase I, Phase II, All-Truth) were coded for subsequent analysis. For Phase I interviews, trained research assistants coded details that were present during the initial chronological narrative that were also *repeated* after the reverse order instruction or structured interview probes were administered, as well as details that were *added* to participant statements after the

instruction or probing questions were implemented. A Total Phase I unique details measure was computed by summing: (i) consistent pre- and post-tactic details, and (ii) new details post-tactic. For participants who were not interviewed during Phase I, the same coding scheme was applied for their Phase II interviews.

For all other participants, Phase II interviews were coded for details that were: (i) repeated between Phase I and Phase II (*consistent* details); (ii) contradictory to details provided during Phase I (*inconsistent-contradiction* details); (iii) added during Phase II but not reported during Phase I (*inconsistent-reminiscent* details); and (iv) failed to be provided during Phase II but were provided at Phase I (*inconsistent-omitted* details). Statements from participants who were interviewed in Phase I and Phase II were coded by two coders for the volume of information provided and the consistency of details that were provided. Inter-rater reliability was high (r's > 0.93 for each described area). We computed the total amount of detail provided at Phase I and Phase II, as well as the amount of consistent details, omitted details, and new details reported across statements.

During the final All-Truth interview, participants were instructed to provide a final truthful statement for both areas of campus they had actually visited during Phase I. These all-truthful statements were then coded for the amount of detail provided for both areas of campus–one that they had rehearsed truthfully in the earlier Phase I and Phase II interviews, and one that they had lied about by describing their activities in an alternate area of campus. We assessed accuracy with respect to participants' freely recalled statements, and with respect to their performance on the cued recall test at the end of Phase II. If participants mentioned a detail they were tasked to remember during their all-truth narrative, the detail was coded as a "1" if it was present and accurate in the statement (e.g., correctly recalling "1926" as the year a fountain was installed). The same was true if participants correctly answered the cued recall question on the final test. A score of "0" for a detail was given for inaccurate details (e.g., an incorrect year), non-specific details (e.g., saying they were told to remember a year, but not providing the year), or when the participant did not mention or said they could not remember the item. Accurate details per area thus ranged from 0 to 6 details, and from these a proportion of accurate details was computed based on the number of details that were mentioned (note: the pattern of results does not change when the proportion of all potential key details are included, rather than just those details mentioned correctly or incorrectly).

3. Results

All materials and data are hosted on OSF.[1] Descriptive statistics from each condition across all measures can be found in Table 1. The following results are distinguished by whether they were pre-registered or exploratory. We first assess the effects of the veracity of a statement and the presence/type of interview tactic used to elicit an initial narrative on participants' ability to correctly recall information learned during the scavenger hunt. Next, we examine the effects of statement veracity and the type of interview technique used on the amount of information provided during initial and subsequent interviews, and then specifically consistent or inconsistent details provided therein. Finally, we explore differences in the amount of detail provided in participants' initial recall statements of each event.

3.1. Pre-registered analyses

3.1.1. Correct recall on final all-truth interview

At the conclusion of Phase II, participants provided a final, truthful account of their activities involving both areas of campus that they visited during the Encoding Phase. This interview allowed us to assess the influence of having previously recalled an event truthfully vs. deceptively. A 3 (*Initial Interview Technique*: Absent, Reverse-Order, Structured Interview) × 2 (*Veracity*: Lie, Truth) ANOVA was conducted on the proportion of accurate details mentioned in participants' all-truth interview statements (see Figure 2). As hypothesized, there was a significant main effect of Veracity such that memory for key details was more accurate for areas of campus that participants had previously truthfully recalled ($M = 0.46$, $SE = 0.03$) than areas of campus they had lied about ($M = 0.33$, $SE = 0.03$), $F(1, 96) = 15.44$, $p < 0.001$, $d = 0.41$ [0.21, 0.62]. Neither the main effect of Interview Technique ($F(2, 96) = 1.13$, $p = 0.33$, $\eta_p^2 = 0.02$) nor the interaction between Initial Interview Technique and Veracity ($F(2, 96) = 1.37$, $p = 0.26$, $\eta_p^2 = 0.03$) were significant. Performance on the cued recall test was similar and is reported on OSF.

3.1.2. Phase I and phase II total details

A 2 (*Initial Interview Technique*: Structured Interview, Reverse-Order) × 2 (*Veracity*: Lie, Truth) × 2 (*Interview Time*: Phase I, Phase II) ANOVA was conducted on the total number of details present in participant's statements. There was a significant main effect of Veracity, $F(1, 68) = 9.28$, $p = 0.003$, $d = 0.37$ [0.13, 0.61]; Interview Time, $F(1, 68) = 67.53$, $p < 0.001$, $d = 0.83$ [0.55, 1.10]; and Interview Technique, $F(1, 68) = 11.62$, $p = 0.001$, $d = 0.82$ [0.33, 1.31]. As expected, people provided more details when truthfully describing their activities ($M = 28.10$, $SE = 1.62$) than when creating false descriptions of their activities ($M = 25.26$, $SE = 1.69$). Further, participants provided more detailed statements during Phase I ($M = 30.67$, $SE = 1.90$) compared to Phase II ($M = 22.69$, $SE = 1.37$). However, participants provided more detailed statements when they were interviewed with a Structured Interview script ($M = 32.08$, $SE = 2.14$) compared with a Reverse Order instruction ($M = 21.28$, $SE = 2.33$), as we predicted. Importantly, the main effects of Interview Time and Initial Interview Technique were qualified by a significant interaction, $F(1, 68) = 44.09$, $p < 0.001$, $\eta_p^2 = 0.39$. While there was a significant decrease in the amount of information recalled from Phase I to Phase II for both conditions, this difference was much greater in the Structured Interview condition ($t(37) = 8.24$, $p < 0.001$, $d = 1.34$ [0.89, 1.77]) than in the Reverse Order condition ($t(31) = 3.95$, $p < 0.001$, $d = 0.70$ [0.31, 1.08]). No other main effects or interactions were significant, F's < 0.25, p's > 0.62.

3.1.3. Between-statement-consistency

A 2 (*Initial Interview Technique*: Reverse-Order, Structured Interview) × 2 (*Veracity*: Lie, Truth) mixed ANOVA was conducted on the number of details *consistently* provided between Phase I and Phase II (see Figure 3, solid gray and black bars). There was a significant main effect of Veracity, $F(1, 68) = 11.37$, $p < 0.001$, $d = 0.41$

1 https://osf.io/atz5h/

TABLE 1 Descriptive statistics for all measures.

	Initial interview absent		Reverse-order		Structured interview	
	Mean	SE	Mean	SE	Mean	SE
Phase II–Final recall accuracy						
Prev Lie	0.33	0.06	0.39	0.06	0.26	0.04
Prev Truth	0.40	0.06	0.51	0.06	0.46	0.05
Phase II–Cued recall accuracy						
Prev Lie	0.58	0.04	0.65	0.05	0.54	0.04
Prev Truth	0.66	0.05	0.71	0.04	0.69	0.03
Phase II–Consistency						
Lie	–	–	20.57	1.11	13.83	1.00
Truth	–	–	23.27	1.18	16.48	1.07
Phase II–Omissions						
Lie	–	–	10.30	1.24	15.62	1.12
Truth	–	–	8.68	1.12	16.88	1.02
Phase II–reminiscence						
Lie	–	–	3.92	0.65	4.86	0.59
Truth	–	–	5.06	0.54	4.64	0.49
Phase I–Total detail						
Lie	–	–	20.88	2.05	37.87	3.47
Truth	–	–	23.22	2.13	40.71	2.91
Phase II–Total detail						
Lie	21.00	1.80	18.88	1.96	23.42	1.99
Truth	25.31	2.07	22.16	2.23	26.32	1.93
Final all-Truth–Total detail						
Prev Lie	13.96	0.90	14.84	1.02	14.92	0.94
Prev Truth	14.37	1.06	15.38	0.93	14.17	0.93

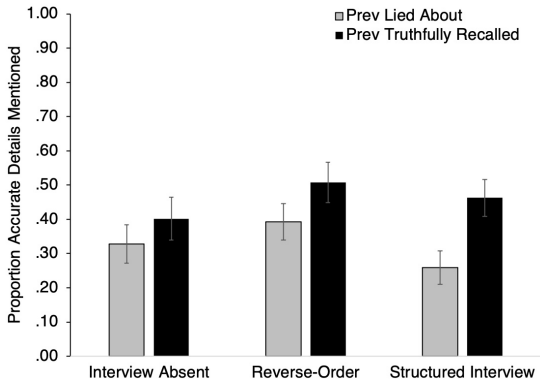


FIGURE 2
Proportion of key details correctly recalled the final all-truth interview.

[0.16, 0.65]). People provided more consistent details between Phase I and Phase II when describing their truthfully rehearsed event ($M = 19.59$, $SE = 1.20$) than when describing their lied-about event ($M = 16.91$, $SE = 1.18$). However, there was no main effect of Interview Technique ($F(1, 68) = 0.73$, $p = 0.40$, $\eta_p^2 = 0.01$, nor an interaction between Veracity and Interview Technique ($F(1, 68) = 0.06$, $p = 0.96$, $\eta_p^2 < 0.01$).

With respect to inconsistency, we examined differences in omissions and new details separately via 2 (*Initial Interview Technique*: Reverse-Order, Structured Interview) × 2 (*Veracity*: Lie, Truth) mixed ANOVAs on the number of *omitted* details (see Figure 3, patterned bars) and the number of *new* details added in Phase II (see Figure 3, open bars). There was a main effect of Interview Technique on the number of details *omitted* from Phase II statements, $F(1, 68) = 47.27$, $p < 0.001$, $d = 1.65$ [1.10, 2.19]. Participants omitted more details from Phase II statements after being interviewed with a Structured Interview in Phase I ($M = 20.16$, $SE = 2.00$) relative to those interviewed with a Reverse Order instruction in Phase I ($M = 4.84$, $SE = 0.51$). However, there was no main effect of Veracity ($F(1, 68) < 0.01$, $p = 0.93$, $\eta_p^2 < 0.01$, nor an interaction between Veracity and Interview Technique, $F(1, 68) = 0.10$, $p = 0.75$, $\eta_p^2 < 0.01$. With respect to *new* details provided during Phase II, there was only a main effect of Interview Technique ($F(1, 68) = 10.10$, $p < 0.01$, $d = 0.76$ [0.27, 1.25]). Participants interviewed with a Structured Interview added more details in their Phase II statements ($M = 5.74$, $SE = 0.62$) than did participants interviewed with a Reverse Order instruction ($M = 3.31$, $SE = 0.40$). Neither the main effect of Veracity ($F(1, 68) = 0.80$, $p = 0.37$,

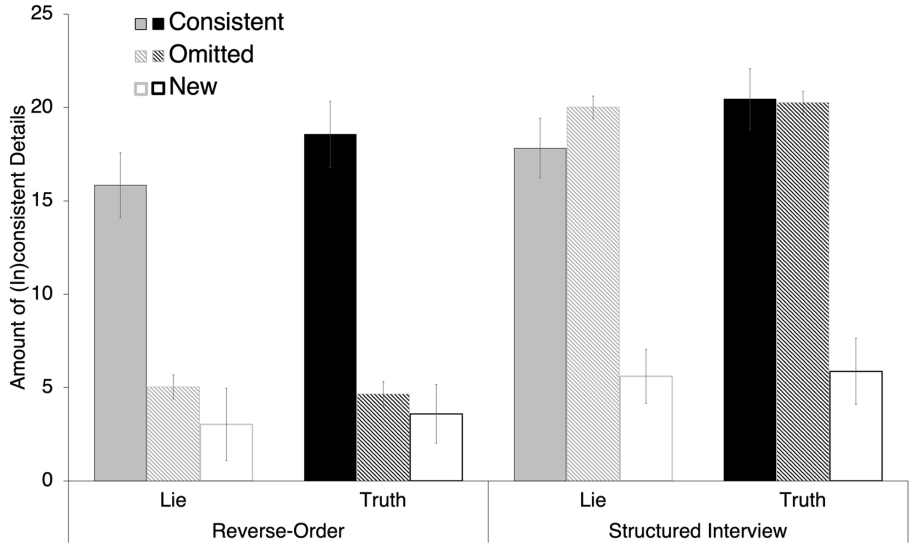


FIGURE 3
Number of details that were consistent or inconsistent between Phase I and Phase II interviews.

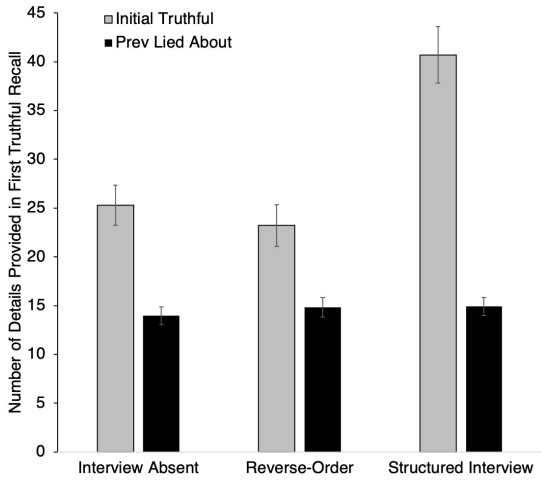


FIGURE 4
Number of details initially truthfully recalled per event area.

$d = 0.11$ [−0.13, 0.34]), nor an interaction between the two ($F(1, 67) = 1.74$, $p = 0.19$, $\eta_p^2 = 0.03$.

3.2. Exploratory analysis

In addition to pre-registered analyses, we also explored how participants not interviewed at Phase I (Initial Interview-Absent) compared to participants who did receive an interview in Phase I with respect to differences in the amount of detail provided for the first time an area of campus was discussed (see Figure 4). For participants who received an initial interview, we examined whether the total amount of detail differed for their initial truthful statement (during the Phase I interview) relative to their truthful statement about the unrehearsed area of campus that they visited (during the Final All-Truth interview). For participants who were not interviewed during Phase I, we compared the amount of detail in their initial truthful statement (during the Phase II interview) to their truthful statement about the area of campus they visited that they did not rehearse previously (during the Final All-Truth interview).

Pairwise analyses were conducted to compare the amount of detail provided for the initial narrative about the previously lied-about event relative to the initial narrative about the previous truthfully rehearsed event for participants in each interview condition. Participants provided significantly more details when initially recalling their truthful event compared to when they truthfully recalled the event that they previously lied about in the Interview Absent condition ($t(26) = 6.80$, $p < 0.001$, $d = 1.31$ [0.78, 1.82]), in the Reverse Order condition ($t(31) = 4.40$, $p < 0.001$, $d = 0.78$ [0.38, 1.17]), and in the Structured Interview condition ($t(36) = 9.55$, $p < 0.001$, $d = 1.57$ [1.08, 2.05]).

4. Discussion

Given the frequency with which investigators repeatedly interview criminal suspects, the current research assessed whether a suspect's choice to lie in an initial interview has consequences for memory accuracy and consistency on a subsequent interview. Here, we examined whether lying affects a suspect's memory with respect to accuracy, as well as how credibility assessment interview techniques (such as the Reverse Order instruction) influence between-statement consistency. Our findings suggest that relative to lying, telling the truth led to better memory for encoded material and more consistent statements across interviews separated by a one-week delay. Further, those interviewed with a Structured Interview were more likely to omit details between two interview statements.

Experiences that participants had truthfully reported in an initial interview were associated with greater detail and were more accurately recalled when compared with those that were initially lied about.

Regardless of whether or how they were interviewed during the initial session, participants provided more spontaneous accurate details in the final truthful interview when they had been previously truthfully interviewed compared to when they had previously lied about the experience. This aligns with previous findings (e.g., Battista et al., 2020; Dianiska and Meissner, 2022), and demonstrates that lying is detrimental to subsequent recall of the truth.

Here, the act of lying required participants to not only refrain from describing one area they visited, but also to create a false description of an area they had not visited. The mnemonic effect of lying seen here may thus be due to a relative lack of rehearsal, as suggested by the MAD framework (Otgaar and Baker, 2018). When participants provided false descriptions about an unvisited area of campus, they did so at the expense of not rehearsing an area of campus that was visited during the scavenger hunt. As a result, people provided less information about the unrehearsed (i.e., lied-about) area of campus when they were later asked to truthfully recall their experience (see Riesthuis et al., 2022b). The fact that the unrehearsed experience was associated with less detail could also be due to a spontaneous inhibition strategy that people may use to facilitate their lie-telling. That is, relative to areas that were truthfully rehearsed, in order to effectively produce a false description of an unexperienced event, people may have attempted to intentionally inhibit information about their activities in the unrehearsed area.

The content of people's statements, both initially and in subsequent interviews, may serve to discriminate lies from truths. In the present experiment, truthfully provided statements about prior experiences were more detailed than experiences that were lied about. This was the case for statements obtained during both initial (Phase I) and delayed (Phase II) interviews. Consistent with prior research it is clear that the level of detail provided about an event can serve as an indicator of veracity (e.g., Evans et al., 2013). However, in the present study, people provided less detail on the Phase II interview for both truthful and deceptive statements after a one-week delay. That is, when providing their deceptive statements, people did not demonstrate the "stability bias" (i.e., similar detail across interviews for lied-about events) that has been observed by others with lengthier delays (i.e., three weeks; Harvey et al., 2017a,b). It may be the case that, at longer delays, the level of detail is a more effective indicator of whether a person is lying or telling the truth.

The nature of these details, such as whether they are consistent across time points, may also be important for the discrimination of lies and truths. Consistency and inconsistency across repeated interviews were considered with respect to four main types of information: repeated, omitted, reminiscent, and contradictory. Opportunities for repeated recall offer truth-tellers an occasion to appear inconsistent, should they provide new information in a subsequent statement. The addition of information that is reminiscent (and therefore inconsistent) may be more likely when people are cued to provide a second statement with a different cue than was used to elicit a prior statement (Gilbert and Fisher, 2006). Liars, on the other hand, may be perceived as suspicious should their statements be inconsistent across interviews and therefore may strategically attempt to maintain their narratives over time. Here, truthfully described activities were associated with a greater proportion of consistent details than were experiences that people lied about.

Asking participants to recall their activities in reverse chronological order actually *improved* between-statement consistency. Specifically, people in the Reverse Order condition omitted fewer details between Phase I and Phase II interviews, compared to people who were asked follow-up probing questions in the Structured Interview condition. Although we expected that participants interviewed with a Reverse Order technique would provide more detailed initial narratives, it was the "tell me more" probing questions in the Structured Interview that led to greater reported details–however, many of these details were subsequently omitted in the Phase II interview. Though accuracy for the details added following these probes could not be assessed for all statements (though other work suggests they may be less accurate than unprompted details; Kontogianni et al., 2020), it is likely that these additional details were peripheral to the primary tasks. For instance, some of these details reflected individuals that they saw (but presumably did not interact with; e.g., "there was a guy with big black glasses" and "I almost ran into a girl"), while others reflected their personal thought processes or observations during the task (e.g., "it was loud in there" and "I was too lazy to scan [a QR code on a flyer in Campus Area A] with my phone"). Therefore, one possibility is that the additional probes in the Structured Interview condition may have prompted less important or less memorable details in the initial interview, leading participants to fail to provide these details during a subsequent interview.

In contrast to expectations, participants were similarly detailed during their Phase II interviews regardless of whether they had been initially interviewed or not. This may be due to participants in the Interview-Absent conditions receiving their cover story information and preparation time immediately preceding their interviews at Phase II. However, this preparation time was needed to equate the instructions to those received by initially interviewed participants.

4.1. Limitations and applied implications

Though it may be possible for truth-tellers to be inconsistent in their repeated recall of an event, our findings suggest that the type of memory cuing afforded during the initial recall episode may be important. Contrary to our expectations, people did not provide more reminiscent details during the Phase II interview when truthfully describing their activities. However, this was likely due to the Phase II interview involving a simple free recall prompt rather than the use of memory-enhancing or varied retrieval approaches. As a result, any reminiscence would have been spontaneous (or self-cued). The use of a memory-enhancing technique, like the Cognitive Interview, has been shown to facilitate the reporting of new details in delayed recall (Odinot et al., 2013; Hope et al., 2014).

Both of the interviewing techniques used to elicit narratives in this study are considered "best practice." The current research did not assess the effect of these best practice techniques in comparison to customary accusatorial tactics, such as those trained in the Reid technique (Inbau et al., 2011; see Meissner et al., 2015). Tactics that are characteristic of the Reid technique include shutting down denials, confronting the suspect with evidence of their guilt, and suggesting scenarios or theories of the crime. In future work, it may be useful to contrast the effects of lying on memory when best practice interview techniques are compared to such guilt-presumptive techniques.

Despite the benefit to some interview outcomes when "best practice" techniques are used (e.g., the diagnosticity of a confession; see Meissner et al., 2012, for a review), such techniques allow a subject to "tell their story" in a way that permits both denials as well as deceptive narratives. In a similar manner, approaches like the

Cognitive Interview can lead to small increases in incorrect details being provided by the subject–though such interviews also lead to large increases in correct details, thereby mitigating the effect on a person's overall accuracy (Memon et al., 2010). Could the provision of deceptive or incorrect information harm subsequent recall? The current data suggest that people who have previously lied are at a disadvantage should they decide at a later point to be truthful and forthcoming with an interviewer. We found that participants provided significantly less detail when they had previously lied about the event. What remains to be examined, however, is whether that harm might be partially or fully ameliorated if memory-enhancing techniques are used to elicit information in a later interview.

To motivate participants to lie convincingly during the experiment, we used a financial incentive based on others' perceptions of their statement. Though participants rated their motivation to be perceived as truthful well above the midpoint of the scale ($M = 5.43$, $SD = 0.99$; on a scale from 1, not at all, to 7, completely), offering a monetary reward for believability might not have adequately motivated someone to lie as they might in an interview. As such, future work should investigate the effects of lying on memory when there is a stronger motivation to lie, such as to avoid punishment or embarrassment (e.g., Riesthuis et al., 2022a).

Finally, given the recent emphasis on increasing the ecological validity of deception experiments (Romeo et al., 2019; Dianiska and Meissner, 2022), participants in the present experiment were permitted to *choose* when to lie and when to tell the truth. Prior to being interviewed, all participants were tasked with choosing one area to tell the truth about and were then given an area of campus to create a lie about their activities. While the paradigm used in the present experiment offers more ecological validity than other lab-based paradigms, it does so at the expense of being able to assess participant's statement accuracy. Given the variability in participants' episodic experiences during the scavenger hunt (e.g., encountering different people and obstacles along the way), we could not assess accuracy. Future studies might involve the inclusion of a confederate or the use of a body camera in the experimental task that would allow for a more natural, yet verifiable, encoding task.

4.2. Conclusion

Taken together, the current findings add further evidence that the act of lying has downstream consequences for the accurate recall of truthfully experienced events. That is, lying about one's experiences led to both less accurate memory for those experiences and less consistent statements. The current data suggest that the act of lying has a detrimental effect on memory for what truthfully occurred. Further, an interviewers' choice of tactic can significantly influence the amount and quality of information provided. The use of a credibility assessment technique, such as the Reverse Order instruction, facilitated between-statement consistency by reducing omissions. In the absence of such a tactic, an interviewers' selection of follow-up topics might, perhaps unintentionally, impede their ability to rely on consistency and the level of detail of a subject's statement as cues to credibility.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://osf.io/atz5h/.

Ethics statement

The studies involving human participants were reviewed and approved by Iowa State University Institutional Review Board (IRB). The patients/participants provided their written informed consent to participate in this study.

Author contributions

RD conceived the original idea and designed the study with contributions from CM, programmed, ran the experiment and analyzed the data, and wrote the primary drafts of the manuscript. CM assisted in analyses and interpretation of the data and provided feedback, and suggested revisions to the manuscript. All authors approved the submitted version of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Anderson, M. C., and Neely, J. H. (1996). "Chapter 8—interference and inhibition in memory retrieval" in *Memory*. eds. E. L. Bjork and R. A. Bjork (Cambridge: Academic Press)

Ballard, P. B. (1913). Obliviscence and reminiscence. *Br. J. Psychol.* 1:82.

Bartlett, F. C. (1932). *Remembering: a study in experimental and social psychology*. Cambridge: Cambridge University Press.

Battista, F., Mangiulli, I., Herter, J., Curci, A., and Otgaar, H. (2020). The effects of repeated denials and fabrication on memory. *J. Cogn. Psychol.* 32, 369–381. doi: 10.1080/20445911.2020.1767626

Battista, F., Mangiulli, I., Riesthuis, P., Curci, A., and Otgaar, H. (2021). Do liars really remember what they lied upon? The impact of fabrication on memory. *Memory* 29, 1076–1090. doi: 10.1080/09658211.2021.1960380

Brandon, S. E., Wells, S., and Seale, C. (2018). Science-based interviewing: information elicitation. *J. Investig. Psychol. Offender Profiling* 15, 133–148. doi: 10.1002/jip.1496

Brewer, N., Potter, R., Fisher, R. P., Bond, N., and Luszcz, M. A. (1999). Beliefs and data on the relationship between consistency and accuracy of eyewitness testimony. *Appl. Cogn. Psychol.* 13, 297–313. doi: 10.1002/(SICI)1099-0720(199908)13:4<297::AID-ACP578>3.0.CO;2-S

Crozier, W. E., Strange, D., and Loftus, E. F. (2017). Memory errors in alibi generation: how an alibi can turn against us. *Behav. Sci. Law* 35, 6–17. doi: 10.1002/bsl.2273

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

Debey, E., Ridderinkhof, R. K., De Houwer, J., De Schryver, M., and Verschuere, B. (2015). Suppressing the truth as a mechanism of deception: Delta plots reveal the role of response inhibition in lying. *Conscious. Cogn.* 37, 148–159. doi: 10.1016/j.concog.2015.09.005

Deeb, H., Vrij, A., Leal, S., Fallon, M., Mann, S., Luther, K., et al. (2022). Sketching routes to elicit information and cues to deceit. *Appl. Cogn. Psychol.* 36, 1049–1059. doi: 10.1002/acp.3989

Dianiska, R. E., Cash, D. K., Lane, S. M., and Meissner, C. A. (2019). "The reciprocal nature of lying and memory: memory confabulation and diagnostic cues to deception" in *The Palgrave handbook of deceptive communication*. ed. T. Docan-Morgan (London: Springer International Publishing)

Dianiska, R. E., and Meissner, C. A. (2022). The effect of lying on memory and metamemory when deception is repeated and volitional. *J. Appl. Res. Mem. Cogn.* 12, 128–140. doi: 10.1037/mac0000036

Evans, J. R., Michael, S. W., Meissner, C. A., and Brandon, S. E. (2013). Validating a new assessment method for deception detection: introducing a psychologically based credibility assessment tool. *J. Appl. Res. Mem. Cogn.* 2, 33–41. doi: 10.1016/j.jarmac.2013.02.002

Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. (2009). Statistical power analyses using G*power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160. doi: 10.3758/BRM.41.4.1149

Fisher, R. P., Vrij, A., and Leins, D. A. (2013). "Does testimonial inconsistency indicate memory inaccuracy and deception? Beliefs, empirical research, and theory" in *Applied issues in investigative interviewing, eyewitness memory, and credibility assessment*. eds. B. S. Cooper, D. Griesel and M. Ternes (New York: Springer)

Geiselman, R. E., and Callot, R. (1990). Reverse versus forward recall of script-based texts. *Appl. Cogn. Psychol.* 4, 141–144. doi: 10.1002/acp.2350040206

Gilbert, J. A. E., and Fisher, R. P. (2006). The effects of varied retrieval cues on reminiscence in eyewitness memory. *Appl. Cogn. Psychol.* 20, 723–739. doi: 10.1002/acp.1232

Granhag, P. A., and Hartwig, M. (2008). A new theoretical perspective on deception detection: on the psychology of instrumental mind-reading. *Psychol. Crime Law* 14, 189–200. doi: 10.1080/10683160701645181

Granhag, P. A., and Strömwall, L. A. (1999). Repeated interrogations –stretching the deception detection paradigm. *Expert. Evid.* 7, 163–174. doi: 10.1023/A:1008993326434

Granhag, P. A., and Strömwall, L. A. (2001). Deception detection based on repeated interrogations. *Leg. Criminol. Psychol.* 6, 85–101. doi: 10.1348/135532501168217

Hartwig, M., Anders Granhag, P., and Strömwall, L. A. (2007). Guilty and innocent suspects' strategies during police interrogations. *Psychol. Crime Law* 13, 213–227. doi: 10.1080/10683160600750264

Harvey, A. C., Vrij, A., Hope, L., Leal, S., and Mann, S. (2017a). A stability bias effect among deceivers. *Law Hum. Behav.* 41, 519–529. doi: 10.1037/lhb0000258

Harvey, A. C., Vrij, A., Leal, S., Hope, L., and Mann, S. (2017b). Deception and decay: verbal lie detection as a function of delay and encoding quality. *J. Appl. Res. Mem. Cogn.* 6, 306–318. doi: 10.1016/j.jarmac.2017.04.002

Hauch, V., Sporer, S. L., Michael, S. W., and Meissner, C. A. (2016). Does training improve the detection of deception? A meta-analysis. *Commun. Res.* 43, 283–343. doi: 10.1177/0093650214534974

Hope, L., Gabbert, F., Fisher, R. P., and Jamieson, K. (2014). Protecting and enhancing eyewitness memory: the impact of an initial recall attempt on performance in an investigative interview: protecting eyewitness memory. *Appl. Cogn. Psychol.* 28, 304–313. doi: 10.1002/acp.2984

Hudson, C. A., Vrij, A., Akehurst, L., and Hope, L. (2019). The devil is in the detail: deception and consistency over repeated interviews. *Psychol. Crime Law* 25, 752–770. doi: 10.1080/1068316X.2019.1574790

Inbau, F., Reid, J., Buckley, J., and Jayne, B. (2011). *Criminal interrogation and confessions*. Burlington: Jones & Bartlett Publishers.

Johnson, M. K., Hashtroudi, S., and Lindsay, D. S. (1993). Source monitoring. *Psychol. Bull.* 114, 3–28. doi: 10.1037/0033-2909.114.1.3

Kassin, S. M., Leo, R. A., Meissner, C. A., Richman, K. D., Colwell, L. H., Leach, A.-M., et al. (2007). Police interviewing and interrogation: a self-report survey of police practices and beliefs. *Law Hum. Behav.* 31, 381–400. doi: 10.1007/s10979-006-9073-5

Kontogianni, F., Hope, L., Taylor, P. J., Vrij, A., and Gabbert, F. (2020). "Tell me more about this…": an examination of the efficacy of follow-up open questions following an initial account. *Appl. Cogn. Psychol.* 34, 972–983. doi: 10.1002/acp.3675

Koriat, A., and Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychol. Rev.* 103, 490–517. doi: 10.1037/0033-295X.103.3.490

Leins, D. A., Fisher, R. P., and Vrij, A. (2012). Drawing on liars' lack of cognitive flexibility: detecting deception through varying report modes: Cognitive flexibility and deception. *Appl. Cogn. Psychol.* 26, 601–607. doi: 10.1002/acp.2837

Masip, J., Blandón-Gitlin, I., Martínez, C., Herrero, C., and Ibabe, I. (2016). Strategic interviewing to detect deception: cues to deception across repeated interviews. *Front. Psychol.* 7:e01702. doi: 10.3389/fpsyg.2016.01702

Meissner, C. A., Kelly, C. E., and Woestehoff, S. A. (2015). Improving the effectiveness of suspect interrogations. *Annu. Rev. Law Soc. Sci.* 11, 211–233. doi: 10.1146/annurev-lawsocsci-120814-121657

Meissner, C. A., Redlich, A. D., Bhatt, S., and Brandon, S. (2012). Interview and interrogation methods and their effects on true and false confessions. *Campbell Syst. Rev.* 8, 1–53. doi: 10.4073/csr.2012.13

Meissner, C. A., Surmon-Böhr, F., Oleszkiewicz, S., and Alison, L. J. (2017). Developing an evidence-based perspective on interrogation: a review of the U.S. government's high-value detainee interrogation group research program. *Psychol. Public Policy Law* 23, 438–457. doi: 10.1037/law0000136

Memon, A., Meissner, C. A., and Fraser, J. (2010). The cognitive interview: a meta-analytic review and study space analysis of the past 25 years. *Psychol. Public Policy Law* 16, 340–372. doi: 10.1037/a0020518

Odinot, G., Memon, A., La Rooy, D., and Millen, A. (2013). Are two interviews better than one? Eyewitness memory across repeated cognitive interviews. *PLoS One* 8:e76305. doi: 10.1371/journal.pone.0076305

Otgaar, H., and Baker, A. (2018). When lying changes memory for the truth. *Memory* 26, 2–14. doi: 10.1080/09658211.2017.1340286

Perfect, T. J., Wagstaff, G. F., Moore, D., Andrews, B., Cleveland, V., Newcombe, S., et al. (2008). How can we help witnesses to remember more? It's an (eyes) open and shut case. *Law Hum. Behav.* 32, 314–324. doi: 10.1007/s10979-007-9109-5

Pickel, K. (2004). When a lie becomes the truth: the effects of self-generated misinformation on eyewitness memory. *Memory* 12, 14–26. doi: 10.1080/09658210244000072

Polage, D. C. (2012). Fabrication inflation increases as source monitoring ability decreases. *Acta Psychol.* 139, 335–342. doi: 10.1016/j.actpsy.2011.12.007

Riesthuis, P., Otgaar, H., Battista, F., and Mangiulli, I. (2022a). Public beliefs on the relationship between lying and memory. *Psychol. Crime Law* 28, 545–568. doi: 10.1080/1068316X.2021.1929979

Riesthuis, P., Otgaar, H., De Cort, A., Bogaard, G., and Mangiulli, I. (2022b). Creating a false alibi leads to errors of commission and omission. *Appl. Cogn. Psychol.* 36, 936–945. doi: 10.1002/acp.3982

Riesthuis, P., Otgaar, H., Mangiulli, I., and de Tauzia, R. (2020). Adopting a fictitious autobiography: fabrication inflation or deflation? *Memory* 28, 741–752. doi: 10.1080/09658211.2020.1771371

Romeo, T., Otgaar, H., Smeets, T., Landstrom, S., and Boerboom, D. (2019). The impact of lying about a traumatic virtual reality experience on memory. *Mem. Cogn.* 47, 485–495. doi: 10.3758/s13421-018-0885-6

Smith, S. M., and Vela, E. (2001). Environmental context-dependent memory: a review and meta-analysis. *Psychon. Bull. Rev.* 8, 203–220. doi: 10.3758/BF03196157

Vieira, K. M., and Lane, S. M. (2013). How you lie affects what you remember. *J. Appl. Res. Mem. Cogn.* 2, 173–178. doi: 10.1016/j.jarmac.2013.05.005

Vredeveldt, A., van Koppen, P. J., and Granhag, P. A. (2014). "The inconsistent suspect: a systematic review of different types of consistency in truth tellers and liars" in *Investigative interviewing*. ed. R. Bull (New York: Springer), 183–207. doi: 10.1007/978-1-4614-9642-7_10

Vrij, A. (2015). *A cognitive approach to lie detection. Detecting deception: current challenges and cognitive approaches* Hoboken: Wiley.

Vrij, A., and Granhag, P. A. (2014). Eliciting information and detecting lies in intelligence interviewing: an overview of recent research. *Appl. Cogn. Psychol.* 28, 936–944. doi: 10.1002/acp.3071

Vrij, A., Hope, L., and Fisher, R. P. (2014). Eliciting reliable information in investigative interviews. *Policy Insights Behav. Brain Sci.* 1, 129–136. doi: 10.1177/2372732214548592

Vrij, A., Leal, S., Granhag, P. A., Mann, S., Fisher, R. P., Hillman, J., et al. (2009). Outsmarting the liars: the benefit of asking unanticipated questions. *Law Hum. Behav.* 33, 159–166. doi: 10.1007/s10979-008-9143-y

Vrij, A., Mann, S. A., Fisher, R. P., Leal, S., Milne, R., and Bull, R. (2008). Increasing cognitive load to facilitate lie detection: the benefit of recalling an event in reverse order. *Law Hum. Behav.* 32, 253–265. doi: 10.1007/s10979-007-9103-y

frontiers | Frontiers in Psychology

TYPE Review
PUBLISHED 16 August 2023
DOI 10.3389/fpsyg.2023.1232753

OPEN ACCESS

EDITED BY
Henry Otgaar,
Maastricht University, Netherlands

REVIEWED BY
Pietro Spataro,
Mercatorum University, Italy
Elisa Krackow,
West Virginia University, United States

*CORRESPONDENCE
Brittany J. Rosendaul
✉ bjr227@cornell.edu

RECEIVED 01 June 2023
ACCEPTED 02 August 2023
PUBLISHED 16 August 2023

CITATION
Rosendaul BJ, Su I-A and Ceci SJ (2023) Normative developmental vs. reverse developmental trends in memory distortion: a framework to investigate the impact of internal and external influences on memory and their relevance to legal decisions.
Front. Psychol. 14:1232753.
doi: 10.3389/fpsyg.2023.1232753

Normative developmental vs. reverse developmental trends in memory distortion: a framework to investigate the impact of internal and external influences on memory and their relevance to legal decisions

Brittany J. Rosendaul*, I-An Su and Stephen J. Ceci

Child Witness and Cognition Lab, Department of Psychology, Cornell University, Ithaca, NY, United States

There are two opposing positions regarding the development of memory: the normative developmental position, and the reverse developmental position. The normative position, which has long been the default presupposition, supports the notion that susceptibility to memory distortion, including false memories, decreases with age. In contrast, the concept of "developmental reversals" supports the notion that susceptibility to memory distortion and false memories increases with age. Each perspective finds support from existing theories as well as from research on endogenous and exogenous sources of influence. In a legal context, having an accurate understanding of the developmental course of false memory can contribute on the one hand to mitigating wrongful convictions and, on the other hand, to appreciating the accuracy of children's statements when warranted. This review aims to integrate the existing literature regarding these seemingly opposite developmental courses and construct a framework outlining the conditions under which we may observe one age trend over the other. This entails an examination of the paradigms that have been invoked to support these competing positions, specifically developmental responses to internal vs. external sources of distortion.

KEYWORDS

memory distortion, developmental reversals, endogenous sources, exogenous sources, false memories, suggestibility, psychology and law

1. Introduction

When it comes to memory, we may intuitively think of a close correlation between accuracy and chronological age, at least until old age when memory declines (Mueller-Johnson and Ceci, 2004; Mueller-Johnson et al., 2007). Children have long been publicly regarded as an unreliable memory source (Wright et al., 2010), made salient by the historical skepticism surrounding child eyewitness testimony dating back to the beginnings of scientific psychology (e.g., Ceci and Bruck, 1993, 1995; Dale et al., 1978; Melton, 1981). Upon asking open-access artificial intelligence to provide "descriptors of children," two of the ten responses refer directly to

children's "active imaginations" (ChatGPT, 2023), suggesting that the skepticism of former times is still with us.

This line of thinking is congruent with the *normative* developmental view of memory; a developmental framework which suggests that as chronological age increases, we anticipate a *decrease* in susceptibility to memory distortion. This normative developmental position has withstood the tests of time and prevailed as the popular consensus among memory developmentalists since the 1970s (see Ceci and Bruck, 1993 for historical review). It is based on an array of paradigms and independent variables (such as recognition, free recall, cued recall, short-term memory savings, release from proactive inhibition, and so on), and it is often associated with age-related changes in specific cognitive, social, and neurobiological factors that drive performance.

Some of the most prominent support for the normative developmental position comes from neurobiological changes with chronological age. Over the course of childhood, neural functioning develops and strengthens, particularly in regions such as the prefrontal cortex that moderate executive functions such as tracking, monitoring, and inhibition. With these neurobiological developments come improvements in working memory and consequently increased overall cognitive functioning (Constantinidis and Klingberg, 2016).

Predating the exponential growth in neurobiology in recent decades, there has been a steady stream of findings documenting cognitive and social influences on memory development. Moreover, endogenous and exogenous factors, including but not limited to strategy development, emotional valence, and suggestive techniques, were shown to be causally related to memory accuracy and resistance to false suggestions. With rare exceptions, this literature demonstrated normative trends of overall memory improvement with chronological age.

Ceci and Bruck (1993) reported that in 15 out of 18 published experiments, there were both age-related decreases in suggestibility and increases in accuracy, with preschoolers showing the poorest overall performance compared to older children, adolescents, and adults. This normative pattern of increasing accuracy with age has been replicated numerous times over the past several decades, both in individual studies and in syntheses of the literature (e.g., Saywitz, 1990; Zaragoza et al., 1992; Ceci et al., 1995; Ceci and Bruck, 1995; Bruck and Ceci, 1999; Gordon et al., 2001; Goodman et al., 2017; Gonzalves et al., 2022). Across four experiments, Ceci et al. (1987) examined mechanisms driving children's heightened vulnerability to external sources of suggestion: misleading post-event information; social forces, such as pressures to conform and receiving praise for wrong responses; and provision of false information by an adult authority figure, each of which were found to yield higher rates of memory error in 3- and 4-year-olds than older children, particularly compared to the old*est* children (12-year-olds).

In two of Ceci et al.'s (1987) experiments, social forces were controlled to parse the true causes of children's suggestibility, and researchers found that accessing the original memory trace was being directly altered by false suggestions. In other words, the experimental design minimized the operation of social forces so that any remaining age-related changes were due to cognitive changes (alteration of the underlying memory trace) rather than desire to please authority figures. This implicates the active role of memory-based mechanisms such as trace attenuation, trace alteration, and retrieval interference. The findings from these studies and many others like them provided early evidence that young children's memories for events—not just their reports of events, but their underlying memories as well—can be distorted as a result of exposure to post-event information. Such studies were taken as evidence of actual memory impairment as opposed to mere report errors. They formed the basis of the normative developmental position.

Many of the factors that have been investigated in the existing literature have been invoked in support of normative developmental age trends. However, particularly in the face of new research, there is evidence that: (1) these factors and paradigms can produce reverse developmental trends, and (2) there are specific factors that may uniquely predispose older populations to memory distortions. The idea that memory becomes *increasingly susceptible* to distortion in tandem with chronological age forms the foundation of *developmental reversals* (Brainerd and Reyna, 1998).

Much of the evidence supporting developmental reversals is more recent than the evidence supporting normative developmental trends (e.g., Brainerd et al., 2008a,b). The specific emphasis on developmental reversals can be traced largely to the contemporary research produced by Brainerd and Reyna's (2001, 2002, 2012) analysis of fuzzy-trace theory. Fuzzy-trace theory, or FTT, suggests that memories are not stored verbatim, but rather through conceptual representations, or "gist traces" (Reyna and Kiernan, 1995). With age, comes the ability to rely on such higher-level cognitive scaffolding. Specifically in terms of memory, this gist retrieval has been shown to coincide with increases in false memories (Brainerd and Reyna, 2012).

When children do not yet possess the skillset or semantic knowledge to create these gist traces to aid the encoding, storage, and retrieval of memories, they must rely more heavily on verbatim memory and are therefore less likely to make gist-based memory errors. For example, the question "Did you drink a Coke at lunch?" may cue verbatim or gist traces, or both. Under FTT, we anticipate that the verbatim memory of having had a Coke at lunch will aid in suppressing false memories (e.g., thinking that you had a Sprite at lunch), whereas gist trace aids in matching semantic features (e.g., remembering that you had a cold, fizzy drink at lunch) which may support false memories of other types of cold fizzy drinks. Both verbatim and gist traces contain potentially accurate information (in the sense that you encoded the experienced event) and therefore do not explicitly contribute to memory distortion. When these conceptual representations are over-relied on, we see the outcome of false memories (Brainerd and Reyna, 2012).

In this paper we begin by synthesizing the developmental findings, going back a half century to the classic work of Flavell and his students and show how these findings continue to be relevant in explaining age trends in memory accuracy. In doing this, we organize the literature into developmental acquisitions that are cognitive and social so to identify pre-existing, non-situationally-dependent internal and societal factors that impact all peoples to some degree. We also briefly address neurobiological developments that moderate these acquisitions such as changes in the prefrontal cortex that subserve monitoring, inhibition, and so on. Following this, we organize this research into endogenous and exogenous forms of suggestibility, with the former (endogenous) referring to internally generated sources, such as when in the absence of any external suggestions a lure spontaneously activates semantically related words that were never encountered during the event or when a witness's internal reveries result in the creation of memory traces as a direct consequence of the

revery as opposed to a direct observation. In contrast, exogenous sources of suggestion refer to factors that are encountered via external influences, such as when an interviewer makes a false suggestion or employs leading questions. We ultimately outline a framework with the intention of explaining and anticipating circumstances under which normative vs. reverse developmental trends may emerge.

2. Cognitive and social considerations

2.1. Knowledge

Knowledge is one of the most significant factors in memory, as it can impact every facet of recollection, including information encoded incorrectly prior to retrieval (e.g., McCutchen, 2000; Dewitt et al., 2012) as well as the retrieval process itself. Within the context of developmental trends, knowledge can cause a chain reaction, so to speak, in which its presence vs. absence can determine the potential for other endogenous and exogenous suggestibility factors. For example, without a knowledge or awareness of social expectations as they pertain to stereotypes, schema, and scripts, these sources of suggestion cannot influence memory.

Although this *may* support normative developmental trends (e.g., when age implicates a greater knowledge of social expectations, which in turn accurately align with an experience and therefore effectively aid in memory storage or retrieval), as the above suggests, neither the possession of knowledge nor its implications are *invariably* age normative. In some situations, a young child may possess greater knowledge than an older child or adult, such as for the characters of their favorite show or the layout of their school. Aspects of familiarity and increased knowledge such as these further supplements a demonstration of developmental reversals. Simultaneously, research has demonstrated that the *lack* of knowledge also affects the capacity to support developmental reversals.

For example, Ceci et al. (2007) implemented the use of multidimensional scaling (MDS) and studied 4- and 9-year-olds' susceptibility to suggestion in the context of a story about a class trip to a zoo where they observed various animals. When there was a dimension available to categorize the animals that only the older children possessed (e.g., predators, arachnids, avians, etc.), they were *more likely* to misreport that they had seen non-observed animals characterized by these dimensions (e.g., other predators or arachnids they had not actually seen) than were younger children who did not possess these dimensions in the first place, hence never encoded them and never employed them during retrieval.

Such findings suggest that knowledge is capable of both supporting accurate recall as well as impeding it, and as we will demonstrate, these opposing outcomes are each compatible with both the normative and reverse developmental trends; developmental outcomes in both directions are related to children's knowledge representations and the direction can be predicted by an understanding of their representations. The organization and interconnectedness of the knowledge structure can have a substantial impact on developmental outcomes. When younger children's representation of knowledge is more elaborate than older children's, they tend to be more likely to make associative errors, whereas the reverse is true when older children's knowledge is more elaborate—which is far more common.

2.2. Scripts

Another focus of memory researchers has been on the understanding and application of *scripts*. Scripts are temporally organized general knowledge structures that depict the sequence of normally occurring events in their proper temporal order (Fivush et al., 1992). These scripts allow us to form expectations regarding the world around us, which leads to inferential reasoning and gap-filling. The ability to accurately organize and depend on scripts develops in tandem with chronological age, but this does not mean that it ineluctably supports the normative developmental position as we will show.

In some cases, we do in fact see demonstrations of normative development. For example, in cases where a vignette follows the script accurately and the latter is used efficiently to support memory, we see a close positive correlation between chronological age and memory accuracy. As an illustration, Pillemer (1992) exposed 3.5- and 4.5-year-olds attending a preschool program to a fire drill evacuation due to a teacher who burned popcorn. The presence of a script for fire drills varied as a function of children's age. The older children had all experienced a fire drill in the past, but the younger children had not. Teachers and children exited the building and waited outside while local police and firefighters turned the alarm off and cleared the building for reentry.

Seven years later, these children were interviewed again and asked about the details of that day. The older children were significantly more accurate in their reporting of the events of the fire evacuation. The younger preschoolers' recollections were not guided by a script or causal mechanism, and as a result their stories and event narratives were more frequently devoid of a causal explanation and structure. Older preschoolers, however, paid greater attention to causality which therefore increased the temporal coherence of their narratives (Pillemer, 1992).

When older children and adults *over*-rely on their scripts, however, it can cause a developmental reversal. This can occur when an event does not accurately or completely fit into one's script; in such a case, memory can be altered by the expectations that preceded. For example, consistent with over-reliance on what usually happens when a child visits the doctor, Ornstein et al. (1998) found 6-year-olds to be more likely to accurately report that they had their heart checked with a stethoscope at a medical appointment than 4-year-olds who did not possess as elaborate of a script for doctor visits. However, when an event fails to mimic a script, it can lead to false recollections, such as claiming that their heart had been checked with a stethoscope when this had not been part of the appointment. Hence, scripts (and more broadly the interconnectedness of knowledge) can lead to both normative and reverse trends depending on the characteristics of the to-be-remembered event and their conformity with expectations based on knowledge structures like scripts and schema.

2.3. Metamemory

Metamemory—or the introspective understanding of how memory functions and the ability to patrol the workings of one's memory in order to monitor and regulate memory effectiveness—is another factor affiliated with age-related trends in memory accuracy. Researchers have charted age trends in the metamemorial processes

involved in regulating memory and the role these play in recall and recognition memory accuracy (for early treatments, see Flavell et al., 1970; Kreutzer et al., 1975; Flavell and Wellman, 1977; Wellman, 1978). Studies by Flavell et al. (1970) have demonstrated normative age trends after evaluating young children's abilities to estimate the number of words they can recall from a list, assess if and when they have committed a list to memory (i.e., "Tell us when you have memorized all of the words in the list"), recognize factors that could impede memory retrieval (e.g., studying for a test in a noisy room), and gauge the current status of their memory. Older children and adolescents possess a greater introspective awareness that enables them to better monitor their own memories and consequently intervene with appropriate strategies (see Schneider, 1999, for a review of age differences in metamemory knowledge).

As children begin to develop an overall mastery of metacognition and its affiliated processes, their overall memory performance improves (Schneider and Pressley, 1997). This is the essence of the normative developmental position. Researchers have documented a number of contextual factors that influence the efficacy of metamemorial processes. Namely, Ceci et al. (2010) found that the nature of the mental representation (how elaborately structured semantic knowledge is) influences the efficacy of metacognition. They found that when younger children's representations are elaborate (e.g., of cartoon characters from shows they frequently watch), their metacognitive awareness is significantly enhanced, whereas when their representation is impoverished vis-à-vis older children—which is usually the case—their metacognition is less efficient. This is an illustration of an important principle: knowledge not only directly influences the recollective process, but it also moderates the efficiency of underlying processes that support memory such as strategies, metacognition, and monitoring.

2.4. Arousal

Every experience invokes some degree of "stress," or *arousal*, best defined within this context as the degree of stimulation in terms of excitement, from low to high, produced while encoding stimuli (Gomes et al., 2013). Previously indistinguishable within literature, arousal is often examined in tandem with the *valence*, or amenity of a stimulus, which we address later in this review as a separate exogenous source of suggestion. The relationship between "emotion," or "mood"—typically a combination of arousal *and* valence—and memory has been investigated in the past (e.g., Gardner, 1932; Redmount, 1959), but it is only within the past several decades that arousal and valence have been analyzed separately (Tellegen, 1985; Brainerd et al., 2008a,b), and even more recently that the relationship *between* the two has been examined (Brainerd et al., 2008a,b; Brainerd, 2018).

The literature has long held and continues to maintain that heightening arousal, valence, and/or "emotion" contributes to increased memorability for events relative to neutral events (Bradley et al., 1992), but simultaneously an increased risk to *inaccurate* memories (Gardner, 1932; Brainerd et al., 2008a,b). As specific effects of low vs. high arousal circumstances have been further investigated, studies have suggested that high arousal stimuli may encourage reliance on gists over verbatim memory (Brainerd et al., 2008a,b; Bookbinder and Brainerd, 2016). This suggests that similarly to the patterns of FTT, when gist-representations are over-relied upon, susceptibility to false memories will increase.

Developmentally, we know that children are less capable of emotional regulation (see Thompson, 1991 for a review), which may suggest normative developmental trends in memory accuracy. Compared to children, adults have significantly greater emotional regulation skills, and may therefore be less susceptible to allowing heightened arousal to produce false memories. However, if trends similar to those of FTT are truly reflected in the context of arousal, we may again observe developmental reversals as age increases one's ability to create gist-representations and therefore *over-rely* on these associations.

2.5. Stereotypes

Depending on the situation, social factors such as stereotypes or security of attachment also have the potential to produce normative or reverse developmental memory trends. For example, Shapiro and Brooks (2018) found that recall of younger children (6-year-olds) showed higher rates of accuracy when thieves exhibited gender-role inconsistent characteristics than did older children for whom the stereotype about thieves being males was firmly established. However, when young children are preemptively exposed to stereotype information, they often demonstrate even higher rates of error than older children because of a combination of the stereotype and the lack of countervailing strategies. Leichtman and Ceci (1995) illustrated this normative trend by introducing preschoolers to a man named "Sam Stone" and presenting false stereotypic information about Sam's clumsiness. The children were then asked suggestive questions such as "When Sam Stone got that [teddy] bear dirty, did he do it on purpose or was it an accident?," and "Remember when Sam Stone ripped the book? Did he rip it on purpose, or on accident?," (p. 577) when he had not made any messes or ripped any books. The youngest children (3- and 4-year-olds) produced many more stereotype-congruent false memories than did their older peers. This is because they lacked the ability to "source" their memories; their "memories" were not the result of retrieving actual observations of Sam Stone, but rather merely reporting what was congruent with a stereotype of being clumsy.

Thus, the effect of stereotypes on suggestibility is similar to the effect of knowledge or scripts; all can elevate memory when they are relevant, but they also can lower recall when they are misplaced. When the development of stereotype-knowledge itself follows a normative trend and young children have not yet acquired it, then younger children's memories cannot be distorted when expectations and experience do not align; however, their memories also cannot be *served* when expectation and experience *would* accurately align, thereby neither exhibiting clear normative nor reverse trends. In contrast, once a child has successfully acquired stereotype-knowledge, as demonstrated in Ceci et al. (1995), young children may over-rely on their expectations derived from the stereotype and normative trends may emerge.

2.6. Attachment style

The relationship that a child has with its parent(s), with a particular emphasis on maternal attachment, has previously been

identified as one of the few individual difference variables that consistently predicts suggestibility (see Bruck and Melnyk, 2004, for a review). The work of Chae et al. (2014, 2021) reported that younger children were no more susceptible to false memories/suggestibility than older children if two factors were present: (1) the context was distressing (employment of the "Strange Situation," in which the parent leaves the child alone in an unfamiliar room for 5 min before being reunited), and (2) the child has a secure attachment to their mother. The idea behind this is that context (distressing vs. non-distressing settings) as well as individual differences in temperament and personality (attachment style) influence memory accuracy. This may be particularly relevant within the context of forensic interviews, where children may be alone with an unfamiliar interviewer under particularly stressful conditions for an extended period of time.

3. Endogenous sources of suggestion

3.1. Source misattribution

Another factor in play regarding memory accuracy is the ability to monitor where information came from. *Source monitoring* refers to the processes involved in making attributions about the origins of memories and beliefs (Johnson et al., 1993). A large literature documents that with some exceptions source monitoring ability increases from early childhood through adolescence and young adulthood, then declines at approximately age 50, thus resulting in an inverted U-shaped developmental function (Foley et al., 1983; Johnson et al., 1993; Fraser Parker, 1995). Especially in young children, this can take the form of misidentifying where they heard or learned information as well as difficulties distinguishing between real, experienced events vs. imagined events (e.g., Ackil and Zaragoza, 1995; Fraser Parker, 1995; Poole and Lindsay, 1995). This can have important legal consequences such as when a witness "remembers" having personally experienced something that someone else had told them about, or a witness confuses something enacted in therapy with having actually experienced it (Loftus, 1997).

As was the case with strategy development, myriad of contextual factors influences the efficacy of source monitoring, including several that moderate age differences, such as the familiarity of the interviewer (Quas et al., 2000). Specifically in a forensic context, interviews and conversations with police, attorneys, judges, juries, social workers, therapists, parents, friends, peers, teachers, and so on, can extend over many months (Ceci and Bruck, 1995). Therefore, it is important to understand children's capacity to segregate conversations and maintain accurate source tracking so experts can be sure that children are aware of whether their statements are the result of direct observation vs. something told to them by a parent or former interviewer.

3.2. Inferential reasoning

Age differences in inferential reasoning can also play a role in memory accuracy. An extensive history dating into the 1970s bears on developmental differences in drawing inferences, gap-filling, and backward causal attributions (working backward from an event to its cause). As we have noted, associations between presented items and non-presented lures typically place older individual's memory at a disadvantage (i.e., claiming they saw or heard semantically related lures), resulting in reversed age trends because younger children often lack knowledge of the associations, thus are not misled by them. For example, if participants are shown a photograph that induces them to make a false backward causal inference (e.g., after viewing a photo of a waiter mopping water at a table where a customer was seated, they later erroneously infer that they saw a photo of the customer spilling the glass of water), older children and adolescents will have more such false memories than younger children given their stronger causal knowledge structures and proficiency in back-filling (Lyons et al., 2010). However, in some cases inferential reasoning does not explicitly lure subjects into false information (e.g., generalizing that the water must have been spilled by *someone* for it to be there, despite not witnessing the incident).

3.3. Strategy acquisition

One of the most researched areas of memory development concerns strategy development, i.e., deliberately deployed procedures that are enacted to achieve a mnemonic end, such as rehearsal or creating reminders (Wellman, 1988). Nearly universally, these strategies are associated with age-normative trends. Older children and adults are more conscious of the value of implementing these strategies and are more capable of using them spontaneously, efficiently, and effectively in tandem with memory. However, despite the usefulness of these strategies, additional factors may counter or even reverse their beneficial effect. Here we analyze three of the most well-researched strategies for supporting memory—rehearsal, organization/clustering, and elaboration.

Rehearsal entails repeating (verbally or mentally) the item(s) to be remembered. The ability to implement rehearsal as a memory-support strategy develops over early childhood but appears to change with age. Young children have the ability to repeat words during a memory task (Flavell, 1966), but they do not appear to fully internalize the benefits of using this strategy to support memory until sometime between 2nd and 6th grade (Justice, 1985). Utilizing rehearsal strategies will often result in normative developmental trends although countervailing factors that are present in the same task are capable of attenuating or reversing the positive effect of strategies like rehearsal. For example, a task that would normally benefit from the use of rehearsal and therefore produce a normative trend because older individuals are more likely to have acquired it, may contain countervailing factors such as semantic associations, which work against older individuals, and could consequently cause reverse outcomes.

The *organization/classification* strategy refers to the grouping of items to be remembered into meaningful clusters or categories (e.g., fruits, animals, farm states). Recognition of the categorical structure of lists reduces the burden on memory by providing natural retrieval cues, i.e., the organizational structure itself. This strategy continues to demonstrate normative developmental trends, with organizational strategies not being implemented consistently until approximately age 8 (Best and Ornstein, 1986) despite the fact that even preschoolers

demonstrate the capability to organize on the basis on semantic meaning such as grouping all fruits or animals (Corsale and Ornstein, 1980). But as one can infer from the aforementioned discussion on knowledge, this strategy can also impede memory and ultimately yield reverse developmental trends when the same semantic knowledge underpinning organizing items also leads to semantic confusions such as the false belief that a non-presented item had been presented.

Elaboration refers to the action of making visual or verbal connections between the items to be remembered or between these items and salient objects. Establishing interrelatedness between objects or ideas serves to establish meaningful connections during encoding, which can serve as memory cues during retrieval. Again, we find that this strategy can be deployed by young children when explicitly instructed to do so, however, like rehearsal, spontaneous use of it does not appear until adolescence (Pressley and Levin, 1977). Naturally, we may anticipate seeing normative developmental trends emerge as the ability to spontaneously implement elaborative techniques develops further. As one of many examples, Beuhring and Kee (1987) found age-related increases in performance on paired-associate tasks between 5th and 12th graders. Moreover, they concluded that 96 percent of this improvement could be accounted for by increased use of the elaboration strategy.

Conversely, one of the most prolific demonstrations of elaboration has been via the Deese, Roediger, and McDermott (DRM) task, which has formed the bedrock of evidence supporting developmental reversals. The DRM standardly entails presenting participants with lists of "target" words (e.g., mad, fear, hate, rage, temper) associated with a non-present "lure" word (anger) followed by recognition and recall tests. Manipulation of factors such as strength of the associated words, speed of presenting words, length of word list, time between list exposure and testing, implementation of memory strategies such as rehearsal, and so on, and their impact on rates of false memories have been previously explored (e.g., Hancock et al., 2003; Watson et al., 2004; Cann et al., 2011; Pardilla-Delgado and Payne, 2017).

A large body of literature investigating the DRM has long reflected the positive linear relationship between likelihood of mistakenly recalling the non-present lure word and age (e.g., Norman and Schacter, 1997; Tun et al., 1998; Balota et al., 1999; Holliday et al., 2011). These trends are likely reflective of the processes affiliated with FTT, and more specifically, age-related trends in spontaneous vs. suggestion-induced false memories. FTT and the DRM have both been associated with the production of *spontaneous* false memories, or memories produced without suggestion from the external environment (Otgaar et al., 2013).

This entire field of research is based in the idea of associative activation, or the triggering of related concepts in one's mind based on exposure to an initial stimulus (Collins and Loftus, 1975; Otgaar et al., 2019). Again tying back to earlier discussion, with age typically comes the skillset of interconnecting ideas, increased knowledge, and greater ability to rely on elaboration skills; thus the strengthening of associative activation. It is when these methods coincide accurately and with experience and aid in memory storage and retrieval that normative developmental trends emerge. However, it is when expectations are broken and age coincides with an *over*-reliance on these associations that we observe developmental reversals.

4. Exogenous sources of suggestion

4.1. Valence

Another variable often studied in the context of the DRM is *valence*. As previously stated, valence refers to the amenity of a stimulus, or its *pleasantness*, ranging from positive to neutral to negative (Gomes et al., 2013). The DRM has been consistently used for examining effects of valence and/or arousal, largely given the numerous standardized lists established by experts [e.g., the Affective Norms for English Words (Bradley and Lang, 1999)]. As discussed earlier, increased arousal has been shown to heighten memorability of an event, but also potentially heighten reliance on gist traces over verbatim memories.

In contrast, valence has been shown to demonstrate a U-shaped relationship to memory, where higher valence, positive *or* negative, is shown to trigger greater memorability than neutral events (Bradley and Lang, 1999). As literature has moved further yet from seeing valence and arousal as one and begun to investigate the relationship between them, researchers have proposed that this U-shaped relationship may be attributable to the rise in arousal that results from increased valence, positive or negative (Brainerd and Bookbinder, 2019). We thereby continue to investigate age trends akin to those discussed in the context of arousal and anticipate that, situationally, normative or reverse developmental trends may emerge.

4.2. Suggestive techniques

Suggestive techniques include a large class of endogenous practices and exogenous procedures such as (mis)leading questions, imagery inductions, repetitions, false assertions, scripts, forced confabulations, forced-choice questions, and so on. Evidence of techniques such as these date back over a century, and their implications in suggestibility have been replicated across both individual studies and literature syntheses ever since (see Ceci and Bruck, 1993, for a review).

Of these techniques, misleading questioning is both one of the oldest on record and one of the most conceptually salient. For example, asking "Was the man's hat black?" when the man was not wearing a hat is an explicitly misleading question. However, later research indicated that forensic interviewers need not make explicit suggestions in order to influence memory. Implicit suggestive strategies such as stereotype-guided recollection (Leichtman and Ceci, 1995), naturally occurring conversations between peers, or overhearing adult conversations (see Principe and Schindewolf, 2012, for a review) were also shown to produce robust misinformation effects, especially in preschoolers. The effects of this were shown to linger, with children continuing to reveal signs of distortion up to 1 year after exposure to false or misleading information (e.g., Peterson et al., 2001; London et al., 2009).

4.3. Suggestion-induced false memories

Our emphasis on evaluating children's memory capabilities does not imply that adults are exempt from memory errors under

suggestive questioning. Dating back to the 1990s with the work of Loftus and her colleagues, there is evidence that memories can not only be altered, but completely created and implanted (e.g., Loftus, 1993, 1997, 2003). This literature continued to expand, examining a number of ways to implant false memories, such as through word lists (e.g., the DRM) and cognitive tasks (e.g., the Brown-Peterson task), in which experts were able to implant false memories (e.g., Roediger and McDermott, 1995; Dodson and Schacter, 2001; Ghetti et al., 2002). The various suggestive techniques and factors previously discussed, such as suggestive questioning techniques have repeatedly been demonstrated to impact the prevalence of false memories across age groups (e.g., Ackil and Zaragoza, 1998; Garven et al., 1998). The overarching trends and likelihood of obtaining false memories do, however, remain based on those of the factors influencing them.

5. Legal relevance

The most direct implication of suggestibility and its impact on memory within the legal system can be seen in eyewitness testimony. Much of the research that has been done on memory and its legal implications has been motivated by cases in which children are called to present eyewitness testimony about criminal and custodial events (Ceci and Bruck, 1995; Poole and Lamb, 1998; Gordon et al., 2001; Ornstein and Haden, 2001).

Eyewitness testimonies are based on the autobiographical memory of a person who is alleged to have witnessed or participated in a crime. Already, it is crucial to be mindful of the myriad sources of suggestion aforementioned, particularly given the customarily negatively valenced and highly arousing nature of crime. In addition to the factors previously discussed that may impact memory prior to forensic interviewing, the interviews themselves may influence memory as well. Specifically, the strategies and techniques employed by law enforcement and social service officials to obtain information or a confession can be coercive or deliberately embellished with false information.

Existing or past interrogation strategies such as the Reid technique have been known to prey on one's susceptibility to memory distortion and exacerbate the likelihood of eliciting a false confession (Kozinski, 2017). As demonstrated in research by Loftus and Pickrell (1995) and Loftus (1996, 1997), false memories are relatively simple to implant, and further studies such as those by Otgaar and his associates (e.g., Otgaar et al., 2021) have shown that increased levels of suggestibility yield higher rates of false confessions. This underscores the necessity to understand how chronological age corresponds with trends in memory development.

Given the influence of the factors discussed throughout this paper, there is no clear indication that the legal testimony of young children need be discarded altogether. It may prove advantageous to take extra precautions when interviewing young, especially preschool-aged, children, considering that age-related differences *do* exist regarding susceptibility to suggestion. But the data indicate that children still are capable of retrieving memories with a great deal of accuracy (Ornstein et al., 1992). Moreover, as we have argued throughout this paper, the possibility of reverse developmental trends underscores the inadvisability of automatically assuming that children's testimony is inferior to that of adults in all circumstances.

6. Conclusion

Experts have historically assumed the normative developmental position, believing that as age increases, susceptibility to memory distortions decreases. The underdevelopment of children's neurobiological architecture involved in source monitoring, tracking, and response inhibition; the underdevelopment of certain cognitive functions (e.g., strategies, associative knowledge, and inferential reasoning); and myriad social influences on young children (e.g., desire to please authority figures, yes-bias, insecure attachment, and peer conformity) undermines their report accuracy. However, this is not to say that there are not cases in which children—even young children—demonstrate superior memory performance to adults. As we have shown, there are times when older individuals' superior semantic knowledge and inferential skills will taint their recollections in ways that do not jeopardize younger children whose lack of knowledge insulates them.

Thus, our literature review reveals that memory trends do not ineluctably follow normative patterns, but rather that there are factors and circumstances that *have the potential* to yield reverse trends; notably, when older individuals possess greater knowledge and comprehension of the world around them that can, in certain contexts, reduce memory accuracy. Older individuals' greater ability to optimize memory-support strategies and greater knowledge can, depending on the specifics of the situation, result in greater memory accuracy, but as demonstrated, this is not always the case—particularly when the greater knowledge leads to false associations. Thus, some of the very acquisitions that lead to accurate recollection can lead to inaccurate recollection when they promote spreading activation of non-observed events.

Overall, we see evidence that when exogenous sources of suggestibility are controlled for and memory aids are employed correctly, age trends are neither clear nor consistent with regards to susceptibility to suggestion or accuracy. The primary dilemma historically has been the lack of distinction between children's true memory accuracy and their reporting accuracy, largely resulting from a lack of control for suggestive factors. As we have gained insight into an extensive list of endogenous and exogenous sources of suggestibility and now further expounded on the conditions under which these factors are subject to influencing developmental trends, the primary responsibility of continuing research will be to investigate how these factors impact one another.

Given the nature of human memory and its variability, susceptibility to memory distortion is context-specific. A task for future researchers will be to formalize the boundary conditions, which will require a dedicated program of research. The ways in which the factors that we have identified interact across every juncture of the memory process illuminates *contingent* age-related trends. The numerosity of these factors, their diversity within the existing literature, and the complexities of their interactions pose a challenge to creating a formal quantitative model that leads to specific expectations. As additional research regarding the sources of memory influence unfolds, it may become easier to create more dynamic and explicit models for predicting age-related trends in memory.

The goal of this review was to synthesize existing studies in such a way that developmental contrasts can be seen across the literature as a whole. In doing so, we have proposed a preliminary outline that can explain and predict when normative vs. reverse developmental outcomes will be observed. This review supports the notion that

children, including young children, should not be automatically discredited within the legal setting purely on the basis of their age. When children are the victims of crime or are the only eyewitness, their testimony can be extremely valuable, and no evidence exists that they are invariably inefficacious. However, to harness children's potential, it is important that we utilize this information and ensure that participants in the legal system (social workers, law enforcement officers, therapists, attorneys) are not contributing to endogenous or exogenous sources of suggestibility.

Author contributions

SC conceived and conceptualized the idea. BR wrote the manuscript with support from I-AS and SC. BR, I-AS, and SC contributed intellectually to the content within this article, as well as reviews and revisions. I-AS organized external affairs (e.g., funding). All authors contributed to the article and approved the submitted version.

Funding

Publishing fees were endowed in-part by the Cornell Open Access Publishing (COAP) Fund.

Acknowledgments

The authors would like to thank the COAP Fund for their financial contributions to this project. We extend our gratitude to the extensive list of colleagues who have supported this work through stimulating discourse and words of encouragement, especially Charles Brainerd.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ackil, J. K., and Zaragoza, M. S. (1995). Developmental differences in eyewitness suggestibility and memory for source. *J. Exp. Child Psychol.* 60, 57–83. doi: 10.1006/jecp.1995.1031

Ackil, J. K., and Zaragoza, M. S. (1998). Memorial consequences of forced confabulation: age differences in susceptibility to false memories. *Dev. Psychol.* 34, 1358–1372. doi: 10.1037/0012-1649.34.6.1358

Balota, D. A., Cortese, M. J., Duchek, J. M., Adams, D., Roediger, H. L., McDermott, K. B., et al. (1999). Veridical and false memories in healthy older adults and in dementia of the ALZHEIMER'S type. *Cogn. Neuropsychol.* 16, 361–384. doi: 10.1080/026432999380834

Best, D. L., and Ornstein, P. A. (1986). Children's generation and communication of mnemonic organizational strategies. *Dev. Psychol.* 22, 845–853. doi: 10.1037/0012-1649.22.6.845

Beuhring, T., and Kee, D. W. (1987). Developmental relationships among metamemory, elaborative strategy use, and associative memory. *J. Exp. Child Psychol.* 44, 377–400. doi: 10.1016/0022-0965(87)90041-5

Bookbinder, S. H., and Brainerd, C. J. (2016). Emotion and false memory: the context–content paradox. *Psychol. Bull.* 142, 1315–1351. doi: 10.1037/bul0000077

Bradley, M. M., Greenwald, M. K., Petry, M. C., and Lang, P. J. (1992). Remembering pictures: pleasure and arousal in memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 18, 379–390. doi: 10.1037/0278-7393.18.2.379

Bradley, M. M., and Lang, P. J., (1999). Affective norms for English words (ANEW): instruction manual and affective ratings. Technical report C-1, The Center for Research in Psychophysiology.

Brainerd, C. J. (2018). The emotional-ambiguity hypothesis: a large-scale test. *Psychol. Sci.* 29, 1706–1715. doi: 10.1177/0956797618780353

Brainerd, C. J., and Bookbinder, S. H. (2019). The semantics of emotion in false memory. *Emotion* 19, 146–159. doi: 10.1037/emo0000431

Brainerd, C. J., and Reyna, V. F. (1998). Fuzzy-trace theory and Children's false memories. *J. Exp. Child Psychol.* 71, 81–129. doi: 10.1006/jecp.1998.2464

Brainerd, C. J., and Reyna, V. F. (2001). "Fuzzy-trace theory: dual processes in memory, reasoning, and cognitive neuroscience" in *Advances in child development and behavior*. eds. H. W. Reese and R. Kail (San Diego, CA, US: Academic Press), 41–100.

Brainerd, C. J., and Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Curr. Dir. Psychol. Sci.* 11, 164–169. doi: 10.1111/1467-8721.00192

Brainerd, C. J., and Reyna, V. F. (2012). Reliability of Children's testimony in the era of developmental reversals. *Dev. Rev.* 32, 224–267. doi: 10.1016/j.dr.2012.06.008

Brainerd, C. J., Reyna, V. F., Ceci, S. J., and Holliday, R. E. (2008a). Understanding developmental reversals in false memory: reply to Ghetti (2008) and Howe (2008). *Psychol. Bull.* 134, 773–777. doi: 10.1037/0033-2909.134.5.773

Brainerd, C. J., Stein, L., Silveira, R., Rohenkohl, G., and Reyna, V. (2008b). How does negative emotion cause false memories? *Psychol. Sci.* 19, 919–925. doi: 10.1111/j.1467-9280.2008.02177.x

Bruck, M., and Ceci, S. J. (1999). The suggestibility of CHILDREN'S memory. *Annu. Rev. Psychol.* 50, 419–439. doi: 10.1146/annurev.psych.50.1.419

Bruck, M., and Melnyk, L. (2004). Individual differences in children's suggestibility: a review and synthesis. *Appl. Cogn. Psychol.* 18, 947–996. doi: 10.1002/acp.1070

Cann, D. R., McRae, K., and Katz, A. N. (2011). False recall in the Deese–Roediger–McDermott paradigm: the roles of gist and associative strength. *Q. J. Exp. Psychol.* 64, 1515–1542. doi: 10.1080/17470218.2011.560272

Ceci, S. J., and Bruck, M. (1993). Suggestibility of the child witness: a historical review and synthesis. *Psychol. Bull.* 113, 403–439. doi: 10.1037/0033-2909.113.3.403

Ceci, S. J., and Bruck, M., (1995). *Jeopardy in the courtroom: A scientific analysis of children's testimony*. American Psychological Association, Washington.

Ceci, S. J., Bruck, M., and Rosenthal, R. (1995). Children's allegations of sexual abuse: forensic and scientific issues: a reply to commentators. *Psychol. Public Policy Law* 1, 494–520. doi: 10.1037/1076-8971.1.2.494

Ceci, S. J., Fitneva, S. A., and Williams, W. M. (2010). Representational constraints on the development of memory and metamemory: a developmental–representational theory. *Psychol. Rev.* 117, 464–495. doi: 10.1037/a0019067

Ceci, S. J., Kulkofsky, S., Klemfuss, J. Z., Sweeney, C. D., and Bruck, M. (2007). Unwarranted assumptions about Children's testimonial accuracy. *Annu. Rev. Clin. Psychol.* 3, 311–328. doi: 10.1146/annurev.clinpsy.3.022806.091354

Ceci, S. J., Ross, D. F., and Toglia, M. P. (1987). Suggestibility of children's memory: Psycholegal implications. *J. Exp. Psychol. Gen.* 116, 38–49. doi: 10.1037/0096-3445.116.1.38

Chae, Y., Goodman, G. S., Larson, R. P., Augusti, E.-M., Alley, D., VanMeenen, K. M., et al. (2014). Children's memory and suggestibility about a distressing event: the role of children's and parents' attachment. *J. Exp. Child Psychol.* 123, 90–111. doi: 10.1016/j.jecp.2014.01.005

Chae, Y., Goodman, G. S., Wang, Y., Goodman, M., McWilliams, K., Shaver, P. R., et al. (2021). Parents' attachment orientation, interviewers' support, and children's memory for a mildly distressing event. *Memory* 29, 1384–1395. doi: 10.1080/09658211.2021.1988644

ChatGPT (2023). WWW Document, Available at: https://chat.openai.com (Accessed 16 April, 2023).

Collins, A. M., and Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychol. Rev.* 82, 407–428. doi: 10.1037/0033-295X.82.6.407

Constantinidis, C., and Klingberg, T. (2016). The neuroscience of working memory capacity and training. *Nat. Rev. Neurosci.* 17, 438–449. doi: 10.1038/nrn.2016.43

Corsale, K., and Ornstein, P. A. (1980). Developmental changes in children's use of semantic information in recall. *J. Exp. Child Psychol.* 30, 231–245. doi: 10.1016/0022-0965(80)90060-0

Dale, P., Loftus, E., and Rathbun, L. (1978). The influence of the form of the question on the eyewitness testimony of preschool children. *J. Psycholinguist. Res.* 7, 269–277. doi: 10.1007/BF01068110

DeWitt, M. R., Knight, J. B., Hicks, J. L., and Ball, B. H. (2012). The effects of prior knowledge on the encoding of episodic contextual details. *Psychon. Bull. Rev.* 19, 251–257. doi: 10.3758/s13423-011-0196-4

Dodson, C. S., and Schacter, D. L. (2001). If I had said it I would have remembered it: reducing false memories with a distinctiveness heuristic. *Psychon. Bull. Rev.* 8, 155–161. doi: 10.3758/BF03196152

Fivush, R., Kuebli, J., and Clubb, P. A. (1992). The structure of events and event representations: a developmental analysis. *Child Dev.* 63, 188–201. doi: 10.2307/1130912

Flavell, J. H. (1966). Role-taking and communication skills in children. *Young Child.* 21, 164–177.

Flavell, J. H., Friedrichs, A. G., and Hoyt, J. D. (1970). Developmental changes in memorization processes. *Cogn. Psychol.* 1, 324–340. doi: 10.1016/0010-0285(70)90019-8

Flavell, J. H., and Wellman, H. M. (1977). "Metamemory" in *Perspectives on the development of memory and cognition*. eds. R. V. Kail and J. W. Hagen (Hillsdale, NJ: Erlbaum).

Foley, M. A., Johnson, M. K., and Raye, C. L. (1983). Age-related changes in confusion between memories for thoughts and memories for speech. *Child Dev.* 54, 51–60. doi: 10.2307/1129860

Fraser Parker, J. (1995). Age differences in source monitoring of performed and imagined actions on immediate and delayed tests. *J. Exp. Child Psychol.* 60, 84–101. doi: 10.1006/jecp.1995.1032

Gardner, D. S. (1932). Perception and memory of witnesses. *Cornell L. Q.* 18, 391–409.

Garven, S., Wood, J. M., Malpass, R. S., and Shaw, J. S. (1998). More than suggestion: the effect of interviewing techniques from the McMartin preschool case. *J. Appl. Psychol.* 83, 347–359. doi: 10.1037/0021-9010.83.3.347

Ghetti, S., Qin, J., and Goodman, G. S. (2002). False memories in children and adults: age, distinctiveness, and subjective experience. *Dev. Psychol.* 38, 705–718. doi: 10.1037/0012-1649.38.5.705

Gomes, C. F. A., Brainerd, C. J., and Stein, L. M. (2013). Effects of emotional valence and arousal on recollective and nonrecollective recall. *J. Exp. Psychol. Learn. Mem. Cogn.* 39, 663–677. doi: 10.1037/a0028578

Gonzalves, L., Chae, Y., Wang, Y., Widaman, K. F., Bederian-Gardner, D., Goodman-Wilson, M., et al. (2022). Children's memory and suggestibility years later: age, distress, and attachment. *Appl. Cogn. Psychol.* 36, 1034–1048. doi: 10.1002/acp.3988

Goodman, G. S., Jones, O., and McLeod, C. (2017). Is there consensus about Children's memory and suggestibility? *J. Interpers. Violence* 32, 926–939. doi: 10.1177/0886260516657358

Gordon, B. N., Baker-Ward, L., and Ornstein, P. A. (2001). Children's testimony: a review of research on memory for past experiences. *Clin. Child Fam. Psychol. Rev.* 4, 157–181. doi: 10.1023/A:1011333231621

Hancock, T., Hicks, J., Marsh, R., and Ritschel, L. (2003). Measuring the activation level of critical lures in the Deese-Roediger-McDermott paradigm. *Am. J. Psychol.* 116, 1–14. doi: 10.2307/1423332

Holliday, R. E., Brainerd, C. J., and Reyna, V. F. (2011). Developmental reversals in false memory: now you see them, now you don't! *Dev. Psychol.* 47, 442–449. doi: 10.1037/a0021058

Johnson, M. K., Hashtroudi, S., and Lindsay, D. S. (1993). Source monitoring. *Psychol. Bull.* 114, 3–28. doi: 10.1037/0033-2909.114.1.3,19931101

Justice, E. M. (1985). Categorization as a preferred memory strategy: developmental changes during elementary school. *Dev. Psychol.* 21, 1105–1110. doi: 10.1037/0012-1649.21.6.1105

Kozinski, W. (2017). The Reid interrogation technique and false confessions: a time for change. *Seattle J. Soc. Just.* 16, 301–346.

Kreutzer, M. A., Leonard, C., Flavell, J. H., and Hagen, J. W. (1975). An interview study of Children's knowledge about memory. *Monogr. Soc. Res. Child Dev.* 40, 1–60. doi: 10.2307/1165955

Leichtman, M. D., and Ceci, S. J. (1995). The effects of stereotypes and suggestions on preschoolers' reports. *Dev. Psychol.* 31, 568–578. doi: 10.1037/0012-1649.31.4.568

Loftus, E. F. (1993). The reality of repressed memories. *Am. Psychol.* 48, 518–537. doi: 10.1037/0003-066X.48.5.518

Loftus, E. F. (1996). Memory distortion and false memory creation. *Bull. Am. Acad. Psychiatry Law* 24, 281–295.

Loftus, E. F. (1997). Creating false memories. *Sci. Am.* 277, 70–75. doi: 10.1038/scientificamerican0997-70

Loftus, E. F. (2003). Make-believe memories. *Am. Psychol.* 58, 867–873. doi: 10.1037/0003-066X.58.11.867

Loftus, E. F., and Pickrell, J. E. (1995). The formation of false memories. *Psychiatr. Ann.* 25, 720–725. doi: 10.3928/0048-5713-19951201-07

London, K., Bruck, M., and Melnyk, L. (2009). Post-event information affects Children's autobiographical memory after one year. *Law Hum. Behav.* 33, 344–355. doi: 10.1007/s10979-008-9147-7

Lyons, K. E., Ghetti, S., and Cornoldi, C. (2010). Age differences in the contribution of recollection and familiarity to false-memory formation: a new paradigm to examine developmental reversals. *Dev. Sci.* 13, 355–362. doi: 10.1111/j.1467-7687.2009.00889.x

McCutchen, D. (2000). Knowledge, processing, and working memory: implications for a theory of writing. *Educ. Psychol.* 35, 13–23. doi: 10.1207/S15326985EP3501_3

Melton, G. B. (1981). Children's competency to testify. *Law Hum. Behav.* 5, 73–85. doi: 10.1007/BF01048573

Mueller-Johnson, K., and Ceci, S. (2004). Memory and suggestibility in older adults: live event participation and repeated interview. *Appl. Cogn. Psychol.* 18, 1109–1127. doi: 10.1002/acp.1078

Mueller-Johnson, K., Toglia, M. P., Sweeney, C. D., and Ceci, S. J. (2007). The perceived credibility of older adults as witnesses and its relation to ageism. *Behav. Sci. Law* 25, 355–375. doi: 10.1002/bsl.765

Norman, K. A., and Schacter, D. L. (1997). False recognition in younger and older adults: exploring the characteristics of illusory memories. *Mem. Cognit.* 25, 838–848. doi: 10.3758/BF03211328

Ornstein, P. A., Gordon, B. N., and Larus, D. M. (1992). Children's memory for a personally experienced event: implications for testimony. *Appl. Cogn. Psychol.* 6, 49–60. doi: 10.1002/acp.2350060103

Ornstein, P. A., and Haden, C. A. (2001). Memory development or the development of memory? *Curr. Dir. Psychol. Sci.* 10, 202–205. doi: 10.1111/1467-8721.00149

Ornstein, P. A., Merritt, K. A., Baker-Ward, L., Furtado, E., Gordon, B. N., and Principe, G. (1998). Children's knowledge, expectation, and long-term retention. *Appl. Cogn. Psychol.* 12, 387–405. doi: 10.1002/(SICI)1099-0720(199808)12:4<387::AID-ACP574>3.0.CO;2-5

Otgaar, H., Howe, M. L., Muris, P., and Merckelbach, H. (2019). Associative activation as a mechanism underlying false memory formation. *Clin. Psychol. Sci.* 7, 191–195. doi: 10.1177/2167702618807189

Otgaar, H., Howe, M. L., Peters, M., Sauerland, M., and Raymaekers, L. (2013). Developmental trends in different types of spontaneous false memories: implications for the legal field. *Behav. Sci. Law* 31, 666–682. doi: 10.1002/bsl.2076

Otgaar, H., Schell-Leugers, J. M., Howe, M. L., Vilar, A. D. L. F., Houben, S. T. L., and Merckelbach, H. (2021). The link between suggestibility, compliance, and false confessions: a review using experimental and field studies. *Appl. Cogn. Psychol.* 35, 445–455. doi: 10.1002/acp.3788

Pardilla-Delgado, E., and Payne, J. D. (2017). The Deese-Roediger-McDermott (DRM) task: a simple cognitive paradigm to investigate false memories in the laboratory. *J. Vis. Exp.* 119:54793. doi: 10.3791/54793

Peterson, C., Moores, L., and White, G. (2001). Recounting the same events again and again: children's consistency across multiple interviews. *Appl. Cogn. Psychol.* 15, 353–371. doi: 10.1002/acp.708

Pillemer, D. B. (1992). "Preschool children's memories of personal circumstances: the fire alarm study" in *Affect and accuracy in recall: Studies of "flashbulb" memories*. eds. E. Winograd and U. Neisser (Cambridge, UK: Cambridge University Press), 121–137.

Poole, D. A., and Lamb, M. E., (1998). *Investigative interviews of children: A guide for helping professionals*. American Psychological Association, Washington, DC, US.

Poole, D. A., and Lindsay, D. S. (1995). Interviewing preschoolers: effects of nonsuggestive techniques, parental coaching, and leading questions on reports of nonexperienced events. *J. Exp. Child Psychol.* 60, 129–154. doi: 10.1006/jecp.1995.1035

Pressley, M., and Levin, J. R. (1977). Developmental differences in subjects' associative-learning strategies and performance: assessing a hypothesis. *J. Exp. Child Psychol.* 24, 431–439. doi: 10.1016/0022-0965(77)90089-3

Principe, G. F., and Schindewolf, E. (2012). Natural conversations as a source of false memories in children: implications for the testimony of young witnesses. *Dev Rev Spec Iss Child Witness Res* 32, 205–223. doi: 10.1016/j.dr.2012.06.003

Quas, J. A., Schaaf, J. M., Alexander, K. W., and Goodman, G. S. (2000). "Do you really remember it happening or do you only remember being asked about it happening?" in *Children's source monitoring in forensic contexts., in: Children's source monitoring* (Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers), 197–226.

Redmount, R. S. (1959). The psychological basis of evidence practices: memory. *J Crim Law Criminol Police Sci* 50, 249–264. doi: 10.2307/1141041

Reyna, V. F., and Kiernan, B. (1995). Children's memory and metaphorical interpretation. *Metaphor Symbol Activ* 10, 309–331. doi: 10.1207/s15327868ms1004_5

Roediger, H. L., and McDermott, K. B. (1995). Creating false memories: remembering words not presented in lists. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 803–814. doi: 10.1037/0278-7393.21.4.803

Saywitz, K. (1990). "The child as witness: experimental and clinical considerations" in *Through the eyes of the child: Obtaining self-reports from children and adolescents*. ed. A. M. La Greca (Boston: Allyn & Bacon), 329–367.

Schneider, W. (1999). "The development of metamemory in children" in *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application*. eds. D. Gopher and A. Koriat (Cambridge, MA, US: The MIT Press), 487–514.

Schneider, W., and Pressley, M., (1997). *Memory development between two and twenty, 2nd ed, memory development between two and twenty, 2nd* Edn. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US.

Shapiro, L. R., and Brooks, E. (2018). Effects of cognitive schemas on children's testimony for a simulated juvenile crime. *J. Appl. Dev. Psychol.* 57, 1–15. doi: 10.1016/j.appdev.2018.05.002

Tellegen, A. (1985). "Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report" in *Anxiety and the anxiety disorders*. eds. A. H. Tuma and J. D. Maser (Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc), 681–706.

Thompson, R. A. (1991). Emotional regulation and emotional development. *Educ. Psychol. Rev.* 3, 269–307. doi: 10.1007/BF01319934

Tun, P. A., Wingfield, A., Rosen, M. J., and Blanchard, L. (1998). Response latencies for false memories: gist-based processes in normal aging. *Psychol. Aging* 13, 230–241. doi: 10.1037/0882-7974.13.2.230

Watson, J. M., Mcdermott, K. B., and Balota, D. A. (2004). Attempting to avoid false memories in the Deese/Roediger—McDermott paradigm: assessing the combined influence of practice and warnings in young and old adults. *Mem. Cognit.* 32, 135–141. doi: 10.3758/BF03195826

Wellman, H. M. (1978). Knowledge of the interaction of memory variables: a developmental study of metamemory. *Dev. Psychol.* 14, 24–29. doi: 10.1037/0012-1649.14.1.24

Wellman, B. (1988). Structural analysis: from method and. *Soc Struct Netw Approach* 15:19.

Wright, D. B., Hanoteau, F., Parkinson, C., and Tatham, A. (2010). Perceptions about memory reliability and honesty for children of 3 to 18 years old. *Leg. Criminol. Psychol.* 15, 195–207. doi: 10.1348/135532508X400347

Zaragoza, M. S., Dahlgren, D., and Muench, J. (1992). "The role of memory impairment in Children's suggestibility" in *Development of long-term retention*. eds. M. L. Howe, C. J. Brainerd and V. F. Reyna (New York, NY: Springer), 184–216.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 24 August 2023
DOI 10.3389/fpsyg.2023.1201674

OPEN ACCESS

EDITED BY
Fabiana Battista,
University of Bari Aldo Moro, Italy

REVIEWED BY
Laura Allen,
University of Minnesota Twin Cities, United States
Erik Mac Giolla,
University of Gothenburg, Sweden

*CORRESPONDENCE
Rachel O'Donnell
✉ rachelod@iastate.edu

RECEIVED 06 April 2023
ACCEPTED 25 July 2023
PUBLISHED 24 August 2023

CITATION
O'Donnell R, Chan JCK, Foster JL and Garry M (2023) Experimental and meta-analytic evidence that source variability of misinformation does not increase eyewitness suggestibility independently of repetition of misinformation.
Front. Psychol. 14:1201674.
doi: 10.3389/fpsyg.2023.1201674

Experimental and meta-analytic evidence that source variability of misinformation does not increase eyewitness suggestibility independently of repetition of misinformation

Rachel O'Donnell[1]*, Jason C. K. Chan[1], Jeffrey L. Foster[2] and Maryanne Garry[3]

[1]Memory, Law, and Education Laboratory, Psychology Department, Iowa State University, Ames, IA, United States, [2]Department of Security Studies and Criminology, Macquarie University, Sydney, NSW, Australia, [3]School of Psychology, The University of Waikato, Hamilton, New Zealand

Considerable evidence has shown that repeating the same misinformation increases its influence (i.e., repetition effects). However, very little research has examined whether having multiple witnesses present misinformation relative to one witness (i.e., source variability) increases the influence of misinformation. In two experiments, we orthogonally manipulated repetition and source variability. Experiment 1 used written interview transcripts to deliver misinformation and showed that repetition increased eyewitness suggestibility, but source variability did not. In Experiment 2, we increased source saliency by delivering the misinformation to participants via videos instead of written interviews, such that each witness was visibly and audibly distinct. Despite this stronger manipulation, there was no effect of source variability in Experiment 2. In addition, we reported a meta-analysis ($k = 19$) for the repeated misinformation effect and a small-scale meta-analysis ($k = 8$) for the source variability effect. Results from these meta-analyses were consistent with the results of our individual experiments. Altogether, our results suggest that participants respond based on retrieval fluency rather than source-specifying information.

KEYWORDS

eyewitness memory, misinformation, repetition, source variability, eyewitness suggestibility, misinformation effect

Introduction

Research on the misinformation effect (when exposure to misleading information harms memory performance) has contributed greatly to the understanding of the fallibility of human memory. Despite its replicability, most of the research in the misinformation literature has used variants of the same three-phase paradigm, which consists of (i) participants witnessing an event, (ii) being introduced to misinformation, and (iii) taking a memory test. Most studies using this paradigm provided misinformation to participants using a single source (e.g., participants might be introduced to misinformation by reading a narrative purportedly written by a professor; Zaragoza et al., 2007; Berkowitz and Loftus, 2018). But crimes are often witnessed

by multiple people, so eyewitnesses may be introduced to misinformation multiple times and/or through multiple sources (Clark and Wells, 2008; Skagerberg and Wright, 2008). For example, co-witnesses to a crime might discuss the details of the event with each other, during which incorrect information might be introduced, and the misinformation might be repeated by the same or other co-witnesses. Given the important role that eyewitnesses play in criminal investigations, it is crucial to understand how an eyewitness' memory may be influenced by receiving the same piece of misinformation multiple times and through more than one source (e.g., from multiple people). Although some reports have shown that repeated exposure to misinformation can exacerbate its influence (Mitchell and Zaragoza, 1996; Walther et al., 2002; Ecker et al., 2011; Bright-Paul and Jarrold, 2012; Schwarz et al., 2016; Ecker et al., 2020), very little research has independently examined the effects of repetition and source variability on eyewitness suggestibility (Mitchell and Zaragoza, 1996; Foster et al., 2012).

The purpose of the current study was to examine how source variability and repetition of misinformation influence eyewitness suggestibility. Source variability was defined as the number of people who delivered the misinformation, and repetition was defined as the number of presentations of the same misinformation. Below, we first review the literature regarding the effect of repetition of misinformation on suggestibility, we then review the literature on memory conformity that pertains to source variability, and we lastly review the previous studies that have investigated both repetition and source variability of misinformation (Mitchell and Zaragoza, 1996; Foster et al., 2012).

Repeated exposure to misinformation

In general, repeated exposure to misinformation increases the misinformation effect (Mitchell and Zaragoza, 1996; Zaragoza et al., 2007; Ecker et al., 2011; Bright-Paul and Jarrold, 2012; Foster et al., 2012). This repetition effect has been observed across different participant populations (see Mitchell and Zaragoza, 1996; Bright-Paul and Jarrold, 2012) and is thought to occur as a result of increased processing fluency or increased belief in the truthfulness of the misinformation (see Arkes et al., 1991; Hassan and Barber, 2021).

The illusory truth effect (Hasher et al., 1977; Dechêne et al., 2010) suggests that repeating misinformation might increase its believability. In the illusory truth paradigm, participants are asked to rate a series of plausible statements for truthfulness ("Lithium is the lightest of all metals"). The typical finding is that repetition increases ratings of truth. There are several predominant explanations for the illusory truth effect, but the source dissociation hypothesis and the processing fluency hypothesis are most relevant to this study. The first hypothesis proposes that successive repetitions increase the processing fluency of an item, and because truth and fluency are highly correlated, people tend to use fluency as a marker for truthfulness (Arkes et al., 1991; Hassan and Barber, 2021). The second hypothesis proposes that repetition increases a statement's credibility because participants mistakenly attribute a prior presentation of the statement to an independent, outside source (Arkes et al., 1991; Roggeveen and Johar, 2002). Both the fluency and source dissociation hypotheses have received empirical support (Begg et al., 1992; Roggeveen and Johar, 2002; Henderson et al., 2021), and the mechanisms underlying each should apply regardless of whether participants are judging the truthfulness of correct statements or misinformation. Together, existing data suggest that repetition of (mis)information should increase eyewitness suggestibility.

Memory conformity, credibility, and source variability

In contrast to the voluminous literature on repetition effects (Mitchell and Zaragoza, 1996; Zaragoza et al., 2007; Ecker et al., 2011; Bright-Paul and Jarrold, 2012; Foster et al., 2012), far less research has investigated whether source variability of misinformation might influence eyewitness suggestibility (Mitchell and Zaragoza, 1996; Foster et al., 2012), but data in the memory conformity and credibility literatures can provide a basis for predictions about the effects of source variability. Memory conformity studies, unlike most misinformation studies (which are typically carried out in solitary circumstances), were originally intended to investigate how participants conform to responses made by others (social influences). In some memory conformity studies, participants receive misinformation from a confederate posing as a co-witness, and participants often mistakenly report that misinformation on a later memory test (Reysen, 2007; Goodwin et al., 2013; Thorley, 2013). Memory conformity is often studied in the context of a single co-witness, but some research has shown that participants exhibited greater conformity when misinformation was provided by two or more co-witnesses (Ost et al., 2008; Jack et al., 2014). Similarly, when the same misinformation was delivered by multiple witnesses, it was judged more convincing than when it was provided by one witness (Lindsay et al., 1986). Extrapolating from these findings, presenting misinformation from multiple sources might increase eyewitness suggestibility.

A serious problem with the above-cited studies and many others in the memory conformity literature (Walther et al., 2002; Vrij et al., 2005; Mojtahedi et al., 2018) is that they have all confounded source variability with repetition, such that when misinformation was delivered by multiple people (increased sources), it was also repeated in each successive presentation (increased repetition). These studies, therefore, do not offer insight regarding whether the effect of group size occurs because of repetition or source variability.

Relatedly, much of the work in the credibility and misinformation effect literature has shown that eyewitnesses are more susceptible to misinformation when it is presented by a more credible source than a less credible one (Dodd and Bradshaw, 1980; Smith and Ellsworth, 1987). Of particular relevance is a study conducted by Park et al. (2017). In this study, participants read fictitious crime vignettes and then made punitive judgments for the suspects and provided confidence for these judgments. Participants were then given a chance to reconsider their judgments after being provided with the average decision of other mock jurors. Importantly, Park et al. manipulated the group size of the jury and found that participants were more likely to yield to the judgment of a putatively larger group than a smaller group. This finding suggests that participants might have regarded a decision made with more sources as one with greater consensus and credibility. Taken together, the results from the aforementioned literatures (i.e., illusory truth effect, memory conformity, group size, and credibility) suggest that when multiple witnesses provide misinformation,

participants might be particularly susceptible to misinformation because of an increased perception of consensus.

Studies that investigated repetition and source variability independently

To our knowledge, only one study to date has examined the effects of both source variability and repetition of misinformation on eyewitness memory (Foster et al., 2012). In Foster et al.'s study, participants watched a short video (~ 6 min) in which an electrician stole several items from a client's house. Following a brief filler task, participants read three reports labeled as the transcript of a police interview, a written police interview, and the transcript of a follow-up interview. Participants were informed that the reports had been created by interviewing other participants in a previous experiment. To manipulate source information, each interview transcript was labeled with a witness identifier. In the one-witness (*1W*) condition, the same identifier (e.g., 9) appeared on all three transcripts; in the three-witness (*3W*) condition, different identifiers (e.g., 5, 9, 16) appeared on each transcript. Participants in the repeated-misinformation (*3X*) condition read three misleading transcripts, in which every piece of misinformation was presented once in each transcript, for a total of three presentations per misinformation. Participants in the nonrepeated-misinformation (*1X*) condition read one misleading transcript and two control transcripts, with the misleading transcript presented either first or last. Within the transcripts, each critical item (e.g., a black or blue cap, depending on the video version) was either misleading (a blue cap was *incorrectly* described as black and vice versa) or neutral (mentioning the cap without describing its color). In summary, misinformation was presented in one of four ways –one exposure via a single source (1X-1W), one exposure via three sources (1X-3W), three exposures via a single source (3X-1W), or three exposures via three sources (3X-3W). After reading the three transcripts, participants took a two-alternative forced choice (2AFC) recognition test (in which participants must choose one response option), with the correct answer and the misinformation serving as the response options. Foster and colleagues found that repetition, but not the number of sources of misinformation, reduced eyewitness memory accuracy.

Mitchell and Zaragoza (1996) also investigated a similar question, but their study did not specifically examine source variability. Here, participants viewed a short police training film and then answered 12 questions, with misinformation embedded in statements before some of the questions. Participants were presented with each set of misinformation zero, one, and three times, with each presentation occurring in a different modality (i.e., via printed paper, via audiotape, *and* via videotape) or the same modality (i.e., via printed paper, via audiotape, *or* via videotape). Finally, participants took a source memory test. Like Foster et al. (2012), Mitchell and Zaragoza found that repeated exposure to the same suggestions increased source misattributions relative to a single exposure. However, unlike Foster et al. when the misinformation was presented three times, participants in the mixed modality condition (which arguably produced more varied sources) made significantly more misattributions than those in the single modality condition. This finding demonstrated that a context manipulation – enacted via presentation modality – increased participants' suggestibility independently of repetition.

It is not clear what contributed to the discrepancies regarding the effects of misinformation presentation context between Foster et al. (2012) and Mitchell and Zaragoza (1996), but one possibility is that Foster et al. varied context via misinformation sources (such as the number of witnesses) whereas Mitchell and Zaragoza varied context via modality. The latter method might have made the context manipulation more salient to participants, thereby enhancing its effects. In particular, the source variability manipulation in Foster et al. – by marking the cover sheet of each interview transcript with a different numeric identifier – might have been too subtle. Specifically, it is possible that participants might not have paid attention to the witness identifier when they read the interview transcripts. If this were the case, participants in the three-witness condition would not remember that they had read transcripts allegedly produced by three different people, thereby rendering the source variability manipulation ineffective. Even if participants had attended the cover page, the written reports did not differ in any perceptually obvious ways, so it might be difficult for participants to distinguish the sources. In two preregistered experiments, we sought to further investigate the effects of repetition and source variability on eyewitness suggestibility. After attempting to conceptually replicate Foster et al.'s study in our Experiment 1, we aimed to boost the salience of our source manipulation in an ecologically realistic manner in Experiment 2.

The current experiments

The goal of the present study was to examine the extent to which source variability and repetition of misinformation influence eyewitness suggestibility. Both experiments were preregistered on the Open Science Framework (OSF), and our experimental materials and data are available at https://osf.io/9zpfk/?view_only=f95ed70720c742d48296fa3b92891ed7. In addition to the two experiments, we also conducted two non-preregistered meta-analyses to further examine the influence of our independent variables (repetition and source variability) on the misinformation effect. We report the results of these meta-analyses at the end of our results section before the General Discussion. To briefly preview, only the current studies and Foster et al. have independently examined the influence of source variability on the misinformation effect, so the source variability meta-analysis included data from only those studies.

Experiment 1 was an attempted conceptual replication of Foster et al.'s Experiment 1[1] Foster et al. (2012) using novel materials. We hypothesized that Experiment 1 would replicate the results of Foster et al., such that the repetition manipulation (three presentations of misinformation relative to one) would decrease participants' response accuracy, but the source variability manipulation (three sources of misinformation relative to one) would not.

In Experiment 2, we attempted to create a more salient source variability manipulation. To this end, we presented the interviews as videos, rather than written transcripts, with three different actors. Some research has suggested that misinformation delivered "directly" – by providing social cues like appearance and mannerisms – creates

1 Instead of eyewitness memory, Experiment 2 in Foster et al. concerned jury decision making, so it is not relevant to our research purpose.

a stronger misinformation effect than misinformation delivered "indirectly" (one that does not provide social cues; Gabbert et al., 2004; Blank et al., 2013). Although the delivery in this case was not done in person, the videos provided rich source-specifying information about each witness (e.g., they looked different, sounded different, and had different mannerisms), and participants could draw on these distinctive, source-specifying details to distinguish the sources. Therefore, we hypothesized that participants would be more suggestible when they received misinformation from multiple people (via videos) relative to one person, especially when the misinformation was repeated.

The design and procedure of both experiments were modeled after Foster et al. (2012). In our Experiment 1, all participants viewed a video and then read three interview transcripts, and misinformation was presented once (1X) or three times (3X), either from one witness (1W) or from three witnesses (3W). In the 1X condition, the misinformation was presented in only one interview, and this interview was presented either first or last. In addition to replicating these conditions from Foster et al. we also created an *additional* 1X condition that distributed the misinformation throughout all three interviews and termed these the *1X distributed* conditions so that every interview presented misinformation, regardless of whether one or three witnesses provided misinformation. Therefore, each experiment had six conditions–(i) 1W-3X, (ii) 3W-3X, (iii) 1W-1X, (iv) 3W-1X, (v) 1W-1X distributed, and (vi) 3W-1X distributed (see Table 1). Figure 1 depicts distribution of the critical items visually.

See Figure 2 for an illustration of the procedure. In Experiment 1, participants first watched an encoding event that depicted a robbery and then read three interview transcripts. The interviews were formatted as an initial interview, a follow-up interview, and a deposition excerpt. Modeling after Foster et al. (2012), each interview had a cover page that described when the interview occurred along with a large, handwritten code that indicated who provided the interview (e.g., 9 for the 1W condition, and 5, 9, and 16 for the 3W condition). Finally, participants took a 2AFC recognition test. The design and procedure of Experiment 2 were the same as Experiment 1, except that each interview was presented as a video featuring different actresses to ensure that the source differences were obvious to participants.

Experiment 1

Participants

A power analysis was conducted to determine sample size. The estimated effect size of misinformation repetition was $d = 0.64$ based on data from Foster et al. (2012). Because Foster et al. did not report a significant effect of source variability on suggestibility, we chose the smallest effect size of interest ($d = 0.25$). We conducted a power analysis for comparison of a main effect, with a Cohen's d of 0.25 and power of 0.50 (one-tailed, $\alpha = 0.05$). The minimum sample size per group was 88 (or 44 per condition), so we aimed to collect data from 264 participants. Note that this sample size provided 0.99 power to detect the repetition main effect of $d = 0.64$ in a two-tailed test at alpha = 0.05. Participants were undergraduate students from Iowa State University who participated for course credit.

TABLE 1 Distribution of misleading claims in each condition.

Condition	Distribution of misleading claims
1W-3X	One witness made the same six misleading claims in each of the three interviews
3W-3X	Three different witnesses made the same six misleading claims in each of the three interviews
1W-1X	One witness made six misleading claims in only one interview
3W-1X	Three witnesses, one of whom made six misleading claims in only one interview
1W-1X Distributed	One witness made two misleading claims in each of the three interviews
3W-1X Distributed	Three witnesses each made two misleading claims in each of the three interviews

Underlines indicate conditions that were in Foster et al. (2012).

All data were collected online via Qualtrics due to COVID-19. A total of 310 participants completed Experiment 1,[2] but data from 43 participants were excluded from analysis (see Table 2 for exclusions and demographic information). The exclusion criteria were preregistered before data collection. Most participants were excluded based on their responses to the survey at the conclusion of the experiment, in which participants self-reported their proficiency in English, if they took the experiment seriously, edited the video in any way, had seen the encoding event before, or experienced any technical issues during the experiment. Other participants were excluded based on their responses to attention checks (Captcha, participation in the filler activities) or survey metadata (duration, devices). The final sample included 267 participants.

Design

In addition to the six between-subjects conditions (illustrated in Table 1), item type (misleading or neutral) was manipulated within-subjects. For the 1X conditions, all of the critical items were presented in either the first or last interview (loading: first, last).

Materials and procedure

The experiment contained five phases (see Figure 2). In Phase 1, participants watched a 20-min excerpt of an episode from the Canadian television show *Flashpoint* (season 1 episode 5). In the video, a former security guard named George attempted to rob the bank where he was employed. The police were called to the scene, and Sergeant Gregory Parker negotiated with George, but George threatened to kill the hostages. The video ended after the police rescued all the hostages except for the bank manager Ruth.

In Phase 2, participants completed a 15-min filler task in which they worked on two Sudoku puzzles (see OSF page for materials), and

2 An additional 112 participants began the experiment but never completed it (91% of these participants completed only 1–2% of the experiment).

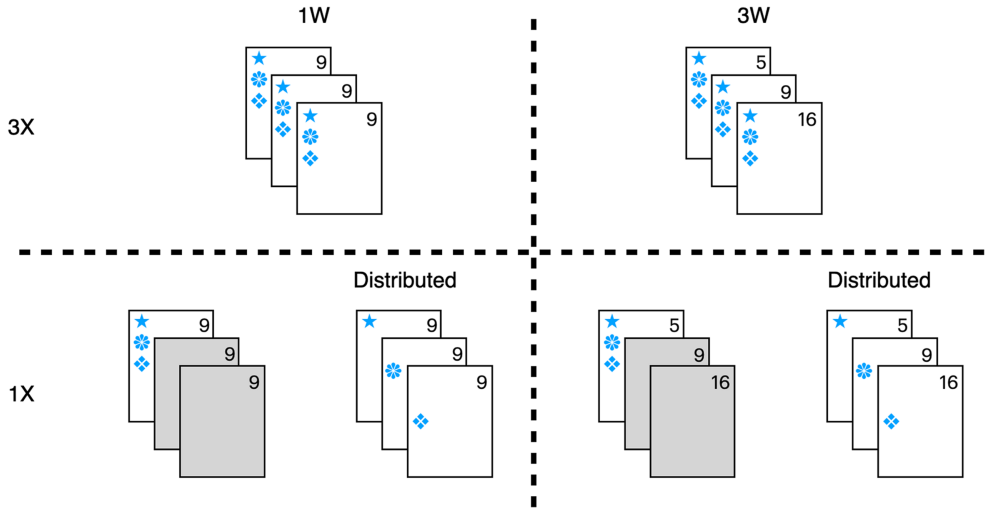


FIGURE 1
Illustration of critical items contained in each interview type and condition. Control interviews, with no misleading items, are indicated in gray. Each symbol represents a different piece of misinformation. In the 3X condition, each piece of misinformation appeared in all three interviews. In the 1X condition, all misinformation appeared only once and in a single interview. In the 1X-Distributed condition, all misinformation appeared only once, but the misinformation was distributed across three interviews.

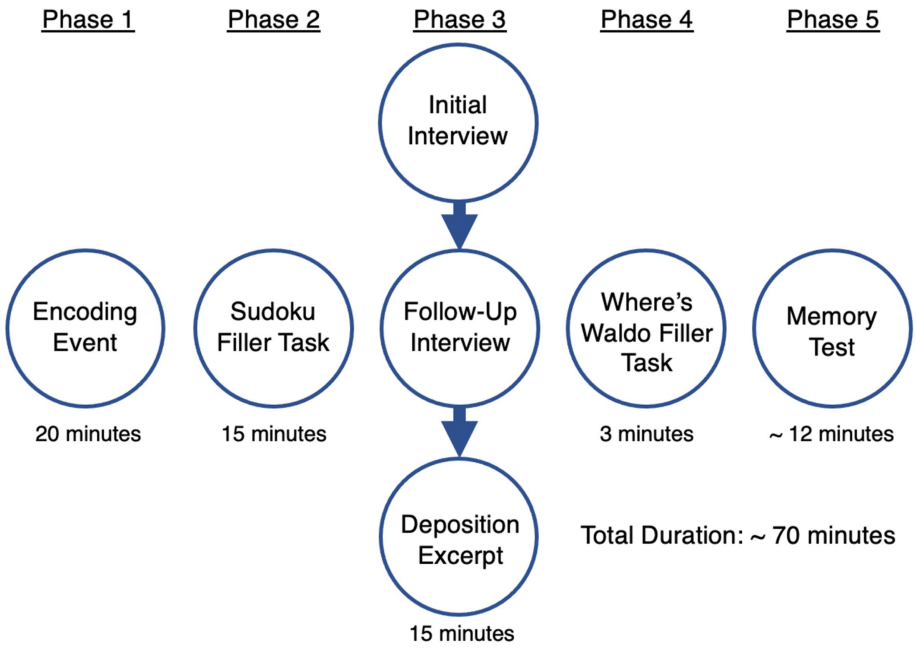


FIGURE 2
Illustration of procedure for experiments 1 and 2. The figure shows the estimated completion time for Experiment 1. Experiment 2's estimated completion time was ~90 min due to the length of the interview videos.

participants were automatically advanced to Phase 3 after the task.

In Phase 3, participants read three interviews. They were informed that a highly trained experimenter had conducted the interviews with participants from a previous experiment. The interviews were presented as an initial interview, a follow-up interview, and a simulated deposition excerpt. The interview transcripts were presented sequentially without breaks. Reading of the interviews was self-paced, but participants were required to spend at least 3 min on each interview. Each interview had a cover page with the handwritten eyewitness identifier (see the top panel of Figure 3) and the day the interview was conducted ("Day of Event," "Day After Event," and "Two Days After Event").

In the 3X condition, all 12 critical items appeared in each of the three interview transcripts. In the 1X condition, all critical items appeared in only the first or last interview. In the 1X

10.3389/fpsyg.2023.1201674

TABLE 2 Number of excluded participants, participants per condition, and demographic information in Experiments 1 and 2.

	E1	E2
Reason for exclusion		
Completed experiment in more than one session	16	9
Did not complete filler tasks (Sudoku, Where's Waldo)	8	8
Edited, paused, or rewatched encoding event or interviews	8	26
Self-reported being not serious or not alert during experiment	5	19
Previously seen encoding event (within last six months)	3	8
Self-reported taking notes during encoding event or interviews	2	3
Duration of experiment exceeded 2 h	1	10
Experienced technical issues (i.e., W-Fi connection)	–	6
Self-reported low English language proficiency	–	4
Completed the study on a mobile device	–	3
Did not agree to the conditions on the consent form	–	3
Participants retained per condition		
1W-3X	44	45
3W-3X	45	45
1W-1X	44	45
3W-1X	46	45
1W-1X Distributed	44	44
3W-1X Distributed	44	44
Ethnicity		
White or Caucasian	82%	79%
Hispanic or Latinx	5%	6%
East Asian	4%	3%
Black or African American	2%	3%
South/Southeast Asian	2%	3%
West Asian/Middle Eastern	2%	2%
Native Hawaiian or Other Pacific Islander	1%	1%
Other	1%	2%
Chose not to respond	< 1%	< 1%
Gender		
Female	57%	57%
Male	42%	42%
Other	< 1%	1%

distributed condition, the critical items were spread across the three interviews, with each interview presenting four different items (see Figure 1). For example, the name of the bank, "City Central," was a critical item. The *misleading* version of this critical detail named the bank "City Towers," whereas the *neutral* version omitted the bank's name. The assignment of each critical detail as misleading or neutral was counterbalanced across participants. Each interview also contained 12 filler items that were presented either three times in the 3X condition or once in the 1X conditions. These filler items were included so that the memory test queried both items that were presented correctly and incorrectly in the interviews rather than querying only omitted (neutral) or incorrect (misled) items.

Pilot testing was conducted to ensure the critical items produced a significant misinformation effect. Here, participants completed a condensed version of Experiment 1 without the main manipulations of repetition and source variability ($N = 73$ total participants in two rounds of pilot testing). The single interview contained 14 critical items (misleading or neutral) and 14 filler items. After an item analysis was conducted, two critical items were removed (low misinformation effect, < 5%), and two filler items were removed (ceiling performance, > 97%). The remaining 12 items produced an average misinformation effect of 21% ($d = 1.12$). Additional details about pilot testing can be found on the OSF project page.

Phase 4 included a 3-min "Where's Waldo" filler task. Participants searched for the cartoon character Waldo (see OSF page for materials) in four pictures.

In Phase 5, participants completed a 24-question 2AFC recognition test (see OSF page for all questions). Twelve questions queried the 12 critical items (where the answer choices were either the correct answer or the misinformation), and 12 questions queried the filler items (where the answer choices were either the correct answer or an incorrect foil). Four filler questions with the highest accuracy ($M = 0.89$) in the pilot were always presented at the beginning of the test so that it would not be perceived as too difficult. The order of the remaining questions and answer choices (for all questions) was randomized. Participants also rated their confidence for each question on a scale from 1 to 5, with 1 meaning not at all confident and 5 meaning very confident. At the conclusion of the study, all participants completed a brief demographic questionnaire with manipulation check questions.[3] Following completion of the questionnaire, all participants were debriefed.

Results and discussion

We first report results of the same analyses as Foster et al. (2012) to determine whether we successfully replicated their recognition results (a 2x2x2 repeated measures ANOVA), and then report follow up *t*-tests to examine both the effect of repeating misinformation and multiple sources of misinformation on eyewitness accuracy. We report these analyses for the conditions that most closely replicated Foster et al. first, and then those that controlled for the distribution of misinformation. Finally, we conducted confidence-accuracy calibration analysis to further examine how item type (misled items vs. neutral items) influenced that relationship (this analysis was not conducted in Foster et al. and as such, is not the target of the replication).

We first reported the 3X vs. 1X comparison (the conditions most similar to Foster et al.) and then the 3X vs. 1X distributed comparison. An independent samples t-test showed that in the 1X condition, participants' recognition accuracy did not differ significantly

3 Participants reported how many witnesses they believed had been interviewed and rated the credibility, accuracy, and consistency of each witness on a 1–5 Likert scale. Unfortunately, in hindsight, the questions were poorly phrased. Participants clearly did not interpret the questions correctly, as indicated by their answers to the question "How many eyewitnesses were interviewed?" ($M = 8.39$, $SD = 5.89$, range = 0–32).

10.3389/fpsyg.2023.1201674

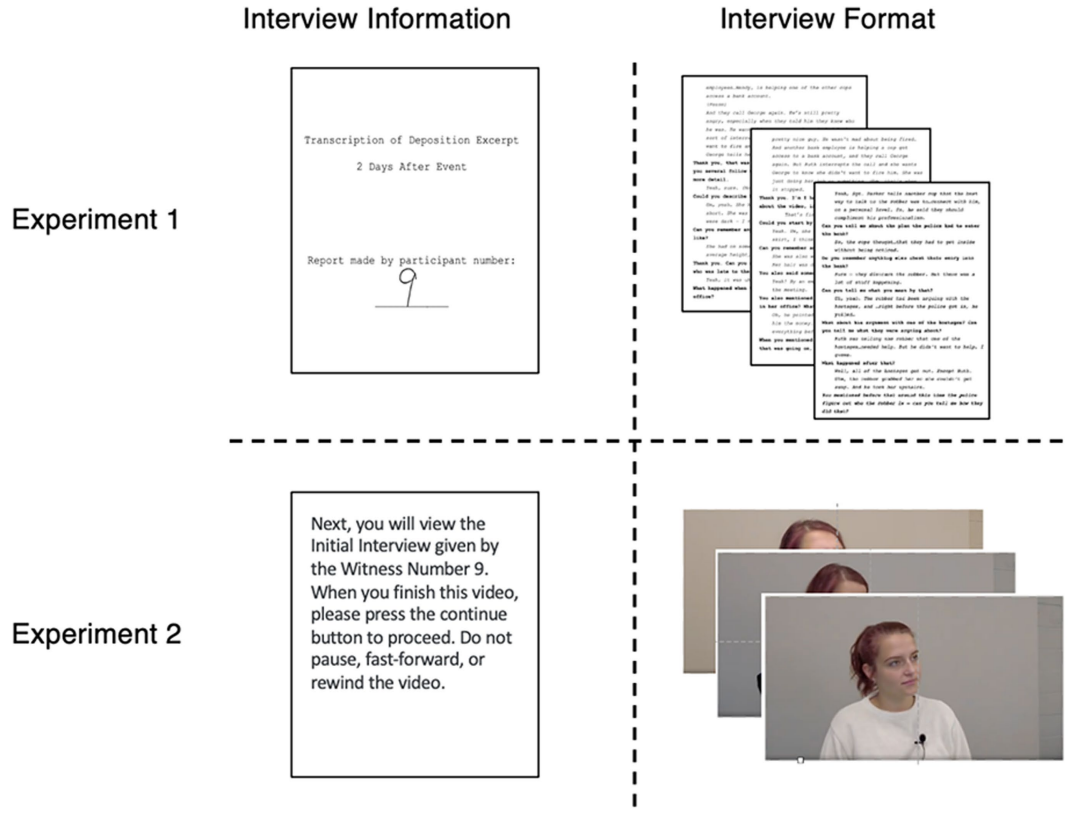


FIGURE 3
Illustration of interview information and format for Experiments 1 and 2.

regardless of whether misinformation was presented in the first or third interview (all $ts < 0.92$, $ps > 0.113$), so the remaining analyses were collapsed across this variable. The selection rates for the filler items are displayed in Table 3.

The effect of repetition and source variability on accuracy

Replication of Foster et al.'s conditions

The most important findings are shown in Figure 4. Replicating the main findings from Foster et al. (2012), repetition of misinformation reduced participants' recognition accuracy, but having three witnesses present misinformation did not affect recognition accuracy relative to one witness. The repetition results can be seen by comparing the left panel (3X) to the middle panel (1X) of Figure 4, and the source variability results can be seen by comparing the first (1W) to the second pair (3W) of bars within each panel.

The above impressions were realized in the results of a 2(repetition: 1X, 3X) × 2(source variability: 1W, 3W) × 2(item type: neutral, misled) repeated measures ANOVA with recognition accuracy (hit rate) as the dependent variable (Figure 4). This ANOVA revealed a main effect of item type, $F(1, 175) = 78.90$, $p < 0.001$, $d = 0.66$, which showed that participants' accuracy was lower for misled items ($M = 0.51$, $SD = 0.21$) than neutral items ($M = 0.69$, $SD = 0.20$) – a misinformation effect. There was also a main effect of repetition, $F(1, 175) = 6.89$, $p = 0.009$, $d = 0.20$, a nonsignificant effect of source, $F(1, 175) = 0.14$, $p = 0.713$, $d = 0.03$, and a nonsignificant interaction between item type and repetition, $F(1, 175) = 3.51$, $p = 0.063$, $\eta_p^2 = 0.02$. All other effects were not significant, $Fs < 1.00$, $ps > 3.19$.

TABLE 3 Recognition performance for filler items in Experiments 1 and 2 per condition.

	Accuracy
Experiment 1	
1W-3X	0.91 (0.12)
3W-3X	0.93 (0.11)
1W-1X	0.93 (0.13)
3W-1X	0.89 (0.13)
1W-1X Distributed	0.88 (0.14)
3W-1X Distributed	0.87 (0.13)
Experiment 2	
1W-3X	0.89 (0.12)
3W-3X	0.89 (0.14)
1W-1X	0.85 (0.18)
3W-1X	0.88 (0.14)
1W-1X Distributed	0.84 (0.15)
3W-1X Distributed	0.81 (0.16)

Standard deviations are represented in parentheses.

Following Foster et al. (2012), we assessed the effects of repeating misinformation on eyewitness accuracy in separate t-tests. Critically,

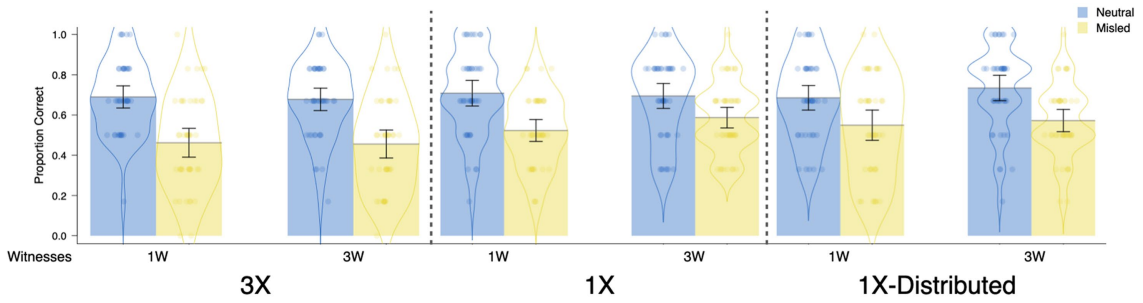


FIGURE 4
Proportion correct by item type in Experiment 1 as a function of repetition and sources. Each dot represents the data of an individual participant. Jitter was introduced to disperse the data points horizontally for visualization purposes. The violin element displays data density.

repetition of misinformation reduced recognition accuracy for the misled items ($M_{1X} = 0.56$, $M_{3X} = 0.46$), $t(177) = 0.60$, $p = 0.002$, $d = 0.47$, but not for the neutral items ($M_{1X} = 0.70$, $M_{3X} = 0.68$), $t(177) = 0.60$, $p = 0.547$, $d = 0.09$. In contrast to these results, having three witnesses deliver misinformation did not reduce recognition accuracy relative to having one witness deliver the same misinformation (accuracy for misled items: $M_{1W} = 0.49$, $M_{3W} = 0.52$, $t[177] = 0.94$, $p = 0.350$, $d = -0.14$; accuracy for neutral items: $M_{1W} = 0.70$, $M_{3W} = 0.69$, $t[177] = 0.44$, $p = 0.662$, $d = 0.07$).

Comparisons that controlled for the distribution of misinformation across interviews

Overall, the same conclusions as above were reached when we distributed the nonrepeated misinformation across all three interview transcripts (rather than presenting them in a single interview), such that repetition, but not source variability, increased the misinformation effect.

The following comparisons included the 1W-1X distributed, 1W-3X, 3W-1X distributed, and 3W-3X conditions. We again conducted a 2 × 2 × 2 ANOVA (Figure 4). There was again a main effect of item type, $F(1, 173) = 82.29$, $p < 0.001$, $d = 0.68$, which showed that participants were less accurate for misled items ($M = 0.51$, $SD = 0.23$) compared to neutral items ($M = 0.70$, $SD = 0.19$). There was also a main effect of repetition, $F(1, 173) = 7.16$, $p = 0.008$, $d = 0.20$, a nonsignificant effect of source, $F(1, 173) = 0.31$, $p = 0.576$, $d = 0.04$, and a nonsignificant interaction between item type and repetition, $F(1, 173) = 3.38$, $p = 0.068$, $\eta_p^2 = 0.02$. All other main effects and interactions were nonsignificant, $Fs < 0.89$, $ps > 0.348$.

We again replicated the key results in Foster et al. (2012). Specifically, repetition of misinformation reduced participants' accuracy for the misled items ($M_{1X} = 0.56$, $M_{3X} = 0.46$), $t(175) = 3.03$, $p = 0.003$, $d = 0.46$, but not for the neutral items ($M_{1X} = 0.71$, $M_{3X} = 0.68$), $t(175) = 0.90$, $p = 0.369$, $d = 0.14$. Moreover, the source variability manipulation did not influence participants' accuracy for both misled items, $M_{1W} = 0.51$, $M_{3W} = 0.51$, $t(175) = 0.22$, $p = 0.824$, $d = -0.03$, and neutral items, $M_{1W} = 0.69$, $M_{3W} = 0.71$, $t(175) = 0.62$, $p = 0.535$, $d = -0.09$.

Across both comparisons in Experiment 1, we replicated the critical pattern of results found in Foster et al. (2012), such that repeating the same piece of misinformation three times reduced participants' accuracy relative to presenting misinformation only once. In addition, we also found no effect of source variability – participants were no less accurate when they read interview transcripts marked as coming from three witnesses as opposed to a single witness.

Confidence-accuracy calibration by item type

To examine whether the relationship between confidence and accuracy varied by item type (for the critical items), we conducted a multilevel logistic regression analysis. We did not anticipate that either of the independent variables would affect the confidence-accuracy relationship, so we did not include them in the model. The multilevel model included data from all participants. Response accuracy (0 and 1) served as the dependent variable, and we regressed this variable on confidence, item type, and their interaction as fixed effects factors. The intercept was allowed to vary across participants as a random effects factor. We did not include any additional random effects factors because the model failed to converge when they were added.

The most important result was a significant interaction between confidence and item type, $B = 0.32$, $SE = 0.06$, $z = 5.767$, $p < 0.001$, such that the confidence-accuracy relationship was much stronger for neutral items, $B = 0.49$, $SE = 0.04$, $z = 12.27$, $p < 0.001$, than for misled items, $B = 0.17$, $SE = 0.04$, $z = 4.42$, $p < 0.001$ – a pattern that is readily apparent in Figure 5. In fact, when viewing the observed data points, the confidence-accuracy relationship for the misled item was essentially flat, with participants performing at close to chance level across the entire confidence range. In contrast, as participants' confidence rose, so did their recognition accuracy for the neutral items. Therefore, encountering misinformation severely undermined the diagnosticity of eyewitness confidence (Chan et al., 2022).

Experiment 2

The goal of Experiment 2 was to address the potential concern that the source variability manipulation in Experiment 1 – via verbal instructions and a digit written on the cover page of the written transcript – was too weak to reveal an effect. To this end, we attempted to provide participants with more obvious source-specifying information in Experiment 2 by showing video interviews featuring different actresses.

10.3389/fpsyg.2023.1201674

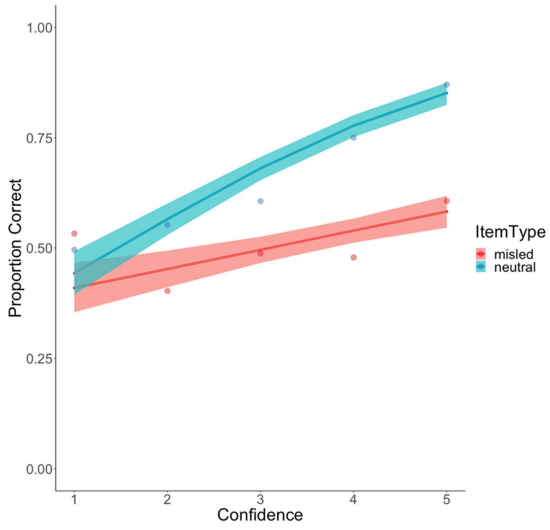


FIGURE 5
Confidence accuracy calibration as a function of item type in Experiment 1. Data points represent observed probabilities, and bands represent 0.95 CI for the fitted multilevel regression lines.

Participants

As in Experiment 1, we aimed to collect data from 44 participants per condition. A total of 367 participants completed the experiment, but data from 99 participants were removed based on the exclusion criteria listed in our preregistration (see Table 2). An additional 144 participants began the experiment but never completed it (~ 83% of these participants completed 0–1% of the experiment). The final sample contained 268 participants.

Materials and procedure

The materials and procedure for Experiment 2 were identical to Experiment 1 except for the interview videos. Each video depicted an interviewer and an interviewee having a conversation. The scripts for the videos were identical to those in Experiment 1, with the addition of natural pauses and vocalized fillers (e.g., uhm, like, okay) to increase realism. Before each interview, participants were shown a statement that contained the witness identifier and the day on which the interview took place. At all times, the videos showed either the interviewee or the interviewer (see Figure 6). Three women acted as interviewees, and their order of appearance across the interviews was counterbalanced across participants. Because participants were told that the interviews occurred across three days, all actresses wore different clothes for each interview.

Results and discussion

Accuracy for the filler questions is presented in Table 3. Because accuracy did not significantly differ in the 1X condition regardless of whether the misinformation was presented in the first or third interview (all $ts < 1.04$, $ps > 0.304$), all analyses reported below were collapsed across this variable.

The effect of repetition and source variability on accuracy

Replication of Foster et al.'s conditions

Figure 7 shows the critical findings from Experiment 2. Replicating Experiment 1, repetition of misinformation reduced participants' accuracy. However, contrary to our expectations (given that we increased the salience of manipulation), we did not find an effect of source variability. That is, there was no difference in participants' recognition accuracy regardless of whether they received misinformation from three people or from one person.

In other words, a repeated measures ANOVA (the same as in Experiment 1) demonstrated a main effect of item type, $F(1, 176) = 54.39$, $p < 0.001$, $d = 0.55$, which revealed a misinformation effect, such that participants were less accurate for misled items ($M = 0.58$, $SD = 0.24$) than for neutral items ($M = 0.72$, $SD = 0.17$). Moreover, there was an interaction between item type and repetition, $F(1, 176) = 4.45$, $p = 0.036$, $\eta_p^2 = 0.03$. All other main effects and interactions were not significant, $Fs < 2.57$, $ps > 0.111$.

The interaction between item type and repetition demonstrated that we again replicated the critical results of Foster et al. (2012), such that repetition of misinformation reduced participant's accuracy on the misled items ($M_{1X} = 0.61$, $M_{3X} = 0.54$), $t(178) = 2.29$, $p = 0.023$, $d = 0.34$, but not on the neutral items ($M_{1X} = 0.72$, $M_{3X} = 0.73$), $t(178) = 0.20$, $p = 0.843$, $d = 0.03$. In addition, although we expected an effect of source variability with the more powerful manipulation in this experiment, increasing the variability of sources did not lead to any significant differences in participants' recognition accuracy for both misled ($M_{1W} = 0.58$, $M_{3W} = 0.57$), $t(178) = 0.38$, $p = 0.702$, $d = 0.06$, and neutral items ($M_{1W} = 0.72$, $M_{3W} = 0.73$), $t(178) = 0.37$, $p = 0.712$, $d = 0.06$. We further examined the implications of this null effect in the General Discussion.

Comparisons that controlled for the distribution of misinformation across interviews

The right side of Figure 7 shows the critical results for these conditions. Overall, contrary to our expectations, we did not find an effect of either the repetition or source variability manipulations. That is, participants' recognition accuracy was not influenced by either repetition or source variability.

In other words, when we conducted the analysis to examine the effects of repetition and source variability on accuracy for the 3X conditions and the 1X distributed conditions (see the right side of Figure 7), the analysis only revealed a misinformation effect, $F(1, 174) = 89.99$, $p < 0.001$, $d = 0.71$, and a surprising interaction between repetition and source variability, $F(1, 174) = 4.35$, $p = 0.039$, $\eta_p^2 = 0.02$. We caution against overinterpreting this interaction given that (i) we did not predict it, and (ii) the effect was small, and (iii) this interaction collapsed across item type, which was the most influential variable. But perhaps most importantly, repetition and item type did not interact, $F(1, 174) = 0.35$, $p = 0.553$, $\eta_p^2 = <0.01$. All other main effects and interactions were nonsignificant, $Fs < 1.56$, $ps > 0.213$.

Contrary to our expectations, we did not find that repetition of misinformation decreased recognition accuracy for the misled items ($M_{1X} = 0.57$, $M_{3X} = 0.54$), $t(176) = 1.21$, $p = 0.229$, $d = 0.18$, and as expected, repetition of misinformation did not influence accuracy for the neutral items ($M_{1X} = 0.74$, $M_{3X} = 0.73$), $t(176) = 0.65$, $p = 0.518$, $d = 0.10$. In addition, we again found that the source variability

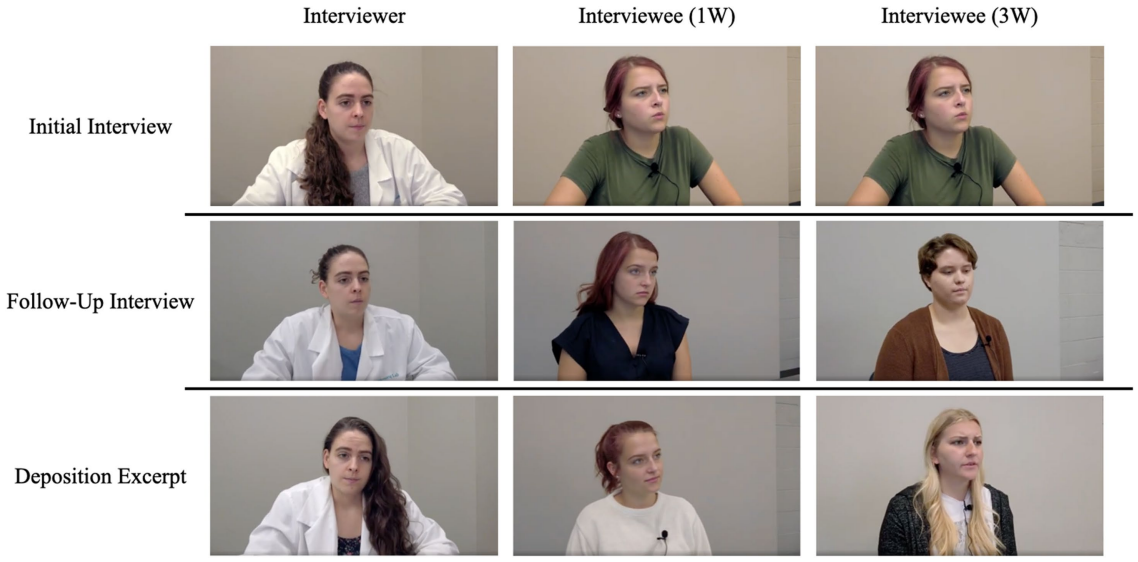


FIGURE 6
Depiction of interviewer and interviewees for 1W and 3W conditions by interview type. The interviewer remained the same across interviews. In the 1W condition, participants viewed the same witness wearing different clothes for each interview. In the 3W condition, participants viewed three different witnesses. The order of witnesses was randomized and counterbalanced.

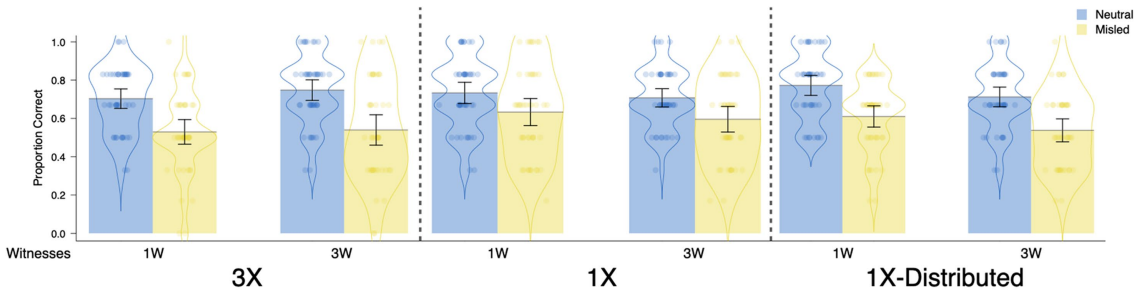


FIGURE 7
Proportion correct by item type in Experiment 2 as a function of repetition and sources. Each dot represents the data of an individual participant. Jitter was introduced to disperse the data points horizontally for visualization purposes. The violin element displays data density.

manipulation did not influence recognition accuracy for both misled ($M_{1W} = 0.57$, $M_{3W} = 0.54$), $t(176) = 0.94$, $p = 0.348$. $d = 0.14$, and neutral items ($M_{1W} = 0.74$, $M_{3W} = 0.73$), $t(176) = 0.27$, $p = 0.786$, $d = 0.04$.

Confidence-accuracy calibration by item type

In Experiment 1, we found that misinformation flattened the confidence-accuracy relationship in participants' responses. We conducted the same multilevel logistic regression for the data in Experiment 2 and found a similar pattern of results. Specifically, there is a significant interaction between confidence and item type, $B = 0.15$, $SE = 0.06$, $z = 2.63$, $p = 0.008$. The confidence-accuracy relationship was stronger for neutral items, $B = 0.44$, $SE = 0.04$, $z = 10.67$, $p < 0.001$, than for misled items, $B = 0.29$, $SE = 0.04$, $z = 7.18$, $p < 0.001$, although this difference was not as dramatic as that in Experiment 1 (see Figure 8). When we examined the observed data, it was clear that recognition performance remained close to chance for the *misled items* across confidence levels 1 to 4, but participants achieved substantially better performance when they reached the highest level of confidence. In contrast, for the *neutral items*, recognition accuracy was consistently above chance and rose with increasing confidence.

Overall, in Experiment 2, we replicated the pattern of results shown in Foster et al. (2012), namely, that repetition of misinformation harmed recognition accuracy. But in the 3X vs. 1X distributed comparison, we did not replicate this effect. It is unclear why participants in Experiment 2's 1X distributed condition selected the misinformation at a higher rate than was typical compared to the other 1X conditions in this study. With no better explanation, we believe this result can be attributed to a sampling error. Moreover, despite increasing the salience of the source manipulation, presenting misinformation from one or three witnesses did not influence participants' accuracy. Finally, we consistently demonstrated that the confidence-accuracy relationship was well-calibrated for neutral items, but the introduction of misinformation flattened this relationship.

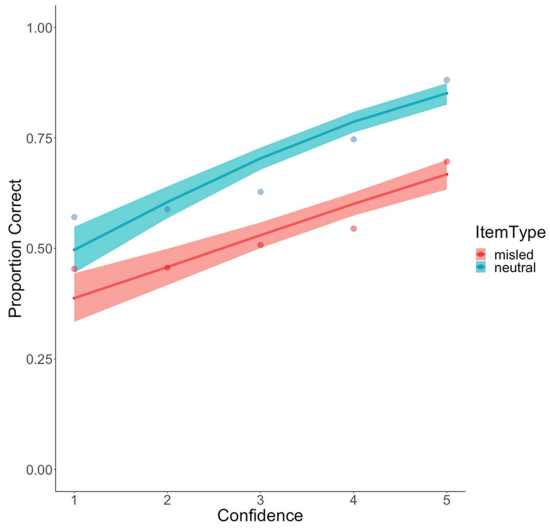


FIGURE 8
Confidence accuracy calibration as a function of item type in Experiment 2. Data points represent observed probabilities, and bands represent 0.95 CI for the fitted multilevel regression lines.

Does repeating misinformation increase its influence? A meta-analysis

Repetition effects have a long history in memory research (since Ebbinghaus, 1964, originally published in 1885) and have been studied quite extensively in the literature on erroneous memory. For example, research on imagination inflation (e.g., Garry et al., 1996; Thomas et al., 2003) and ironic effects of repetition (e.g., Jacoby, 1999; Benjamin, 2001) showed that repetition can drive false remembering. As we have reviewed in the Introduction, several studies have shown that repetition of misinformation can increase its influence. In the current study, three of our four comparisons revealed a significant repetition effect. Overall, we found that repetition increased the misinformation effect, although, as described previously, one comparison in Experiment 2 produced a null effect. To further contextualize the repetition effect in the misinformation literature, we examined extant studies that have used repetition manipulations to provide a meta-analytic estimate of the effect size.

The literature of repetition effects on memory is enormous. It is therefore important to define and constrain the criteria for inclusion to make this meta-analysis feasible. To be faithful to the *misinformation effect design*, we included only studies in which the *to-be-rejected* (i.e., misleading) materials were repeated after a neutral encoding phase (that did not include misinformation). In addition, studies were excluded if the encoding phase was interactive – that is, if a participant engaged with the experimenter during the back-and-forth phases. These studies mainly included children as the participants, presumably to keep participants engaged during the encoding event. Together, our selection criteria excluded DRM or inference-driven false memory studies (Benjamin, 2001; McDermott and Chan, 2006), in which the to-be-rejected items were never presented for encoding. We also excluded studies that repeated the to-be-remembered (rather than to-be-rejected) items during encoding, such as studies that demonstrated the illusory truth effect (Arkes et al., 1991; Hassan and Barber, 2021), repetition effects in verbal learning (Howe, 1970, 1972; Melton, 1970; Jacoby, 1978; Hintzman, 2010), studies that demonstrated the imagination inflation effect (Garry et al., 1996; Goff and Roediger, 1998; Thomas et al., 2003), or studies that use a repeated retrieval (rather than repeated encoding) procedure to induce false memories (Poole and White, 1991; Shaw et al., 1995; Memon and Vartoukian, 1996; Roediger et al., 1996; Hauer et al., 2007; Henkel, 2007, 2008; Hershkowitz and Terner, 2007; La Rooy et al., 2010; Chan and Langley, 2011).

To find articles, we searched PsycINFO, PsycArticles, PubMed, and Google Scholar with the following search terms: "misinformation effect AND repetition," "repeated (or repetition of) misinformation," and "misinformation AND repetition." We also searched for studies that either cited or were cited by Mitchell and Zaragoza (1996) or Foster et al. (2012). Among the search results, we included only studies that used the misinformation effect procedure and involved a repetition manipulation (typically once vs. two or three presentations). Finally, we included only studies that contained enough information to calculate effect sizes (accuracy under one presentation of misinformation vs. accuracy under multiple presentations of misinformation).

In total, 19 effect sizes from 12 studies were included in this meta-analysis. Data from the current experiments were collapsed across the 1X variable for both Experiments 1 and 2 to avoid over-representing data from the 3X group.[4] Studies included in the meta-analysis are marked with an asterisk in the Reference section. Most included studies either directly reported an effect size of repetition or reported enough information for an effect size to be derived. For one study (Mitchell and Zaragoza, 1996), standard deviation was not reported (resulting in not enough information to calculate an effect size), so we imputed their standard deviation based on the remaining studies in the meta-analysis. We used the "meta" package in R to conduct the meta-analysis. As we anticipated heterogeneity between studies, we used a random-effects model to pool effect sizes. In addition, we used the restricted maximum likelihood estimator to calculate the heterogeneity variance τ^2 and the Knapp-Hartung adjustments to calculate confidence interval around the pooled effect.

Figure 9 shows a forest plot of this meta-analysis, with the random effects model producing a moderate repetition effect, $g = 0.38$ [0.18, 0.59], $p < 0.001$. A majority of the sampled studies showed a positive repetition effect,[5] and only two effect sizes were negative. We conducted an Egger's test, $t(17) = -2.98$, $p = 0.01$, (in R using the metabias function) to examine asymmetry and found some evidence of asymmetry in the funnel plot, although this did not necessarily indicate publication bias. That is, given the negative

4 We chose this method to prevent the data from the 3X condition from being included twice. If we had instead treated the data from each condition separately, we would have computed an effect for the 1X vs. 3X comparison and another effect size for the 1X distributed vs. 3X comparison.

5 One might wonder whether Bright-Paul and Jarrold (2012) is an outlier. An analysis conducted without that study resulted in a slightly smaller, but still significant effect size of g=0.34 [0.17, 0.51], p<0.001.

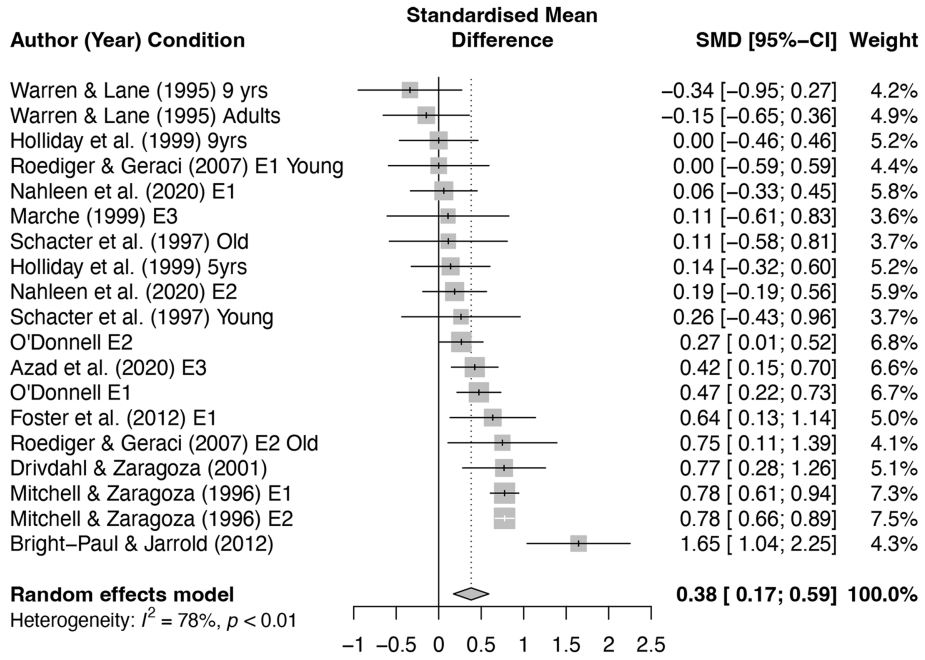


FIGURE 9
A forest plot of repetition effect sizes for misinformation effect studies.

result from the above analysis, it is clear from examining the funnel plot (see Supplementary material, Figure 1) that the data trended in the opposite direction of publication bias, as demonstrated by the concentration of data in the bottom left of the plot. Note that this meta-analysis was not meant to be exhaustive but was designed to provide a broader, quantitative examination of the extant effect sizes regarding repetition and the misinformation effect. A note about the existing literature: perhaps due to practical constraints, most studies had compared a single presentation of misinformation to either two or three presentations, so we currently do not know if the repetition effect is monotonically positive, if it would reach an asymptote, or if it would take on an inverted U function, such that repeating the same misinformation "too many" times would reduce its influence, similar to how being exposed to "too many" pieces of misinformation can increase people's resistance to the misinformation (Pena et al., 2017). Further research is needed to address this important question.

A meta-analysis of source variability

In the current study, we failed to observe a source variability effect. The null effect in Experiment 2 was particularly notable because we expected that making clear that three different people (instead of a single person) provided the same misinformation would strengthen its influence. One argument that could be raised is that the null effect of Experiment 2, despite our best effort to make the witnesses clearly different, could still have been the result of a weak source manipulation. That is, one might argue that we could further increase the distinctiveness of the sources by further distinguishing the witnesses in terms of gender, age, and other obvious characteristics. Although each of the confederates was female and around the same age (18–22), it is very unlikely that they were indistinguishable. Each confederate had a different hair color (red, blond, and brown) and style, all three used different mannerisms, and each interviewee had a different voice. Thus, it is unlikely that the null effect of source variability is attributable to a weak source manipulation.

We believe that a more likely argument is that source variability by itself does not significantly influence accuracy or suggestibility. Across four experimental comparisons involving misled items, we did not find any significant effects of the source manipulation. Notably, these null effects occurred regardless of whether each piece of misinformation was presented only once (which allowed us to examine the influence of source variability on its own) or three times (which allowed us to examine the effect of source variability in the context of repeated misinformation). In addition, when we conducted a meta-analysis using the data from both the current study and Foster et al. (2012), a fixed effects model showed that the *source variability effect* was essentially nil, $g = -0.01$ [−0.18, 0.15], $p = 0.856$ (See Figure 10). For this meta-analysis, we did not conduct an Egger's test, given that there were not enough studies to conduct this test and that there was only one published study included. However, the funnel plot (see Supplementary material, Figure 2) does not show evidence of asymmetry. We used the same package in R to conduct this analysis but used a fixed-effects model, given that both studies used largely the same design. In addition, we applied *ad hoc* variance correction, given the small Hartung-Knapp variance estimate. Furthermore, the current experiments demonstrated the same null effect shown in Foster et al. (2012) with greater statistical power. Altogether, the data from the current study and Foster et al. suggest that source variability had little to no discernible influence on participants' memory, at least in conditions similar to those tested here.

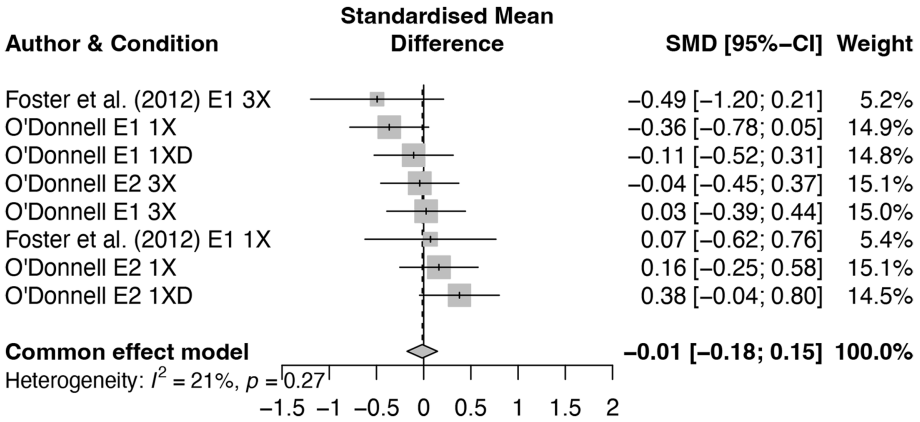


FIGURE 10
A forest plot of the source variability effect in the current study and Foster et al. (2012) by repetition condition.

General discussion

In two experiments, we examined whether multiple presentations of the same misinformation and increasing the number of people who provide the same misinformation would affect people's suggestibility. Overall, both experiments provided results that replicated Foster et al., such that repetition, but not source variability, increased the misinformation effect. One exception to this pattern was that repetition did not significantly harm performance when comparing the 3X against the 1X distributed condition in Experiment 2, which we attributed to a sampling error. Moreover, across both experiments, participants exhibited a strong, positive confidence-accuracy relationship for the neutral items, but exposure to misinformation severely depressed this relationship, such that only the most confident responses exceeded chance level accuracy.

In the remainder of the General Discussion, we briefly review the results of our repetition manipulation, then discuss the potential reasons why there were no differences in accuracy when participants read or watched interviews from three witnesses compared to one witness, and finally, consider the applied implications of these findings.

The effect of repeated misinformation on suggestibility

In Experiment 1, participants read three interview transcripts that either introduced them to misinformation once or three times. In Experiment 2, the interview transcripts were formatted as videos, and participants were again introduced to misinformation either once or three times. Across both experiments, the effect of repetition was reliable but modest (g = 0.47 in Experiment 1 and g = 0.27 in Experiment 2). Indeed, the repetition effect was only significant in one of two comparisons in Experiment 2. It should be noted that the current studies were powered to detect a larger effect size than we found. However, the meta-analysis demonstrated a significant, moderate effect of repetition among studies in the misinformation literature that used a similar paradigm. Moreover, although the current studies demonstrated a smaller effect size than that of Foster et al. (g = 0.64), a sampling of the literature shows a variety of effect sizes, and indeed, our effect sizes fell squarely inside the 0.95 CI of the meta-analytic effect of 0.18 to 0.59.

A practical note is that our manipulation of repetition, and indeed the repetition manipulation implemented in most misinformation studies to date, is relatively weak. We had participants read/watch someone reproduce the same piece of misinformation up to three times within a 10–20 min span. In actual criminal investigations, an eyewitness might be exposed to the same piece of misinformation on far more occasions stretched across a much longer interval (Barry et al., 2017). Given that spaced presentations enhance learning relative to massed presentations (Cepeda et al., 2006), real-life eyewitnesses who are repeatedly exposed to the same piece of misinformation across a longer time interval than is typical in laboratory settings might demonstrate a greater repetition effect. Future research should examine the influence of varying number of repetitions and the intervals between repetition of misinformation on eyewitness suggestibility.

Does varying misinformation sources increase its influence?

In the present experiments, – and in Foster et al. (2012) – increasing the number of misinformation sources did not affect participants' suggestibility. In addition, a meta-analysis of the data from both the current studies and Foster et al. demonstrated that, among eight effect sizes, the effect of source variability was essentially nil. Of course, source variability might yet produce an effect in future studies with heretofore unexamined variables (such as with even more eyewitnesses who deliver the misinformation, with a source memory test, with different participant populations, or if participants discussed the details of the encoding event with confederates).

Assuming that source variability does not normally affect eyewitness memory, what might explain this null effect? Foster et al. (2012) theorized that people might not account for the number of eyewitnesses who make a statement during memory retrieval. Rather, they rely on the fluency or familiarity of the information they are retrieving without recalling where the information came from or how many sources contributed to the information. This idea is supported

by other research (Thomas et al., 2010; Chan et al., 2012), including studies that showed that perception of consensus is driven primarily by the fluency of the retrieved information (Weaver et al., 2007; Schwarz et al., 2021). By this logic, information that is given once by three different sources would produce a similar feeling of consensus as information that is repeated three times by the same source. Finally, across six experiments, Weaver et al. demonstrated that repeatedly presenting one piece of information gave participants a sense of consensus for that information, even though only one person presented it. In these experiments, participants read focus group opinion statements. In Experiments 1 and 5, participants either read one opinion from one person, a similar opinion three times from the same person, or a similar opinion from three different people (participants were told these opinions were sampled from a focus group of five people). Experiments 2a, 2b, 3, and 4 contained only the first two conditions listed above. After reading the statements, participants were asked to estimate the opinions of the focus group on a 1–7 Likert scale. Participants consistently perceived an opinion to have more support when it was repeated by the same person, despite knowing that only one person was providing the opinion (as each opinion had a name attached to it). This result suggests that participants relied on fluency when making consensus judgments. As such, it could be that the null effect of source variability was a result of participants ascribing similar feelings of consensus to information that was presented three times by one person and information that was presented once each by three different people.

Additionally, Foster et al. (2012) argued that different mechanisms might contribute to the credibility of information from one source compared to information from three sources. Specifically, when one witness repeats a claim, the person may be judged as more credible (than a person who does not repeat a claim) because of the consistency exhibited *across instances* (Chan et al., 2017; Smelter and Calvillo, 2020). When three witnesses make the same claim, it may be judged as credible because of the consistency exhibited *across individuals* (Lindsay et al., 1986). Thus, a claim that has been repeated by a single witness may be perceived as more accurate because the *witness* is deemed consistent, whereas perceptions of accuracy in a claim that is made by three different witnesses may be judged as more accurate because *the claim itself* is viewed as more credible than a claim that has not been uttered by multiple witnesses. Ultimately, however, it is the repetition, not variation of sources, that increases the credibility of a claim.

The lack of a source variability effect suggests that participants might not have used source-specifying information during their retrieval, but it is also possible that this null effect occurred because participants were not given explicit instructions to use source specifying information. A notable difference between the results of Foster et al.'s (2012) and Mitchell and Zaragoza (1996), aside from the differences in the materials and the nature of the source manipulation,[6] was that Mitchell and Zaragoza gave participants a source discrimination test, whereas Foster et al. and (the current experiments) gave participants a 2AFC recognition test. Therefore, it is possible that a source variability effect might only surface when participants were forced to consider source-specifying information, such as when they were told explicitly to retrieve contextual information or when they were provided with explicit warnings. Future research can address this possibility by manipulating both source variability and retrieval requirements.

Applied implications

From an applied perspective, our results indicate that an individual who hears a piece of misinformation repeatedly would be more likely to report that misinformation than an individual who hears the same piece of misinformation only once, and this is true regardless of how many sources repeatedly present the misinformation. More broadly, an intriguing implication of these findings is that attempts to retract or debunk misinformation should avoid including the misinformation (e.g., correcting the misinformation without explicitly restating it), given that repetitions can increase suggestibility. Research on the continued influence effect has examined this possibility with somewhat mixed results. Several researchers have found that including the details of the misinformation in a retraction can have a backfire effect, such that people often falsely remember the information being corrected as true (Skurnik et al., 2005; Nyhan and Reifler, 2010; Peter and Koch, 2016). However, more recent research showed that retractions that included the misinformation were more effective at reducing the continued influence effect than retractions that did not mention the misinformation (Ecker et al., 2017, 2020). A possible explanation for these discrepant findings is that timing of the correction matters. In studies that did not demonstrate a backfire effect, participants read statements or new articles and received a correction (with or without a reminder of the fake news) *after a delay* (Ecker et al., 2017, 2020). In contrast, in the studies that demonstrated a backfire effect, participants usually read statements or news articles with a truth verification *simultaneously*.

The current experiments showed that participants were generally quite adept at judging their response accuracy, as they demonstrated a positive confidence-accuracy relationship for the neutral items (see Figures 5, 8), which is consistent with recent findings in the literature (Wixted et al., 2018). But perhaps more remarkable was the much flatter confidence-accuracy curves for the misled items. In line with Wixted et al.'s argument, the confidence-accuracy relationship was flatter when it was contaminated by misinformation. In both experiments, participants exhibited near-chance performance for these items, and accuracy only exceeded chance at the highest level of confidence. This poor confidence-accuracy relationship shows the evidence-contaminating power of misleading suggestions and replicates recent findings that showed that, in the absence of a warning, misinformation can damage *both the accuracy* of eyewitness memory reports *and the diagnosticity* of eyewitness confidence (Chan et al., 2022).

Conclusion

In two preregistered, high-powered experiments, we attempted to conceptually replicate Experiment 1 of Foster et al. (2012) and determine whether both repetition and source variability influence

6 Note again, however, Mitchell and Zaragoza did not manipulate the number of witnesses who provided the same piece of misinformation; rather, they manipulated presentation modality.

eyewitness suggestibility. In three of the four comparisons, we demonstrated a significant effect of repetition,[7] and our small-scale meta-analyses provided further evidence that repetition of misinformation can exacerbate eyewitness suggestibility.

In contrast to the effect of repetition, we found no effect of source variability. Although one might suggest that this null effect was the result of a weak manipulation, we argue here that our manipulation produced obvious differences across the three interviewees in an ecologically realistic manner. We therefore conclude that, as Foster et al. (2012) did, it is repetition of the misinformation that increases an eyewitness's suggestibility, not the number of people who provide the misinformation.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://osf.io/9zpfk/?view_only=f95ed70720c742d48296fa3b92891ed7.

Ethics statement

The studies involving human participants were reviewed and approved by the Institutional Review Board at Iowa State University. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

7 Note that in all four comparisons, repetition reduced recognition accuracy for the misled items quantitatively.

Author contributions

RO'D and JC conceived the research idea. RO'D adapted the master's thesis (O'Donnell, 2022), which was written to fulfill the requirements of her master's degree during her Doctoral Program in Psychology at Iowa State University, collected and analyzed the data for both experiments under the guidance of JC, and wrote the full first draft of the manuscript. JC provided revisions. MG and JF provided consultations on the project, suggestions regarding the design and procedure, and comments on the manuscript. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1201674/full#supplementary-material

References

*References that were included in the meta-analysis.

Arkes, H. R., Boehm, L. E., and Xu, G. (1991). Determinants of judged validity. *J. Exp. Soc. Psychol.* 27, 576–605. doi: 10.1016/0022-1031(91)90026-3

Barry, J., Lindsay, L., Begley, T., Edwards, D., and Cardoret, C. (2017). Criminal court of the city of New York 2016 annual report. Available at: https://www.nycourts.gov/LegacyPDFs/COURTS/nyc/criminal/2016-Annual-Report-Final.pdf (Accessed December, 2022).

Begg, I. M., Anas, A., and Farinacci, S. (1992). Dissociation of processes in belief: source recollection, statement familiarity, and the illusion of truth. *J. Exp. Psychol. Gen.* 121, 446–458. doi: 10.1037/0096-3445.121.4.446

Benjamin, A. S. (2001). On the dual effects of repetition on false recognition. *J. Exp. Psychol. Learn. Mem. Cogn.* 27, 941–947. doi: 10.1037/0278-7393.27.4.941

Berkowitz, S., and Loftus, E. F. (2018). "Misinformation in the courtroom" in *Finding the truth in the courtroom: Dealing with deception, lies, and memories*. eds. H. Otgaar and M. L. Howe (Oxford: Oxford University Press), 11–30.

Blank, H., Ost, J., Davies, J., Jones, G., Lambert, K., and Salmon, K. (2013). Comparing the influence of directly vs. indirectly encountered post-event misinformation on eyewitness remembering. *Acta Psychol.* 144, 635–641. doi: 10.1016/j.actpsy.2013.10.006

*Bright-Paul, A., and Jarrold, C. (2012). Children's eyewitness memory: repeating post-event misinformation reduces the distinctiveness of a witnessed event. *Memory* 20, 818–835. doi:10.1080/09658211.2012.708345

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., and Rohrer, D. (2006). Distributed practice in verbal recall tasks: a review and quantitative synthesis. *Psychol. Bull.* 132, 354–380. doi: 10.1037/0033-2909.132.3.354

Chan, J. C. K., and Langley, M. M. (2011). Paradoxical effects of testing: retrieval enhances both accurate recall and suggestibility in eyewitnesses. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 248–255. doi: 10.1037/a0021204

Chan, J. C. K., Manley, K. D., and Lang, K. (2017). Retrieval-enhanced suggestibility: a retrospective and a new investigation. *J. Appl. Res. Mem. Cogn.* 6, 213–229. doi: 10.1016/j.jarmac.2017.07.003

Chan, J. C. K., O'Donnell, R., and Manley, K. D. (2022). Warning weakens retrieval-enhanced suggestibility only when it is given shortly after misinformation: the critical importance of timing. *J. Exp. Psychol. Appl.* 28, 694–716. doi: 10.1037/xap0000394

Chan, J. C. K., Wilford, M. M., and Hughes, K. L. (2012). Retrieval can increase or decrease suggestibility depending on how memory is tested: the importance of source complexity. *J. Mem. Lang.* 67, 78–85. doi: 10.1016/j.jml.2012.02.006

Clark, S. E., and Wells, G. L. (2008). On the diagnosticity of multiple-witness identifications. *Law Hum. Behav.* 32, 406–422. doi: 10.1007/s10979-007-9115-7

Dechêne, A., Stahl, C., Hansen, J., and Wänke, M. (2010). The truth about the truth: a meta-analytic review of the truth effect. *Personal. Soc. Psychol. Rev.* 14, 238–257. doi: 10.1177/1088868309352251

Dodd, D. H., and Bradshaw, J. M. (1980). Leading questions and memory: pragmatic constraints. *J. Verbal Learn. Verbal Behav.* 19, 695–704. doi: 10.1016/S0022-5371(80)90379-5

*Drivdahl, S. B., and Zaragoza, M. S. (2001). The role of perceptual elaboration and individual differences in the creation of false memories for suggested events. *Appl. Cognit. Psychol.* 15, 265–281. doi:10.1002/acp.701

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology*. New York City, NY: Рипол Классик.

Ecker, U. K. H., Hogan, J. L., and Lewandowsky, S. (2017). Reminders and repetition of misinformation: helping or hindering its retraction? *J. Appl. Res. Mem. Cogn.* 6, 185–192. doi: 10.1037/h0101809

Ecker, U. K. H., Lewandowsky, S., and Chadwick, M. (2020). Can corrections spread misinformation to new audiences? Testing for the elusive familiarity backfire effect. *Cogn. Res. Princ. Implic.* 5:41. doi: 10.1186/s41235-020-00241-6

Ecker, U. K. H., Lewandowsky, S., Swire, B., and Chang, D. (2011). Correcting false information in memory: manipulating the strength of misinformation encoding and its retraction. *Psychon. Bull. Rev.* 18, 570–578. doi: 10.3758/s13423-011-0065-1

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

Gabbert, F., Memon, A., Allan, K., and Wright, D. B. (2004). Say it to my face: examining the effects of socially encountered misinformation. *Leg. Criminol. Psychol.* 9, 215–227. doi: 10.1348/1355325041719428

Garry, M., Manning, C. G., Loftus, E. F., and Sherman, S. J. (1996). Imagination inflation: imagining a childhood event inflates confidence that it occurred. *Psychon. Bull. Rev.* 3, 208–214. doi: 10.3758/BF03212420

Goff, L. M., and Roediger, H. L. (1998). Imagination inflation for action events: repeated imaginings lead to illusory recollections. *Mem. Cogn.* 26, 20–33. doi: 10.3758/BF03211367

Goodwin, K. A., Kukucka, J. P., and Hawks, I. M. (2013). Co-witness confidence, conformity, and eyewitness memory: an examination of normative and informational social influences. *Appl. Cognit. Psychol.* 27, 91–100. doi: 10.1002/acp.2877

Hasher, L., Goldstein, D., and Toppino, T. (1977). Frequency and the conference of referential validity. *J. Verbal Learn. Verbal Behav.* 16, 107–112. doi: 10.1016/S0022-5371(77)80012-1

Hassan, A., and Barber, S. J. (2021). The effects of repetition frequency on the illusory truth effect. *Cogn. Res. Princ. Implic.* 6:38. doi: 10.1186/s41235-021-00301-5

Hauer, B. J. A., Wessel, I., Merckelbach, H., Roefs, A., and Dalgleish, T. (2007). Effects of repeated retrieval of central and peripheral details in complex emotional slides. *Memory* 15, 435–449. doi: 10.1080/09658210701312085

Henderson, E. L., Simons, D. J., and Barr, D. J. (2021). The trajectory of truth: a longitudinal study of the illusory truth effect. *J. Cogn.* 4:29. doi: 10.5334/joc.161

Henkel, L. A. (2007). The benefits and costs of repeated memory tests for young and older adults. *Psychol. Aging* 22, 580–595. doi: 10.1037/0882-7974.22.3.580

Henkel, L. A. (2008). Maximizing the benefits and minimizing the costs of repeated memory tests for older adults. *Psychol. Aging* 23, 250–262. doi: 10.1037/0882-7974.23.2.250

Hershkowitz, I., and Terner, A. (2007). The effects of repeated interviewing on children's forensic statements of sexual abuse. *Appl. Cognit. Psychol.* 21, 1131–1143. doi: 10.1002/acp.1319

Hintzman, D. L. (2010). How does repetition affect memory? Evidence from judgments of recency. *Mem. Cogn.* 38, 102–115. doi: 10.3758/MC.38.1.102

*Holliday, R. E., Douglas, K. M., and Hayes, B. K. (1999). Children's eyewitness suggestibility: memory trace strength revisited. *Cogn. Dev.* 14, 443–462 doi: 10.1016/S0885-2014(99)00014-3

Howe, M. J. (1970). Repeated presentation and recall of meaningful prose. *J. Educ. Psychol.* 61, 214–219. doi: 10.1037/h0029102

Howe, M. J. A. (1972). Repeated presentation and retention of meaningful information. *Psychol. Rep.* 31, 840–842. doi: 10.2466/pr0.1972.31.3.840

Jack, F., Zydervelt, S., and Zajac, R. (2014). Are co-witnesses special? Comparing the influence of co-witness and interviewer misinformation on eyewitness reports. *Memory* 22, 243–255. doi: 10.1080/09658211.2013.778291

Jacoby, L. L. (1978). On interpreting the effects of repetition: solving a problem versus remembering a solution. *J. Verbal Learn. Verbal Behav.* 17, 649–667. doi: 10.1016/S0022-5371(78)90393-6

Jacoby, L. L. (1999). Ironic effects of repetition: measuring age-related differences in memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 3–22. doi: 10.1037/0278-7393.25.1.3

La Rooy, D., Katz, C., Malloy, L. C., and Lamb, M. E. (2010). Do we need to rethink guidance on repeated interviews? Psychol. *Public Policy Law* 16, 373–392. doi: 10.1037/a0019909

Lindsay, R. C. L., Lim, R., Marando, L., and Cully, D. (1986). Mock-juror evaluations of eyewitness testimony: a test of metamemory hypotheses. *J. Appl. Social. Psychol.* 16, 447–459. doi: 10.1111/j.1559-1816.1986.tb01151.x

*Marche, T. A. (1999). Memory strength affects reporting of misinformation. *J. Exp. Child Psychol.* 73, 45–71. doi: 10.1006/jecp.1998.2489

McDermott, K. B., and Chan, J. C. K. (2006). Effects of repetition on memory for pragmatic inferences. *Mem. Cogn.* 34, 1273–1284. doi: 10.3758/BF03193271

Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *J. Verbal Learn. Verbal Behav.* 9, 596–606. doi: 10.1016/S0022-5371(70)80107-4

Memon, A., and Vartoukian, R. (1996). The effects of repeated questioning on young children's eyewitness testimony. *Br. J. Psychol.* 87, 403–415. doi: 10.1111/j.2044-8295.1996.tb02598.x

*Mitchell, K. J., and Zaragoza, M. S. (1996). Repeated exposure to suggestion and false memory: the role of contextual variability. *J. Mem. Lang.* 35, 246–260. doi:10.1006/jmla.1996.0014

Mojtahedi, D., Ioannou, M., and Hammond, L. (2018). The dangers of co-witness familiarity: investigating the effects of co-witness relationships on blame conformity. *J. Police. Crim. Psych.* 33, 316–326. doi: 10.1007/s11896-018-9254-4

*Nahleen, S., Strange, D., and Takaran, M. K. T. (2020). Does emotional or repeated misinformation increase memory distortion for a trauma analogue event? *Psychol. Res.* 85, 2453–2465. doi: 10.1007/s00426-020-01409-x

Nyhan, B., and Reifler, J. (2010). When corrections fail: the persistence of political misperceptions. *Polit. Behav.* 32, 303–330. doi: 10.1007/s11109-010-9112-2

O'Donnell, R. (2022). Does source variability of misinformation influence eyewitness suggestibility independently of repetition of misinformation? [Master's thesis, Iowa State University]. ProQuest Dissertations and Theses Global.

Ost, J., Ghonouie, H., Cook, L., and Vrij, A. (2008). The effects of confederate influence and confidence on the accuracy of crime judgements. *Acta Psychol.* 128, 25–32. doi: 10.1016/j.actpsy.2007.09.007

Park, S. A., Goïame, S., O'Connor, D. A., and Dreher, J.-C. (2017). Integration of individual and social information for decision-making in groups of different sizes. *PLoS Biol.* 15:e2001958. doi: 10.1371/journal.pbio.2001958

Pena, M. M., Klemfuss, J. Z., Loftus, E. F., and Mindthoff, A. (2017). The effects of exposure to differing amounts of misinformation and source credibility perception on source monitoring and memory accuracy. *Psychol. Conscious. Theory Res. Pract.* 4, 337–347. doi: 10.1037/cns0000137

Peter, C., and Koch, T. (2016). When debunking scientific myths fails (and when it does not): the backfire effect in the context of journalistic coverage and immediate judgments as prevention strategy. *Sci. Commun.* 38, 3–25. doi: 10.1177/1075547015613523

Poole, D. A., and White, L. T. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Dev. Psychol.* 27, 975–986. doi: 10.1037/0012-1649.27.6.975

Reysen, M. B. (2007). The effects of social pressure on false memories. *Mem. Cogn.* 35, 59–65. doi: 10.3758/BF03195942

*Roediger, H. L., and Geraci, L. (2007). Aging and the misinformation effect: a neuropsychological analysis. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 321–334 doi: 10.1037/0278-7393.33.2.321

Roediger, H. L., Jacoby, J. D., and McDermott, K. B. (1996). Misinformation effects in recall: creating false memories through repeated retrieval. *J. Mem. Lang.* 35, 300–318. doi: 10.1006/jmla.1996.0017

Roggeveen, A. L., and Johar, G. V. (2002). Perceived source variability versus familiarity: testing competing explanations for the truth effect. *J. Consum. Psychol.* 12, 81–91. doi: 10.1207/S15327663JCP1202_02

*Schacter, D. L., Koutstaal, W., Johnson, M. K., Gross, M. S., and Angell, K. E. (1997). False recollection induced by photographs: a comparison of older and younger adults. *Psychol. Aging* 12, 203–215. doi: 10.1037/0882-7974.12.2.203

Schwarz, N., Jalbert, M., Noah, T., and Zhang, L. (2021). Metacognitive experiences as information: processing fluency in consumer judgment and decision making. *Consum. Psychol. Rev.* 4, 4–25. doi: 10.1002/arcp.1067

Schwarz, N., Newman, E., and Leach, W. (2016). Making the truth stick & the myths fade: lessons from cognitive psychology. *Behav. Sci. Policy* 2, 85–95. doi: 10.1353/bsp.2016.0009

Shaw, J. S., Bjork, R. A., and Handal, A. (1995). Retrieval-induced forgetting in an eyewitness-memory paradigm. *Psychon. Bull. Rev.* 2, 249–253. doi: 10.3758/BF03210965

Skagerberg, E. M., and Wright, D. B. (2008). The prevalence of co-witnesses and co-witness discussions in real eyewitnesses. *Psychol. Crime Law* 14, 513–521. doi: 10.1080/10683160801948980

Skurnik, I., Yoon, C., Park, D. C., and Schwarz, N. (2005). How warnings about false claims become recommendations. *J. Consum. Res.* 31, 713–724. doi: 10.1086/426605

Smelter, T. J., and Calvillo, D. P. (2020). Pictures and repeated exposure increase perceived accuracy of news headlines. *Appl. Cogn. Psychol.* 34, 1061–1071. doi: 10.1002/acp.3684

Smith, V. L., and Ellsworth, P. C. (1987). The social psychology of eyewitness accuracy: misleading questions and communicator expertise. *J. Appl. Psychol.* 72, 294–300. doi: 10.1037/0021-9010.72.2.294

Thomas, A. K., Bulevich, J. B., and Chan, J. C. K. (2010). Testing promotes eyewitness accuracy with a warning: implications for retrieval enhanced suggestibility. *J. Mem. Lang.* 63, 149–157. doi: 10.1016/j.jml.2010.04.004

Thomas, A. K., Bulevich, J. B., and Loftus, E. F. (2003). Exploring the role of repetition and sensory elaboration in the imagination inflation effect. *Mem. Cogn.* 31, 630–640. doi: 10.3758/BF03196103

Thorley, C. (2013). Memory conformity and suggestibility. *Psychol. Crime Law* 19, 565–575. doi: 10.1080/1068316X.2011.648637

Vrij, A., Pannell, H., and Ost, J. (2005). The influence of social pressure and black clothing on crime judgements. *Psychol. Crime Law* 11, 265–274. doi: 10.1080/10683160410001680780

Walther, E., Bless, H., Strack, F., Rackstraw, P., Wagner, D., and Werth, L. (2002). Conformity effects in memory as a function of group size, dissenters and uncertainty. *Appl. Cognit. Psychol.* 16, 793–810. doi: 10.1002/acp.828

*Warren, A., and Lane, P. (1995). "Effects of timing and type of questioning on eyewitness accuracy and suggestibility," in *Memory and testimony in the child witness*. eds. M. S. Zaragoza, J. R. Graham, R. Hirschman, G. C. N. Hall and Y. S. Ben-Porath (London: Sage Publications, Inc.).

Weaver, K., Garcia, S. M., Schwarz, N., and Miller, D. T. (2007). Inferring the popularity of an opinion from its familiarity: a repetitive voice can sound like a chorus. *J. Pers Soc Psychol.* 92, 821–833.

Wixted, J. T., Mickes, L., and Fisher, R. P. (2018). Rethinking the reliability of eyewitness memory. *Perspect. Psychol. Sci.* 13, 324–335. doi: 10.1177/1745691617734878

Zaragoza, M. S., Belli, R. F., and Payment, K. E. (2007). "Misinformation effects and the suggestibility of eyewitness memory" in *Do justice and let the sky fall: Elizabeth Loftus and her contributions to science, law, and academic freedom*. eds. M. Garry and H. Hayne (Mahwah: Lawrence Erlbaum Associates Publishers).

frontiers | Frontiers in Psychology

TYPE Review
PUBLISHED 28 August 2023
DOI 10.3389/fpsyg.2023.1081528

OPEN ACCESS

EDITED BY
Iris Blandon-Gitlin,
California State University, Fullerton,
United States

REVIEWED BY
Stefanie Sharman,
Deakin University, Australia
Zhiwei Liu,
Sichuan University of Science and Engineering,
China

*CORRESPONDENCE
Henry Otgaar
✉ henry.otgaar@kuleuven.be;
✉ henry.otgaar@maastrichtuniversity.nl

RECEIVED 27 October 2022
ACCEPTED 31 July 2023
PUBLISHED 28 August 2023

CITATION
Otgaar H, Mangiulli I, Battista F and Howe ML (2023) External and internal influences yield similar memory effects: the role of deception and suggestion.
Front. Psychol. 14:1081528.
doi: 10.3389/fpsyg.2023.1081528

External and internal influences yield similar memory effects: the role of deception and suggestion

Henry Otgaar[1,2]*, Ivan Mangiulli[1,3], Fabiana Battista[3] and Mark L. Howe[2,4]

[1]Leuven Institute of Criminology (LINC), Catholic University of Leuven, Leuven, Belgium, [2]Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Netherlands, [3]Department of Education, Psychology, and Communication Sciences, University of Bari Aldo Moro, Bari, Italy, [4]Department of Psychology, City, University of London, London, United Kingdom

In legal cases, testimonies can become contaminated because of an amalgam of external and internal influences on memory. It is well-established that external influences (e.g., suggestive interviews) can hurt memory. However, less focus has been placed on the impact of internal influences (e.g., lying) on memory. In the current review, we show that the available evidence suggests that both external and internal influences exert similar effects on memory. That is, we review studies showing that suggesting non-occurrences and suggesting non-experiences can lead to omission errors and false memories, respectively. Likewise, these memory effects are also observed when focusing on internal influences. That is, false denials, feigning amnesia and fabrication have been shown to affect memory in terms of forgetting (i.e., omissions) and false memories (i.e., commissions). Also, we show that both external and internal influences can lead to changes in the belief that an event occurred. We argue that in legal cases, triers of fact should concentrate on whether both types of influences might have affected testimonial accuracy in witnesses, victims, and suspects.

KEYWORDS

lying, suggestion, forgetting, false memory, cognitive dissonance

How internal and external influences can yield similar memory effects: the role of deception and suggestion

What witnesses, victims, and suspects can accurately remember about their experiences is oftentimes a crucial issue in the court. The reason is straightforward. In many legal cases, objective evidence such as fingerprints or DNA samples is lacking (Howe and Knott, 2015). In these cases, triers of fact need to base their legal decision making on the memorial record of witnesses, victims, and suspects. Although triers of fact strive for memory reports containing a high degree of accuracy, this is frequently not what happens. Specifically, people might misremember details of, or even entire, autobiographical events. When triers of fact (e.g., judges) deem such statements to be authentic, miscarriages of justice might prevail (see Howe et al., 2018).

In the (legal) psychological literature and case reports, memory failures are often depicted as arising from external sources such as the use of suggestive interviewing techniques by the police (e.g., Otgaar et al., 2019b). However, a new branch of research has

shown that memory failures can also occur because of internal influences such as lying about a crime (see for a review, Otgaar and Baker, 2018; Battista and Otgaar, 2022).

What we will show in the current review is that these external and internal factors oftentimes exert similar effects on memory, implying that similar mechanisms might underpin these memory effects. By assembling empirical research on these themes, our review will focus specifically on two factors, namely suggestion (i.e., suggesting non-experience/non-occurrences) and lying (fabrication, feigning amnesia, false denial) as external and internal influences, respectively. The reason why we focused on these influences is that they share important similarities in how they target memories. For one thing, suggesting and fabricating non-experiences both imply the invention of details and/or events that never occurred, while suggesting non-occurrence and feigning amnesia/falsely denying details both relate to rebutting experiences that did occur (see also **Figure 1**). Finally, we will show how relevant and informative these findings are for legal proceedings in which memory failures might play a pivotal role determining a verdict.

What are memory failures?

In the eyewitness memory field, much attention has been given to the following two memory failures: Omission and commission errors (e.g., Wright and Loftus, 1998; Loftus, 2005). When errors of omission take place, people who have experienced an event, fail to report (parts of) the experience. A notable example of omission errors is forgetting. Although forgetting is a normal memory phenomenon, it is often marshaled as an important memory failure (Fawcett and Hulbert, 2019). For example, forgetting is seen as a form of cognitive decline more likely to occur when getting older (e.g., Fraundorf et al., 2019). Moreover, forgetting important details of, say, a traumatic experience (e.g., sexual abuse) can be seen as a memory failure (e.g., Bell and Loftus, 1988; Ernberg et al., 2018; Shaw and Loftus, 2020) because less detailed statements can be unduly regarded as less accurate (e.g., Curci et al., 2020). However, forgetting can also be beneficial as, for instance, it aids in the facilitation of cognitive functioning (Nørby, 2015; Fawcett and Hulbert, 2019). Furthermore, a failure to report even a traumatic event (e.g., sexual abuse) does not entail an entire inability to remember it (McNally, 2005; Otgaar et al., 2019b).

Several theoretical notions have been proposed to account for the occurrence of forgetting. Classic theories of forgetting have involved principles of decay and interference (Fawcett and Hulbert, 2019). Decay theory states that forgetting takes place due to the "wasting effects of time" (McGeoch, 1932, p. 354). Most evidence, however, points to the idea that interference is the main source of forgetting (Wixted, 2004, 2005). Interference can be broadly differentiated into proactive and retroactive interference. Proactive interference refers to forgetting occurring due to prior learning affecting the retention of current information. Retroactive interference, instead, involves the negative impact of new information on previously encoded information. Research suggests that retroactive interference is the most likely candidate to explain forgetting (Wixted, 2004, 2005).

Forgetting can also be exerted intentionally (e.g., Anderson and Green, 2001; Macleod, 2012). For example, in the directed forgetting paradigm (word list variant) (Macleod, 2012), participants receive two word lists and they are instructed to forget one word list while remembering the other one. When participants are asked to recall all of the words they were presented, they typically remember fewer words from the list that had to be forgotten than the list that had to be remembered (e.g., Bjork, 1989; Bjork and Bjork, 1996; Macleod, 1999, 2012; Conway et al., 2000). Furthermore, in the Think/No Think method (Anderson and Green, 2001), participants are trained on several unrelated word pairs (e.g., ordeal-roach). Next, participants are reminded of these word pairs by being presented with a cue word (e.g., ordeal) and for each cue, one of two instructions is provided: Participants have to either recall the associated item (e.g., roach) ("think" instruction) or have to not think of the associated response ("no-think"). During the last phase, participants receive all cue words and are asked to come up with the associated words. The general finding is a memory impairment for "no-think" items compared with memory for items that were only presented during the first and last phase (Anderson and Green, 2001). This memory suppression effect has also been conceptually replicated in other studies (e.g., Joormann et al., 2005; Bergström et al., 2007).

Moreover, forgetting can also occur because of the inhibitory processes that occur in retrieval. That is, retrieval of practiced information causes the suppression of unpracticed related information (i.e., retrieval-induced forgetting effect or RIF; Anderson et al., 1994). In the RIF paradigm, participants are asked to learn a set of category item pairs (e.g., fruit–apple and drink-gin) and then are instructed to practice half of the studied pairs from half of the categories (e.g., fruit–apple). Finally, participants are asked to recall all words they can remember from the first phase. The typical finding is that the unpracticed items (e.g., pear) from the practiced categories (e.g., fruit) are more poorly recalled than unpracticed items (e.g., gin) from unpracticed categories (e.g., drink).

Although several studies have been conducted on forgetting, there are also a plethora of studies focusing on errors of commission which are instances in which people either remember events or details that were not experienced or remember them differently as compared with what they really experienced. In this category of errors there are also false memories[1] that can be elicited spontaneously (i.e., spontaneous false memories) or because of external suggestion (i.e., suggestive false memories). Here too, false memories are commonly regarded as dangerous memory failures in the courtroom as they can lead to false accusations and miscarriages of justice (e.g., Otgaar et al., 2022a,b). To be more specific, the main contributing factor of wrongful convictions - for around 70%–is false testimony, wherein eyewitnesses have misidentified an innocent suspect during a line-up (Innocence Project, 2020). Another example concerns people who falsely remember having been abused, something that can lead to false accusations (e.g., Otgaar et al., 2019b). However, just like forgetting, the production of false memories can be seen as an integral part of a normal and adaptive memory system. Indeed, having false memories can

1 In the current paper, we use the term "false memories" as referring to remembering non-experienced details/events and referring to instances in which details of events are remembered differently (also called memory distortions). Also, we use the terms omissions and forgetting, and commissions and false memories interchangeably in the current manuscript.

sometimes even be beneficial in that it can aid in prospective problem-solving (Howe, 2011; Otgaar et al., 2015; Howe et al., 2017).

Theories that explain the occurrence of false memories are, for example, source monitoring framework (SMF; Johnson et al., 1993), fuzzy-trace theory (FTT; Brainerd et al., 2008), and associative-activation theory (AAT; Howe et al., 2009). According to the source monitoring framework, people make attributions about the sources of their memories (Johnson et al., 1993). When mental representations contain a high degree of memory qualities usually associated with correct recollections, people frequently attribute these mental representations to memories of experienced events. However, when these representations contain qualities such as cognitive operations (e.g., thoughts), they are more likely to be attributed to imagination or reasoning processes. False memories originate from source monitoring errors when, for example, people attribute a mental representation to a memory for an experienced event while that experience was actually suggested by someone else (e.g., a police officer; Johnson et al., 1993).

Fuzzy-trace theory stipulates that when experiencing an event, two independent memory traces are stored (Brainerd et al., 2008): Gist and verbatim traces. Verbatim traces are involved in the storage of specific details of an experience (e.g., remembering that the color of the jacket of a bank robber was red), whereas gist traces refer to the storage of the underlying meaning of an event (e.g., remembering that a bank was robbed). According to FTT, verbatim traces fade more rapidly than gist traces, making people more reliant on gist traces over time. False memories are assumed to occur when people rely on such gist traces.

Associative-activation theory uses the principle of spreading activation to explain the formation of false memories (Howe et al., 2009; Otgaar et al., 2019a). According to the tenets of AAT, when experiencing an event, this experience leads to a spread of activation through a memory network containing nodes (e.g., memories) of related experiences. This spreading activation would also activate related nodes of events that were not experienced leading to false memories. False memories are especially likely to occur when spreading activation runs rapidly and automatically through a network and when relations between nodes are strong.

In the current review, we will focus on how such memory failures can arise. Specifically, the center of our discussion lies between the impact of a specific set of external and internal influences on both forgetting (i.e., omission) and false memory production (i.e., commission). It is relevant to stress here that myriad forms of external and internal influences exist and an exhaustive review of all of these influences is beyond the scope of the current review. Here, therefore, our focus is on a certain selection of these influences, keeping the following three considerations in mind. First, our discussion will concentrate on influences in which there was an (externally or internally) "active" overt attempt to affect memory. Second, we will describe influences that are often discussed in the legal realm. Third, we discuss these influences in tandem because the available research suggests that they exert similar effects on memory. We will both discuss relevant research conducted with adult and child samples to show how these influences might affect memory.

External influences on memory

Several studies have underlined the robust effect of suggestion on memory. Suggestion is called an external influence because oftentimes it originates from an external source (e.g., police officer, therapist, etc.). An abundance of studies using a variety of paradigms (e.g., misinformation paradigm, memory conformity, etc.) have shown that external suggestion can taint someone's memory (e.g., Loftus, 2005). Importantly, external suggestion can take two forms. People can suggest that details/events were present while actually they were not. However, people can also falsely suggest that certain details/events were not experienced, while in fact they were. These different variants of suggestion can lead to differential effects on memory (e.g., Merckelbach et al., 2007; Af Hjelmsäter et al., 2008; Wright et al., 2009; Otgaar et al., 2010a; Frenda et al., 2011; Scoboria et al., 2017; Azad et al., 2022; Rassin, 2022). We will now describe how these different forms of suggestion can lead to very specific memory effects (see also **Figure 1**).

Suggesting non-experience can lead to false memories

Several methods have been devised to investigate how suggestion of non-experienced details/events can impact memory. One of the most studied ones is the misinformation paradigm (Loftus, 2005). Basically, this paradigm follows a three-stage procedure. First, participants are presented with some stimuli (e.g., video of a burglary) or are involved in an interactive event (e.g., a science demonstration). Then, participants receive misinformation in the form of, for example, an eyewitness account containing false details (e.g., that the burglar stole jewellery while money was really stolen). Finally, a memory test is provided in which participants have to state which details they can still recollect. The misinformation effect refers to the finding that suggested false details are often reported by participants as having occurred during the first phase (Frenda et al., 2011).

Another paradigm used to externally engender entire false autobiographical experiences is the false memory implantation paradigm (Loftus and Pickrell, 1995). In this paradigm, participants are asked to report what they can still remember about events that ostensibly happened to them in their childhood. The important manipulation is that one of the events is false (i.e., being lost in a mall), being fabricated by the experimenters. After multiple interviews, during which the researchers suggested participants had experienced the false event, the canonical finding is that about 30% of participants fall prey to the suggestion and report having experienced such false event (Scoboria et al., 2017). Researchers have successfully implanted a wide array of false autobiographical events that share characteristics with events such as sexual abuse. For example, researchers have succeeded to implant painful events (e.g., being bitten by a dog; Porter et al., 1999), shameful events (e.g., swimming trousers falling off during swimming; Otgaar et al., 2021), and events that allegedly occurred more than once (Calado et al., 2021).

An additional way to evoke false memories is the memory conformity paradigm (Wright et al., 2009). This paradigm has been

used with three variants. In the first one, pairs of participants are presented with stimuli (e.g., picture of a desk). Participants are under the impression that they are witnessing the same stimuli, but each participant is viewing a slightly different version of the stimuli. For example, one participant might see a pen on the desk, while the other participant is presented with a desk without a pen. After the encoding phase, participants have to recall the stimuli collaboratively. What happens here is that participants will (unintentionally) suggestively influence each other's statements. During a final memory test, participants have to individually report what they can still remember concerning the stimuli.

In the second variant, group of participants are presented with the stimuli and, then, they discuss such stimuli. However, some participants are confederates of the experimenter who provide misleading information as actual elements of the stimuli. Finally, in the last variant, participants are simply presented with information of what was said by co-participants. Overall, these studies found that participants report having seen details that were actually suggested/discussed by the other participants (Wright et al., 2009).

Collectively, these paradigms largely show that external suggestion can lead to the production of false memories. Several theoretical explanations exist to explain how suggestion can foment false memories creation, and one of the most popular ones is the source monitoring framework, such that false memories due to external suggestion are basically source monitoring errors (e.g., Lindsay and Johnson, 2000).

Suggesting non-occurrence can lead to omission errors

Apart from suggesting that an event or detail was experienced while it was not, the reverse can also take place. That is, one can suggest that something was *not* experienced, while it actually was. Studies using this variant of suggestion have shown that this can lead to omission errors or failures to report experienced events (e.g., Pezdek and Roe, 1995; Wright et al., 2001; Merckelbach et al., 2007; Af Hjelmsäter et al., 2008; Otgaar et al., 2010a; Azad et al., 2022).

Compared with work on suggesting non-experiences, empirical research focusing on the suggestion of non-occurrence is quite limited. To study this, researchers have simply tweaked the usual false memory methods and focused on suggesting non-occurrences instead of suggesting non-experiences. For example, in one of the first of these studies focusing on suggesting non-occurrences conducted by Pezdek and Roe (1995), 4- and 10-year old children were touched in a specific way (e.g., hand on the children's shoulder) or not touched at all. Children were told that a different touch, a new touch, or no touch at all had happened. Of relevance to the current discussion is the condition in which children were touched but were told that nothing occurred. The authors found that children were not likely to accept the suggestion that no touch occurred and, hence, did not demonstrate significant more omission errors in their memory reports than children who did not receive the suggestion that no touch occurred.

Other studies, however, have shown that suggestion of non-occurrences *can* lead to omission errors. In two experiments, Wright et al. (2001) showed that post-event information suggesting that event was not experienced could make the memory concerning that event less accessible. In their experiments, participants saw certain stimuli (e.g., a restaurant scene depicted in slides). After this, they were again provided with these stimuli, but a critical scene (e.g., waitress taking an order) was omitted. Participants were instructed to use these stimuli to generate a story (Experiment 1) or imagine a scene (Experiment 2). During a final memory task, the important result was that the post-event omission led people not to report the critical scene in free recall and recognition. This effect has also been demonstrated when children were involved as participants (Williams et al., 2002).

In addition, recent work has extended our prior understanding of the memory consequences of suggestions of non-occurrence by further investigating this issue in a sample of adults (Azad et al., 2022). In three studies, Azad et al. (2022) asked participants to watch a video (i.e., child kidnapping case) and then exposed them to suggestions of non-occurrence once (Studies 1 and 3) or multiple times (Studies 2 and 3). In a final stage, participants' memory for the video was tested. Interestingly, single suggestions of non-occurrence did not make participants prone to omissions, but they did find that repeated suggestions of non-occurrence led participants to omit video-related information.

The just-mentioned studies used a rather subtle manipulation to induce omission errors. In Otgaar et al.'s (2010a) study, younger (4–5-year-olds) and older children (9–10-year-olds) had to remove three pieces of clothing of a puppet. In one condition, it was suggested to the children that they actually removed two pieces of clothing. This was done using a verbal suggestion and false evidence (putting one piece on the puppet again without the child noticing it). The authors found that although children initially claimed to take off three pieces of clothing, after the suggestion, a significant minority of children reported to have only removed two pieces of clothing (Otgaar et al., 2010a; see for similar results, Merckelbach et al., 2007; Af Hjelmsäter et al., 2008). Moreover, in a second study, Otgaar et al. (2010b) asked children to erroneously report that they only removed two pieces of clothing. This group had to complete a choice reaction time task consisting of pictures of different types of clothing. Their instruction was to indicate whether they removed these pieces of clothing or not. The primary result was that children made significantly more errors for removed pieces of clothing that they failed to report than for those they had not removed.

Contrary to the formation of false memories, little attention has been paid to the mechanisms underpinning omission errors. Pezdek and Roe (1995) referred to terms such as "erasing" memories thereby implying that the suggested memory is gone or –to use a less dramatic connotation– has become inaccessible to retrieval processes. However, it has also been shown that at least for omission errors in children, erasure was not a viable candidate to explain the failure to report experienced events (Otgaar et al., 2010b). Thus, a more promising explanation for omission errors is that suggesting non-occurrences does not impact the recollection of experienced events, but the belief that a particular event occurred. Alternatively, omission errors might simply refer to failures to report remembered information.

Suggestion (of non-experience and non-occurrence) can lead to false beliefs and non-believed memories

Previous research has mainly focused on the impact of suggesting non-experiences and non-occurrences on false memories and omission errors. However, recent research shows that these suggestions can even have more subtle effects on true and false memories. A recent surge of experimentation has shown that suggestion can impact the belief that an event occurred rather than the recollection of an event (e.g., Mazzoni et al., 2010; Clark et al., 2012; Otgaar et al., 2014b; Li et al., 2020). Believing that an event took place and recollecting an event are two different concepts contributing to the phenomenology of remembering. Belief refers to trusting that an event occurred, while recollection refers to re-experiencing an event including vivid images concerning this event (Scoboria et al., 2004, 2014). For perhaps most of our memories, we are prone to believe that an event occurred *and* to have vivid recollections concerning that event. However, for certain events, believing and recollecting are detached from each other. For example, people believe that they were born, but have no recollection of that event. Interestingly, for certain experiences, people have vivid recollections of an event, but no longer believe in the occurrence of that particular event. This latter type of memory has been called non-believed memories (Mazzoni et al., 2010; see for a review, Otgaar et al., 2014b). This counterintuitive memory phenomenon has stirred an abundance of research as it might clarify how suggestion can shape memory.

For example, it has been suggested that before a false memory for an event can be evoked, the event should be first considered plausible and then a belief that the event has occurred should have been formed (e.g., Mazzoni et al., 2001). This implies that suggesting non-experiences might also lead to false beliefs. This is indeed what research has been showing. For example, Scoboria et al. (2012)–using the false memory implantation paradigm–found that false suggestions increased false beliefs of a non-experienced event. Furthermore, Scoboria et al. (2017) performed a mega-analysis on false memory implantation studies and it was found that participants often expressed high belief in the occurrence of the falsely suggested event.

However, the reverse can occur as well. That is, when suggestion is provided about non-occurrences, belief can be affected as well. An increasing body of research is showing that suggesting non-occurrences can lead to reductions in belief and even end up in non-believed memories. In two experiments, Otgaar et al. (2013) used the false memory implantation paradigm to induce false memories of a hot air balloon ride. Adults (Experiment 1) and children (Experiment 2) were suggestively told that they experienced a hot air balloon ride when they were younger during multiple suggestive interviews. Importantly, when they were debriefed about true nature of the study, they were asked whether they still believed in the occurrence of the false event and still had recollections concerning the false event. The principal finding was that a significant proportion of subjects (13% for adults and 15% for children) developed non-believed memories after debriefing. That is, they still had a memory of going in a hot air balloon, but no longer believed that the event happened. Follow-up studies have confirmed the suggestion that non-occurrences can lead to belief changes and result in non-believed memories when (1) different paradigms are used (e.g., Clark et al., 2012), (2) suggestion is provided on true and false memories (e.g., Scoboria et al., 2012; Mazzoni et al., 2014), and (3) children and adults are tested (e.g., Otgaar et al., 2017). In addition, more recently, in two studies, Li et al. (2020) tested whether non-believed memories can also be reported for bizarre events in the standard imagination inflation paradigm. They asked participants to perform or imagine both simple familiar actions and bizarre actions. After 1 day, participants were invited to imagine simple actions of which some were new actions and some were actions performed the day before. After a week, participants completed a memory task and, when some actions were (correctly or incorrectly) recognized as performed, they were negatively challenged (i.e., participants were told that the action was not performed). The authors found that challenging actions that participants claimed to have performed decreased beliefs in these actions and led to the production of non-believed memories both for bizarre actions and familiar actions.

To recap, external influences can affect memory in different forms. When someone is suggestively told that an event or detail was experienced, while in fact this was not, a multitude of studies shows that such suggestions can facilitate the formation of false beliefs and false memories. Furthermore, when someone is suggestively told that a certain event did not occur while in fact it did, research indicates that it can lead to omission errors, belief reductions, and even non-believed memories. We now turn our attention to influences that are exerted internally and how they might contaminate memory performance.

Internal influences on memory

An increasing body of research is currently showing that deception is a powerful internal influence that can affect memory (e.g., Otgaar and Baker, 2018; Paige et al., 2022; Vo et al., 2022). Vrij (2008) defined lying as a "a successful or unsuccessful deliberate attempt, without forewarning, to create in another a belief which the communicator considers to be untrue" (p. 5). According to this definition, lying is exerted intentionally *by the one* exercising the lie. This is relevant because self-generation might lead to stronger memory contamination because it could be speculated that such self-generation makes the lie also more personally relevant (e.g., Howe et al., 2013; Wang et al., 2019; but see also Pezdek et al., 2009). Even though lying occurs on an almost daily basis in everyday life (Riesthuis et al., 2022a), this behavior is legally relevant because it is often exerted by suspects, victims, and witnesses (e.g., Vrij, 2008; Otgaar and Baker, 2018; Verigin et al., 2019).

Several deceptive strategies can be exercised and evidence is accruing that different forms of deception can lead to different memory effects. Otgaar and Baker (2018) argued that these differential memory effects might be caused by differences in cognitive resources that are needed to exercise certain types of lies (see also Battista et al., 2021a,c). In this section, we will focus on the memory effects of three types of lying: Fabrication, false denials, and feigning amnesia (see also **Figure 1**).

Fabrication can lead to false memories (and forgetting)

In legal contexts, fabrication is a common phenomenon. For example, perpetrators might willingly distort the truth or invent an entire story in order to mislead the police. Still witnesses and victims might come up with false information while answering police interviews, such as accusatorial interviews (Garven et al., 1998; Meissner et al., 2014). One of the most often-used methods to examine the impact of (self-generated information) fabrications on memory is the forced confabulation paradigm (Ackil and Zaragoza, 1998). In the first study using this paradigm, children and adults viewed a clip from a movie. After viewing the movie, participants were instructed to answer some questions concerning the movie. In the forced confabulation group, participants were told to provide an answer to every question and guess if they did not know the answer. By contrast, participants in the control condition only had to answer questions of which they were sure they knew the answer and were instructed not to guess. Importantly, participants were presented with questions about details that were presented in the movie, but were also presented with questions about non-presented details (e.g., a question about what was stolen when actually nothing was stolen). One week later, all participants received a source memory task. Specifically, they were asked whether they talked about certain details the week before and whether they saw these details in the video. The most important finding was that participants who were in the forced confabulation group claimed to have seen their own confabulations in the movie. In other words, forced confabulations led to the production of false memories for the confabulated responses.

Subsequent research has extended this work and, for example, showed that forcing participants to fabricate entire events (instead of details) can also generate false memories (e.g., Pickel, 2004; Chroback and Zaragoza, 2008). For example, Chroback and Zaragoza (2008) had participants view a clip from a movie (i.e., Looking for Miracles). Two days later, participants had to answer several interview questions of which some referred to false events. Participants were explicitly instructed to provide an answer to every question and guess if they did not know the answer. One week and 8 weeks after viewing the movie clip, participants received a recognition and recall test, respectively. Although false memory formation was limited after 1 week, after 8 weeks, participants claimed to have seen their own forced confabulations nearly 50% of the time.

Furthermore, apart from using the forced confabulation paradigm, other related research has also shown that self-generated fabrications can lead to false memories. In fact, Pickel (2004) showed that participants who fabricated misinformation themselves started to falsely remember this misinformation as being true. Schreiber et al. (2001) showed that when children were instructed to speculate about what objects could do, after a 5–6-month delay, children formed false answers to what these objects could do. Specifically, in their study, children received atypical actions for common objects (e.g., throwing a knife away). One week later, children were asked to speculate what else these objects could do (e.g., "What else can could he have done with a knife?"). The researchers found that inviting children to speculate could lead to false answers of these speculations at follow-up memory tests.

In short, lying, in the form of fabrication, can lead to the formation of false memories and this effect seems also to be not mitigated or exacerbated by other factors (e.g., incentive to lie, cognitive resources, personality traits) as shown in some recent experiments (Battista et al., 2021b,c, 2023; Riesthuis et al., 2022b; Battista et al., under review[2]). All these studies suggest that one explanation for this effect is that, just as false memories, fabrications can result in source monitoring errors because the fabrications appear phenomenologically similar to memories of experienced events.

Meanwhile, there is some limited evidence demonstrating that fabrication can also engender forgetting effects as well. For example, Pickel (2004) not only showed that self-generated misinformation was misremembered but that it also led participants to remember less about the target stimulus. A similar finding was observed by Riesthuis et al. (2022b) who found that creating a false alibi not only generated false memories but also resulted in omission errors. A possible interpretation for why fabrication led to forgetting is because the act of fabrication prevented participants to rehearse the experienced stimuli. This lack of rehearsal might have led to the forgetting of details concerning the event (see also Pickel, 2004; Riesthuis et al., 2022b).

False denials can lead to forgetting and false memories

A simpler deceptive strategy than fabrication is falsely denying that an experienced event unfolded. There is a vast literature showing that offenders of violent crimes (e.g., homicide, sexual abuse) oftentimes falsely deny that they committed a criminal act (e.g., Henning et al., 2005; Watson et al., 2016). Furthermore, false denials have been mentioned as one of several strategies that victims use to cope with sexually abusive experiences (Romeo et al., 2018; Ahern et al., 2019; Bücken et al., 2022c). A fundamental question here is whether the act of false denials might have memory impairing effects.

Research is amassing revealing that false denials can lead to omission errors. In the first study on this issue, Vieira and Lane (2013) instructed participants to study pictures of different objects (e.g., teacup). After this, participants received studied and unstudied objects and had to tell the truth or deny seeing these objects. The consequence was that for certain objects, they falsely denied studying these objects. Following this, participants were presented with a source memory test. Of relevance for the current discussion was the finding that participants forgot having falsely denied certain objects.

Otgaar et al. (2014a) found similar memory effects of false denials. In their experiment, they adapted the forced confabulation paradigm and added a false denial condition. Specifically, in their experiments, children (6–8- and 10–12-year-olds) and adults viewed a video and then received a memory test about details of that video. After this, participants were invited to lie (i.e., falsely deny or fabricate) or tell the truth about what seen in the video. Of importance for the current discussion are the false denial and

2 Battista, F., Otgaar, H., Riesthuis, P., and Mangiulli, I. (under review). *Lying on misleading information: False confirmation leads to fabricated memories.*

control conditions. After the memory test, participants in the false denial condition had to falsely deny seeing certain details while control participants had to tell the truth. One week later, participants received a source memory test in which they were asked whether they talked about certain details and whether they saw certain details in the video. The most interesting finding was that participants in the false denial forgot they had talked about certain details which in fact they did. This memory impairing effect of false denials has been dubbed denial-induced forgetting (Otgaar et al., 2016b).

After this first demonstration, the denial-induced forgetting effect has been observed using various stimuli like pictures (Otgaar et al., 2016b), virtual reality (Romeo et al., 2019) or daily life actions (Li et al., 2022a,b), and memory tasks (i.e., recognition and recall) (Otgaar et al., 2018). Taken together, false denials have been shown to lead to omission errors and especially omissions errors for details that were discussed rather than to the forgetting of the event.

Although some studies have found that false denials - in specific circumstances–might undermine our memory for the event, this work is limited. For instance, Battista et al. (2020) asked participants to repeatedly deny certain details while denying other details only once. They demonstrated that when details were denied four times, correct recall levels were lower than when details were denied once. Still, a detrimental effect of false denials on memory for the original event was also found by Romeo et al. (2019) in a study in which they tested the mnemonic impact of falsely denying emotional events. Similarly, another recent study (Battista et al., 2021a) found that when the false denials strategy requires a high involvement of cognitive resources to be employed, it can also result in a forgetting effect for the event (but see also Li and Liu, 2021).

Recent experimentation has shifted attention to the question whether false denials might also affect false memory production. The reasoning here was as follows. If false denials lead to omission errors then, based on theories such as FTT and AAT, such omission errors should affect the risk of false memory production. AAT would, for example, predict that when omission errors occur, activation will spread less to neighboring nodes thereby reducing the production of false memories (Howe et al., 2009). Evidence for this was found by Otgaar et al. (2020). They showed participants lists containing associatively-related words (e.g., tears, sorrow, grief) linked to a non-presented theme word (i.e., cry). After the encoding phase, half of the participants had to falsely deny seeing these words, while the other half had to tell the truth. During a final memory task, participants who had to deny created fewer false memories than truth-tellers (Experiment 1).

Recently, Bücken et al. (2022a,b), however, demonstrated that false denials can increase people's willingness to go along with false information. In one of their studies (Bücken et al., 2022b), participants viewed a video of a car crash and following this, half of them falsely denied that certain details were in the video while others had to tell the truth. After 1-week participants received misinformation concerning what happened during the interview and the car crash. False denials increased susceptibility to misinformation concerning the interview.

Scholars suggested that a lack of rehearsal might be a possible mechanism to explain the mnemonic consequences of false denials (Otgaar and Baker, 2018). Nevertheless, there are some recent indications that inhibition could be the mechanism underpinning the denial-induced forgetting effect and, occasionally, a forgetting of the event (Otgaar et al., 2020). The rationale here is that during the act of denial, retrieval of the-to-remembered event is temporarily inhibited leading to forgetting effects.

Feigning amnesia can lead to forgetting and false memories

A deceptive strategy that is also oftentimes used by offenders of violent crimes is pretending to suffer from memory loss for such event (e.g., Cima et al., 2002; Pyszora et al., 2003; Jelicic, 2018). Offenders claim amnesia for several reasons such as obstructing police investigations and interfering with legal proceedings (Tysse, 2005; Tysse and Hafemeister, 2006). In general, prevalence data show that about 30% of offenders who have committed violent crimes claim memory loss (see for a review, Mangiulli et al., 2021). Like other deceptive strategies such as fabrication and false denials, feigning amnesia has been shown to exert memory undermining effects.

Specifically, an increasing corpus of studies have shown that feigning amnesia can foster omission errors (e.g., Christianson and Bylin, 1999; Bylin and Christianson, 2002; Van Oorsouw and Merckelbach, 2004; Sun et al., 2009; Mangiulli et al., 2018a,b, 2019a,b). In the first study of this kind (Christianson and Bylin, 1999), participants were presented with a description of a crime and had to imagine being the offender of that particular crime. During a memory test, one group was instructed to feign memory loss for the crime while another group had to report the same event truthfully. One week later, all participants had to respond truthfully during a final memory test. The central finding was that those participants who feigned amnesia remembered fewer details (i.e., omissions) than truthful responders. Since then, research has replicated this effect using different stimuli (e.g., Van Oorsouw and Merckelbach, 2004; Sun et al., 2009; Mangiulli et al., 2018a,b). The memory undermining effect of feigning amnesia has been mostly attributed to a lack of rehearsal (but see for an alternative explanation, Mangiulli et al., 2019b). Indeed, studies that included a third group that was only tested after a delay showed that those feigning amnesia and participants in this delayed control did not statistically differ from each other in terms of memory performance. The reason for this is because both groups were less likely to rehearse the stimuli than the honest control group (Bylin and Christianson, 2002; Van Oorsouw and Merckelbach, 2004).

Apart from the finding that feigning amnesia can lead to omission errors, studies have also revealed that it can engender false memory creation. The explanation behind this is that feigning amnesia does not involve a single concrete deceptive strategy (see Mangiulli et al., 2020, 2021). That is, people who choose to feign amnesia can do so by just claiming memory loss ("*I do not remember*"), but also by adding fictitious details to their amnesic claim ("*I cannot remember because I was somewhere else during the crime*"). Interestingly, research in which participants were specifically instructed not to just deny the experience (and thus potentially fabricate details) demonstrated that they had elevated levels of false memories (Van Oorsouw and Giesbrecht, 2008). Recent research has shown that those feigning amnesia who decided to omit information were the ones with the lowest memory performance while those who also added false details in their feigned account for a crime reported the highest amount of commission errors (Mangiulli et al., 2019b, 2020).

However, when Mangiulli et al. (2020) examined whether feigning amnesia would also increase the risk of reporting misinformation, no evidence was observed. Thus, although the act of simulating amnesia can lead to errors of omission and commission, it does not seem to increase people's susceptibility to external pressure.

Lying can lead to changes in belief

Lying not only leads to false memories and omission errors, but there is some research showing that it can also affect the belief that an event took place (Polage, 2004, 2012, 2019; Romeo et al., 2018). For example, Polage (2004) asked participants to rate the likelihood that certain events happened to them before the age of ten. Approximately 2 weeks later, participants had to falsely claim to an experimenter that they experienced an event that they previously rated as unlikely to have happened them. One week later, participants had to truthfully rate the same events and indicate how likely it was that these events happened to them before the age of ten. In general, participants rated the events as less likely to have occurred to them after lying about them: An effect called fabrication deflation. However, what was also found was that a small subset of participants (10–16% in two studies) were more likely to claim that the events happened to them after lying about them, which Polage (2012) referred to as fabrication inflation effect. Interestingly, individual differences might play a role in this fabrication inflation effect as there is some preliminary evidence showing that high levels of dissociative experience might be positively related to the fabrication inflation effect (Polage, 2012).

So, it seems that fabrication might lead to increases in the belief that a non-experienced event occurred (but see also Riesthuis et al., 2020). Recent research has also focused on whether false denials might lead to decreases in the belief that an event occurred. Otgaar et al. (2016b) examined the denial-induced forgetting effect and compared a group that had to falsely deny that certain details were experienced and a group that was falsely suggested by an experimenter that certain details were not experienced. Decreases in belief were not found for the false denial group. However, Polage (2019) used a similar methodology in her fabrication inflation work and included a false denial condition and she did find that false denials led to decreases in the belief that events were experienced. Furthermore, Romeo et al. (2018) showed that feigning amnesia led to decreases in belief as well as recently Li et al. (2022b) found that when people mix different type of lies (i.e., false denial and fabrication) their beliefs for the occurrence of event-related details can decrease.

Taken together, the work on internal influences and memory, and more specifically the work on the impact of lying on memory, has shown that lying has differential effects on memory. Fabrication has been shown to lead to increases in belief in the occurrence of the self-generated information and false memories, while false denials have primarily been found to lead to omission errors. In addition, depending on which specific strategy is used, feigning amnesia has been found to lead to belief changes, omission errors, and false memories (see also **Figure 1**).

Future perspectives

Despite one being other-generated and the other self-generated, it is evident throughout this review that external and internal influences oftentimes exert similar effects on memory. For example, as displayed above, suggesting false experiences as well as self-generating false information (and feigning amnesia) can lead to false memories. Moreover, suggesting non-occurrence and falsely denying or feigning amnesia for experienced event can both lead

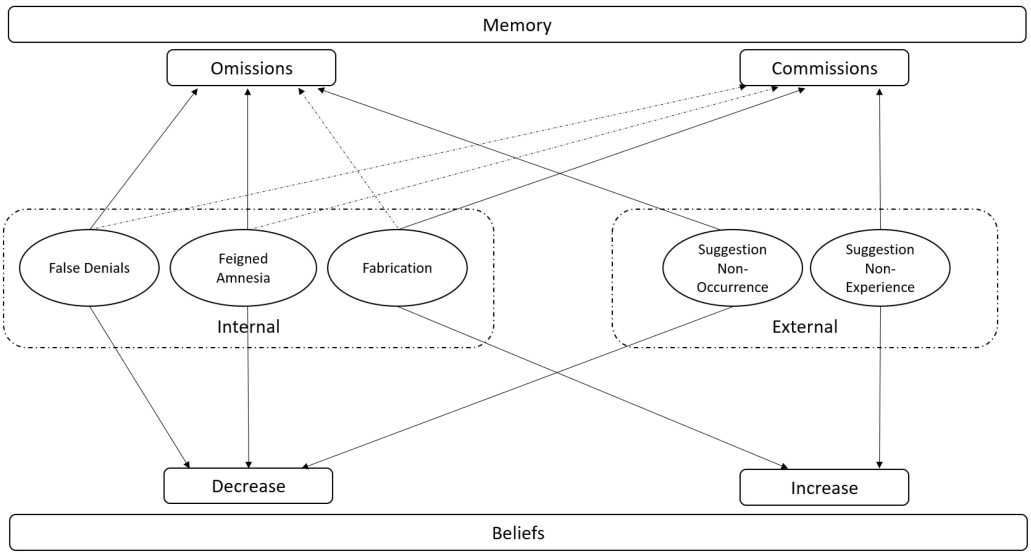


FIGURE 1
Schematic overview on how internal and external influences affect memory. In the figure, it can be seen that there are different internal and external influences and that they both can lead to changes in memory and belief. Please note that dotted lines indicate that only some studies detected a relationship between the specific factor and memory outcome.

to forgetting. Finally, it appears that both types of influences can impact the belief in the truth value of an experienced event.

Having established the effects of both external and internal influences on memory, a timely question arises. This question–which likely could orient future research paths–concerns examining whether these effects are (partially) driven by a common mechanism. We propose that one such a mechanism could be cognitive dissonance (i.e., a displeasing psychological state caused by a mismatch between two or more elements in a cognitive structure; Festinger, 1957, 1962). That is, dissonance is thought to play an essential role in whether belief is reduced or not when suggestion is provided about non-occurrences (Scoboria et al., 2014). According to the model postulated by Scoboria and Henkel (2020), when someone is suggestively told that their memory is incorrect, dissonance arises. Such dissonance can be at an interpersonal or intrapersonal level. Concerning interpersonal dissonance, the idea is that people would evaluate the costs and benefits of (dis)agreeing with the suggestion. If people agree with the suggestion, a reduction in belief might take place. On an intrapersonal level, instead, people would evaluate the suggestion with their own memory (e.g., if the suggestion pertains to a memory which is vague). Here too, if the suggestion is accepted, it is likely that belief in the occurrence of an event will be undermined. Evidence of this possibility comes also from recent studies investigating a possible relationship between non-believed memories and memory distrust (i.e., people's distrust toward their own memories) (Zhang et al., 2022a,b). These studies found support for a positive association between memory distrust and non-believed memories such that memory distrust was assumed to be a reason why people reduce beliefs in the occurrence of events. In addition, the idea that cognitive dissonance can play a crucial role in how internal and external influences can affect our memory comes from work–albeit limited–showing that dissonance is also related to the production of false memories and lying (e.g., Merckelbach and Merten, 2012; Rodriguez and Strange, 2015). For example, Rodriguez and Strange (2014) had participants make an easy or difficult choice between two smartphones. Following this, participants were instructed to remember their decision experience. Participants receiving the difficult choice experienced cognitive dissonance and were more likely to misremember their initial decisions than participants receiving the easy choice.

Cognitive dissonance might also be related to when people lie and then come forward with the truth. That is, some lies (e.g., false denials) might be displeasing if they are exercised under pressure and hence, create a mismatch with a memory for an experienced event. Therefore, based on the proposition that cognitive dissonance is assumed to play a role in in how external and internal influences affect memory, several specific future research enterprises and predictions can be postulated. For example, if dissonance plays a role in how false denials lead to omission errors, then the following might be expected: Omission errors would be more likely to occur when, for example, the memory of the experienced event is weak because of high intrapersonal dissonance. The reason is because when dissonance takes place, people will simply try to resolve it. So, people would only agree to the false denial of the event if the denial does not conflict too much with their own experience. This means that when people have difficulties in remembering an event, the act of false denials will more likely be accepted, hence leading to omission errors. A possible way to empirically test this idea in future studies is by having participants experience an event and then assigning some of them to a group that has to immediately deny experiencing the event, while others have to falsely deny experiencing the event a week later (i.e., delayed group). The prediction would be that the latter group will have a weaker memory performance for the event than the other one, making it more susceptible to intrapersonal dissonance. This, in turn, would lead the false denials to robust memory undermining effects.

Beyond the idea of testing a possible effect of dissonance, there are also other routes that might be fruitful to explore. For example, one interesting avenue is to examine when external or internal influences affect memory and/or belief. Based on earlier models and research (e.g., Mazzoni et al., 2001; Scoboria et al., 2004, 2014), the idea is that people first form a belief that an event happened and after that a recollection of an event is created. This work has also shown that beliefs are more malleable than recollections (e.g., Otgaar et al., 2014b). A critical question for future experimentations is to investigate whether manipulating the levels of the impact of external/internal influences can divergently affect beliefs and recollections. That is, there might be some forms of dose-response relationship in that weaker forms of external/internal influences (e.g., subtle external suggestion using misinformation in an eyewitness testimony) are more likely to affect belief, while stronger forms of external/internal influences (e.g., a policeman providing harsh suggestive interviewing tactics) are more likely to target recollection.

One might also wonder whether the observed effects of lying on memory are perhaps due to fact that participants were "instructed" to lie, while in real life settings, witnesses, victims, and suspects frequently choose to lie. An imperative question is to empirically test the proposition that "instructed" lies have different effects on memory than "voluntary" lies. Although limited, recent research has examined whether the volitional act of lying has different effects on people's memory than when they are instructed to lie. Interestingly, these studies observed that similar memory undermining effects are detected when participants can freely choose to deceive, thereby suggesting that the act of lie is the determining factor in the observed memory effects (Dianiska and Meissner, 2022; Li et al., 2022a; Riesthuis et al., 2022d). Of course, future research could increase the knowledge base in this area and attempt to replicate these recent studies.

Another important avenue for research could be investigating what are the memory consequences caused by the interplay of different influences. Indeed, in the current review, we have focused on how lying and external suggestions can taint memory and result in forgetting and false memories. However, it is important to be cognizant of the fact that such memory failures can also arise because these influences might well work in tandem (e.g., Mangiulli et al., 2020; Bücken et al., 2022a,b; see text footnote 2). The investigation of such interactions might result into a more all-encompassing understanding on how different types of influences impact memory.

Legal implications

Wrongful convictions can be caused by memory failures. For example, suggestive therapeutic sessions can lead to false memories

of sexual abuse leading to false accusations and miscarriages of justice (Howe et al., 2018; Otgaar et al., 2019b). Also, data from the American Innocence Project has revealed that about 70% of wrongful convictions were the result of eyewitness misidentification (Innocence Project, 2020; see also Wells and Olson, 2003; Saks and Koehler, 2005). Such misidentifications, which may lead innocent people to be imprisoned, are memory failures and they have received a wealth of empirical attention within the psychological and legal realm. Importantly, such false positives are sometimes regarded as more serious than false negatives (acquitting a guilty person), an adage also known as the Blackstone ratio (Blackstone, 1765; see also de Keijser et al., 2014).

However, likely because of sentiments as the Blackstone ratio, other memory impairments (e.g., forgetting) perhaps did not receive so much attention as the former. However, omission errors and not believing that a certain event took place can have egregious consequences in police investigations and legal cases. For example, when witnesses are unable to remember how a certain criminal experience exactly unfolded, it might become difficult for the police to find a suspect, wasting unnecessary resources. Also, if a victim of abuse expresses low belief that the abusive event truly happened, an accusation might not be taken seriously and would hinder fact-finding in a criminal investigation. But also suggesting non-experiences to victims, witnesses, and suspects might lead to misidentifications and false confessions (e.g., Zajac and Henderson, 2009; Frenda et al., 2011; Scherr et al., 2020), such as when the perpetrators silence their victims claiming that nothing happened (e.g., Shepp et al., 2019).

Interestingly, we have additionally shown that memory failures such as forgetting and false memories can not only be prompted by external influences, but can also be initiated by means of internal influences. What we have shown is that false denials (and feigning amnesia too) can result into forgetting and decreases in belief, while fabrication (and as well as feigning amnesia in certain circumstances) can boost false beliefs and false memory formation. Collectively, this work has demonstrated that lying can exert similar effects on memory as external influences. However, from a practical perspective, research on how lying affects memory is still limited (see also Battista and Otgaar, 2022). The issue of lying has often been examined, but this examination is predominantly in the context of deception detection (e.g., Granhag et al., 2015). Although research in the area of deception detection sometimes uses principles of memory (e.g., recognition) to detect concealed knowledge (e.g., Verschuere et al., 2011), the work described here ascribes causal effects of lying on memory, clearly demonstrating that the act of lying can have deteriorating effects on memory. This pattern of results can be relevant for different professionals working in the legal arena. For example, memory researchers working as expert witnesses are often asked to estimate the reliability of testimonies of witnesses, victims, and suspects (Otgaar et al., 2017). What such expert witnesses basically do is to evaluate whether statements might have been affected by, for example, suggestive interviewing techniques. However, so far, the impact of internal influences on testimonial accuracy is not clear yet. To give a case example, it is common that victims of sexual abuse falsely deny being abused and only after a certain period of time come forward with the truth (Magnusson et al., 2017). Memory experts who are asked to evaluate the reliability of this victim's statement might now also note that false denials can have detrimental effects on memory too. Similarly, legal professionals (e.g., police officers), who know that fabrication can result into false memories, would dismiss the use of coercive and accusatorial interrogations that make the interviewee more likely to come up with false information for the forgotten crime-related details (Garven et al., 1998).

On a related note, one issue that has to be at the foreground concerning the legal implications of the reviewed work is to what extent effects observed in experiments on external and internal influences on memory are meaningful and practically relevant. That is, although the reviewed literature shows that these influences can negatively impact memory, a basic but forthright question is whether these findings have any practical meaning. So, experiments conducted in this area should establish certain effect sizes and the question is whether such effect sizes bear any relevance in actual legal cases. This question can only be answered if, as a field, we agree to some extent on which effect sizes are of relevance in practical settings. In other words, the problem is understanding to what extent the achieved results of psychological studies are sufficiently informative and can bring a significant contribution for legal professionals' practice (Riesthuis et al., 2022c). Specifically, this question is related to what the smallest effect size is of interest in these memory experiments (see Lakens et al., 2018). That is, we recently argued that memory experiments should contain elements that can generate effects of interest for the (legal) field (see Otgaar et al., 2022b). If we consider what the smallest effect size is of interest for research on external and internal influences on memory, our argument is that even when such influences (e.g., false denial, suggestion non-experiences) lead to increases or decreases of only one (falsely) remembered detail, this might be of high value to the legal field. This is vital because even one remembered (or forgotten) detail can be determining, for instance, in the reconstruction of the crime (see also Mangiulli et al., 2021; Riesthuis et al., 2022b). Taken together, establishing which effects are of interest concerning the impact of external and internal influences on memory might lead to stronger experiments to demonstrate such effects. This might also reveal whether internal or external influences evince larger effects on memory and which one might be more practically relevant. Hence, we believe that if memory researchers are planning new studies and want to conduct an *a priori* power analysis, they should estimate which effect size is needed to establish an effect of interest. If the field agrees that effect sizes such as the one explained above (i.e., increase/decrease one remembered/forgotten detail) are of practical interest, stronger memory experiments can be built.

Concluding remarks

In what way can memory be shaped? In the present review, we have demonstrated that external and internal influences can exert similar effects on memory. Specifically, we showed that forgetting and false memories can arise when people are exposed to suggestion of non-occurrences and non-experiences, respectively. Similarly, such forgetting and false memories can also be produced when people falsely deny, feign amnesia for, and fabricate events. Furthermore, we have demonstrated that these influences can also lead to amplifications and reductions in the belief that an event occurred. We speculated that focusing on whether cognitive dissonance might be a centerpiece mechanism will likely

engender novel research on how internal and external influences can shape memory.

Author contributions

HO conceived the project and wrote the manuscript. MH, IM, and FB wrote the manuscript and critically revised it. All authors contributed to the article and approved the submitted version.

Funding

The current manuscript has been supported by a C1 (Application number: C14/19/013) and FWO Research Project grant (Application number: G0D3621N; Research Foundation—Flanders, Belgium) awarded to HO.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Ackil, J. K., and Zaragoza, M. S. (1998). Memorial consequences of forced confabulation: Age differences in susceptibility to false memories. *Dev. Psychol.* 34, 1358–1372. doi: 10.1037/0012-1649.34.6.1358

Af Hjelmsäter, E. R., Granhag, P. A., Strömwall, L. A., and Memon, A. (2008). The effects of social influence on children's memory reports: The omission and commission error asymmetry. *Scand. J. Psychol.* 49, 507–513. doi: 10.1111/j.1467-9450.2008.00679.x

Ahern, E. C., Hershkowitz, I., Lamb, M. E., Blasbalg, U., and Karni-Visel, Y. (2019). Examining reluctance and emotional support in forensic interviews with child victims of substantiated physical abuse. *Appl. Dev. Sci.* 23, 227–238. doi: 10.1080/10888691.2017.1387057

Anderson, M. C., Bjork, R. A., and Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 1063–1087.

Anderson, M. C., and Green, C. (2001). Suppressing unwanted memories by executive control. *Nature* 410, 366–369. doi: 10.1038/35066572

Azad, T., Lindsay, D. S., and Zaragoza, M. S. (2022). Can suggestions of non-occurrence lead to claims that witnessed events did not happen? *J. Gen. Psychol.* 149, 349–370. doi: 10.1080/00221309.2020.1860889

Battista, F., Curci, A., Mangiulli, I., and Otgaar, H. (2021a). What can we remember after complex denials? The impact of different false denials on memory. *Psychol. Crime Law* 27, 914–931. doi: 10.1080/1068316X.2020.1865956

Battista, F., Mangiulli, I., Riesthuis, P., Curci, A., and Otgaar, H. (2021c). Do liars really remember what they lied upon? The impact of fabrication on memory. *Memory* 29, 1076–1090. doi: 10.1080/09658211.2021.1960380

Battista, F., Otgaar, H., Mangiulli, I., and Curci, A. (2021b). The role of executive functions in the effects of lying on memory. *Acta Psychol.* 215:103295. doi: 10.1016/j.actpsy.2021.103295

Battista, F., Lanciano, T., Curci, A., Mirandola, C., and Otgaar, H. (2023). I lie because I am good at: Psychopathic traits do not influence the effects of fabrication on memory. *J. Cogn. Psychol.* doi: 10.1080/20445911.2023.2245603

Battista, F., Mangiulli, I., Herter, H., Curci, A., and Otgaar, H. (2020). The effects of repeated denials and fabrication on memory. *J. Cogn. Psychol.* 32, 369–381. doi: 10.1080/20445911.2020.1767626

Battista, F., and Otgaar, H. (2022). Research on the effects of lying on memory: A scientometric analysis and a call for new studies. *Front. Psychol.* 13:837265. doi: 10.3389/fpsyg.2022.837265

Bell, D. E., and Loftus, E. F. (1988). Degree of detail of eyewitness testimony and mock juror judgments. *J. Appl. Soc. Psychol.* 18, 1171–1192. doi: 10.1111/j.1559-1816.1988.tb01200.x

Bergström, Z. M., Velmans, M., de Fockert, J., and Richardson-Klavehn, A. (2007). ERP evidence for successful voluntary avoidance of conscious recollection. *Brain Res.* 1151, 119–133.

Bjork, E. L., and Bjork, R. A. (1996). Continuing influences of to-be-forgotten information. *Conscious. Cogn.* 5, 176–196. doi: 10.1006/ccog.1996.0011

Bjork, R. A. (1989). "Retrieval inhibition as an adaptive mechanism in human memory," in *Varieties of memory and consciousness: Essays in honor of Endel Tulving*, eds H. L. Roediger and F. I. M. Craik (Hillsdale, NJ: Erlbaum), 309–330.

Blackstone, W. (1765). *Commentaries on the Laws of England*. Oxford: Clarendon Press.

Brainerd, C. J., Reyna, V. F., and Ceci, S. J. (2008). Developmental reversals in false memory: A review of data and theory. *Psychol. Bull.* 134, 343–382. doi: 10.1037/0033-2909.134.3.343

Bücken, C. A., Otgaar, H., London, K., Riesthuis, P., Battista, F., and Mangiulli, I. (2022c). 'Nothing happened': Legal implications of false denials among abused children. *Child Abuse Rev.* 32, e2791. doi: 10.1002/car.2791

Bücken, C. A., Mangiulli, I., and Otgaar, H. (2022a). Simulating denial increases false memory rates for abuse unrelated information. *Behav. Sci. Law* 40, 433–451. doi: 10.1002/bsl.2566

Bücken, C. A., Mangiulli, I., Uzun, S., and Otgaar, H. (2022b). False denials increase false memories for trauma-related discussions. *Memory* 30, 1158–1171. doi: 10.1080/09658211.2022.2094964

Bylin, S., and Christianson, S. Å (2002). Characteristics of malingered amnesia: Consequences of withholding vs. distorting information on later memory of a crime event. *Legal Criminol. Psychol.* 7, 45–61. doi: 10.1348/135532502168379

Calado, B., Luke, T. J., Connolly, D. A., Landström, S., and Otgaar, H. (2021). Implanting false autobiographical memories for repeated events. *Memory* 29, 1320–1341. doi: 10.1080/09658211.2021.1981944

Christianson, S. A., and Bylin, S. (1999). Does simulating amnesia mediate genuine forgetting for a crime event? *Appl. Cogn. Psychol.* 13, 495–511. doi: 10.1002/(SICI)1099-0720(199912)13:6<495::AID-ACP615>3.0.CO;2-0

Chroback, Q. M., and Zaragoza, M. S. (2008). Inventing stories: Forcing witnesses to fabricate entire fictitious events leads to freely reported false memories. *Psychon. Bull. Rev.* 15, 1190–1195. doi: 10.3758/PBR.15.6.1190

Cima, M., Merckelbach, H. L. G. J., Nijman, H., Knauer, E., and Hollnack, S. (2002). I can't remember your honor: Offenders who claim amnesia. *German J. Psychiatry* 5, 24–34.

Clark, A., Nash, R. A., Fincham, G., and Mazzoni, G. (2012). Creating non-believed memories for recent autobiographical events. *PLoS One* 7:32998. doi: 10.1371/journal.pone.0032998

Conway, M. A., Harries, K., Noyes, J., Racsma'ny, M., and Frankish, C. R. (2000). The disruption and dissolution of directed forgetting: Inhibitory control of memory. *J. Memory Lang.* 43, 409–430. doi: 10.1006/jmla.2000.2706

Curci, A., Lanciano, T., Curtotti, D., and Sartori, G. (2020). Lessons for the courtroom from the study of Flashbulb memory: An integrative review. *Memory* 28, 441–449. doi: 10.1080/09658211.2020.1727522

de Keijser, J. W., de Lange, E. G., and van Wilsem, J. A. (2014). Wrongful convictions and the Blackstone ratio: An empirical analysis of public attitudes. *Punish. Soc.* 16, 32–49. doi: 10.1177/1462474513504800

Dianiska, R. E., and Meissner, C. A. (2022). The effect of lying on memory and metamemory when deception is repeated and volitional. *J. Appl. Res. Mem. Cogn.* 12, 128–140. doi: 10.1037/mac0000036

Ernberg, E., Magnussen, M., Landström, S., and Tidefors, I. (2018). Court evaluations of young children's testimony in child sexual abuse cases. *Legal Criminol. Psychol.* 23, 176–191. doi: 10.1111/lcrp.12124

Fawcett, J. M., and Hulbert, J. C. (2019). The many faces of forgetting. Toward a constructive view of forgetting in everyday life. *J. Appl. Res. Memory Cogn.* 9, 1–18. doi: 10.1016/j.jarmac.2019.11.002

Festinger, L. (1957). *A theory of cognitive dissonance*. Standford, CA: Standford University Press. doi: 10.1515/9781503620766

Festinger, L. (1962). *A theory of cognitive dissonance*. Standford, CA: Stanford university press. doi: 10.1038/scientificamerican1062-93

Fraundorf, S. H., Hourihan, K. L., Peters, R. A., and Benjamin, A. S. (2019). Aging and recognition memory: A meta-analysis. *Psychol. Bull.* 145, 339–371. doi: 10.1037/bul0000185

Frenda, S. J., Nichols, R. M., and Loftus, E. F. (2011). Current issues and advances in misinformation research. *Curr. Dir. Psychol. Sci.* 20, 20–23. doi: 10.1177/0963721410396620

Garven, S., Wood, J. M., Malpass, R. S., and Shaw, J. S. (1998). More than suggestion: The effect of interviewing techniques from the McMartin preschool case. *J. Appl. Psychol.* 83, 347–359. doi: 10.1037/0021-9010.83.3.347

Granhag, P. A., Vrij, A., and Verschuere, B. (eds) (2015). *Detecting deception. Current challenges and cognitive approaches*. Chichester: Wiley. doi: 10.1002/9781118510001

Henning, K., Jones, A. R., and Holdford, R. (2005). "I didn't do it, but if I did I had a good reason": Minimization, denial, and attributions of blame among male and female domestic violence offenders. *J. Family Viol.* 20, 131–139. doi: 10.1007/s10896-005-3647-8

Howe, M. L. (2011). The adaptive nature of memory and its illusions. *Curr. Dir. Psychol. Sci.* 20, 312–315. doi: 10.1177/0963721411416571

Howe, M. L., and Knott, L. M. (2015). The fallibility of memory in judicial processes: Lessons from the past and their modern consequences. *Memory* 23, 633–656. doi: 10.1080/09658211.2015.1010709

Howe, M. L., Knott, L. M., and Conway, M. A. (2018). *Memory and Miscarriages of Justice*. Abingdon: Routledge. doi: 10.4324/9781315752181

Howe, M. L., Threadgold, E., Norbury, J., Garner, S., and Ball, L. J. (2013). Priming children's and adults' analogical problem solutions with true and false memories. *J. Exp. Child Psychol.* 116, 96–103. doi: 10.1016/j.jecp.2013.03.006

Howe, M. L., Threadgold, E., Wilkinson, S., Garner, S. R., and Ball, L. J. (2017). "The positive side of memory illusions: New findings about how false memories facilitate reasoning and problem solving," in *False and distorted memories*, eds R. A. Nash and J. Ost (Abingdon: Psychology Press), 127–140.

Howe, M. L., Wimmer, M. C., Gagnon, N., and Plumpton, S. (2009). An associative-activation theory of children's and adults' memory illusions. *J. Memory Lang.* 60, 229–251. doi: 10.1016/j.jml.2008.10.002

Innocence Project (2020). *We work to free the innocent, prevent wrongful convictions, and create fair, compassionate, and equitable systems of justice for everyone*. New York, NY: Innocence Project.

Jelicic, M. (2018). Testing claims of crime-related amnesia. *Front. Psychiatry* 9:617. doi: 10.3389/fpsyt.2018.00617

Johnson, M. K., Hashtroudi, S., and Lindsay, D. S. (1993). Source monitoring. *Psychol. Bull.* 114, 3–28. doi: 10.1037/0033-2909.114.1.3

Joormann, J., Hertel, P. T., Brozovich, F., and Gotlib, I. H. (2005). Remembering the good, forgetting the bad: Intentional forgetting of emotional material in depression. *J. Abnorm. Psychol.* 114:640.

Lakens, D., Scheel, A. M., and Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Adv. Methods Pract. Psychol. Sci.* 1, 259–269. doi: 10.1177/2515245918770963

Li, C., Wang, J., and Otgaar, H. (2020). Creating nonbelieved memories for bizarre actions using an imagination inflation procedure. *Appl. Cogn. Psychol.* 34, 1277–1286. doi: 10.1002/acp.3722

Li, Y., and Liu, Z. (2021). Involvement modulates the effects of deception on memory in daily life. *Front. Psychol.* 12:4639. doi: 10.3389/fpsyg.2021.756297

Li, Y., Liu, Z., and Liu, X. (2022a). Who did I lie to that day? Deception impairs memory in daily life. *Psychol. Res.* 86, 1763–1773. doi: 10.1007/s00426-021-01619-x

Li, Y., Liu, Z., and Liu, X. (2022b). More lies lead to more memory impairments in daily life. *Front. Psychol.* 13:822788. doi: 10.3389/fpsyg.2022.822788

Lindsay, D. S., and Johnson, M. K. (2000). False memories and the source monitoring framework: Reply to Reyna and Lloyd (1997). *Learn. Ind. Differ.* 12, 145–161. doi: 10.1016/S1041-6080(01)00035-8

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

Loftus, E. F., and Pickrell, J. E. (1995). The formation of false memories. *Psychiatr. Ann.* 25, 720–725. doi: 10.3928/0048-5713-19951201-07

Macleod, C. M. (1999). The item and list methods of directed forgetting: Test differences and the role of demand characteristics. *Psychon. Bull. Rev.* 6, 123–129. doi: 10.3758/BF03210819

Macleod, C. M. (2012). Directed forgetting. *Encycloped. Sci. Learn.* 993–995. doi: 10.1007/978-1-4419-1428-6_1084

Magnusson, M., Ernberg, E., and Landström, S. (2017). Preschoolers' disclosures of child sexual abuse: Examining corroborated cases from Swedish courts. *Child Abuse Neglect* 70, 199–209. doi: 10.1016/j.chiabu.2017.05.018

Mangiulli, I., Lanciano, T., Jelicic, M., van Oorsouw, K., Battista, F., and Curci, A. (2018a). Can implicit measures detect source information in crime-related amnesia? *Memory* 26, 1019–1029. doi: 10.1080/09658211.2018.1441421

Mangiulli, I., Van Oorsouw, K., Curci, A., Merckelbach, H., and Jelicic, M. (2018b). Feigning amnesia moderately impairs memory for a mock crime video. *Front. Psychol.* 6:625. doi: 10.3389/fpsyg.2018.00625

Mangiulli, I., Lanciano, T., van Oorsouw, K., Jelicic, M., and Curci, A. (2019a). Do reminders of the crime reverse the memory-undermining effects of simulating amnesia? *Memory Cogn.* 47, 1375–1385. doi: 10.3758/s13421-019-00939-z

Mangiulli, I., Van Oorsouw, K., Curci, A., and Jelicic, M. (2019b). Retrieval-induced forgetting in the feigning amnesia for a crime paradigm. *Front. Psychol.* 10:928. doi: 10.3389/fpsyg.2019.00928

Mangiulli, I., Otgaar, H., Curci, A., and Jelicic, M. (2020). An experimental investigation of the misinformation effect in crime-related amnesia claims. *Appl. Cogn. Psychol.* 34, 1092–1100. doi: 10.1002/acp.3697

Mangiulli, I., Riesthuis, P., and Otgaar, H. (2021). The memory-undermining effect of simulated crime-related amnesia and its legal implications: A review. *Psychol. Injury Law* 15, 213–226. doi: 10.1007/s12207-021-09441-x

Mazzoni, G., Clark, A., and Nash, R. A. (2014). Disowned recollections: Denying true experiences undermines belief in occurrence but not judgments of remembering. *Acta Psychol.* 145, 139–146. doi: 10.1016/j.actpsy.2013.11.007

Mazzoni, G., Scoboria, A., and Harvey, L. (2010). Nonbelieved memories. *Psychol. Sci.* 21, 1334–1340. doi: 10.1177/0956797610379865

Mazzoni, G. A., Loftus, E. F., and Kirsch, I. (2001). Changing beliefs about implausible autobiographical events: A little plausibility goes a long way. *J. Exp. Psychol.* 7:51. doi: 10.1037/1076-898X.7.1.51

McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychol. Rev.* 39, 352–370. doi: 10.1037/h0069819

McNally, R. J. (2005). Debunking myths about trauma and memory. *Can. J. Psychiatry* 50, 817–822. doi: 10.1177/070674370505001302

Meissner, C. A., Redlich, A. D., Michael, S. W., Evans, J. R., Camilletti, C. R., Bhatt, S., et al. (2014). Accusatorial and information-gathering interrogation methods and their effects on true and false confessions: A meta-analytic review. *J. Exp. Criminol.* 10, 459–486. doi: 10.1007/s11292-014-9207-6

Merckelbach, H., and Merten, T. (2012). A note on cognitive dissonance and malingering. *Clin. Neuropsychol.* 26, 1217–1229. doi: 10.1080/13854046.2012.710252

Merckelbach, H., Van Roermund, H., and Candel, I. (2007). Effects of collaborative recall: Denying true information is as powerful as suggesting misinformation. *Psychol. Crime Law* 13, 573–581. doi: 10.1080/10683160601160679

Nørby, S. (2015). Why forget? On the adaptive value of memory loss. *Perspect. Psychol. Sci.* 10, 551–578. doi: 10.1177/1745691615596787

Otgaar, H., and Baker, A. (2018). When lying changes memory for the truth. *Memory* 26, 2–14. doi: 10.1080/09658211.2017.1340286

Otgaar, H., Candel, I., Smeets, T., and Merckelbach, H. (2010a). 'You didn't take Lucy's skirt off': The effect of misleading information on omissions and commissions in children's memory reports. *Legal Criminol. Psychol.* 15, 229–241. doi: 10.1348/135532509X471951

Otgaar, H., Meijer, E. H., Giesbrecht, T., Smeets, T., Candel, I., and Merckelbach, H. (2010b). Children's suggestion-induced omission errors are not caused by memory erasure. *Conscious. Cogn.* 19, 265–269. doi: 10.1016/j.concog.2009.09.015

Otgaar, H., Curci, A., Mangiulli, I., Battista, F., Rizzotti, E., and Sartori, G. (2022a). A court ruled case on therapy-induced false memories. *J. For. Sci.* 67, 2122–2129. doi: 10.1111/1556-4029.15073

Otgaar, H., Howe, M. L., and Dodier, O. (2022b). What can expert witnesses reliably say about memory in the courtroom? *For. Sci. Int.* 3:100106. doi: 10.1016/j.fsiml.2022.100106

Otgaar, H., de Ruiter, C., Howe, M. L., Hoetmer, L., and van Reekum, P. (2017). A case concerning children's false memories of abuse: Recommendations regarding expert witness work. *Psychiatry Psychol. Law* 24, 365–378. doi: 10.1080/13218719.2016.1230924

Otgaar, H., Howe, M. L., Smeets, T., and Wang, J. (2016b). Denial-induced forgetting: False denials undermine memory, but external denials undermine belief. *J. Appl. Res. Memory Cogn.* 5, 168–175. doi: 10.1016/j.jarmac.2016.04.002

Otgaar, H., Howe, M. L., Brackmann, N., and Smeets, T. (2016a). The malleability of developmental trends in neutral and negative

memory illusions. *J. Exp. Psychol.* 145, 31–55. doi: 10.1037/xge0000127

Otgaar, H., Howe, M. L., Mangiulli, I., and Bücken, C. (2020). The impact of false denials on forgetting and false Memory. *Cognition* 202:104322. doi: 10.1016/j.cognition.2020.104322

Otgaar, H., Howe, M. L., Memon, A., and Wang, J. (2014a). The development of differential mnemonic effects of false denials and forced confabulations. *Behav. Sci. Law* 32, 718–731. doi: 10.1002/bsl.2148

Otgaar, H., Scoboria, A., and Mazzoni, G. (2014b). On the existence and implications of nonbelieved memories. *Curr. Dir. Psychol. Sci.* 23, 349–354. doi: 10.1177/0963721414542102

Otgaar, H., Howe, M. L., Patihis, L., Merckelbach, H., Lynn, S. J., Lilienfeld, S. O., et al. (2019b). The return of the repressed: The persistent and problematic claims of long-forgotten trauma. *Perspect. Psychol. Sci.* 14, 1072–1095. doi: 10.1177/1745691619862306

Otgaar, H., Howe, M. L., Muris, P., and Merckelbach, H. (2019a). Associative activation as a mechanism underlying false memory formation. *Clin. Psychol. Sci.* 7, 191–195. doi: 10.1177/2167702618807189

Otgaar, H., Howe, M. L., van Beers, J., van Hoof, R., Bronzwaer, N., and Smeets, T. (2015). The positive ramifications of false memories using a perceptual closure task. *J. Appl. Res. Mem. Cogn.* 4, 43–50. doi: 10.1016/j.jarmac.2014.12.001

Otgaar, H., Moldoveanu, G., Melis, V., and Howe, M. L. (2021). A new method to implant false autobiographical memories: Blind implantation. *J. Appl. Res. Mem. Cogn.* [Epub ahead of print]. doi: 10.1037/mac0000028

Otgaar, H., Romeo, T., Ramakers, N., and Howe, M. L. (2018). Forgetting having denied: The "amnesic" consequences of denial. *Memory Cogn.* 46, 520–529. doi: 10.3758/s13421-017-0781-5

Otgaar, H., Scoboria, A., and Smeets, T. (2013). Experimentally evoking nonbelieved memories for childhood events. *J. Exp. Psychol.* 39:717. doi: 10.1037/a0029668

Paige, L. E., Wolf, J. M., and Gutchess, A. (2022). Evaluating heart rate variability as a predictor of the influence of lying on memory. *Memory* 30, 785–795. doi: 10.1080/09658211.2020.1849307

Pezdek, K., Lam, S. T., and Sperry, K. (2009). Forced confabulation more strongly influences event memory if suggestions are other-generated than self-generated. *Legal Criminol. Psychol.* 14, 241–252. doi: 10.1348/135532508X344773

Pezdek, K., and Roe, C. (1995). The effect of memory trace strength on suggestibility. *J. Exp. Child Psychol.* 60, 116–128. doi: 10.1006/jecp.1995.1034

Pickel, K. (2004). When a lie becomes the truth: The effects of self-generated misinformation on eyewitness memory. *Memory* 12, 14–26. doi: 10.1080/09658210244000072

Polage, D. C. (2004). Fabrication deflation? The mixed effects of lying on memory. *Appl. Cogn. Psychol.* 18, 455–465. doi: 10.1002/acp.995

Polage, D. C. (2012). Fabrication inflation increases as source monitoring ability decreases. *Acta Psychol.* 139, 335–342. doi: 10.1016/j.actpsy.2011.12.007

Polage, D. C. (2019). Liar, liar: Consistent lying decreases belief in the truth. *Appl. Cogn. Psychol.* 33, 527–536. doi: 10.1002/acp.3489

Porter, S., Yuille, J. C., and Lehman, D. R. (1999). The nature of real, implanted, and fabricated memories for emotional childhood events: Implications for the recovered memory debate. *Law Hum. Behav.* 23, 517–537. doi: 10.1023/A:1022344128649

Pyszora, N. M., Barker, A. F., and Kopelman, M. D. (2003). Amnesia for criminal offences: A study of life sentence prisoners. *J. For. Psychiatry* 14, 475–490. doi: 10.1080/14789940310001599785

Rassin, E. (2022). Suggested false memories of a non-existent film: Forensically relevant individual differences in the *crashing memories* paradigm. *Memory* 30, 1205–1211. doi: 10.1080/09658211.2022.2085750

Riesthuis, P., Otgaar, H., Battista, F., and Mangiulli, I. (2022a). Public beliefs on the relationship between lying and memory. *Psychol. Crime Law* 28, 545–568. doi: 10.1080/1068316X.2021.1929979

Riesthuis, P., Otgaar, H., De Cort, A., Bogaard, G., and Mangiulli, I. (2022b). Creating a false alibi leads to errors of commission and omission. *Appl. Cogn. Psychol.* 36, 936–945. doi: 10.1002/acp.3982

Riesthuis, P., Mangiulli, I., Broers, N., and Otgaar, H. (2022c). Expert opinions on the smallest effect size of interest in false memory research. *Appl. Cogn. Psychol.* 36, 203–215. doi: 10.1002/acp.3911

Riesthuis, P., Otgaar, H., Hope, L., and Mangiulli, I. (2022d). Registered report: The effects of incentivized lies on memory. *Appl. Cogn. Psychol.* 36, 1150–1161. doi: 10.1002/acp.3986

Riesthuis, P., Otgaar, H., Mangiulli, I., and de Tauzia, R. (2020). Adopting a fictitious autobiography: Fabrication inflation or deflation? *Memory* 28, 741–752. doi: 10.1080/09658211.2020.1771371

Rodriguez, D. N., and Strange, D. (2014). Dissonance-induced false memories: Evidence from a free-choice paradigm. *J. Cogn. Psychol.* 26, 571–579. doi: 10.1080/20445911.2014.925459

Rodriguez, D. N., and Strange, D. (2015). False memories for dissonance inducing events. *Memory* 23, 203–212. doi: 10.1080/09658211.2014.881501

Romeo, T., Otgaar, H., and Landström, S. (2018). Coping with Sexual Abuse. *Psychol. Res. Urban Soc.* 1, 46–54. doi: 10.7454/proust.v1i1.11

Romeo, T., Otgaar, H., Smeets, T., Landstrom, S., and Boerboom, D. (2019). The impact of lying about a traumatic virtual reality experience on memory. *Memory Cogn.* 47, 485–495. doi: 10.3758/s13421-018-0885-6

Saks, M. J., and Koehler, J. J. (2005). The coming paradigm shift in forensic identification science. *Science* 309, 892–895. doi: 10.1126/science.1111565

Scherr, K. C., Redlich, A. D., and Kassin, S. M. (2020). Cumulative disadvantage: A psychological framework for understanding how innocence can lead to confession, wrongful conviction, and beyond. *Perspect. Psychol. Sci.* 15, 353–383. doi: 10.1177/1745691619896608

Schreiber, N., Wentura, D., and Bilsky, W. (2001). "What else could he have done?" Creating false answers in child witnesses by inviting speculation. . *J. Appl. Psychol.* 86, 525–532. doi: 10.1037/0021-9010.86.3.525

Scoboria, A., and Henkel, L. (2020). Defending or relinquishing belief in occurrence for remembered events that are challenged: A social-cognitive model. *Appl. Cogn. Psychol.* 34, 1243–1252. doi: 10.1002/acp.3713

Scoboria, A., Jackson, D. L., Talarico, J., Hanczakowski, M., Wysman, L., and Mazzoni, G. (2014). The role of belief in occurrence within autobiographical memory. *J. Exp. Psychol.* 143:1242. doi: 10.1037/a0034110

Scoboria, A., Mazzoni, G., Kirsch, I., and Relyea, M. (2004). Plausibility and belief in autobiographical memory. *Appl. Cogn. Psychol.* 18, 791–807. doi: 10.1002/acp.1062

Scoboria, A., Wade, K. A., Lindsay, D. S., Azad, T., Strange, D., Ost, J., et al. (2017). A mega-analysis of memory reports from eight peer-reviewed false memory implantation studies. *Memory* 25, 146–163. doi: 10.1080/09658211.2016.1260747

Scoboria, A., Wysman, L., and Otgaar, H. (2012). Credible suggestions affect false autobiographical beliefs. *Memory* 20, 429–442. doi: 10.1080/09658211.2012.677449

Shaw, E. V., and Loftus, E. F. (2020). Punishing the crime of forgetting. *J. Appl. Res. Mem. Cogn.* 9, 24–28. doi: 10.1016/j.jarmac.2019.11.003

Shepp, V., O'Callaghan, E., and Ullman, S. E. (2019). Interactions with offenders post-assault and their impacts on recovery: A qualitative study of sexual assault survivors and support providers. *J. Aggress. Maltreat. Trauma* 29, 725–747. doi: 10.1080/10926771.2019.1660443

Sun, X., Punjabi, P. V., Greenberg, L. T., and Seamon, J. G. (2009). Does feigning amnesia impair subsequent recall? *Mem. Cogn.* 37, 81–89. doi: 10.3758/MC.37.1.81

Tysse, J. E. (2005). Note: The right to an imperfect trial – Amnesia, malingering, and competency to stand trial. *Wm. Mitchell L. Rev.* 32:353.

Tysse, J. E., and Hafemeister, T. L. (2006). Amnesia and the determination of competency to stand trial. *Dev. Mental Health Law* 25, 65–80.

Van Oorsouw, K., and Giesbrecht, T. (2008). Minimizing culpability increases commission errors in a mock crime paradigm. *Legal Criminol. Psychol.* 13, 335–344. doi: 10.1348/135532507X228539

Van Oorsouw, K., and Merckelbach, H. (2004). Feigning amnesia undermines memory for a mock crime. *Appl. Cogn. Psychol.* 18, 505–518. doi: 10.1002/acp.999

Verigin, B. L., Meijer, E. H., Bogaard, G., and Vrij, A. (2019). Lie prevalence, lie characteristics and strategies of self-reported good liars. *PLoS One* 14:e0225566. doi: 10.1371/journal.pone.0225566

Verschuere, B., Spruyt, A., Meijer, E. H., and Otgaar, H. (2011). The ease of lying. *Cons. Cogn.* 20, 908–911.

Vieira, K. M., and Lane, S. M. (2013). How you lie affects what you remember. *J. Appl. Res. Mem. Cogn.* 2, 173–178. doi: 10.1016/j.jarmac.2013.05.005

Vo, T. V. A., Gunderson, C. A., and Ten Brinke, L. (2022). How deception and believability feedback affect recall. *Memory* 30, 706–714. doi: 10.1080/09658211.2021.1883064

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. Chichester: John Wiley and Sons.

Wang, J., Otgaar, H., Howe, M. L., and Zhou, C. (2019). A self-reference false memory effect in the DRM paradigm: Evidence from Eastern and Western samples. *Mem. Cogn.* 47, 76–86. doi: 10.3758/s13421-018-0851-3

Watson, S., Harkins, L., and Palmer, M. (2016). The experience of deniers on a community sex offender group program. *J. For. Psychol. Pract.* 16, 374–392. doi: 10.1080/15228932.2016.1219219

Wells, G. L., and Olson, E. A. (2003). Eyewitness testimony. *Annu. Rev. Psychol.* 54, 277–295. doi: 10.1146/annurev.psych.54.101601.145028

Williams, S. J., Wright, D. B., and Freeman, N. H. (2002). Inhibiting children's memory of an interactive event: The effectiveness of a cover-up. *Appl. Cogn. Psychol.* 16, 651–664. doi: 10.1002/acp.821

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annu. Rev. Psychol.* 55, 235–269. doi: 10.1146/annurev.psych.55.090902.141555

Wixted, J. T. (2005). A theory about why we forget what we once knew. *Curr. Dir. Psychol. Sci.* 14, 6–9. doi: 10.1111/j.0963-7214.2005.00324.x

Wright, D. B., and Loftus, E. F. (1998). How misinformation alters memories. *J. Exp. Child Psychol.* 71, 155–164.

Wright, D. B., Loftus, E. F., and Hall, M. (2001). Now you see it; now you don't: Inhibiting recall and recognition of scenes. *Appl. Cogn. Psychol.* 15, 471–482. doi: 10.1002/acp.719

Wright, D. B., Memon, A., Skagerberg, E. M., and Gabbert, F. (2009). When eyewitnesses talk. *Curr. Dir. Psychol. Sci.* 18, 174–178. doi: 10.1111/j.1467-8721.2009.01631.x

Zajac, R., and Henderson, N. (2009). Don't it make my brown eyes blue: Co-witness misinformation about a target's appearance can impair target-absent line-up performance. *Memory* 17, 266–278. doi: 10.1080/09658210802623950

Zhang, Y., Battista, F., Thissen, D., Otgaar, H., Wang, J., and Jelicic, M. (2022a). Examining the associations between nonbelieved memories and memoryd, self-esteem, and rumination. *Psychol. Conscious.* [Epub ahead of print]. doi: 10.1037/cns0000344

Zhang, Y., Otgaar, H., and Wang, J. (2022b). Memory distrust is related to memory errors, self-25 esteem, and personality. *Appl. Cogn. Psychol.* 36, 283–292. doi: 10.1002/acp.3917

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 22 November 2023
DOI 10.3389/fpsyg.2023.1212709

OPEN ACCESS

EDITED BY
Ivan Mangiulli,
University of Bari Aldo Moro, Italy

REVIEWED BY
Richard S. John,
University of Southern California, United States
Dara Mojtahedi,
University of Huddersfield, United Kingdom

*CORRESPONDENCE
Lauren Knott
✉ lauren.knott.1@city.ac.uk

RECEIVED 26 April 2023
ACCEPTED 01 November 2023
PUBLISHED 22 November 2023

CITATION
Shah D and Knott L (2023) The detrimental effects of delay on the endorsement of misleading details for emotionally salient events.
Front. Psychol. 14:1212709.
doi: 10.3389/fpsyg.2023.1212709

The detrimental effects of delay on the endorsement of misleading details for emotionally salient events

Datin Shah and Lauren Knott*

Department of Psychology, City, University of London, London, United Kingdom

Previous research has shown that the exposure to misleading information continues its detrimental effect on memory over time for negatively arousing events. However, research has also shown that both high-and low-arousing negative events are vulnerable to distortion from misinformation. Therefore, the present study set out to explore the impact of retention interval on memory for negative (arousing and non-arousing) and neutral events in the misinformation paradigm. Participants were presented with a negative high-arousing, a negative low-arousing, and a neutral scene, and exposed to misleading information for central and peripheral aspects of each scene. Recognition memory for scene details was measured 10 min after misinformation exposure and again after one week. We found that, regardless of the type of detail, the effect of misinformation persisted over time for the negative-arousing event but disappeared one week later for the negative low-arousing and neutral events. The results are explained in relation to adaptive function and theories of source monitoring. The findings of this study provide important forensic implications, especially when we consider the arousing nature of crimes.

KEYWORDS

misinformation paradigm, retention interval, memory, emotion, arousal

1 Introduction

False memories occur when one recalls an entirely new experience that never occurred or incorrectly recalls details of an experienced event (Roediger and McDermott, 1995). The distortion of memory has become an increasingly prominent focus of cognitive psychological research, especially with its implications to the applied setting. Understanding the fundamental nature of a reconstructive memory system and the memory errors that can occur from it is of paramount importance to the legal field. We should strive to understand factors that may contaminate eyewitness testimony. To this end, over the past four to five decades, researchers have set out to understand the factors affecting errors in recollection and the mechanisms that cause them (Zhang et al., 2021).

When it comes to research examining the formation of false memories, one of the key laboratory paradigms used is Loftus' misinformation paradigm (Loftus et al., 1978). In the standard three-stage paradigm, participants are first presented with an event (e.g., in the form of a slide show, video, or a staged event). Thereafter, some participants receive misleading information about the event, typically embedded into a questionnaire or a written narrative. Finally, memory is tested for original event details. The "misinformation effect" occurs when participants falsely report the misleading information in their memory reports as being part of

the original event. Since the original research, many factors have been shown to influence misremembering. For example, studies have shown that the size of the misinformation effect increases over longer retention intervals (e.g., Frost, 2000; Frost et al., 2002; Mudd and Govern, 2004). Frost et al. (2002) argued that the reduced number of source cues available after a long delay can make participants more susceptible to misleading information, thereby increasing misinformation errors. In contrast factors such as warning conditions, if given prior to the misinformation presentation, can lead to a decrease in susceptibility (see Loftus, 2005) as we are more vigilant to post-event information discrepancies.

Although the misinformation effect has been used in numerous studies, investigating various key factors, it seems that little research has focused on the influence of emotion on the susceptibility to misinformation. In the legal field, eyewitnesses will be questioned about events that will inevitably be emotionally arousing, particularly serious crimes (e.g., an assault, a theft, or murder) therefore the impact of emotion on misinformation warrants a detailed investigation. Interestingly, the manner in which negative events are encoded and later retrieved elicits conflicting views from the field. For example, research has shown that events containing negative emotional detail are better remembered compared to those containing neutral detail (Hamann, 2001; Talmi et al., 2007). When positive and negative images (taken from the International Affective Picture System [IAPS]; Lang et al., 2008) are shown at study, negative images are better remembered at test (Charles et al., 2003). Similar findings are shown for emotional words, although emotionally arousing taboo words are better recalled when neutral and negative items are matched for relatedness (MacKay et al., 2004; Buchanan et al., 2006). This emotional enhanced effect has been shown immediately after study (Murty et al., 2010; Talmi and McGarry, 2012), and is thought to be primarily due to the attraction of attention during encoding (Sommer et al., 2008; Talmi, 2013), but also over a period of delay. The latter has been attributed to consolidation consistent with the Emotional Synaptic Tagging Hypothesis (Bergado et al., 2011; McReynolds and McIntyre, 2012), with greater activity in the amygdala, hippocampus, and parahippocampus, in addition to visual, prefrontal, and parietal areas (LaBar and Cabeza, 2006; Murty et al., 2010; Dolcos et al., 2012).

However, research has also shown that emotions can impair memory for certain details by producing an emotional memory narrowing effect (e.g., Kaplan et al., 2012). This is a phenomenon whereby one remembers information that is central to an emotional event but has poorer memory for peripheral or background information about the event (Kaplan et al., 2012). According to Easterbrook's (1959) cue-utilisation theory, an individual has a limited number of cues that they can process at any one time. Therefore, as the arousal of an emotional event increases, attention narrows to the most central/arousing aspects of the event and away from the peripheral/background information (Heuer and Reisberg, 1990; Christianson et al., 1991; Steblay, 1992).

The narrowing effect may be specific to negatively arousing stimuli (e.g., Waring and Kensinger, 2009; Yegiyan and Yonelinas, 2011; Van Damme and Smets, 2014). In a review of evidence on the effects of emotion, Kensinger (2009) showed that a narrowed attentional scope to central/specific details was associated with negative emotion but a broader attentional scope was associated with positive emotion. According to the affect-as-information theory (e.g., Schwarz and Clore, 1983), positive emotion indicates a safe and unproblematic situation that does not require the need for increased attention to specific details, thereby resulting in broader information processing. In contrast, negative emotion suggests a problem that must be dealt with, thus there is a greater need to focus on relevant information within the environment, resulting in narrow item-specific processing.

So it appears that emotionally negative events cause an enhanced memory effect, although potentially leading to impaired memory for certain details. Paradoxically, studies have also shown that negative events are susceptible to distortion from misleading information. Porter et al. (2003) examined whether the effects of misinformation exposure varied with the emotionality of photographic scenes. They found that the endorsement of "major misinformation" (a major peripheral object non-existent in the picture) was most common with negative scenes than with positive and neutral scenes. In addition, Porter et al. (2008) asked participants to try and recall "widely publicised" positively-valenced and negatively-valenced public events. Half of the events were fictitious. It was found that recollection was greater for true-negative than for true-positive events, and greater for false-negative than for false-positive events. Similar findings are true of children recalling emotional false memories too (Otgaar et al., 2008).

Van Damme and Smets (2014) were the first to manipulate the effects of both valence and arousal on suggestibility. The emotional nature of an event can be described by means of (at least) two dimensions: valence and arousal (e.g., Russell, 1980, 2003). They presented participants with high-and low-arousing positive, negative, and neutral photographs. Half of the participants were later exposed to misleading central and peripheral details. They found that, regardless of prior exposure to misinformation, participants were less accurate and endorsed more misleading information for peripheral details associated with the negative events (both high and low in arousal). This indicated that *negative valence* narrowed attention. High arousal improved memory for correct central details, and both negative valence and high arousal inhibited control participants' tendency to endorse false central detail, however, this effect disappeared with previous exposure to misinformation. Van Damme and Smets suggested that the main parts of the negative scenes may act as attention magnets (i.e., a salient or distinctive part that captures one's attention; Laney et al., 2003) and that this level of memory narrowing may have been due to the activation of goals associated with the negative emotion. That is, the narrowing effect occurs towards details that are goal-relevant (i.e., the goal-relevance approach; Levine and Edelstein, 2009).

Studies such as those presented above use retrieval tasks with only a short delay. What impact does negative emotion have on the misinformation effect over a longer delay? In veridical memory research, we know that memory for emotional stimuli remains stable or improves over time (e.g., LaBar and Phelps, 1998; Sharot and Phelps, 2004; Wang, 2014; for a meta-analysis, see Park, 2005). In addition, central details seem to benefit most from a lower rate of decay (e.g., Christianson and Loftus, 1987). From an evolutionary perspective, being able to remember an arousing experience over time can help an individual prepare for similar events, and guide future behaviour to approach or avoid such situations (Porter and Peace, 2007; Van Damme and Smets, 2014). To the best of our knowledge, only one study has manipulated testing interval and misinformation

exposure to examine their effect on susceptibility to misinformation for emotional events. Porter et al. (2010) presented participants with positive and negative emotional images. Misinformation was introduced to half of the participants with a retrieval task that took place immediately and either 1 week or 1 month later. Regardless of event emotion, they found that overall accuracy for misleading details was lower for misled participants than for nonmisled participants across all retrieval intervals, and misled participants showed a greater reduction in accuracy from 1 week to 1 month compared to control participants. However, negative images (compared to positive) were associated with a greater susceptibility to major misleading details, a pattern found at both immediate and delayed retrieval sessions. Thus, relative to positive emotion, negative emotion heightens suggestibility at least for major misinformation, and this persists over time.

Porter et al. (2008, 2010) argued that negative information is better retained in memory over time but is also vulnerable to distortion from misleading information (*paradoxical negative emotion* hypothesis; Porter et al., 2008). Remembering information from negative events can help individuals to avoid or deal with future dangers (Porter and Peace, 2007). However, negative events are also susceptible to distortion. This has been explained as an adaptive need to retain relevant information concerning negative events from trustworthy sources to ensure one is prepared for future related dangers. Porter et al. (2010) argued that major details indicate a significant change in one's recollection, thus constituting valuable information that may serve a greater benefit in the future. Consequently, at least for Porter *et al*'s study, major details associated with negative events were more likely to be incorporated into one's memory reports.

Source monitoring failure may also be used to explain these findings. Source misattributions can most often occur when there are similarities between the original information and the post-event information (Johnson et al., 1993; Mitchell and Johnson, 2000). When participants process the post-event information, they may mentally reconstruct the original event or engage in active rehearsal, thus increasing the overlap between the two sources of information (e.g., in sensory/perceptual characteristics) and strengthening the post-event information (Zaragoza and Mitchell, 1996; Mitchell and Johnson, 2000). Source confusion may be worse for negative high-arousing events relative to neutral and emotionally low-arousing events. Negative high arousing events have been shown to benefit memory consolidation of negative emotional details through the activation of the amygdala and hippocampus (e.g., McGaugh, 2000; Dolcos et al., 2005). Thus, it is plausible to assume that mental visualisations of the post-event information would be more vivid and better integrated into memory for the original event, especially if the availability of source cues fades with time (Frost et al., 2002). We may assume that this would not be the case for negative low-arousing events, although this has yet to be examined.

1.1 Present study

The aim of the present study was to examine the impact of delayed retrieval and exposure to misinformation on memory for emotionally negative and neutral events, and central and peripheral aspects of these events. We manipulated arousal in negative emotional images, with a neutral image comparison across a period of delay. Negative events regardless of the level of arousal have been shown to be better remembered than neutral information (e.g., Kensinger and Corkin, 2004), but also be susceptible to misinformation (Van Damme and Smets, 2014). We aimed to explore whether retention interval and misinformation exposure differentially impacted misinformation for high and low-arousing negative events. In addition, memory for central details of negatively arousing events may persist over time more than peripheral details (Christianson and Loftus, 1987; Burke et al., 1992). Central details from negative events and high-arousing events have shown to be vulnerable to prior exposure to misinformation (Van Damme and Smets, 2014), though its effect over time is yet to be seen. Thus, we aimed to systematically study the impact of delayed retrieval on susceptibility to misinformation for central and peripheral aspects of negative events. Based on previous research, we predicted that for the negatively arousing event, the magnitude of the misinformation effect for central details would be similar over time, but would increase for peripheral details. As for the negative low-arousing and neutral events, the misinformation effect for central and peripheral details would increase over time. Finally, we were keen to replicate Porter *et al*'s (2010) findings but with a test for different details at immediate and delayed testing sessions. This would eliminate any concern regarding repeat testing with the same memory test (see Porter et al., 2010). This could affect the interpretation of the memory reports if participants contaminate memory for the event images with test responses from a previous test condition. Considering the above, we believe that the present study is the first to examine the impact of delayed retrieval and exposure to misinformation for central and peripheral details for emotionally negative (both high and low in arousal) and neutral images.

2 Methods

2.1 Participants

Forty-eight participants (age: $M = 35.35$, $SD = 14.60$, age range = 18–60; sex: 32 females & 16 males) took part in both sessions of the study in return for course credits or a small fee. The study was conducted online. An *a priori* power analysis using MorePower 6.0 (Campbell and Thompson, 2012) indicated that a sample size of between 32 and 80 was adequate to detect medium to large interaction effect (see Paz-Alonso et al., 2013 for similar design and effect size) with a power of 0.80. The participants had English as their first language, normal or corrected-to-normal vision, and were not colour-blind. Participants were recruited via online participant recruitment platforms (Sona and Prolific). City, University of London's Psychology Research Ethics Committee approved the study and ethical principles were followed.

2.2 Design

There were two experimental designs: one for minor details and one for major details. See Figure 1 for a visual representation of the two designs. All participants saw both types of details. For minor

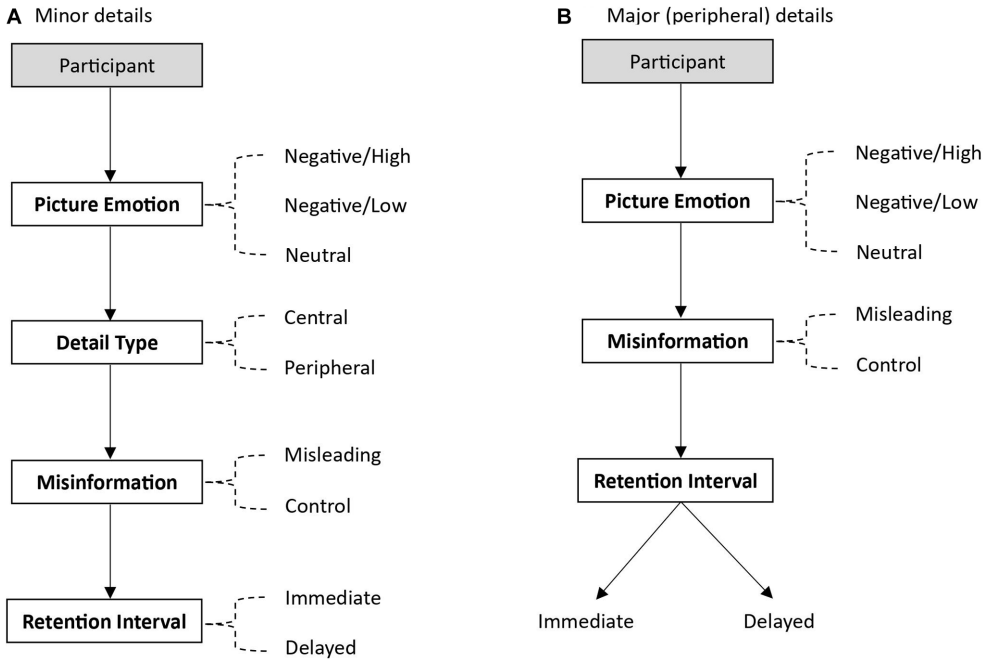


FIGURE 1
Diagrams representing the experimental design for minor details **(A)** and major details **(B)**. To illustrate the complex designs, the arrows go through the names of the independent variables, with dashed lines indicating all the conditions associated with the variables that participants take part in. For diagram A, the design was fully within-subjects. That is, participants were presented with minor details for all combinations of the levels across the four experimental factors. For diagram B, the design was mixed, with Retention Interval being between subjects. Thus, participants were presented with major peripheral details for all combination of the levels across two experimental factors (Picture Emotion & Misinformation) but were tested on these details either in an immediate or a delayed test.

details (i.e., details that contradict what was in the picture), the design was a 3 (picture emotion: negative/high vs. negative/low vs. neutral) x 2 (detail type: central vs. peripheral) x 2 (misinformation: misled vs. control) x 2 (retention interval: immediate vs. delayed) repeated measures design. For *picture emotion*, each participant saw three pictures, and the order was counterbalanced: one negative high-arousing (negative/high), one negative low-arousing (negative/low), and one neutral. The presence of *misinformation* was manipulated using five misleading details (two central and three peripheral details) and five control details (i.e., no misinformation was provided for these details; two central and three peripheral details). The misleading and control details were counterbalanced. For *retention interval*, participants completed a recognition test immediately and one week later. As such, the misleading and control details were split between the immediate and delayed recognition tests and counterbalanced. For major details (i.e., a salient peripheral detail that is not present in the picture), there was only one major misleading detail and one major control detail. Thus, the design was a 3 (picture emotion) x 2 (misinformation) x 2 (retention interval) mixed design, with retention interval as a between-subjects variable. Twenty-four participants were in both immediate and delayed conditions. Porter et al. (2003, 2010) only had one major misleading detail in each picture. These details are considered salient and should be noticeable if present, therefore including more than one suggested major detail could make participants aware of the presence of false information and the purpose of the study. The dependent variable was the false recognition of the incorrect answer in the misleading and control questions.

2.3 Materials

2.3.1 Picture characteristics

Three pictures taken from the International Affective Picture System (IAPS; Bradley and Lang, 2007; Lang et al., 2008) database were used as to-be-remembered events. The negative high-arousing event was an assault scene (IAPS number: 9254; Valence: 2.03; Arousal: 6.04), the negative low-arousing event was a cemetery scene (IAPS number: 9220; Valence: 2.06; Arousal: 4.00), and the neutral event was a restaurant scene (IAPS number: 2593; Valence: 5.80; Arousal: 3.42).

Central and peripheral details were determined using a pilot task. We used a similar approach to Van Damme and Smets (2014) and Porter et al. (2003), whereby 30 participants (not included in the present study) were asked to draw lines around the central information on each picture. That is, participants were asked to circle on the pictures the area(s) with "the main information that is directly connected to the event, or gist of the event, depicted in the scene" (Christianson, 1992; Luna and Martín-Luengo, 2018). Participants had no time limit for this task. Details were considered central if they fell within the area(s) of the picture, or peripheral if they fell outside of the area(s), by at least 70% of the participants. The central areas judged by the participants included the main characters and objects that were part of the event depicted in the scenes. As such, the central details taken from within these areas included, for example, details about the colour/pattern of clothing worn by the main persons in the scenes, and the type and descriptive features of main objects (e.g., a gravestone). Everything outside of the enclosed lines was considered

peripheral information (e.g., the type and number of objects such as streetlamps, the colour of background objects, descriptive aspects of background people) Assessing such details is in line with previous research examining central and peripheral details (e.g., Van Damme and Smets, 2014; Luna and Martin-Luengo, 2018; Jobson et al., 2023).

2.3.2 Post-event questionnaire

The post-event questionnaire (titled "Perception Questionnaire" for the participants) consisted of 10 Yes/No questions about each picture (30 in total). For the questionnaire, we chose ten critical details. Eight of the critical details were *minor* details and two were *major* details. The minor details were selected from the pilot study, whereby four of the details were central (i.e., fell within the central area) and four were peripheral (i.e., fell outside the central area). The major details were *only* peripheral details (thus no central major details were examined). Following Porter et al. (2003), a major detail was defined as a person, animal, or a major object that is falsely suggested to be present in the pictures. Although it is not possible to define the size of the detail since the major details do not exist, in a similar manner to Porter and colleagues, we considered that most, if not all, participants would notice this salient information if present.

For each critical detail, we created a misleading question and a control question. The phrasing of the control questions was kept as similar as possible to the misinformation question except that the misinformation was omitted or the detail was mentioned in a neutral form. An example question with a minor (central) misleading detail concerning the colour of the woman's top was, "Did you see that the woman's **brown** top was long-sleeved?" [whereas in fact the top was black; the detail in **bold** was removed in the control version of the question]. Thus, minor misinformation contradicted the details in the pictures. An example question with a major misleading detail concerning the presence of a bird was, "Behind the injured man sitting on the right, did you see the hedge **that had a large pigeon on it**?" [whereas in fact there was no pigeon; the text in **bold** was removed in a control version of the question]. Thus, salient major misinformation added details in the peripheral area. All critical details were never the direct focus of the question; rather, they added extra information in the question.

For each participant, half of the critical details (two central minor, two peripheral minor, and one peripheral major) were misleading, and the remaining half were controls. Thus, the post-event questionnaire contained five misleading questions and five control questions. To counterbalance the combination of detail type and misinformation, two versions of the questionnaire were created. Misleading details in Version A were control details in Version B, and control details in Version A were misleading details in Version B. Therefore, each critical detail served equally often as a misleading and control detail.

2.3.3 Memory test

Recognition memory for the pictures was assessed using 14 two-alternative forced-choice questions per picture. Since the participants were tested both immediately and one week later, two recognition tests were constructed, whereby the 14 questions per picture were split between the two tests. In both Test One and Test Two, two questions probed memory for previously suggested misleading minor details (one central and one peripheral), two questions probed memory for non-suggested control minor details (one central and one peripheral), and two questions probed memory for non-leading details (one central and one peripheral) not previously suggested to all participants. In addition, Test One further included two questions probing memory for the major details (one misleading and one control). Overall, Test One consisted of eight questions per picture and Test Two consisted of six questions per picture. The order of the tests was counterbalanced, such that half of the participants received Test One in the first session and Test Two in the second session. Therefore, for minor details, there was only one question in each test for each combination of detail type and misinformation. For major details, the two questions for major details appeared only in Test One, thus participants were either tested on major critical details in session one or session two (i.e., between-subjects).

For the misleading questions, the two response alternatives were a correct detail (consistent with the picture), and a misleading detail (consistent with the questionnaire). The same response alternatives were used for control questions targeting those details that were misleading for half of the participants. For both the control and non-leading questions, a correct detail and a novel foil were possible answers. An example of a misleading and control question asked during the recognition test is the following: "What colour was the top worn by the woman?" along with two response options: (1) Black [correct] or (2) Brown [misleading/control]. An example of a recognition question targeting a major misleading peripheral detail is the following: "Was there a pigeon in the picture?" along with two response options: (1) No, there was no pigeon [correct] or (2) Yes, there was a large pigeon [misleading/control]. In both tests, participants were instructed to select one of the response alternatives based on their own memory for the pictures. The questions and response alternatives were presented in random order. If they did not know the answer, they were told to make their best guess.

2.3.4 Mood ratings

Pictures may invoke mood changes. Research has shown that mood may impact suggestibility (Forgas et al., 2005; Van Damme and Seynaeve, 2013; Zhang et al., 2021). For example, Forgas et al. (2005) found that positive mood increased misinformation susceptibility whereas negative mood inhibited the endorsement of misinformation. Zhang et al. (2021) found that a positive mood increased misinformation endorsement for neutral scenes. To ensure that there is no confounding effect of a person's mood on the outcome of the results, we collected participants' mood ratings at different points during the experiment using Self-Assessment Manikin (SAM; Bradley and Lang, 1994) scales to check for any significant mood changes between sessions. Mood was assessed immediately before picture encoding (session 1) and before each recognition test (sessions 1 and 2).

2.4 Procedure

See Figure 2 for a visual overview of the study's procedure. Participants took part in two sessions. In session one, participants first provided informed consent and then completed the first SAM scale. Thereafter, participants were told that they will be shown some pictures for 30 s each. They were instructed to "Please look at each picture as if you unexpectedly witness the event." Preceding each picture was a fixation cross for two seconds. The presentation order of the three pictures was counterbalanced. Once all three

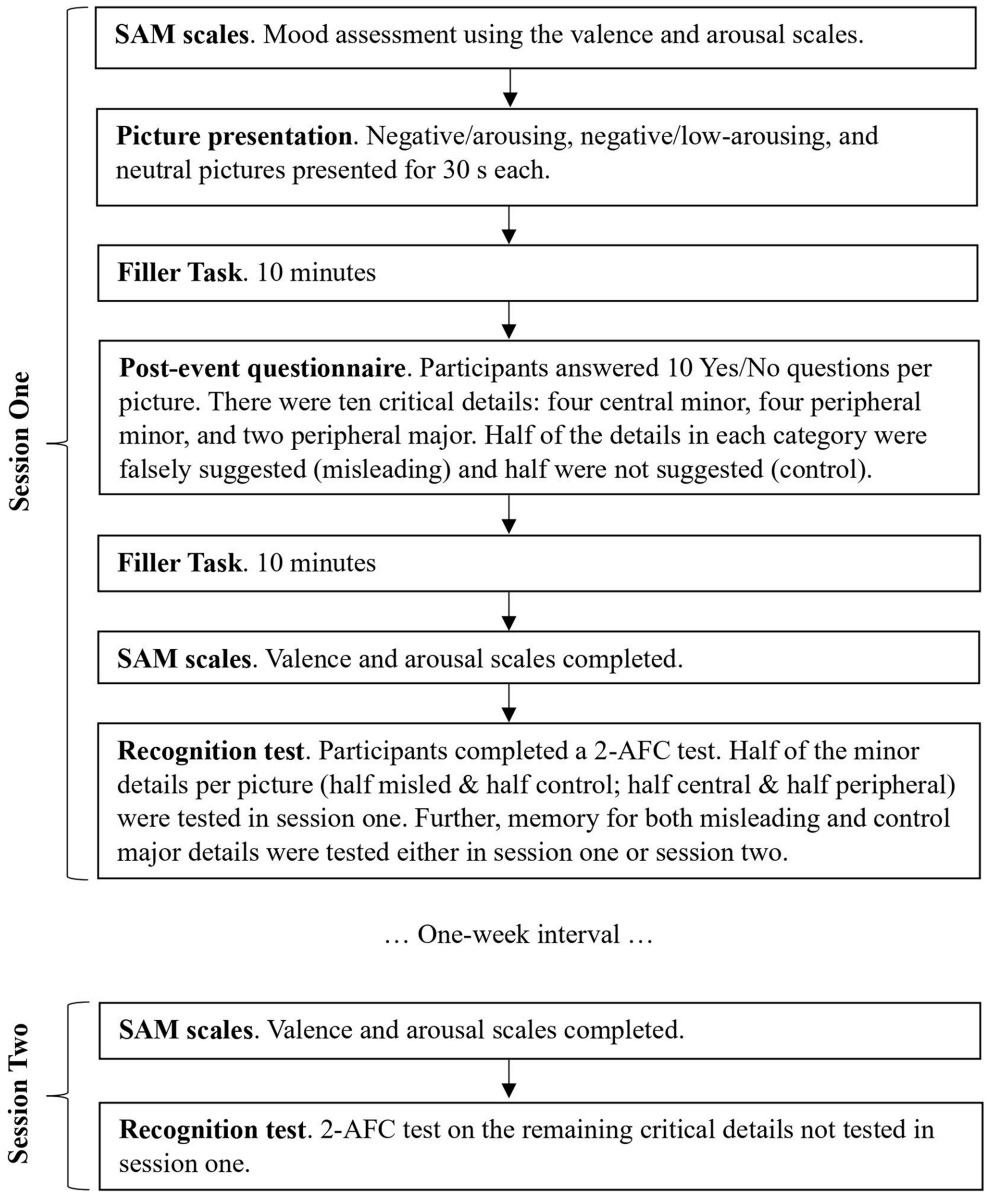


FIGURE 2
Overview of experimental procedures.

pictures had been presented, there was a 10-min interval during which time participants completed unrelated filler tasks (i.e., mathematical problems and unrelated anagrams). Thereafter, participants completed the post-event questionnaire in which half of the questions suggested misleading information. The participants were not warned about potential discrepancies between the information in the questions and the picture. The order of the sets of questions about each picture followed the picture presentation order at the encoding stage. After the post-event phase, there was another 10-min interval during which time participants completed reasoning problems. Following this, all participants completed the SAM questionnaire again and the first recognition test. Whether participants received test one or test two in this session depended on the counterbalancing condition that they were randomly assigned to. Before finishing, participants were falsely told that the second session in one week would involve a new set of pictures and they would rate these pictures on two emotional dimensions (valence and arousal). This instruction was used in an attempt to reduce the likelihood of rehearsal in the interim.

Exactly one week later, participants were sent a link for the second part of the study. The link was sent in the morning and participants had until 9 pm on the same day to complete the second part. They first completed the SAM questionnaire to assess their current mood state. Thereafter, they were given the second recognition test. Participants who received test one or two in session one completed test two or test one in the second session, respectively. After completing the recognition test, participants provided demographic information and a debriefing.

3 Results

Two participants were removed from all analyses due to failing more than one attention check.[1] The final sample consisted of 46 participants (age: $M = 35.48$, $SD = 14.63$, age range = 18–60; sex: 30 females & 16 males). For the analysis of major misinformation, there remained 22 participants in the immediate condition and 24 in the delayed condition. The answers in the recognition test were coded dichotomously reflecting false recognition (i.e., correct answer = 0, incorrect answer = 1). The main analyses were conducted on binary false responses to minor and major critical details. An alpha level of 0.05 was used for all statistical tests.

3.1 Mood check

To check whether there were any significant changes to participants' mood between three points in the experiment (Time 1: start of session one; Time 2: immediately before the recognition test of session one; Time 3: start of session two), One-way ANOVAs were conducted on valence and arousal scores separately. Of interest is the difference between Time 1 and Time 3, and between Time 2 and Time 3, since the former represents the start of each session, and the latter represents participants' mood before each recognition test. No difference in valence scores was found between Time 1 and Time 3 ($p = 1.00$) and between Time 2 and Time 3 ($p = 0.149$) and no significant differences in arousal were found between Time 1 and Time 3 ($p = 0.203$) and between Time 2 and Time 3 ($p = 0.220$).

3.2 False recognition

The data represented binary responses (0 = correct, and 1 = incorrect). Since log-linear cannot analyse within-subjects data with complex designs, the data were analysed using Generalised Estimating Equations (GEE; Zeger and Liang, 1986). GEE, an extension of the Generalized Linear Model, is an approach that allows for the analysis of repeated measurements and non-normally distributed data. The false recognition responses to misleading and control details were analysed using GEE with a Binomial distribution and log link function.[2] The repeated factors in the model were picture emotion (negative/high vs. negative/low vs. neutral), detail type (central vs. peripheral), misinformation (misled vs. control), and retention interval (immediate vs. delayed). See Table 1 for means and standard deviations. Post-hoc tests of significant interactions were Bonferroni corrected. Effect sizes for mean differences were estimated using Cohen's d with the interpretation as follows: small = 0.2, medium = 0.5, and large = 0.8. The means reported in-text are estimated marginal means along with their respective standard deviations. There was a significant main effect of misinformation, Wald $\chi^2(1, N = 46) = 35.74$, $p < 0.001$, $d = 1.13$, and detail type, Wald $\chi^2(1, N = 46) = 4.50$, $p = 0.034$, $d = 0.39$. False recognition was significantly higher for misleading details ($M = 0.47$, $SD = 0.16$) compared to control details ($M = 0.30$, $SD = 0.14$) and for central details ($M = 0.41$, $SD = 0.15$) compared to peripheral details ($M = 0.35$, $SD = 0.16$). There was also a significant retention interval x misinformation interaction, Wald $\chi^2(1, N = 46) = 9.74$, $p = 0.002$, and a picture emotion x retention interval x misinformation interaction that approached significance (see Figure 3), Wald $\chi^2(1, N = 46) = 5.92$, $p = 0.052$. There were no further main effects (Wald χ^2's < 5.02, $ps > 0.081$), two-way interactions (Wald χ^2's < 2.86, $ps > 0.239$), three-way interactions (Wald χ^2's < 2.78, $ps > 0.133$), and a four-way interaction (Wald $\chi^2 = 1.24$, $p = 0.537$). Because the three-way interaction approached significance and was of interest to our aim to understand the impact of misinformation on memory for negative emotional events over time, we explored this further.

For the negative/high picture, there was a significant main effect of misinformation, Wald χ^2 $(1, N = 46) = 9.51$, $p = 0.002$, $d = 0.67$, but no significant effect of retention interval ($p = 0.458$) nor interaction ($p = 0.507$), suggesting that the size of the misinformation effect was similar at both immediate and delayed sessions, and no differences in the false recognition of misleading and control details over time. For the negative/low picture, there was a significant main effect of misinformation, Wald χ^2 $(1, N = 46) = 20.18$, $p < 0.001$, $d = 0.93$, but not retention interval, Wald χ^2 $(1, N = 46) = 2.86$, $p = 0.091$, $d = 0.35$. However, there was a significant interaction, Wald χ^2 $(1, N = 46) = 9.54$, $p = 0.002$. A misinformation effect was found at immediate testing (misleading: $M = 0.61$, $SD = 0.39$, control: $M = 0.21$, $SD = 0.28$), Wald $\chi^2(1, N = 46) = 36.76$, $p < 0.001$, $d = 1.18$, but not at delayed testing (misleading: $M = 0.32$, $SD = 0.33$, control: $M = 0.28$, $SD = 0.31$), Wald $\chi^2(1, N = 46) = 0.19$, $p = 0.661$, $d = 0.12$. There appears to be a decrease in false recognition of the misleading details over time. For the neutral picture, there was a significant main effect of misinformation, Wald $\chi^2(1, N = 46) = 6.70$, $p = 0.010$, $d = 0.58$, but not retention interval, Wald $\chi^2(1, N = 46) = 0.01$, $p = 0.905$, $d = 0.04$. However, there was a significant interaction, Wald $\chi^2(1, N = 46) = 10.97$, $p < 0.001$. Similar to the negative/low picture, a misinformation effect was found at immediate testing (misleading: $M = 0.54$, $SD = 0.35$, control: $M = 0.25$, $SD = 0.31$), Wald $\chi^2(1, N = 46) = 17.72$, $p < 0.001$, $d = 0.88$, but not at delayed testing (misleading: $M = 0.36$, $SD = 0.29$, control: $M = 0.40$, $SD = 0.35$), Wald $\chi^2(1, N = 46) = 0.43$, $p = 0.514$, $d = 0.12$. It appears that misinformation continued to influence memory performance over time for the high-arousing negative event, but for the low-arousing events, there was no significant negative impact of misinformation on memory over time; in fact, false recognition of the misleading details decreased over time. Although detail type did not interact with this effect, Table 1 suggests that this was more apparent in the peripheral compared to central detail type.

1 Typically used for studies conducted online to ensure engagement. There were two attention checks during the picture presentation stage but between stimulus, and one during the post-event stage.

2 For the purposes of comparability and consistency with relevant previous research (e.g., Porter et al., 2010; Van Damme and Smets, 2014; Jobson et al., 2023), we also performed an ANOVA on binary data. Previous research has conducted an ANOVA on such binary data (e.g., Porter et al., 2010 and Peace and Constantin, 2016, on major misinformation data; Sutherland and Hayne, 2001,). By also using a similar statistical approach, this can help to determine whether the effects found in previous research are also found in the present study. The findings from the ANOVA analysis were similar to those obtained using GEE, and are provided in full as a supplementary material for those interested. Both analyses lead to similar outcomes.

TABLE 1 Descriptive statistics for the false recognition of misleading and control details as a function of picture emotion, detail type, misinformation, and retention interval.

Retention interval	Immediate testing				Delayed testing			
Misinformation	Misleading		Control		Misleading		Control	
	M	SD	M	SD	M	SD	M	SD
Central details								
Negative/High	0.59	0.50	0.33	0.47	0.46	0.50	0.41	0.50
Negative/Low	0.67	0.47	0.22	0.42	0.35	0.48	0.39	0.49
Neutral	0.59	0.50	0.22	0.42	0.43	0.50	0.39	0.49
Peripheral details								
Negative/High	0.52	0.51	0.35	0.48	0.50	0.51	0.26	0.44
Negative/Low	0.54	0.50	0.20	0.40	0.28	0.46	0.17	0.38
Neutral	0.50	0.51	0.28	0.46	0.28	0.46	0.41	0.50

M and SD refer to Mean and Standard Deviation, respectively.

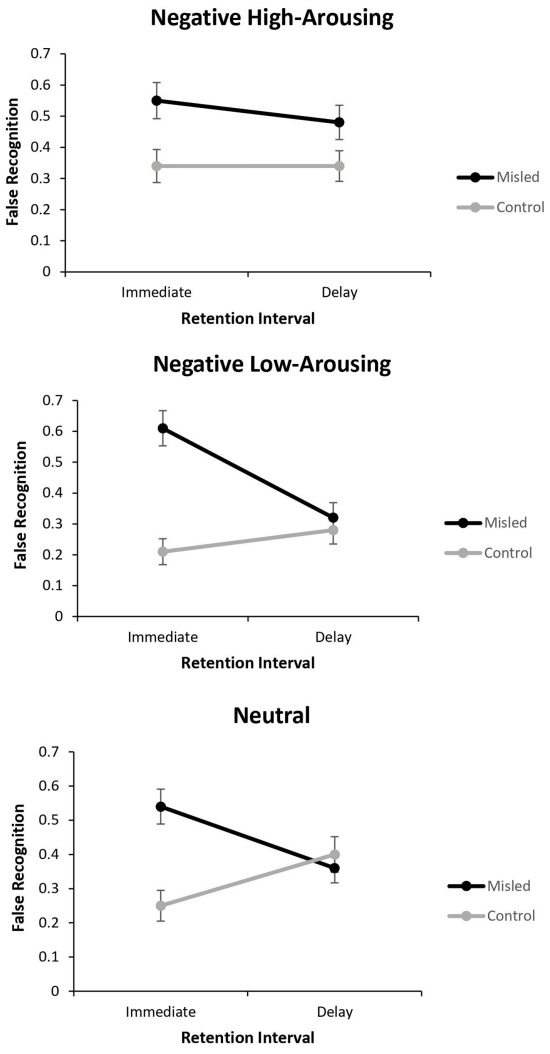


FIGURE 3
Line graphs showing the proportion of false recognition (out of two binary questions) of the critical minor details for each picture as a function of Retention Interval and Misinformation (Error bars represent the standard error).

Based on Porter et al. (2003, 2010) and Van Damme and Smets (2014), differences in the endorsement of the major misleading details across negative and neutral pictures over time were investigated. To do so, the factors picture emotion, misinformation, and retention interval, with between-subjects on the last factor, were submitted to a Generalised Estimating Equation analysis. See Table 2 for means and standard deviations. There was a significant misinformation effect, Wald $\chi^2(1, N=46)=16.96$, $p<0.001$, $d=0.80$. Furthermore, there was also a significant misinformation x retention interval interaction, Wald $\chi^2(1, N=46)=10.31$, $p=0.001$. At immediate testing, false endorsement rates were higher for misleading major details ($M=0.53$, $SD=0.35$) compared to control major details ($M=0.13$, $SD=0.21$), Wald $\chi^2(1, N=22)=29.07$, $p<0.001$, $d=1.39$. However, this misinformation effect was no longer significant at delayed testing (misleading: $M=0.33$, $SD=0.32$, control: $M=0.28$, $SD=0.25$), Wald $\chi^2(1, N=24)=0.45$, $p=0.501$, $d=0.17$. There were no further significant main effects (Wald χ^2's <0.69, $ps>0.710$), two-way interactions (Wald χ^2's <2.17, $ps>0.339$), and a three-way interaction (Wald $\chi^2=0.35$, $p=0.839$).

4 Discussion

Although extensive research has examined factors that increase and decrease the extent to which we are susceptible to misleading information regarding event recall, there are still questions to be answered regarding the impact of affective factors and how they influence memory distortion. The present study aimed to explore the impact of delayed retrieval and susceptibility to deception for negative/high arousal, negative/low arousal and neutral events, and based on previous emotion memory literature, whether there would be differential effects on memory distortion for central and peripheral details (Kaplan et al., 2012). Although previous research has examined the impact of delay on valanced stimuli, arousal was high for negative and positive images (Porter et al., 2010). To understand the role of arousal on memory distortion over time, participants were presented with a negative high-arousing, negative low-arousing, and neutral scene, followed by exposure to misleading central and peripheral details. Recognition memory was measured shortly after misinformation exposure and one week later.

TABLE 2 Descriptive statistics for the false recognition of major details as a function of picture emotion, misinformation, and retention interval.

Retention interval	Immediate testing				Delayed testing			
Misinformation	Misleading		Control		Misleading		Control	
	M	SD	M	SD	M	SD	M	SD
Negative/High	0.59	0.50	0.09	0.29	0.33	0.48	0.25	0.44
Negative/Low	0.45	0.51	0.18	0.40	0.25	0.44	0.29	0.46
Neutral	0.55	0.51	0.14	0.35	0.42	0.50	0.29	0.46

M and SD refer to Mean and Standard Deviation, respectively.

For the negative high-arousing event, we found, regardless of detail type, a misinformation effect that persisted over time. The magnitude of this effect was medium. Such a finding fits with the paradoxical negative emotion (Porter et al., 2008) hypothesis. This predicts that negative information will be remembered well over time, but can be associated with a greater susceptibility to distorting misleading information relative to other emotional events. This is because retaining memory of negative arousing events can be of adaptive significance (Porter et al., 2008) but it is also adaptive to incorporate all relevant information about negative events from trustworthy sources to further prepare for and/or avoid similar "dangerous" events in the future (Porter et al., 2008). Consistent with this, we found continued susceptibility to misinformation for the negative arousing events over time.

Such outcomes may also be explained based on source confusion. Post-event misinformation associated with the negative arousing event may have a strong memory trace and be more integrated into the original event, making source monitoring difficult. When answering the post-event questions, participants likely engage in the reconstruction of the original event and the rehearsal and visualisation of the misleading information (Johnson et al., 1993). This increases the overlap between memory for the original event and memory for the post-event information, consequently increasing source confusion. This has been empirically demonstrated in previous research (e.g., Dobson and Markham, 1993; Zaragoza and Lane, 1994). Misleading information about the negative low-arousing and neutral events can also be accompanied by mental visualisations. However, since arousal has been shown to benefit memory consolidation of negative information through the activations of the amygdala and hippocampus (e.g., McGaugh, 2002; Dolcos et al., 2005), it is plausible to assume that mental visualisations of the post-event information would be vivid, better integrated into memory for the original event, and better remembered over time. Consequently, misinformation may continue to affect memory for a negative arousing event due to source confusion, especially if the availability of source cues fades with time (Frost et al., 2002).

For the negative low-arousing and neutral events, the effect of misinformation at immediate testing (with a large effect size) disappeared after a delay. This was driven by a significant reduction in the recognition of misleading details after one week. Using an activation-based explanation (e.g., Source of Activation Confusion Model; Ayers and Reder, 1998), we could argue that such an effect occurs because the concept's strength (e.g., original or misleading detail) decays over time. Therefore, a stronger activation of the recently presented misinformation relative to a weaker activation of the original detail may lead to source misattribution errors of the activated concept. When testing after a short interval, misinformation receives more activation than the original detail because of its recent exposure, thus the original detail is less likely to be retrieved. At delayed testing, however, memory traces for both details are weaker, but the strength of the misinformation item is roughly equivalent to or below that of the original item's strength (Ayers and Reder, 1998; Lustig et al., 2004). The misleading information has a less distortive effect on memory at one week because its recency advantage is reduced and is thus less accessible to memory. Therefore, the original detail receives more activation and is subsequently retrieved (Ayers and Reder, 1998).

Overall, the reduction in the endorsement of misleading information associated with the negative low-arousing and neutral scenes after one week may be due to the reduced accessibility of the misleading information and greater activation of the original information over time. One could ask why this would not be the case for the negative high-arousing event. However, there are at least two possible reasons for why the spontaneous recovery of the original information did not occur for the high-arousing event. First, as mentioned earlier, the processing of misleading information and its integration within the original event may be stronger through the reconstruction of the original event. Thus, it is plausible to assume that there would be greater source confusions associated with the negative high-arousing event, particularly as we have suggested that cues fade over time (Frost et al., 2002). Second, high arousing information specifically benefits from long-term consolidation (e.g., Kensinger and Corkin, 2004). It may be that the visualisation of the post-event information with the negative-arousing event increases emotional arousal, thereby enhancing the encoding and consolidation of the misleading information and memory over time. Together, misleading information continues to interfere with memory for the negatively arousing event, thereby preventing an increase in correct recognition after a period of delay.

According to our analysis, the effect of retention interval and misinformation on memory for negative and neutral events did not significantly vary for central and peripheral details. Although note the contribution of peripheral details to the reported three-way interaction and therefore the need for future research to continue examining central and peripheral memory. Research has shown that negative events in general cause memory narrowing and that the presence of misinformation increases susceptibility to central misinformation (Van Damme and Smets, 2014). In addition, central information in an arousing event may specifically benefit from long-term consolidation (Christianson, 1992; Park, 2005). Based on these previous findings, we rationalised that retention interval could affect memory for central and peripheral misleading details for different emotional events. Although our findings did not support this rationale, previous misinformation studies have reported mixed results regarding the

effect of emotion on memory for central and peripheral misinformation (see Sharma et al., 2022, for a review). This could be attributed to methodological variations between the studies (e.g., the type of memory test, and the way central and peripheral details are determined). Our findings support Porter et al. (2003), who found no significant difference across emotional and neutral scenes. However, future research can determine whether our finding, irrespective of detail type, is a genuine result or an artefact of the study's design/procedure.

Turning briefly to major misinformation. Porter et al. (2010) found that major (peripheral) details associated with moderate-to-high arousing negative events were vulnerable to misinformation, which persisted over time. Although we saw a misinformation effect for major misinformation details at immediate testing with a large effect, this disappeared after a period of delay, and this did not differentiate across emotional picture conditions. We were unable to replicate negative emotion's specific susceptibility to "major misinformation" details. Two limitations should be mentioned. First, as this was treated as a between-participants factor due to methodological constraints, our sample size was low for analysing major misinformation. Second, there are procedural differences between these studies, including the type of test, definitions for central/peripheral details, and images used. The misinformation literature is fraught with procedural differences and understanding the impact of those differences in relation to the impact of emotion on memory distortion is work for future research.

To conclude, we found that misleading information continued to distort memory for a negatively arousing event over time, whereas memory performance improved for the negative low-arousing and neutral events. This has important applied implications for the development of false memories in forensic/legal settings. Eyewitnesses typically experience events that are negatively valenced and highly arousing (e.g., a robbery or an assault). They may also be exposed to misleading information about the events from, for example, other witnesses or the media. Indeed the latter point has some significance regarding the impact of conformity to misinformation from certain sources, such as those with perceived high credibility, intelligence, and authority on misinformation endorsement (e.g., Thorley, 2015; Mojtahedi et al., 2020). Furthermore, eyewitnesses may be asked to recall the event immediately after experiencing it or a few hours to weeks after the event (Neubauer and Fradella, 2011). Our findings highlight the continued detrimental impact of misinformation on memory for a negatively arousing event over time. Interestingly though, if the event is low arousing any impact of misleading information may not have a prolonged effect. Finally, we accept that there are some limitations to the present research. First, witnessing a photograph of a traumatizing scene is different from witnessing a real-life crime. Second, the use of forced-choice recognition tests may not reflect most recollections of real-life events where eyewitnesses are less likely to be forced to respond using a set number of responses (though see Howe et al., 2010). Future research can aim to use open questions to reduce the possible impact of correct guessing (Loftus et al., 1985) and response biases (Zaragoza et al., 1987) through the elimination of written cues, thereby increasing the probability of detecting the misinformation effects. Third, in line with previous misinformation research (e.g., Porter et al., 2003, 2010; Peace and Constantin, 2016), the pictures were chosen based on the normed valence and arousal from the IAPS database. Since these ratings were not collected in the present study, the manipulation of valence and arousal was not directly confirmed. Nonetheless, the current study provides insights into the potential impact of arousing negative events and the influences on susceptibility to misinformation for such events. This appears to be specific to arousing negative details and not negative valenced events in general. Given that our research suggests that people who view a highly disturbing scene are far more prone to incorporate misinformation into their memory relative to other scenes, suggests that despite the level of complexity of the event, it is essential that improper questioning techniques be avoided in practise to reduce the problem of inaccurate testimony.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Psychology Ethics Committee/City, University of London. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

The research design, data collection, analysis, and the initial full draft of the manuscript was conducted by DS. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1212709/full#supplementary-material

References

Ayers, M. S., and Reder, L. M. (1998). A theoretical review of the misinformation effect: predictions from an activation-based memory model. *Psychon. Bull. Rev.* 5, 1–21. doi: 10.3758/BF03209454

Bergado, J. A., Lucas, M., and Richter-Levin, G. (2011). Emotional tagging—a simple hypothesis in a complex reality. *Prog. Neurobiol.* 94, 64–76. doi: 10.1016/j.pneurobio.2011.03.004

Bradley, M. M., and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25, 49–59. doi: 10.1016/0005-7916(94)90063-9

Bradley, M. M., and Lang, P. J. (2007). "The international affective picture system (IAPS) in the study of emotion and attention". in *Handbook of emotion elicitation and assessment* Eds. J. A. Coan and J. J. B. Allen (New York: Oxford University Press) 29–46.

Buchanan, T. W., Etzel, J. A., Adolphs, R., and Tranel, D. (2006). The influence of autonomic arousal and semantic relatedness on memory for emotional words. *Int. J. Psychol.* 61, 26–33. doi: 10.1016/j.ijpsycho.2005.10.022

Burke, A., Heuer, F., and Reisberg, D. (1992). Remembering emotional events. *Mem. Cogn.* 20, 277–290. doi: 10.3758/BF03199665

Campbell, J. I., and Thompson, V. A. (2012). MorePower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis. *Behav. Res. Methods* 44, 1255–1265. doi: 10.3758/s13428-012-0186-0

Charles, S. T., Mather, M., and Carstensen, L. L. (2003). Aging and emotional memory: the forgettable nature of negative images for older adults. *J. Exp. Psychol. Gen.* 132, 310–324. doi: 10.1037/0096-3445.132.2.310

Christianson, S. Å. (1992). Emotional stress and eyewitness memory: a critical review. *Psychol. Bull.* 112, 284–309. doi: 10.1037/0033-2909.112.2.284

Christianson, S. Å., and Loftus, E. F. (1987). Memory for traumatic events. *Appl. Cogn. Psychol.* 1, 225–239. doi: 10.1002/acp.2350010402

Christianson, S. Å., Loftus, E. F., Hoffman, H., and Loftus, G. R. (1991). Eye fixations and memory for emotional events. *J. Exp. Psychol. Learn. Mem. Cogn.* 17:693. doi: 10.1037//0278-7393.17.4.693

Dobson, M., and Markham, R. (1993). Imagery ability and source monitoring: implications for eyewitness memory. *Br. J. Psychol.* 84, 111–118. doi: 10.1111/j.2044-8295.1993.tb02466.x

Dolcos, F., Denkova, E., and Dolcos, S. (2012). Neural correlates of emotional memories: a review of evidence from brain imaging studies. *Psychologia* 55, 80–110. doi: 10.2117/psysoc.2012.80

Dolcos, F., LaBar, K. S., and Cabeza, R. (2005). Remembering one year later: role of the amygdala and the medial temporal lobe memory system in retrieving emotional memories. *Proc. Natl. Acad. Sci.* 102, 2626–2631. doi: 10.1073/pnas.0409848102

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

Forgas, J. P., Laham, S. M., and Vargas, P. T. (2005). Mood effects on eyewitness memory: affective influences on susceptibility to misinformation. *J. Exp. Soc. Psychol.* 41, 574–588. doi: 10.1016/j.jesp.2004.11.005

Frost, P. (2000). The quality of false memory over time: is memory for misinformation "remembered" or "known"? *Psychon. Bull. Rev.* 7, 531–536. doi: 10.3758/BF03214367

Frost, P., Ingraham, M., and Wilson, B. (2002). Why misinformation is more likely to be recognised over time: a source monitoring account. *Memory* 10, 179–185. doi: 10.1080/09658210143000317

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends Cogn. Sci.* 5, 394–400. doi: 10.1016/S1364-6613(00)01707-1

Howe, M. L., Candel, I., Otgaar, H., Malone, C., and Wimmer, M. C. (2010). Valence and the development of immediate and long-term false memory illusions. *Memory* 18, 58–75. doi: 10.1080/09658210903476514

Heuer, F., and Reisberg, D. (1990). Vivid memories of emotional events: the accuracy of remembered minutiae. *Mem. Cogn.* 18, 496–506. doi: 10.3758/BF03198482

Jobson, L., Wade, K. A., Rasor, S., Spearing, E., McEwen, C., and Fahmi, D. (2023). Associations between the misinformation effect, trauma exposure, and symptoms of posttraumatic stress disorder and depression. *Memory* 31, 179–191. doi: 10.1080/09658211.2022.2134422

Johnson, M. K., Hashtroudi, S., and Lindsay, D. S. (1993). Source monitoring. *Psychol. Bull.* 114, 3–28. doi: 10.1037/0033-2909.114.1.3

Kaplan, R. L., Van Damme, I., and Levine, L. J. (2012). Motivation matters: differing effects of pre-goal and post-goal emotions on attention and memory. *Front. Psychol.* 3:404. doi: 10.3389/fpsyg.2012.00404

Kensinger, E. A. (2009). Remembering the details: effects of emotion. *Emot. Rev.* 1, 99–113. doi: 10.1177/1754073908100432

Kensinger, E. A., and Corkin, S. (2004). Two routes to emotional memory: distinct neural processes for valence and arousal. *Proc. Natl. Acad. Sci.* 101, 3310–3315. doi: 10.1073/pnas.0306408101

LaBar, K. S., and Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nat. Rev. Neurosci.* 7, 54–64. doi: 10.1038/nrn1825

LaBar, K. S., and Phelps, E. A. (1998). Arousal-mediated memory consolidation: role of the medial temporal lobe in humans. *Psychol. Sci.* 9, 490–493. doi: 10.1111/1467-9280.00090

Laney, C., Heuer, F., and Reisberg, D. (2003). Thematically induced arousal in naturally occurring emotional memories. *Appl. Cogn. Psychol.* 17, 995–1004. doi: 10.1002/acp.951

Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual* Gainesville, FL: University of Florida.

Levine, L. J., and Edelstein, R. S. (2009). Emotion and memory narrowing: a review and goal-relevance approach. *Cognit. Emot.* 23, 833–875. doi: 10.1080/02699930902738863

Loftus, E. F. (2005). Planting misinformation in the human mind: a 30-year investigation of the malleability of memory. *Learn. Mem.* 12, 361–366. doi: 10.1101/lm.94705

Loftus, E. F., Miller, D. G., and Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *J. Exp. Psychol. Hum. Learn. Mem.* 4, 19–31.

Loftus, E. F., Schooler, J. W., and Wagenaar, W. A. (1985). The fate of memory: comment on McCloskey and Zaragoza. *J. Exp. Psychol. General* 114, 375–380. doi: 10.1037/0096-3445.114.3.375

Luna, K., and Martín-Luengo, B. (2018). The (limited) effect of emotional arousal in the regulation of accuracy in eyewitness memory. *Psicológica J* 39, 1–24. doi: 10.2478/psicolj-2018-0001

Lustig, C., Konkel, A., and Jacoby, L. L. (2004). Which route to recovery? Controlled retrieval and accessibility bias in retroactive interference. *Psychol. Sci.* 15, 729–735. doi: 10.1111/j.0956-7976.2004.00749.x

MacKay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., and Dyer, J. R. (2004). Relations between emotion, memory and attention: evidence from taboo stroop, lexical decision, and immediate memory tasks. *Mem. Cogn.* 32, 474–487. doi: 10.3758/BF03195840

McGaugh, J. L. (2000). Memory--a century of consolidation. *Science* 287, 248–251. doi: 10.1126/science.287.5451.248

McGaugh, J. L. (2002). Memory consolidation and the amygdala: a systems perspective. *Trends Neurosci.* 25, 456–461. doi: 10.1016/s0166-2236(02)02211-7

McReynolds, J. R., and McIntyre, C. K. (2012). Emotional modulation of the synapse. *Rev. Neurosci.* 23, 449–461. doi: 10.1515/revneuro-2012-0047

Mitchell, K. J., and Johnson, M. K. (2000). "Source monitoring: attributing mental experiences" in *The Oxford handbook of memory*. eds. E. Tulving and F. Craik (New York: Oxford University Press), 179–195.

Mojtahedi, D., Ioannou, M., and Hammond, L. (2020). Intelligence, authority and blame conformity: co-witness influence is moderated by the perceived competence of the information source. *J. Police Crim. Psychol.* 35, 422–431. doi: 10.1007/s11896-019-09361-2

Mudd, K., and Govern, J. M. (2004). Conformity to misinformation and time delay negatively affect eyewitness confidence and accuracy. *N. Am. J. Psychol.* 6, 227–238.

Murty, V. P., Ritchey, M., Adcock, R. A., and LaBar, K. S. (2010). fMRI studies of successful emotional memory encoding: a quantitative meta-analysis. *Neuropsychologia* 48, 3459–3469. doi: 10.1016/j.neuropsychologia.2010.07.030

Neubauer, D., and Fradella, H. (2011). *America's courts and the criminal justice system*. Belmont, CA: Wadsworth.

Otgaar, H., Candel, I., and Merckelbach, H. (2008). Children's false memories: easier to elicit for a negative than a neutral event. *Acta Psychol.* 128, 350–354. doi: 10.1016/j.actpsy.2008.03.009

Park, J. (2005). Effect of arousal and retention delay on memory: a meta-analysis. *Psychol. Rep.* 97, 339–355. doi: 10.2466/pr0.97.2.339-355

Paz-Alonso, P. M., Goodman, G. S., and Ibabe, I. (2013). Adult eyewitness memory and compliance: effects of post-event misinformation on memory for a negative event. *Behav. Sci. Law* 31, 541–558. doi: 10.1002/bsl.2081

Peace, K. A., and Constantin, K. M. (2016). Misremembering events: emotional valence, psychopathic traits, and the misinformation effect. *J. Police Crim. Psychol.* 31, 189–199. doi: 10.1007/s11896-015-9178-1

Porter, S., Bellhouse, S., McDougall, A., Ten Brinke, L., and Wilson, K. (2010). A prospective investigation of the vulnerability of memory for positive and negative emotional scenes to the misinformation effect. *Can. J. Behav. Sci.* 42, 55–61. doi: 10.1037/a0016652

Porter, S., and Peace, K. A. (2007). The scars of memory: a prospective, longitudinal investigation of the consistency of traumatic and positive emotional memories in adulthood. *Psychol. Sci.* 18, 435–441. doi: 10.1111/j.1467-9280.2007.01918.x

Porter, S., Spencer, L., and Birt, A. R. (2003). Blinded by emotion? Effect of the emotionality of a scene on susceptibility to false memories. *Can. J. Behav. Sci.* 35, 165–175. doi: 10.1037/h0087198

Porter, S., Taylor, K., and Ten Brinke, L. (2008). Memory for media: investigation of false memories for negatively and positively charged public events. *Memory* 16, 658–666. doi: 10.1080/09658210802154626

Roediger, H. L., and McDermott, K. B. (1995). Creating false memories: remembering words not presented in lists. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 803–814. doi: 10.1037/0278-7393.21.4.803

Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol.* 39, 1161–1178. doi: 10.1037/h0077714

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychol. Rev.* 110, 145–172. doi: 10.1037/0033-295X.110.1.145

Schwarz, N., and Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: informative and directive functions of affective states. *J. Pers. Soc. Psychol.* 45, 513–523. doi: 10.1037/0022-3514.45.3.513

Sharma, P. R., Wade, K. A., and Jobson, L. (2022). A systematic review of the relationship between emotion and susceptibility to misinformation. *Memory* 1-21:2120623. doi: 10.1080/09658211.2022.2120623

Sharot, T., and Phelps, E. A. (2004). How arousal modulates memory: disentangling the effects of attention and retention. *Cogn. Affect. Behav. Neurosci.* 4, 294–306. doi: 10.3758/CABN.4.3.294

Sommer, T., Glascher, J., Moritz, S., and Buchel, C. (2008). Emotional enhancement effect of memory: removing the influence of cognitive factors. *Learn. Mem.* 15, 569–573. doi: 10.1101/lm.995108

Steblay, N. M. (1992). A meta-analytic review of the weapon focus effect. *Law Hum. Behav.* 16, 413–424. doi: 10.1007/BF02352267

Sutherland, R., and Hayne, H. (2001). Age-related changes in the misinformation effect. *J. Exp. Child Psychol.* 79, 388–404. doi: 10.1006/jecp.2000.2610

Talmi, D. (2013). Enhanced emotional memory: cognitive and neural mechanisms. *Curr. Dir. Psychol. Sci.* 22, 430–436. doi: 10.1177/0963721413498893

Talmi, D., Luk, B. C., McGarry, L. M., and Moscovitch, M. (2007). The contribution of relatedness and distinctiveness to emotionally-enhanced memory. *J. Mem. Lang.* 56, 555–574. doi: 10.1016/j.jml.2007.01.002

Talmi, D., and McGarry, L. M. (2012). Accounting for immediate emotional memory enhancement. *J. Mem. Lang.* 66, 93–108. doi: 10.1016/j.jml.2011.07.009

Thorley, C. (2015). Blame conformity: innocent bystanders can be blamed for a crime as a result of misinformation from a young, but not elderly, adult co-witness. *PLoS One* 10:e0134739. doi: 10.1371/journal.pone.0134739

Van Damme, I., and Seynaeve, L. (2013). The effect of mood on confidence in false memories. *J. Cogn. Psychol.* 25, 309–318. doi: 10.1080/20445911.2013.769440

Van Damme, I., and Smets, K. (2014). The power of emotion versus the power of suggestion: memory for emotional events in the misinformation paradigm. *Emotion* 14, 310–320. doi: 10.1037/a0034629

Wang, B. (2014). Effect of time delay on recognition memory for pictures: the modulatory role of emotion. *PLoS One* 9:e100238. doi: 10.1371/journal.pone.0100238

Waring, J. D., and Kensinger, E. A. (2009). Effects of emotional valence and arousal upon memory trade-offs with aging. *Psychol. Aging* 24, 412–422. doi: 10.1037/a0015526

Yegiyan, N. S., and Yonelinas, A. P. (2011). Encoding details: positive emotion leads to memory broadening. *Cognit. Emot.* 25, 1255–1262. doi: 10.1080/02699931.2010.540821

Zaragoza, M. S., and Lane, S. M. (1994). Source misattributions and the suggestibility of eyewitness memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 934–945. doi: 10.1037//0278-7393.20.4.934

Zaragoza, M. S., and Mitchell, K. J. (1996). Repeated exposure to suggestion and the creation of false memories. *Psychol. Sci.* 7, 294–300. doi: 10.1111/j.1467-9280.1996.tb00377.x

Zaragoza, M. S., McCloskey, M., and Jamis, M. (1987). Misleading postevent information and recall of the original event: Further evidence against the memory impairment hypothesis. *J Exp Psychol Learn Mem Cogn.* 13, 36–44. doi: 10.1037//0096-3445.114.1.1

Zeger, S., and Liang, K. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 42, 121–130. doi: 10.2307/2531248

Zhang, W., Cowan, G., Colombo, M., Gross, J., and Hayne, H. (2021). Emotional content of the event but not mood influences false memory. *Appl. Cogn. Psychol.* 35, 1418–1426. doi: 10.1002/acp.3872

frontiers | Frontiers in Psychology

TYPE Review
PUBLISHED 07 December 2023
DOI 10.3389/fpsyg.2023.1268757

OPEN ACCESS

EDITED BY
Fabiana Battista,
University of Bari Aldo Moro, Italy

REVIEWED BY
Rafaela Tavares Seixas,
University of Porto, Portugal
Adriano Schimmenti,
Kore University of Enna, Italy

*CORRESPONDENCE
Olivier Dodier
✉ olivier.dodier@unimes.fr

†These authors have contributed equally to this work

RECEIVED 31 July 2023
ACCEPTED 13 November 2023
PUBLISHED 07 December 2023

CITATION
Dodier O, Barzykowski K and Souchay C (2023) Recovered memories of trauma as a special (or not so special) form of involuntary autobiographical memories.
Front. Psychol. 14:1268757.
doi: 10.3389/fpsyg.2023.1268757

Recovered memories of trauma as a special (or not so special) form of involuntary autobiographical memories

Olivier Dodier[1*†], Krystian Barzykowski[2†] and Céline Souchay[3]

[1]APSY-V Laboratory, Department of Psychology, Literature, Languages and History, University of Nîmes, Nîmes, France, [2]Applied Memory Research Laboratory, Faculty of Philosophy, Institute of Psychology, Jagiellonian University, Kraków, Poland, [3]Laboratoire de Psychologie et Neurocognition, LPNC CNRS 5105, Université Grenoble Alpes, Grenoble, France

Recovered memories of trauma are memories of traumatic events experienced generally during childhood, but of which the persons were unaware until they retrieved it. Legal decisions are sometimes based on such recovered memories, the validity of which is often questioned. Yet, people can recover genuine traumatic memories of childhood abuse. In this paper, we present and further discuss the idea that recovered traumatic memories can be interpreted in the context of the autobiographical memory framework. Specifically, we argue that recovered memories may be accessed after exposure to incidental cues that initiate unexpected spontaneous memory retrieval. Thus, we relate the recovered memory phenomenon to involuntary autobiographical memories and argue that it is an example of highly stressful, emotionally negative, and intense involuntary memories that were yet never recalled. This novel, evidence-based perspective leads us to reconsider the examination of the validity of eyewitness testimony as a continuum ranging from the least valid form (i.e., memories recovered in highly suggestive context facilitating its factitious reconstruction) to the most valid form (i.e., memories that were triggered by cues without any person's voluntary engagement), and this in relation with how internal (e.g., age and internal cue) or external (e.g., suggestion in therapy, suggestion during investigative interview, and contextual cue) factors may influence memory retrieval. Finally, we propose several recommendations for experts that may be useful in assessing the validity of a testimony based on recovered memories.

KEYWORDS

recovered memory, involuntary autobiographical memory, false memory, traumatic memory, trauma, retrieval cue

1 Introduction

In this paper, we present and further discuss the idea that memory processes can contribute to the explanation of (at least) some instances of the phenomenon of recovered memories of traumatic events. This type of memories was the subject of heated debate during the "memory wars" in the 1990s (Crews, 1995). On the one hand, therapists considered that it was possible for traumatic memories to be repressed and thus pushed outside the boundaries of consciousness, before returning in their original form, notably through therapeutic methods (e.g., Bass and Davis, 1988). On the other hand, skeptics, mainly experimental psychologists, but also several

clinical psychologists and psychiatrists, considered that such recovered memories were probably false memories suggested by third parties, generally psychotherapists (in this case memories of events that never took place, thus rejecting the hypothesis of traumatic repression, Loftus, 1993; Holmes, 1994; Lilienfeld and Loftus, 1999). The "memory wars" controversy came to a head when it reached the courtroom, where people were accused of committing sexual abuse based on memories recovered in psychotherapy (Loftus, 1993). This led to quarrels between experts, with supporters of repression on a side and skeptics explaining these accusations by induced false memories on the other. These disputes unfortunately turned an initially scientific debate into a popular one (see detailed examples in Loftus, 1993; see also examples of false memories allegedly induced by therapists in Kaplan and Manicavasagar, 2001).

Recent work has shown that memory wars still rage on (Dodier, 2019; Otgaar et al., 2019; Battista et al., 2023), with similar opposing camps. However, it appears that the debates have shifted towards neuroscientific approaches, with some authors (e.g., Markowitsch and Staniloiu, 2013; Chechko et al., 2018; Dimitrova et al., 2021) suggesting the existence of brain biomarkers of dissociative amnesia (another way of calling repression, Otgaar et al., 2019, 2023), while others have criticized the validity and lack of homogeneity of the results found in studies using brain imaging (Otgaar et al., in press). Closely associated with the topic addressed in our article, the debate has also shifted somewhat to something other than a repressed vs. false memory dichotomy. Specifically, while repression (or dissociative amnesia) is still criticized for its validity (Otgaar et al., 2019; Patihis et al., 2019; Battista et al., 2023; Pope et al., 2023), other explanations besides false memories are being put forth to explain recovered memories of trauma that may occur outside of a therapeutic setting. For example, some memories may not be fully encoded, due for example to stress that may limit the integration of certain information (Deffenbacher et al., 2004), or the use of substances such as alcohol or drugs (see Kloft et al., 2021; Segura et al., 2021). Another example is that in some cases the recovered memories of trauma are in fact continuous memories (i.e., memories of events that people feel they have always known occurred) that have only been reinterpreted as abuse with the time and maturity to understand the event (McNally and Geraerts, 2009; McNally, 2023). In this case, the event would be experienced in a non-traumatic way during encoding (because children are too young to understand the event, especially when it is of a sexual nature), with time the individuals do not think about it, before exposure to a contextual cue allows for the involuntary retrieval of an autobiographical memory, which can then take on its traumatic nature.

Specifically, our aim is to argue that recovered memories of actual traumatic events (i.e., true memories, as opposed to false memories, in this case memories of events that did not take place) are usually recovered after exposure to incidental cues that initiate unexpected spontaneous memory retrieval. In this context, after (i) a brief review of the literature on recovered memories of trauma, we will (ii) relate this phenomenon to the literature on involuntary autobiographical memories and argue that, at least in some cases, it may be an example of highly stressful, emotionally intense, and extremely negative involuntary memories that were yet never recalled prior to the unexpected memory recovery. Next, we will then (iii) propose that the validity of eyewitness testimonies (focusing on ones that come from recovered memories of trauma) may lie along a continuum ranging from the least valid memories (i.e., memories recovered in highly suggestive context facilitating its factitious reconstruction, e.g., during therapy) to the most valid memories that were triggered by cues without any person's voluntary engagement (e.g., involuntary autobiographical memories retrieved unexpectedly during watching a movie). Finally, we will (iv) propose brief recommendations for expert witnesses that may be useful in assessing the validity of a testimony based on recovered memories of trauma.

2 Recovered memories of trauma

Recovered memories of trauma are memories of a generally stressful and distressing event that a person has, of which he or she was unaware until he or she remembered it (Ost, 2006). For example, in a case described by Dodier et al. (2023), a 16-year-old girl suddenly recovered her memory of sexual abuse by her great uncle when she was 8, after hearing his name in the middle of a discussion. In an interview, she said she did not know the abuse had happened until she retrieved it. Such memories are usually accompanied by a sense of surprise at (re)discovering the facts (Geraerts et al., 2007), to the point where they could lead to significant psychological distress upon recovery (McNally and Geraerts, 2009). Recovered memories, when they are traumatic in nature, can join the reversible feature of dissociative amnesia, as defined by the DSM-5-TR (American Psychiatric Association, 2022).

Traditionally, recovered memories of trauma have been the subject of debate between scientists and clinicians in the fields of research, therapy, and justice, as the validity of such memories is difficult to establish and therefore so is the examination of the likelihood of those memories being false. Indeed, much laboratory work, widely replicated, has highlighted the ease of creating false memories of entire events in people (Scoboria et al., 2017), including criminal events (Shaw and Porter, 2015; but see Wade et al., 2018). These criticisms were made because many cases of recovered memories of trauma occurred in a therapeutic setting, by some clinicians who were convinced that childhood traumas had been repressed, and that it was necessary to recover them in order to heal people. The problem was that the methods dedicated to recovering memories were highly suggestive and resembled in many ways the experimental methods used to access for creating false memories (Ost, 2006). While there is a clear difference between laboratory experiments and false memories in real life (which are usually traumatic), it is important to note that documented cases of suggested false traumatic memories have been reported in the literature (Loftus, 1993; Kaplan and Manicavasagar, 2001; Otgaar et al., 2022).

Clearly, not all recovered memories of trauma are false memories; we know of no memory specialist who would consider every recovered memory of trauma to be necessarily a false memory. Rather, our position is that it depends on the context in which these recovered memories are, in fact, recovered. This is consistent with our central idea that the validity of testimonies should be examined on a continuum, rather than in a purely categorical fashion.

As just mentioned, recovered memories of trauma are traditionally associated with the therapeutic context, to the point where there is a popular belief that it is necessary to recover repressed memories in order to heal various disorders that are believed to be unconscious expressions of childhood trauma (Otgaar et al., 2021). However, it appears that the vast majority of recovered memories of trauma occur

outside of any therapeutic context, and even that individuals are alone when they recover the memory (Dodier and Patihis, 2021). The aforementioned study did point out that such recovered memories alone could result from suggestions (e.g., following discussions or self-documentation on the topic of repressed traumatic memories), but it also raises the question of the possibility that memories of events with traumatic potential can be recovered following exposure to a contextually derived retrieval cue, consistent with Tulving's classic work on the specificity of encoding (e.g., Tulving and Thomson, 1973) and the multiple pathways to access a memory trace (Tulving, 1974).

This is related to the hypothesis put forward by McNally and Geraerts (2009) according to which recovered memories of childhood sexual abuse that actually occurred reflect that the individuals were too young at the time of the event to understand. The memory would then be encoded as a bizarre, confusing and unusual event (Clancy and McNally, 2005/2006), but not as traumatic. Thus, as demonstrated by experimental work on the forgetting curve (Murre and Dros, 2015), the strength of the memory trace of the event would decline. After a period of time that can vary from months to decades, a retrieval cue (e.g., hearing about the perpetrator, returning to the scene) would allow the involuntary and spontaneous retrieval of the memory, which would then be reinterpreted as sexual abuse. This hypothesis was corroborated by Dodier and Patihis (2021) showing that a third of the people claiming to have recovered memories of sexual abuse during childhood specified that they had never really forgotten it, but had reinterpreted it over time.

These different research findings then allow us to consider a new evidence-based approach to explaining recovered memories of trauma. More specifically, in the next section we will develop the encoding and retrieval mechanisms behind involuntary autobiographical memories, and how this work offers a powerful explanatory framework to account for such recovered memories.

3 Involuntary autobiographical memories as a possible framework to understand recovered memories of trauma

Our main idea put forward here is that, under certain circumstances, a memory of a traumatic/unpleasant past episode may simply pop into our mind without any preceding intention (i.e., when one did not try to recall a given memory). Importantly, such involuntary retrieval of a past episode may become a core element (or starting point) of a recovered memory of trauma (e.g., when one spontaneously experiences a past memory of being abused by a given person may start voluntarily thinking about that situation elaborating further on that experience giving a rise to recovered memory of trauma). While there may be several possible mechanisms of recovered memories of trauma (e.g., simple processes such as failure to remember a prior recall of the event and/or forgetting mechanisms; Otgaar et al., 2019), which may not be mutually exclusive, we focus on some instances of recovered memories of trauma as a result of involuntary autobiographical memory retrieval. To this end, we briefly introduce the self-memory system first (Conway and Pleydell-Pearce, 2000), followed by conceptualization of memory retrieval stages (Barzykowski and Mazzoni, 2022; Barzykowski and Moulin, 2023). Then, we discuss possible circumstances under which such recovered memories may be most likely to occur and, as we argue in the present paper, may be even a more valid representation of the event.

3.1 The self-memory system

The ability to remember our personal past; namely, anything that we have witnessed and/or experienced while being self-reflectively aware that a given remembered event belongs to our personal past is called autobiographical memory (e.g., Tulving, 1985; Conway and Pleydell-Pearce, 2000). Recent theories of autobiographical memory acknowledge two broad ways in which such memories can be accessed and which are the result of the presence or the lack of conscious intention (i.e., wanting to recall a given memory): involuntary and voluntary retrieval (Roberts et al., 1994; Berntsen, 1996; Schlagman and Kvavilashvili, 2008; for similarities and differences between involuntary and voluntary memories see also, Barzykowski and Staugaard, 2016, 2018; Barzykowski et al., 2019a). Therefore, each time we want to recall, for instance, a childhood summer holiday at our grandparents with more or less detailed events (e.g., *eating cherries directly from the tree*, etc.), we use our voluntary memories. However, sometimes such memories may come to our mind without any conscious attempt at retrieval, for example, when watching a movie, a memory of *having delicious cherries with grandparents during childhood holiday* may simply enter our mind without being sought-for. While involuntary autobiographical memories were somewhat ignored for several decades (e.g., Miller, 1962/1974), they are now considered as a basic mode of remembering, central to psychological well-being and, importantly, frequently experienced in a daily life (e.g., Berntsen, 2010; Rasmussen and Berntsen, 2011; Uzer et al., 2012). An important result from a naturalistic diary study, among others, is that involuntary memories arise in response to incidental (both internal or external) cues that usually overlap with key features of the memory content (Berntsen, 1998; Schlagman et al., 2007; e.g., *seeing a cherry may trigger a certain past episode of picking and eating them in grandparents' garden*).

According to the influential model proposed by Conway and Pleydell-Pearce (2000; for later modifications see also: Conway, 2008, 2009; Conway and Jobson, 2012), autobiographical memory consists of a hierarchical network of interconnected nodes that differ in terms of their level of specificity. At the top of the network are superordinate levels of important periods of one's life (e.g., *when being a child*), general events (e.g., *holiday*) and common themes (e.g., *summer holiday with grandparents*). At the bottom of the network are stored fragments of events with specific sensory details (e.g., *details experienced when picking/eating cherries directly from the tree*). Higher levels are constituted by such basic and specific memory contents. Importantly, the self-memory system and the ability to remember personal past emerges over the years of cognitive development (e.g., language). For instance, there is a strong relationship between language development and memory showing that the better language skills, the better (more efficiently) autobiographical memory works (e.g., Fivush and Hamond, 1990; Leichtman and Ceci, 1993; Nelson, 1993). This means that over the course of cognitive and language development children are better at taking control over their memory, being able to more efficiently: encode events (knowing what is important, paying attention to the events, elaborating them to be better remembered, understanding an event within a broader

context), store events (as they are stored in relatively stable, organized cognitive schemas, scripts and knowledge nodes/units), retrieve (as they are better at using cues and strategies to recall a given event).

3.2 Conceptualization of memory retrieval stages

Conway and Pleydell-Pearce (2000) suggested that retrieval of a given memory is due to the activation of autobiographical information that spreads across the network. Furthermore, such activation may be elicited by different types of cues leading to either generative or direct retrieval. While generative retrieval is a result of a top-down cognitively controlled search process (i.e., we know what memory we want to recall and we do our best to do so), direct retrieval is thought of circumventing the search process accessing a memory very quickly (and thus is beyond our control as it happens to us rather than we have control over it). Conway and Pleydell-Pearce (2000) argued that fragments of memory representations are constantly activated at the bottom level of the hierarchy by different internal and external cues—i.e., those in the environment, those acted upon by the recognition memory system; but the vast majority of such memories cued by internal and external cues never reach consciousness. While there are several hypotheses of why such memories do not reach the awareness threshold (see for instance Vannucci et al., 2015; Barzykowski et al., 2019b), from the perspective of the present paper it is more important to reflect on the question of how memories are successfully retrieved. As presented in Figure 1, to simplify the act of memory retrieval, it may be conceptualized into four following demonstrative stages (Barzykowski and Mazzoni, 2022; Barzykowski and Moulin, 2023): (1) pre-retrieval stage (e.g., modifying memory accessibility by priming), (2) retrieval stage (forming and developing a memory), (3) post-retrieval stage (becoming aware of a memory and further its processing), and (4) retrieval outcome report stage (giving verbal account and gaining understanding of a memory and past episode) (see Figure 1).

The pre-retrieval stage (1) may be associated with any cognitive processes that either facilitate or impair retrieval. More precisely, during the pre-retrieval stage an individual may be in "*retrieval mode*" in which "*the cognitive system is prepared for or expects memory construction and recollection*" (Conway, 2001, p. 1379). For instance, one may be occasionally (i.e., incidentally and peripherally) exposed to some information more or less directly relating to childhood abuse. This may lead to the effect of priming which, for some memories, may enhance the likelihood that a given memory will be triggered and/or will enter a person's awareness (e.g., seeing children).

The retrieval stage (2) relates to the forming and development of an autobiographical memory. According to the self-memory system, any memory information stored in the autobiographical memory system may be: (a) retrieved automatically without any conscious intention, and/or (b) triggered by internal or external cue, and/or (c) activated by spreading activation mechanism. This means that while we may have an access to memory content we know (or think) exists in our memory (but there are some exceptions, e.g., non-believed memories, see Mazzoni et al., 2010), there may also be instances of retrieval of memories that were either (i) forgotten (inaccessible and/or unavailable but recognized as known and experienced in personal past, e.g., *a memory of cherry picking in my grandparents' garden popped in my mind, I forgot about it but now I remember it*) leading to strong feeling of surprise (as in the proustian-like memories) or (ii) rediscovered (inaccessible and unavailable and, importantly, recognized as not known before as experienced in the personal past, e.g., *I have just remembered a neighbor touching me up when I was picking cherries in my grandparents' garden*). In general, during the retrieval stage (2) a memory is triggered by and/or accessed via a given cue, and it may be either reconstructed, directly retrieved, voluntarily searched or involuntarily recalled, depending on the memory pre-retrieval and retrieval processes involved. It is also worth underlining that such a memory retrieval may be without explicit self-reflection; namely, a given memory might have been formed but one may not be explicitly aware of it yet (something that refers to an "*experiential level of consciousness*"; Baird et al., 2013).

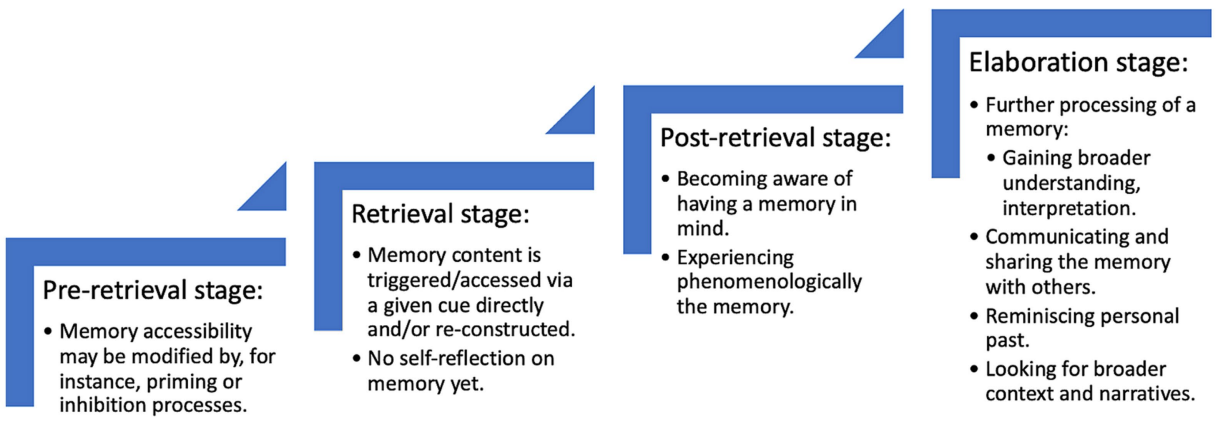


FIGURE 1
The theoretical and demonstrative conceptualization of autobiographical memory retrieval to better disentangle the possible factors operating during a memory retrieval as developed and described by Barzykowski and Mazzoni, 2022 and Barzykowski and Moulin (2023). These stages are thought to be as a dynamic system, where the retrieval process can jump back and forth and even activate several stages at any given time. A helpful analogy is a neural network.

Once the memory is formed, during the post-retrieval phase (3), people may become aware of having the memory in mind. Thus, this stage relates to the ability to, for example, extract autobiographical content from the stream of consciousness to explicitly become aware of having a memory that is autobiographical (this is the level of meta-awareness). At this stage, one is fully aware that an autobiographical memory was actually retrieved but no further understanding or interpretation is done yet. Importantly, becoming aware of an involuntary memory popping into mind is most frequent when one is engaged in an attentionally undemanding activity (e.g., driving a car, washing a dish etc., for experimental studies see Vannucci et al., 2015; Barzykowski and Niedźwieńska, 2018a).

In the last stage (4), the retrieved memory may be shared with others and reported by giving a verbal account of the content. This is a stage during which a broader understanding of a given memory is developed. Depending on the context of that memory (e.g., *memory of sitting on a Santa-Claus laps while having a photo in a shopping mall* vs. *sitting on a lap of a given person while being alone in a bedroom*), one may come to a conclusion that a given memory may be understood as an experience of, for instance, childhood abuse, although it was not encoded/understood in such a manner at the time of the event occurrence. Furthermore, different events may be encoded/remembered with various levels of emotional intensity. As demonstrated in previous studies (e.g., Barzykowski and Staugaard, 2016, 2018; also, studies showing that emotional memories contain more details, e.g., D'Argembeau et al., 2003; Comblain et al., 2005) the general rule is that the more emotionally intense an original event was, the easier a given memory should pass the awareness threshold making ones' self-aware of having a memory of that event (Stage 3). However, in some cases, especially if the event was not fully understood, emotions may arise in response to interpretation of an event made in Stage 4 when a broader understanding of it is gained. Put differently, a child may not fully understand the event of abuse. However, once it is recalled as an adult, the event may be reinterpreted as such given the knowledge one has about what may or may not be defined as an abuse.

3.3 Recovered memories of trauma as a form of involuntary autobiographical memory

We argue that recovered memory of trauma relate to a broader concept than involuntary memory retrieval; namely, while typical memory of a past event is experienced and realized by a rememberer (i.e., one may recall sitting on someone's lap in a bedroom), an involuntary memory may become a recovered memory of trauma when it is interpreted and understood as an abuse. However, there are two questions to be asked; namely, (1) why an event was not remembered earlier so a person is surprised when it is involuntarily recalled for a first time and (2) why and how such a memory is eventually recalled at a given time? These questions relate to two broad threads of encoding (the first thread) and retrieval (the second thread). We briefly elaborate further on these two threads.

The first question regards the reasons why a given event was not remembered earlier so one might not have been aware of experiencing an event in the past until it was spontaneously recovered. For instance, it may be argued that the recovered experience at the time of occurrence was not easily and straightforwardly understood by a rememberer. This challenges the efficient coding and of the memory, making it more prone to be forgotten (if it was indeed remembered in the first place) and/or rendered unavailable (i.e., difficult to access). Thus, a given incident may not be fully processed and encoded within an autobiographical memory base and may therefore be less accessible via a top-down (generative) memory retrieval process. This makes such a memory difficult to be accessed as, in general, the better an event fits an already existing cognitive schema/script, the easier and better it is later remembered (e.g., Pillemer et al., 1994). This was also suggested by McNally and Geraerts (2009); namely, that some instances of childhood abuse may be experienced and encoded as confusing and bizarre, but may not be traumatic. Also, the less frequent the recovered experience is, the lower the likelihood is that the memory about that experience will be accessible and rehearsed. Repeated events of past abuse that follow the same/similar pattern may be better encoded and remembered even if not fully understood (but such an understanding of the situation may be reached over time). At the same time, (non-traumatic) events that happened once may be relatively easily forgotten (i.e., when not having the possibility to rehearse and elaborate on the event). Put differently, it may be relatively likely that a recovered memory of trauma will relate to a single event that might have not been sufficiently encoded (e.g., processed, elaborated, understood), and therefore remains unavailable for a controlled and voluntary retrieval. Yet, a memory could be still retrieved if only automatic processing of cues in the environment detected some conceptual or perceptual overlap with stored memory representations of such long-forgotten event. Therefore, one would not only expect involuntary memories to arise (as we observe in an every-day life) but some of these memories on relatively rare occasions may relate to the personal past that one was not fully aware of (which is also observed in an every-day life). For instance, one may not remember meeting a given person during the conference but then seeing that person may trigger that forgotten memory.

The second question relates to the reasons why a given memory was recently retrieved. A critical issue we develop here is the cue-dependency of involuntary memory retrieval, that is, involuntary memories may be triggered spontaneously by any type of cue even when the rememberer does not expect a given memory to be retrieved. For instance, there may be a higher likelihood of recovering a memory about some past experience if such experience was combined with some attention-catching, unique or focal cue. This is not to say that one has to be self-aware of that stimuli but that that stimuli may be encoded as a somewhat vivid element of that event. The more unique and distinctive the accompanying cue is, the higher the likelihood is that such cue will efficiently trigger a given memory (see for instance Berntsen et al., 2013). As mentioned above (and as suggested by Barzykowski and Moulin, 2023), as our cognitive system automatically matches, as quickly and effortlessly as possible, the contents of mental representations stored in memory with the current contents of perception/attention, a spontaneous retrieval of a past memory is more likely to occur, if the memory contains something unique, distinctive that may trigger that representation. For instance, a unique scent accompanying a given event that might have been encoded/memorized alongside, may thereafter trigger such memory when one is exposed to this scent again. Conversely, if there are no distinctive cues encoded with an event, it may decrease the likelihood of incidental involuntary retrieval. We elaborate on the issue of cues below.

The uniqueness of a cue is somewhat crucial as according to the principle of cue overload, it is most likely that a cue will match several past events. Berntsen (2009) proposed a mechanism of cue-item discriminability, defined as "*how easily a given cue isolates an item*" (Rubin, 1995, p. 151 as cited in Berntsen, 2009, p. 107). That is, the more events that are associated with a particular cue, the less efficient this cue will be in triggering any one of them. This was also demonstrated by Berntsen et al. (2013); namely, that involuntary episodic memories are retrieved more often in response to unique compared to repeated (i.e., associated with many memories) cues, which confirmed the principle of the cue-overload. Therefore, the more unique a cue accompanying the recovered experience is, the higher the likelihood is that such a cue may trigger abruptly and involuntarily a memory of that experience. This does not imply that a less unique cue cannot trigger a recovered memory of trauma. This is particularly evident since any type of cue has the potential to elicit a memory of that event, if only it was present during the recovered experience. For instance, priming processes (i.e., increasing the activation of a memory information by prior encounter with the contents of memory representation) may actually increase such a cue-item discriminability allowing for efficient activation of a particular involuntary memory (e.g., Mace, 2005; Barzykowski and Niedźwieńska, 2018b). It is also possible that on some occasions an environment/surrounding setting may consist of cues that map onto a given past memory or there is a relevant configural, contextual similarity between the current situation and a given past events. While these cues/contexts may not be efficient in triggering a given memory, it is nonetheless possible that some of them may increase the accessibility of that memory. In other words, over time, recurrent exposure to cues might lead to the memory being fully retrieved.

3.4 Conclusion

The idea that recovered memories of trauma may be a form of involuntary autobiographical memory suggests that they may reflect quite authentic events and therefore be very valid. This raises the question of the context of recovery, and more precisely of the variations in the different contexts of retrieval of recovered memories of trauma. Insofar as these may vary, then the validity of the memories may vary because these contexts may be more or less suggestive. Thus, we explore in the following section how the validity of memory reports can be examined in a legal framework. We propose that maintaining an opposition between false and true memories, or between valid and invalid testimonies might be counterproductive. We therefore present in the following section that memory reports in a legal framework should be examined on a continuum of validity, rather than in a category-based approach.

4 Testimony validity as a continuum

Recovered memories of trauma may reflect events that never happened, as well as memories of perfectly genuine events (McNally and Geraerts, 2009). While these two phenomena may be labeled as, respectively, "false" and "true" memories, we propose that the validity of eyewitness testimony should not be viewed solely through the lens of "true" vs. "false memories." The accuracy of a witness's report can be considered in a more balanced way, since, for example, the term false memory is more "a linguistic convenience" than a truly unified phenomenon (Bernstein et al., 2018, p. 161), and can in fact refer as much to an event that never happened, as to elements of an event that actually did happen (in which case, the quantity of false elements can widely vary from one individual to another, in the same individual, and from one event to another).

Another (while related) argument is that there is now a vast body of literature showing that there is no relationship between different forms of false memory. Specifically, sensitivity to one type of false memory (e.g., spontaneous false memories, elicited in particular through the DRM task; Deese, 1959; Roediger and McDermott, 1995; see below for more details) does not predict sensitivity to other types of false memory (e.g., misinformation effect, Loftus, 2005; Loftus and Klemfuss, 2023; or creation of rich false memories, via the lost-in-the-mall paradigm, Loftus and Pickrell, 1995). Bernstein et al. (2018) therefore suggested that different constructs may underlie the different types of false memories. Mazzoni (2002) proposed that false memories could be distinguished by their origin: suggestion-dependent false memories (i.e., misinformation effect, rich false memory creation), and false memories produced by the reconstructive nature of memory (e.g., DRM paradigm, schema-based false memories).

What can be drawn from those arguments is that there might be a variety of memories different in their nature and validity. As a result an eyewitness may be based on a recovered memory of trauma, which may reflect an event that never happened or an event that actually did happen. It may also be a continuous memory and have been more or less distorted by external suggestions and/or a natural (and internal) reconstruction process. We therefore believe that expert witnesses (e.g., forensic psychologists, memory experts, clinical psychologists) should assess the validity of eyewitness reports as a continuum, ranging from the least valid form (i.e., recovered memory of a traumatic event that never occurred) to the most valid form (i.e., involuntary autobiographical recovered memory of a traumatic event that did occurred) (see Figure 2).

4.1 The event did not occur at all

We consider that the least valid form of testimony based on a recovered memory of trauma is when the event did not occur at all. Typically, this type of memory is the result of suggestion by a third party, where several techniques have been used to recover (and, in this case, create) the memory. In the laboratory, the seminal study to highlight the possibility of individuals developing memories of events that never occurred was that of Loftus and Pickrell (1995), traditionally known as the "lost-in-the-mall" study. In this study, the experimenters provided participants with a booklet containing three true autobiographical accounts of their childhood, as well as a false autobiographical account according to which the participant, then a young child, had become lost in a shopping mall. This false event was constructed to be credible (i.e., including true biographical elements). On the basis of this booklet, participants were then asked to indicate what they remembered about each of the events, all presented as true, with the option of specifying that they had no recollection. Logically, in the first interview with the experimenter, the false event was given very little detail. Participants were then interviewed twice more, between which they were invited to think about the events, try to

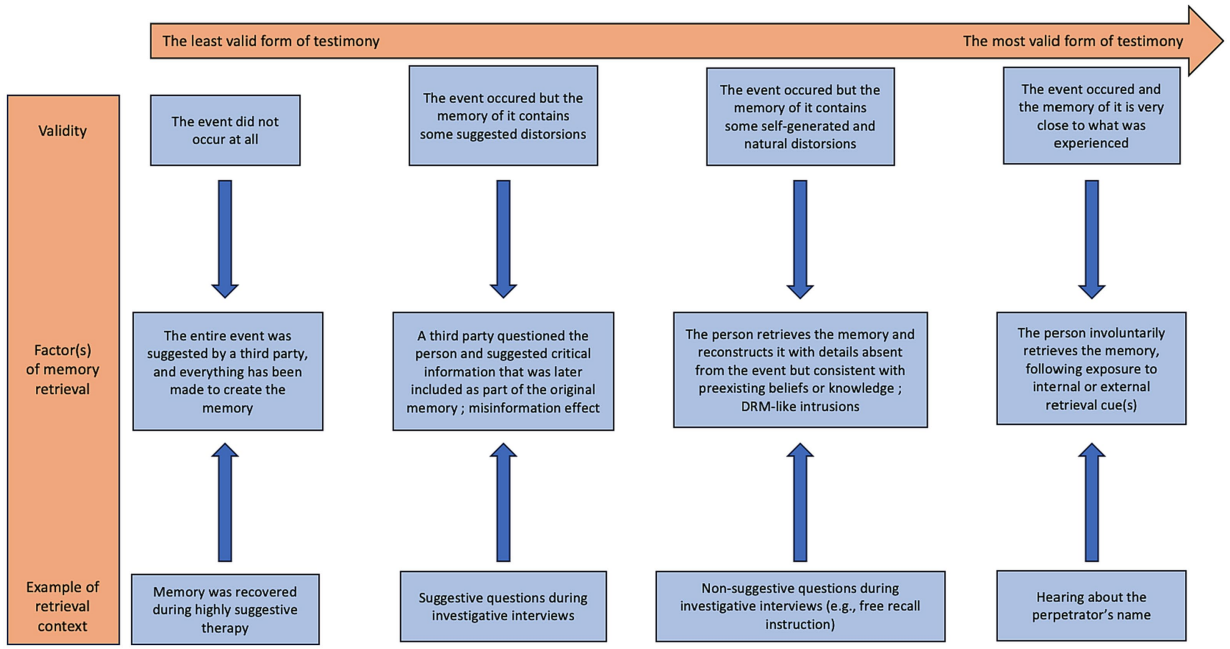


FIGURE 2
Continuum of the validity of memory reports in legal contexts.

remember them and imagine what might have happened. In the experiment, at the last interview, around 25% of participants partially or fully remembered the false event.

While this study has been criticized for its lack of ecological validity (e.g., the events in these studies do not equate with a traumatic event, Pezdek et al., 2006), and statistical issues (Crook and McEwen, 2019), there is evidence to support the conclusion that it is possible to induce detailed memories of events that never happened. Firstly, a very recent preregistered replication, with a larger number ($N = 123$) of participants than in the original study ($N = 24$), showed that 35% of participants recalled getting lost in a shopping mall during childhood, even though this event never occurred (Murphy et al., 2023). Secondly, a meta-analysis (i.e., raw data from several studies are aggregated for a broader analysis) showed that, in studies using this paradigm, around 30% of people came to remember and detail an event they had never experienced (Scoboria et al., 2017).

Also, a number of studies showed that highly suggestive methods could also lead people to recall such events that had never occurred, such as guided imagery (Hyman and Pentland, 1996), dream interpretation (Mazzoni et al., 1999), or hypnosis (Green et al., 1998). Interestingly, it appears that memories recovered in therapy are frequently the result of such methods (Dodier et al., 2019; Patihis and Pendergrast, 2019).

We consider this to be the least valid form of testimony, insofar as everything has been done to create a recovered memory of trauma without knowing or not if the experience of the event actually happened. Thus, by thinking of giving retrieval cues (or by deliberately trying to induce a false memory, see example below), an individual could then suggest a whole scenario to a person that they could endorse, even if it does not correspond to any personally experienced event. In the worst-case scenario, the event simply never happened, and the risk of a miscarriage of justice is greatly increased if no one in the judicial process properly assesses the validity of the testimony.

For example, in 2012, a French psychologist was convicted of moral abuse for deliberately suggesting memories of events that never actually occurred in two patients. It was established that these events had never been experienced by the two victims, as they were intrauterine memories of attempted abortions. By inducing the victims to develop such memories, the psychologist led them to break their family ties and make them dependent on his therapeutic care, guaranteeing him substantial income (for more information, see: https://www.lemonde.fr/societe/article/2012/04/13/le-proces-d-un-therapeute-accuse-d-inventer-de-faux-traumatismes-a-ses-patients_1684943_3224.html, article in French).

It should be noted, however, that these cases represent anecdotal evidence, and that recent work shows that cases of recovered memories in a therapeutic setting represent a minority of cases of recovered memories of traumatic events (Dodier and Patihis, 2021), and that this seems to be independent of the therapy type (Dodier et al., 2019; Patihis and Pendergrast, 2019). There is also anecdotal evidence that recovered memories of traumatic events during therapy may correspond to events that truly occurred (Schimmenti, 2017). This reinforces the need to conceive of the validity of testimonies as a continuum and to be able to explore the precise context in which memories are recovered to give an expert opinion.

4.2 The event occurred but the memory of it contains some suggested distortions

Progressing along the continuum, we arrive at memories relating to events that actually happened, but which contain details that never occurred in the event, and which were suggested by third parties.

When people incorporate into their memory information suggested by others (or by the media) after the event has taken place, this refers to the misinformation effect. Decades of research have demonstrated the robustness of this effect and the extent to which memories are sensitive to external factors and therefore can be reconstructed by including encoded post-event information (Loftus, 2005; Loftus and Klemfuss, 2023). This effect can occur with very explicit suggestions, where a person directly suggests the information that then could be included in the memory (e.g., "so he was holding a knife, right?"). However, it was shown in early work that a simple variation in the violent connotation of a car accident (i.e., use of the conjugated verbs "collided," "bumped," "contacted," "hit," or "smashed" to illustrate a car accident) could significantly alter the proportion of people falsely recalling seeing broken glass in the accident, when there was none (Loftus and Palmer, 1974).

The misinformation effect can arise from multiple sources such as co-witnesses (see, e.g., Hope et al., 2008; Jack et al., 2014), relatives, or police investigators (Loftus, 2005). The latter case is particularly sensitive, as investigators generally have information (or even assumptions) about the course of events and, without training in questioning techniques, can easily come to suggest information (e.g., Launay and Py, 2015; Verkampt et al., 2021).

We consider these memories to be more valid than those in which the event did not occur, because in this case (this will sound trivial), the event did occur, but in a different form to that represented in memory. However, distorted aspects are relative to assumptions made by others, and can sometimes even be the result of confirmation bias, where investigators already have a precise idea of how the event took place, and seek to confirm it by directing their questions (concerning child sexual abuse, see Zhang et al., 2022), and the probability is therefore high that these assumptions are related to critical information that can redirect an investigation. Thus, even without malicious intent, the memory may adopt a form desired by a third party, and thus its validity is highly questionable.

Take, for example, a situation where a person gives evidence of a bank robbery. During the interview, the police officer asks the witness "Was the robber carrying a weapon? A gun?" Regardless of the immediate answer given by the witness, the misinformation effect would be that, during a second interview, the person spontaneously reports that the robber had a firearm, when they did not. In this case, the person would have included in their memory of the event, even though they had really experienced it, the presence of a firearm which did not exist in the event.

4.3 The event occurred but the memory of it contains some self-generated and natural distortions

Since Bartlett's seminal work (Bartlett, 1932) showing how cognitive schemas can modify and adapt memories to bring them a certain coherence, it has been widely accepted that memory is reconstructive, and that in the absence of external suggestions, our memories are comprised of our personally lived experiences, but also of intrusions, often semantic. Work on the DRM paradigm (Deese, 1959; Roediger and McDermott, 1995), which has been widely replicated, confirms this natural aspect of reconstruction. In this paradigm, a list of words (e.g., dream, pillow, yawn, and tired, etc.) is presented and linked to a critical lure (e.g., sleep), which is absent from the list. Typically, during a recall or recognition task, the critical lure is frequently falsely recalled or misrecognized as part of the list.

These distortions correspond to the ordinary functioning of memory and have an adaptive nature, in the sense that they can enable memories to be enriched with relevant information, maintain a certain coherence and can make it easier to plan future events or solve problems (i.e., Schacter, 2012). Their causes are twofold: they result from the activation of knowledge networks at encoding, contributing to their more or less explicit encoding with the event, but also from difficulties in identifying the source of memory encoding (e.g., the event vs. one's own thoughts) (Roediger et al., 2001). Since they result from activations of knowledge networks during encoding, these intrusions are often consistent with the event experienced. While in everyday life, such errors do not seem to pose major problems, the impact is different in the forensic setting, where details can sometimes be of great importance (e.g., remembering the presence of a weapon, when there was none).

Despite this, it appears that these memories can be considered more valid than the memories previously described in this section insofar as (i) they are perfectly natural and they seem to concern everyone, including people with exceptional memory and autobiographical abilities (Paihis et al., 2013), (ii) unlike memories induced by external sources, it is extremely difficult to be able to estimate which details may be distortions or not in the absence of corroborating or contradicting evidence. Moreover, such intrusions may well occur when police interviews are conducted in an entirely appropriate way, without suggestions and with free recall tasks that can increase self-generated retrieval cues of erroneous information (e.g., see work on the cognitive interview, where an increase in errors is generally observed with this tool compared to a control tool; Memon et al., 2010).

It should also be noted that this phenomenon is dependent on internal factors, such as age, a phenomenon known as developmental reversal, according to which children are less sensitive to this type of distortion than adults (and therefore make fewer recall or recognition errors in a DRM task), because their knowledge networks are less mature (Howe et al., 2011). In adults, we also generally observe an increase in these distortions with age (e.g., Colombel et al., 2016).

Consider the example in the previous section, where a witness describes the robbery that they experienced. Assume that the police officer does not suggest the presence of a firearm. Because the "gun" information is semantically consistent with the "robbery" event, it would be possible for the witness to generate and include this information in their memory of the event, which they would have truly experienced, without any external influence. In this case, their prior knowledge (e.g., script) could have led to the spontaneous intrusion of the "firearm" information into the "robbery" event. In the same way, the person could also have included other information semantically linked to the concept of "robbery," such as the fact that the robber was wearing a mask, or that they said certain words such as "nobody moves," etc.

In the case of natural distortions (or suggested distortions, see previous section), the validity of testimonies is greatly reduced because distortions can be very relevant to the investigation and thus redirect it completely (e.g., suggesting a physical characteristic of a suspect, when the perpetrator is someone else) and/or be very consistent with the event, and therefore credible, and potentially aggravate the situation (e.g., presence of a gun when there was none).

4.4 The event occurred and the memory of it is very similar to what was experienced

While there are several methods used to study involuntary autobiographical memories (e.g., semi-structuralized diary methods, questionnaires), a typical laboratory-based method (originally developed by Schlagman and Kvavilashvili, 2008) is built on the observation that involuntary memories are most frequently triggered by easily identifiable cues present in the closes surrounding (most of which are verbal-type, e.g., heard/read words). Thus, in this method participants are engaged in a monotonous vigilance task (detecting seldomly occurring pattern of vertical lines) while exposed to irrelevant word phrases (e.g., birthday party and upsetting conversation etc.) some of which may incidentally trigger involuntary memories. Importantly, participants are also instructed to write down any spontaneously occurring thought and/or memories during performance of the vigilance task. Such laboratory-based procedure provided in recent years experimental way to investigate involuntary memory retrieval under well-controlled conditions. As mentioned earlier, an interesting consequence of involuntary autobiographical memory retrieval, as described in the previous section, is that such memories are activated at the bottom level of the hierarchy by different internal and external cues circumventing the search process (also schemas and cognitive scripts), they may be considered as more valid. For instance, Barzykowski and colleagues (e.g., Barzykowski and Staugaard (2016, 2018; also, Barzykowski et al., 2019a) demonstrated in a well-controlled experimental condition that while some memories may require very little reconstruction and other memories rely more heavily on reconstructive efforts, reconstruction may not be required specifically for memories that are involuntarily retrieved. Therefore, we argue that involuntary autobiographical memories, following exposure to an internal or external retrieval cue, may be most likely to reflect the event as it occurred.

Put differently, it may be that a recovered memory of trauma retrieved in response to an external/internal cue directly relating to something presented during the occurrence of an original event may be the closest approximation (especially when compared to voluntary retrieval) of something that occurred in the past. For instance, Staugaard and Berntsen (2019) provided evidence demonstrating that indeed over time memories for past events become more cue-dependent which may eventually hamper strategic and voluntary retrieval at longer delays. Interestingly, they also showed a somewhat steeper slope of the forgetting curve in the voluntary compared to involuntary memory retrieval (Staugaard and Berntsen, 2019). When discussing their findings, they also suggest the plausible link between recovered memories and involuntary memories as follows (p. 903–904): "*The unexpected activation of dormant memories in response to situational cues is also consistent with some observations in clinical psychology of "recovered memories" of childhood trauma (see, e.g.,* Conway, 1997; Read and Lindsay, 1997, *for reviews). Although the notion of recovered memories is contentious, and although the majority of such recovered memories appear to have been brought about through strategic retrieval attempts in the course of psychotherapy* (*e.g.,* Geraerts et al., 2007), *there are some examples of recovered memories outside of therapeutic settings in response to situational cues (e.g.,* Bendiksen, 1997*), which might be conceptualized as involuntary memories of forgotten events. However, the fact that the present studies used laboratory material without the personal significance and levels of complexity associated with real-life events renders these possibilities highly tentative and speculative*." We fully agree with the authors on this idea, and we also call for more studies on the possible mechanisms underlying such resurfacing of the past events in every day context. As for now this issue and its implications for memory accuracy still requires a robust and thorough discussion. In all cases, we do not argue though that such a memory may never be distorted but that there may be a high likelihood that such memory actually happened, especially if it relates to an event with a distinctive and noticeable cue (see also below).

The presented idea that recovered memories of trauma based on involuntary autobiographical memories may be highly valid is based on the assumption that a cue serves as a way to access a memory information that already exists within the autobiographical memory base in a form that it was encoded rather than launches a reconstructive memory retrieval process. Thus, it is rather unlikely that a cue can readily access information that was not previously presented (encoded or experienced) within the memory system, at least under typical circumstances. However, if only previously suggested, imagined, elaborated, that is, in any way "artificially created," a cue may also trigger such a memory representation that although existing in the memory system remains rather false.

For example, a person who spontaneously recalls, in detail and without any effort at retrieval, having been subjected to violence during their adolescence by their neighbor, on hearing his name during a discussion, could have a memory, at the time of retrieval, in the version closest to what they have encoded. For this reason, it is crucial for a testimony to be collected promptly and in an appropriate manner (i.e., free from suggestion and based on free recall) following the recovery of a memory in such a context. This helps minimize the risks of contamination or excessive reconstructions resulting from numerous retrieval efforts.

Thus, a proper analysis of the context of memory retrieval (e.g., unexpectedly retrieved for the very first time and not elaborated, developed in any way previously) and memory content *per se* (e.g., the correspondence between triggering cue and the memory content) may help in evaluating the memory's validity. We elaborate on this idea in the last section of the article.

Finally, it is also important to highlight and explicitly stipulate that while describing the continuum, we refer to "*the most valid*" memory. However, we do not imply that this memory is 100% accurate or it is not susceptible to distortion, fading, false details or any other mechanism(s) underlying memory erroneous retrieval. We rather would like to argue that such memory may be, in some cases, the closest approximation (compared to voluntary retrieval) to the representation of the original event and the way it was initially encoded. While we present and further develop such an idea, we are fully aware that future studies and more empirical data are needed.

5 Brief recommendations for practice and concluding remarks

Although for reasons of clarity and practicality we may give the impression that we have identified four types of testimony, it is more appropriate to consider them as dimensions that can be placed on a continuum (see Figure 2). Each of these dimensions may overlap with others, and the aim is that expert witnesses assess how valid a testimony is, rather than identifying to which category it belongs. For example, a memory may have been induced in therapy, but concern only an entire part of an event that actually took place in the first place

(e.g., a family meal that actually took place, but where the subsequent acts of sexual abuse are totally suggested by a therapist). It is also quite likely that a memory will include details suggested by an investigator, as well as self-generated and natural distortions. As a final example, a memory of childhood abuse may be recovered spontaneously following exposure to an environmental cue, but may then have been recalled repeatedly with suggestive questions by, for example, family members, friends, and/or investigators.

Most existing tools for discriminating between true and false testimonies have theoretical and practical limitations. For example, the *Criteria-Based Content Analysis* (Steller and Koehnken, 1989; see also Volbert and Steller, 2014 for an update), one of the four parts of the Statement Validity Analysis (Granhag et al., 2015) has focused most of its work on detecting lies and deception, based on the idea that an invented narrative will be qualitatively different from one based on a real memory. Limitations in the tool's validity have been pointed out (Rassin, 2000). When applied to the distinction between true and false memory, it appears that this tool is of limited use (Kulkofsky, 2008; Volbert and Steller, 2014). Another example is the *Reality Monitoring* (RM) framework, which aims to distinguish between internally and externally generated memories (Johnson and Raye, 1981). However, the ambiguity of the definitions of what is internally generated and what is externally generated was highlighted very early on (Johnson et al., 1993). Similarly, the aim was rather to detect deception or lies. Here again, the usefulness of RM in distinguishing between true and false memories seems quite limited (Otgaar et al., 2010). In any case, such tools aimed at identifying markers of truthfulness in testimonies serve to categorize testimonies as true or false, and in our view represent a reductive view of the validity of testimonies.

The role of expert witnesses when evaluating eyewitness testimonies is therefore to be able to weigh the extent to which the event contains, for example, suggested information, in order to place the testimony on the overall dimension of validity. This requires a precise, rigorous and thorough assessment of the context in which a memory has emerged. The following questions can be asked when assessing the validity of a memory report: Is the memory retrieved? If so, how? As a result of discussions? If so, with whom? In what form? After how long a discussion? If not, spontaneously? Under what circumstances? What was the likely cue that activated the memory and led to its retrieval? Etc. If the testimony stems from a continuous memory (i.e., one that has not been suddenly recovered), has the person discussed it with others? If so, who? Were any questions put to the person during these discussions? Which questions? How did the investigators gather their testimony? How many times has the same event been recalled? Over what period of time? Etc.

It is necessary to explore all the elements in a case file in order to make a critical assessment of the retrieval contexts (i) of the memory in general, (ii) if possible, of all the critical information. In this way, expert witnesses will be able to better assess the validity of the testimony in detail, and not by relying on a binary conception of "true vs. false memory," which may certainly simplify the understanding of triers of facts, but which probably reduces too much the form that memories can take, and the underlying mechanisms that enable them to be created. This is a micro-level analysis which, as we believe, can be integrated into more general methods that have been proposed by memory scholars to provide expert reports that are as immune as possible to bias (e.g., Otgaar et al., 2017; Vredeveldt et al., 2022; Arbiyah et al., 2023).

Of course, this work requires the intervention of memory experts with extensive and precise knowledge of how memory works (see Magnussen and Melinder, 2012). If a legal expert does not have this knowledge, it would seem necessary to call upon a memory expert to assist on these critical issues of eyewitness testimony (see fuller arguments in Dodier, 2018; Dodier et al., 2023). However, non-memory experts may sometimes have to give their opinion on memory phenomena, either because memory experts do not exist in the legal system (e.g., this is the case of France's legal system, see Dodier, 2018; Dodier et al., 2023), or because memory phenomena are not at the heart of a legal case. In this respect, we advocate the idea that precise evidence-based guidelines and tools should be constructed and designed for use by expert witnesses (forensic or clinical) who are not memory specialists. In other words, memory specialists should work to create tools that can be used and adapted to the level of expertise of psychologists who carry out forensic examinations.

Although it has been shown that false memories are a "linguistic convenience" (Bernstein et al., 2018, p. 161), that is, the different types of false memories do not correlate particularly well, due to different and specific underlying processes, we find it crucial to examine future research aimed at clearly defining the processual differences behind false memories, but also the overlapping processes and mechanisms. For example, source-monitoring is generally invoked to explain natural distortions of the DRM type, but also the misinformation effect. Precisely identifying the limits and overlaps would make it possible to refine the model of testimony validity that we are advocating. Also, field studies using corroborated versus uncorroborated testimonies (e.g., Geraerts et al., 2007) could provide some external validity to the model.

We hope that this article will continue to bring balance to the debate on recovered memories of trauma, and that it will provide valuable resources for expert witnesses who must give their opinion in court cases based on eyewitness testimony, whether or not they include such recovered memories. Indeed, we can foresee that our contribution could be applied to contexts broader than this mere issue. If the place we give to recovered memories allows us to propose a model of testimony validity, it appears that, as we have just developed, its practical relevance can be found in all contexts of assessment of the validity of testimonies, whether the memories are continuous or recovered.

Author contributions

OD: Writing – original draft. KB: Writing – original draft. CS: Funding acquisition, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. While writing the paper KB was supported by a grant from the National Science Centre, Poland (2019/35/B/HS6/00528). We acknowledge the financial support of Grenoble Alpes University for this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

American Psychiatric Association (2022). *Diagnostic and statistical manual of mental disorders*. 5th Edn. American Psychiatric Association.

Arbiyah, N., Otgaar, H., Sauerland, M., Rassin, E., Maegherman, E., and Merckelbach, H. (2023). The use of alternative scenarios in assessing the reliability of victims' statements. *Psychol. Crime Law*, 1–18. doi: 10.1080/1068316X.2023.2236274

Baird, B., Smallwood, J., Fishman, D. J. F., Mrazek, M. D., and Schooler, J. W. (2013). Unnoticed intrusions: dissociations of meta-consciousness in thought suppression. *Conscious. Cogn.* 22, 1003–1012. doi: 10.1016/j.concog.2013.06.009

Bartlett, F. C. (1932). *Remembering: a study in experimental and social psychology*. Cambridge University Press.

Barzykowski, K., and Mazzoni, G. (2022). Do intuitive ideas of the qualities that should characterize involuntary and voluntary memories affect their classification? *Psychol. Res.* 86, 170–195. doi: 10.1007/s00426-020-01465-3

Barzykowski, K., and Moulin, C. (2023). Are involuntary autobiographical memory and déjà vu natural products of memory retrieval? *Behav. Brain Sci.*, 1–66. doi: 10.1017/S0140525X22002035

Barzykowski, K., and Niedźwieńska, A. (2018a). Involuntary autobiographical memories are relatively more often reported during high cognitive load tasks. *Acta Psychol.* 182, 119–128. doi: 10.1016/j.actpsy.2017.11.014

Barzykowski, K., and Niedźwieńska, A. (2018b). Priming involuntary autobiographical memories in the lab. *Memory* 26, 277–289. doi: 10.1080/09658211.2017.1353102

Barzykowski, K., Niedźwieńska, A., and Mazzoni, G. (2019a). How intention to retrieve a memory and expectation that it will happen influence retrieval of autobiographical memories. *Conscious. Cogn.* 72, 31–48. doi: 10.1016/j.concog.2019.03.011

Barzykowski, K., Radel, R., Niedźwieńska, A., and Kvavilashvili, L. (2019b). Why are we not flooded by involuntary thoughts about past and future? Testing the cognitive inhibition dependency hypothesis. *Psychol. Res.* 83, 666–683. doi: 10.1007/s00426-018-1120-6

Barzykowski, K., and Staugaard, S. R. (2016). Does retrieval intentionality really matter? Similarities and differences between involuntary memories and directly and generatively retrieved voluntary memories. *Br. J. Psychol.* 107, 519–536. doi: 10.1111/bjop.12160

Barzykowski, K., and Staugaard, S. R. (2018). How intention and monitoring your thoughts influence characteristics of autobiographical memories. *Br. J. Psychol.* 109, 321–340. doi: 10.1111/bjop.12259

Bass, E., and Davis, L. (1988). *The courage to heal: a guide for women survivors of child sexual abuse*. Perennial Library/Harper & Row Publishers.

Battista, F., Mangiulli, I., Patihis, L., Dodier, O., Curci, A., Lanciano, T., et al. (2023). A scientometric and descriptive review on the debate about repressed memories and traumatic forgetting. *J. Anxiety Disord.* 97:102733. doi: 10.1016/j.janxdis.2023.102733

Bendiksen, M. (1997). Circumstances and the phenomenology of a recovered memory: A corroborated case study. In D. S. Lindsay and J. D. Read (Eds.). *Recollections of trauma: Scientific evidence and clinical practice (NATO ASI Series*, p. 566). Plenum.

Bernstein, D. M., Scoboria, A., Desjarlais, L., and Soucie, K. (2018). "False memory" is a linguistic convenience. *Psychol. Conscious. Theory Res. Pract.* 5, 161–179. doi: 10.1037/cns0000148

Berntsen, D. (1996). Involuntary autobiographical memories. *Appl. Cogn. Psychol.* 10, 435–454. doi: 10.1002/(SICI)1099-0720(199610)10:5<435::AIDACP408>3.0.CO;2-L

Berntsen, D. (1998). Voluntary and involuntary access to autobiographical memory. *Memory* 6, 113–141. doi: 10.1080/741942071

Berntsen, D. (2009). *Involuntary autobiographical memories*. An introduction to the unbidden past. Cambridge University Press.

Berntsen, D. (2010). The unbidden past: involuntary autobiographical memories as a basic mode of remembering. *Curr. Dir. Psychol. Sci.* 19, 138–142. doi: 10.1177/0963721410370301

Berntsen, D., Staugaard, S. R., and Sørensen, L. M. T. (2013). Why am i remembering this now? Predicting the occurrence of involuntary (spontaneous) episodic memories. *J. Exp. Psychol. Gen.* 142, 426–444. doi: 10.1037/a0029128

Chechko, N., Stickel, S., Kellermann, T., Kirner, A., Habel, U., Fernandez, G., et al. (2018). Progressively analogous evidence of covert face recognition from functional magnetic resonance imaging and skin conductance responses studies involving a patient with dissociative amnesia. *Eur. J. Neurosci.* 48, 1964–1975. doi: 10.1111/ejn.14087

Clancy, S. A., and McNally, R. J. (2005/2006). Who needs repression? Normal memory processes can explain "forgetting" of childhood sexual abuse. *Sci. Rev. Ment. Health Pract.* 4, 66–73.

Colombel, F., Tessoulin, M., Gilet, A.-L., and Corson, Y. (2016). False memories and normal aging: links between inhibitory capacities and monitoring processes. *Psychol. Aging* 31, 239–248. doi: 10.1037/pag0000086

Comblain, C., D'Argembeau, A., and Van der Linden, M. (2005). Phenomenal characteristics of autobiographical memories for emotional and neutral events in older and younger adults. *Exp. Aging Res.* 31, 173–189. doi: 10.1080/03610730590915010

Conway, M. A. (1997). Past and present: Recovered memories and false memories. In M. A. Conway (Ed.), *Recovered memories and false memories* (pp. 150–191). Oxford University Press. doi: 10.1093/med:psych/9780198523864.003.0007

Conway, M. A. (2001). Sensory-perceptual episodic memory and its context: autobiographical memory. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 356, 1375–1384. doi: 10.1098/rstb.2001.0940

Conway, M. A. (2008). "Exploring episodic memory" in *Handbook of episodic memory*. eds. E. Dere, A. Easton, L. Nadel and J. P. Huston, *vol. 18* (Elsevier), 19–29.

Conway, M. A. (2009). Episodic memories. *Neuropsychologia* 47, 2305–2313. doi: 10.1016/j.neuropsychologia.2009.02.003

Conway, M. A., and Jobson, L. (2012). "On the nature of autobiographical memory" in *Understanding autobiographical memory. Theories and approaches*. eds. D. Berntsen and D. C. Rubin (Cambridge University Press), 54–69.

Conway, M. A., and Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychol. Rev.* 107, 261–288. doi: 10.1037/0033-295X.107.2.261

Crews, F. (1995). *The memory wars: Freud's legacy in dispute*. The New York Review of Books.

Crook, L. S., and McEwen, L. E. (2019). Deconstructing the lost in the mall study. *J. Child Custody* 16, 7–19. doi: 10.1080/15379418.2019.1601603

D'Argembeau, A., Comblain, C., and Van der Linden, M. (2003). Phenomenal characteristics of autobiographical memories for positive, negative, and neutral events. *Appl. Cogn. Psychol.* 17, 281–294. doi: 10.1002/acp.856

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *J. Exp. Psychol.* 58, 17–22. doi: 10.1037/h0046671

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., and McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law Hum. Behav.* 28, 687–706. doi: 10.1007/s10979-004-0565-x

Dimitrova, L. I., Dean, S. L., Schlumpf, Y. R., Vissia, E. M., Nijenhuis, E. R. S., Chatzi, V., et al. (2021). A neurostructural biomarker of dissociative amnesia: a hippocampal study in dissociative identity disorder. *Psychol. Med.* 53, 805–813. doi: 10.1017/S0033291721002154

Dodier, O. (2018). The need for memory experts in French courts. *J. Forensic Psychol. Res. Pract.* 18, 158–176. doi: 10.1080/24732850.2018.1444912

Dodier, O. (2019). A bibliometric analysis of the recovered-memory controversy in the 21st Century. *Applied Cognitive Psychology* 33, 571–584. doi: 10.1002/acp.3498

Dodier, O., Morlet-Rivelli, M., and Verkampt, F. (2023). France lags behind in the application of memory science in the courtroom. *J. Appl. Res. Mem. Cogn.* 12, 203–207. doi: 10.1037/mac0000088

Dodier, O., and Patihis, L. (2021). Recovered memories of child abuse outside of therapy. *Appl. Cogn. Psychol.* 35, 538–547. doi: 10.1002/acp.3783

Dodier, O., Patihis, L., and Payoux, M. (2019). Reports of recovered memories of childhood abuse in therapy in France. *Memory* 27, 1283–1298. doi: 10.1080/09658211.2019.1652654

Fivush, R., and Hamond, N. R. (1990). "Autobiographical memory across the preschool years: toward reconceptualising childhood amnesia" in *Knowing and remembering in young children*. eds. R. Fivush and J. A. Hudson (Cambridge: Cambridge University Press), 223–248.

Geraerts, E., Schooler, J. W., Merckelbach, H., Jelicic, M., Hauer, B. J. A., and Ambadar, Z. (2007). The reality of recovered memories: corroborating continuous and discontinuous memories of childhood sexual abuse. *Psychol. Sci.* 18, 564–568. doi: 10.1111/j.1467-9280.2007.01940.x

Granhag, P. A., Vrij, A., and Verschuere, B. (Eds.) (2015). *Detecting deception: current challenges and cognitive approaches* Wiley-Blackwell.

Green, J. P., Lynn, S. J., and Malinoski, P. (1998). Hypnotic pseudomemories, prehypnotic warnings, and malleability of suggested memories. *Appl. Cogn. Psychol.* 12, 431–444. doi: 10.1002/(SICI)1099-0720(199810)12:5<431::AID-ACP516>3.0.CO;2-R

Holmes, D. S. (1994). Is there evidence of repression? Doubtful. *Harv. Ment. Health Lett.* 10, 4–6.

Hope, L., Ost, J., Gabbert, F., Healey, S., and Lenton, E. (2008). "With a little help from my friends…": the role of co-witness relationship in susceptibility to misinformation. *Acta Psychol.* 127, 476–484. doi: 10.1016/j.actpsy.2007.08.010

Howe, M. L., Wimmer, M. C., Gagnon, N., and Plumpton, S. (2011). An associative-activation theory of children's and adults' memory illusions. *J. Mem. Lang.* 60, 229–251. doi: 10.1016/j.jml.2008.10.002

Hyman, I. E. Jr., and Pentland, J. (1996). The role of mental imagery in the creation of false childhood memories. *J. Mem. Lang.* 35, 101–117. doi: 10.1006/jmla.1996.0006

Jack, F., Zydervelt, S., and Zajac, R. (2014). Are co-witness special? Comparing the influence of co-witness and interviewer misinformation on eyewitness reports. *Memory* 22, 243–255. doi: 10.1080/09658211.2013.778291

Johnson, M. K., Hashtroudi, S., and Lindsay, D. S. (1993). Source monitoring. *Psychol. Bull.* 114, 3–28. doi: 10.1037/0033-2909.114.1.3

Johnson, M. K., and Raye, C. L. (1981). Reality monitoring. *Psychol. Rev.* 88, 67–85. doi: 10.1037/0033-295X.88.1.67

Kaplan, R., and Manicavasagar, V. (2001). Is there a false memory syndrome? A review of three cases. *Compr. Psychiatry* 42, 342–348. doi: 10.1053/comp.2001.24588

Kloft, L., Monds, L. A., Blokland, A., Ramaekers, J. G., and Otgaar, H. (2021). Hazy memories in the courtroom: a review of alcohol and other drug effects on false memory and suggestibility. *Neurosci. Biobehav. Rev.* 124, 291–307. doi: 10.1016/j.neubiorev.2021.02.012

Kulkofsky, S. (2008). "Credible but inaccurate: can criterion-based content analysis (CBCA) distinguish true and false memories?" in *Child sexual abuse: Issues and challenges*. ed. M. J. Smith (Nova Science Publishers), 21–42.

Launay, C., and Py, J. (2015). Methods and aims of investigative interviewing of adult witnesses: an analysis of professional practices. *Prat. Psychol.* 21, 55–70. doi: 10.1016/j.prps.2014.11.001

Leichtman, M. D., and Ceci, S. J. (1993). "The problem of infantile amnesia: lessons from fuzzytrace theory" in *Emerging themes in cognitive development*. eds. M. L. Howe and R. Pasnak, *vol. 1* (Foundations New York: Springer), 195–213.

Lilienfeld, S. O., and Loftus, E. F. (1999). A step backward in the recovered memory debate. *Professional Psychology: Research and Practice*, 30, 623. doi: 10.1037/0735-7028.30.6.623.b

Loftus, E. F. (1993). The reality of repressed memories. *Am. Psychol.* 48, 518–537. doi: 10.1037/0003-066X.48.5.518

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

Loftus, E. G., and Klemfuss, J. Z. (2023). Misinformation – past, present, and future. *Psychol. Crime Law*, 1–7. doi: 10.1080/1068316X.2023.2219813

Loftus, E. F., and Palmer, J. C. (1974). Reconstruction of automobile destruction: an example of the interaction between language and memory. *J. Verbal Learn. Verbal Behav.* 13, 585–589. doi: 10.1016/S0022-5371(74)80011-3

Loftus, E. F., and Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals*, 25, 720–725. doi: 10.3928/0048-5713-19951201-07

Mace, J. H. (2005). Priming involuntary autobiographical memories. *Memory* 13, 874–884. doi: 10.1080/09658210444000485

Magnussen, S., and Melinder, A. (2012). What psychologists know and believe about memory: A survey of practitioners. *Applied Cognitive Psychology*, 26, 54–60. doi: 10.1002/acp.1795

Markowitsch, H. J., and Staniloiu, A. (2013). The impairment of recollection in functional amnesic states. *Cortex* 49, 1494–1510. doi: 10.1016/j.cortex.2012.05.020

Mazzoni, G. (2002). Naturally occurring and suggestion-dependent memory distortions: the convergence of disparate research traditions. *Eur. Psychol.* 7, 17–30. doi: 10.1027//1016-9040.7.1.17

Mazzoni, G. A. L., Lombardo, P., Malvagia, S., and Loftus, E. F. (1999). Dream interpretation and false beliefs. *Prof. Psychol. Res. Pract.* 30, 45–50. doi: 10.1037/0735-7028.30.1.45

Mazzoni, G., Scoboria, A., and Harvey, L. (2010). Nonbelieved memories. *Psychol. Sci.* 21, 1334–1340. doi: 10.1177/0956797610379865

McNally, R. J. (2023). The return of repression? Evidence from cognitive psychology. *Top. Cogn. Sci.* Advance online publication. doi: 10.1111/tops.12638

McNally, R. J., and Geraerts, E. (2009). A new solution to the recovered memory debate. *Perspect. Psychol. Sci.* 4, 126–134. doi: 10.1111/j.1745-6924.2009.01112

Memon, A., Meissner, C. A., and Fraser, J. (2010). The cognitive interview: a meta-analytic review and study space analysis of the past 25 years. *Psychol. Public Policy Law* 16, 340–372. doi: 10.1037/a0020518

Miller, G. A. (1962/1974). *Psychology: the science of mental life* Penguin.

Murphy, G., Dawson, C. A., Huston, C., Ballantyne, L., Barrett, E., et al. (2023). Lost in the mall again: a preregistered replication and extension of Loftus & Pickrell (1995). *Memory* 31, 818–830. doi: 10.1080/09658211.2023.2198327

Murre, J. M. J., and Dros, J. (2015). Replication and analysis of Ebbinghaus' forgetting curve. *PLoS One* 10:e0120644. doi: 10.1371/journal.pone.0120644

Nelson, K. (1993). The psychological and social origins of autobiographical memory. *Psychol. Sci.* 4, 7–14. doi: 10.1111/j.1467-9280.1993.tb00548.x

Ost, J. (2006). "Recovered memories" in *Investigative interviewing: rights, research, regulation*. ed. T. Williamson (Willan Publishing), 259–291.

Otgaar, H., Candel, I., Memon, A., and Almerigogna, J. (2010). Differentiating between children's true and false memories using reality monitoring criteria. *Psychol. Crime Law* 16, 555–566. doi: 10.1080/10683160902926166

Otgaar, H., Curci, A., Mangiulli, I., Battista, F., Rizzotti, E., and Sartori, G. (2022). A court ruled case on therapy-induced false memories. *J. Forensic Sci.* 67, 2122–2129. doi: 10.1111/1556-4029.15073

Otgaar, H., de Ruiter, C., Howe, M. L., Hoetmer, L., and van Reekum, P. (2017). A case concerning children's false memories of abuse: recommendations regarding expert witness work. *Psychiatry Psychol. Law* 24, 365–378. doi: 10.1080/13218719.2016.1230924

Otgaar, H., Dodier, O., Garry, M., Howe, M. L., Loftus, E. F., and Lynn, S. J. (2023). Oversimplifications and misrepresentations in the repressed memory debate: A reply to Ross. *Journal of Child Sexual Abuse*, 32, 116–126. doi: 10.1080/10538712.2022.2133043

Otgaar, H., Howe, M., Dodier, O., Lilienfeld, S. O., Loftus, E. F., Lynn, S. J., et al. (2021). Belief in unconscious repressed memory persists. *Perspect. Psychol. Sci.* 16, 454–460. doi: 10.1177/1745691621990628

Otgaar, H., Howe, M. L., Patihis, L., Mangiulli, I., Dodier, O., Huntjens, R., et al. (in press). The neuroscience of dissociative amnesia and repressed memory: premature conclusions and unanswered questions. *Leg. Criminol. Psychol.*

Otgaar, H., Howe, M. L., Patihis, L., Merckelbach, H., Lynn, S. J., Lilienfeld, S. O., et al. (2019). The return of the repressed: the persistent and problematic claims of long-forgotten trauma. *Perspect. Psychol. Sci.* 14, 1072–1095. doi: 10.1177/1745691619862306

Paihis, L., Frenda, S., LePort, A. K. R., and Loftus, E. F. (2013). False memories in highly superior autobiographical memory individuals. *Proc. Natl. Acad. Sci.* 110, 20947–20952. doi: 10.1073/pnas.1314373110

Patihis, L., Otgaar, H., and Merckelbach, H. (2019). Expert witnesses, dissociative amnesia, and extraordinary remembering: response to brand et al. *Psychol. Inj. Law* 12, 281–285. doi: 10.1007/s12207-019-09348-8

Patihis, L., and Pendergrast, M. H. (2019). Reports of recovered memories of abuse in therapy in a large age-representative U.S. national sample: therapy type and decade comparisons. Clinical. *Psychol. Sci.* 7, 3–21. doi: 10.1177/2167702618773315

Pezdek, K., Blandon-Gitlin, I., Lam, S., Hart, R. E., and Schooler, J. W. (2006). Is knowing believing? The role of event plausibility and background knowledge in planting false beliefs about the personal past. *Mem. Cogn.* 34, 1628–1635. doi: 10.3758/BF03195925

Pillemer, D. B., Picariello, M. L., and Pruett, J. C. (1994). Very long-term memories of a salient preschool event. *Appl. Cogn. Psychol.* 8, 95–106. doi: 10.1002/acp.2350080202

Pope, H. G., Schnabel, J., and Hudson, J. I. (2023). Current scientific interest in dissociative amnesia: a bibliometric analysis. *Appl. Cogn. Psychol.* 37, 42–51. doi: 10.1002/acp.4021

Rasmussen, A. S., and Berntsen, D. (2011). The unpredictable past: spontaneous autobiographical memories outnumber autobiographical memories retrieved strategically. *Conscious. Cogn.* 20, 1842–1846. doi: 10.1016/j.concog.2011.07.010

Rassin, E. (2000). Criteria based content analysis: the less scientific road to truth. *Expert. Evid.* 7, 265–278. doi: 10.1023/A:1016627527082

Read, J. D., and Lindsay, D. S. (1997). *Recollections of trauma: Scientific evidence and clinical practice*. Plenum.

Roberts, P., McGinnis, D., and Bladt, L. (1994). The unexpected excursion: passive memories in everyday life. Poster presented at the 3rd practical aspects of memory conference, University of Maryland.

Roediger, H. L., and McDermott, K. B. (1995). Creating false memories: remembering words not presented in lists. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 803–814. doi: 10.1037/0278-7393.21.4.803

Roediger, H. L., Watson, J. M., McDermott, K. B., and Gallo, D. A. (2001). Factors that determine false recall: a multiple regression analysis. *Psychon. Bull. Rev.* 8, 385–407. doi: 10.3758/BF03196177

Rubin, D. C. (1995). *Memory in oral traditions. The cognitive psychology of epic, ballads, and counting-out rhymes* Oxford University Press.

Schacter, D. L. (2012). Adaptive constructive processes and the future of memory. *Am. Psychol.* 67, 603–613. doi: 10.1037/a0029869

Schimmenti, A. (2017). Elena: a case of dissociative identity disorder from the 1920s. *Bull. Meninger Clin.* 81, 281–298. doi: 10.1521/bumc_2017_81_08

Schlagman, S., and Kvavilashvili, L. (2008). Involuntary autobiographical memories in and outside the laboratory: how different are they from voluntary autobiographical memories? *Mem. Cogn.* 36, 920–932. doi: 10.3758/MC.36.5.920

Schlagman, S., Kvavilashvili, L., and Schultz, J. (2007). "Effects of age on involuntary au- tobiographical memories" in *Involuntary memory*. ed. J. H. Mace (Malden, MA: Blackwell), 1–19.

Scoboria, A., Wade, K. A., Lindsay, D. S., Azad, T., Strange, D., Ost, J., et al. (2017). A mega-analysis of memory reports from eight peer-reviewed false memory implantation studies. *Memory* 25, 146–163. doi: 10.1080/09658211.2016.1260747

Segura, I. A., McGhee, J., Della Sala, S., Cowan, N., and Pompéia, S. (2021). A reappraisal of acute doses of benzodiazepines as a model of anterograde amnesia. *Hum. Psychopharmacol. Clin. Exp.* 36:e2774. doi: 10.1002/hup.2774

Shaw, J., and Porter, S. (2015). Constructing rich false memories of committing crime. *Psychol. Sci.* 26, 291–301. doi: 10.1177/0956797614562862

Staugaard, S. R., and Berntsen, D. (2019). Retrieval intentionality and forgetting: how retention time and cue distinctiveness affect involuntary and voluntary retrieval of episodic memories. *Mem. Cogn.* 47, 893–905. doi: 10.3758/s13421-019-00904-w

Steller, M., and Koehnken, G. (1989). "Criteria-based statement analysis" in *Psychological methods in criminal investigation and evidence*. ed. D. C. Raskin (Springer Publishing Company), 217–245.

Tulving, E. (1974). Cue-dependent forgetting. *Am. Sci.* 62, 74–82.

Tulving, E. (1985). Memory and consciousness. *Can. Psychol.* 26, 1–12. doi: 10.1037/h0080017

Tulving, E., and Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychol. Rev.* 80, 352–373. doi: 10.1037/h0020071

Uzer, T., Lee, P. J., and Brown, N. R. (2012). On the prevalence of directly retrieved autobiographical memories. *J. Exp. Psychol. Learn. Mem. Cogn.* 38, 1296–1308. doi: 10.1037/a0028142

Vannucci, M., Pelagatti, C., Hanczakowski, M., Mazzoni, G., and Paccani, C. R. (2015). Why are we not flooded by involuntary auto- biographical memories? Few cues are more effective than many. *Psychological Research* 79, 1077–1085. doi: 10.1007/s00426-014-0632-y

Verkampt, F., Dodier, O., Ginet, M., and Milne, R. (2021). An analysis of the quality of French investigative interviews with children: age of the witnesses does matter. *Police Pract. Res. Int. J.* 22, 1130–1154. doi: 10.1080/15614263.2019.1658581

Volbert, R., and Steller, M. (2014). Is this testimony truthful, fabricated, or based on false memory? Credibility assessment 25 years after Steller and Köhnken (1989). *Eur. Psychol.* 19, 207–220. doi: 10.1027/1016-9040/a000200

Vredeveldt, A., van Rosmalen, E. A. J., van Koppen, P. J., Dror, I. E., and Otgaar, H. (2022). Legal psychologists as experts: guidelines for minimizing bias. *Psychol. Crime Law*, 1–25. doi: 10.1080/1068316X.2022.2114476

Wade, K. A., Garry, M., and Pezdek, K. (2018). Deconstructing rich false memories of committing crime: commentary on Shaw and Porter (2015). *Psychol. Sci.* 29, 471–476. doi: 10.1177/0956797617703667

Zhang, Y., Segal, A., Pompedda, F., Haginoya, S., and Santilla, P. (2022). Confirmation bias in simulated CSA interviews: how abuse comparison assumption influences interviewing and decision-making processes? *Leg. Criminol. Psychol.* 27, 314–328. doi: 10.1111/lcrp.12213

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 17 January 2024
DOI 10.3389/fpsyg.2023.1239139

OPEN ACCESS

EDITED BY
Fabiana Battista,
University of Bari Aldo Moro, Italy

REVIEWED BY
Gianmarco Convertino,
Sapienza University of Rome, Italy
Raffaella Maria Ribatti,
University of Bari Aldo Moro, Italy

*CORRESPONDENCE
Magdalena Kękuś
✉ mkekus@swps.edu.pl

RECEIVED 12 June 2023
ACCEPTED 18 December 2023
PUBLISHED 17 January 2024

CITATION
Kękuś M, Szpitalak M, Polczyk R and Barzykowski K (2024) Online misinformation can distort witnesses' memories. Analysis of co-witness discussions using an online version of the MORI-v technique. *Front. Psychol.* 14:1239139. doi: 10.3389/fpsyg.2023.1239139

Online misinformation can distort witnesses' memories. Analysis of co-witness discussions using an online version of the MORI-v technique

Magdalena Kękuś[1]*, Malwina Szpitalak[2], Romuald Polczyk[2] and Krystian Barzykowski[2]

[1]Faculty of Psychology in Kraków, SWPS University, Kraków, Poland, [2]Institute of Psychology, Jagiellonian University, Kraków, Poland

Introduction: The memory conformity effect occurs when people witness a given incident and then talk to each other about it, and the statement of one person affects the memory account of another person with respect to that incident. The main objectives of this experiment were (1) to examine the effectiveness of a modified version of the MORI-v technique in inducing the memory conformity effect and (2) to investigate how the manner in which participants discuss the observed event influences the magnitude of this effect. In general, the modified online MORI-v technique consists of the following main elements: (1) *original material*, that is, two versions of a short film which are identical except for certain critical details; for example, in one version, a thief puts on a red cap, but in the other version it is black; (2) the *collaborative recognition test*, that is, a discussion about the original material which leads to mutual misinformation; and (3) an *individual recognition test* that checks the effect of the discussion on the memory account of the original material.

Methods: A total of 72 participants (36 pairs) aged 18–54 took part in the research. Participants were tested using the online MORI-v technique: They were familiarized with the original material on their computers at home, and then they talked about it via a video communication app and completed an individual recognition test on their computers. Importantly, the discussions were recorded and analyzed in detail after the experimental session.

Results and discussion: Using the online MORI-v technique, the effect of memory conformity was demonstrated, that is, in the individual recognition test, the proportion of correct answers to questions about discussed details (related to misinformation) was lower than the proportion of correct answers to questions about non-discussed details. It was also demonstrated that if one participant introduced misinformation during the discussion about a particular item and the other did not question it, the latter's answer to that item during the individual recognition test was most often incorrect. However, if one participant introduced misinformation during the discussion about an item and the other questioned it, the latter's answer about that item during the individual recognition test was most often correct.

KEYWORDS

memory conformity effect, memory distortions, co-witness discussion, eyewitness testimonies, MORI technique

10.3389/fpsyg.2023.1239139

1 Introduction

The memory conformity effect occurs when people witness a given incident (e.g., a crime) and then talk to each other about it, and the statement of one person affects the memory account of the other person with respect to that incident (Wright et al., 2000). For instance, a witness, despite seeing a criminal wearing a gray coat, testifies that the coat was brown as this was the information heard from a co-witness. Thus, this phenomenon may contribute to unreliable testimonies, which remain the leading cause of incorrect court decisions (Smith and Cutler, 2013). It should be stressed that such a mistake may have serious consequences, including the conviction of an innocent person or the acquittal of a guilty one (Greene and Loftus, 1984). Consequently, the memory conformity effect is the subject of a lot of research.

Various methods have been used to present participants with misinformation stemming from other persons. In a seminal experiment, Schneider and Watkins (1996) presented a list of words to pairs of participants (Experiment 1) or to pairs in which one member was a researcher's accomplice (Experiment 2). The members of the pairs took turns to respond first. The second responses were strongly biased by the first responses. This effect has been replicated many times (e.g., Wright et al., 2000, 2005; Skagerberg and Wright, 2008, 2009; Thorley and Dewhurst, 2009, Experiment 3).

The above-described procedure consisted of two stages: Presentation of some original material and participants taking turns to respond to questions. An important extension of this paradigm that consists of three stages was proposed by Roediger et al. (2001; they called this paradigm "social contagion"). They studied pairs of people, one of whom was actually a confederate. The pairs watched household scenes, and during the following collaborative recall, the confederate introduced false answers in some cases. This resulted in "recalling" many of the suggested answers in a subsequent individual memory test. The fact that memory reports may be distorted as a result of wrong answers provided previously by a confederate has been replicated many times (e.g., Meade and Roediger, 2002; Wright et al., 2005; Allan and Gabbert, 2008; Zajac and Henderson, 2009; Thorley, 2013; Williamson et al., 2013; Szpitalak et al., 2015; Doughty et al., 2017; Calado et al., 2018, Experiments 1 and 2).

Another design used in memory conformity research consists of having two actual participants listening to or watching slightly different versions of the original material without knowledge of the differences. For example, in research by Gabbert et al. (2007), two participants were seated at computer desks with their backs to one another while they looked at pictures, which actually differed in some details. After a filler task, the participants then "recalled jointly" the pictures; in this way, they often misinformed each other because they were recalling details that were correct for only one of them. In a final performed individual memory test, fewer correct answers were given for critically discussed details. Other methods used to make the pairs of participants believe that they were watching the same video included, for example, having the participants watch the video one at a time "as there was only one monitor" (Bartlett et al., 2021) or separating them with a screen (Gabbert et al., 2003). Similarly, listening to different versions of an audiotape (Oeberst and Seidemann, 2014) and collaboratively recalling it resulted in biased answers in a subsequent individual memory test. In a rare example of an experiment outside the laboratory, Carlucci et al. (2011) had a confederate approach to small groups on a beach and interacted with one group member. Afterward, the experimenter asked the group members to identify the confederate in a target-absent line-up. Group members are more likely to conform to the responses of the person responding first if this person was the group member who had interacted with the confederate.

An interesting procedure in memory conformity research is the MORI technique (*Manipulation of Overlapping Rivalrous Images*; Mori, 2003, 2007), which consists of the simultaneous projection of two movie versions on the same screen. The two versions differ in terms of certain details; for example, in one version, the criminal checks the time on a wristwatch, while in the other, he does so on a wall clock. Thanks to the polarized glasses (which look like regular sunglasses) that are worn by the pair of participants, each of them sees a different movie version while being convinced that they are both actually watching the same movie clip. Thus, the participants sit beside each other, look at the same screen, and are not aware that the other person is watching a different version.

A disadvantage of the MORI technique is that it requires sophisticated technical equipment. To overcome this, Cadavid and Luna (2021) proposed a modification of this procedure called the MORI-v technique. Here, pairs of participants separately get acquainted with the different movie versions displayed on smartphone screens. Thus, polarized glasses are not necessary. Afterward, in order to introduce misinformation via an instant-messaging app, the experimenter sends multiple-choice questions concerning the movie, and the participants discuss them by chatting on the app. In the end, participants undergo a virtual individual recognition test on smartphones (Cadavid and Luna, 2021; Experiment 2). It should be stressed that regardless of which of the aforementioned procedures is used, the memory conformity effect is reliably replicated. Therefore, it is reasonable to ask about the mechanisms of this phenomenon, that is, the reasons why a given person succumbs to misinformation from a co-witness.

In the relatively scarce studies of the mechanisms of memory conformity, the classification proposed by Wright et al. (2009) is used most often. It is partly based on the distinction proposed by Deutsch and Gerard (1955) in the context of social influence and describes three processes which may cause the memory conformity effect. The first one is *normative influence*, which results from the need for social acceptance and avoiding the costs of disagreeing in social situations. It occurs because a person does not agree with information provided by an interlocutor but does not disclose their own opinion to avoid confrontation (Asch, 1956). It basically involves comparing the costs of disagreeing with the costs of making an error (Wright et al., 2010). This was confirmed in a study by Baron et al. (1996, Experiment 1), who manipulated the cost of an error: Two times as many participants yielded to misinformation when they were told that the experiment consisted of pilot data compared to participants who thought that the data would be used by police and courts and that the most accurate participants would be given a monetary prize. Also, Skagerberg and Wright (2008, 2009) showed that the perceived power of the partner influences memory conformity; this effect can be explained, for example, in terms of normative influence.

The second possible mechanism underlying memory conformity is *informational influence*. It refers to a situation in which a person succumbs to misinformation provided by an interlocutor as they are certain the interlocutor is right and is the source of more accurate

information. While normative influence involves comparing the costs of disagreeing with the cost of errors, the informational impact is connected with comparing the relative likelihood of the other person being correct versus oneself being correct (Wright et al., 2009). In contrast to normative influence, which could be expected to be reduced to zero when the participants answer in private (as there are no costs of disagreeing in this situation), the informational impact may be present in this context as there is no social influence.

Informational influence can be expected to be related to the perceived accuracy of one's own memory (as well as the partner's), the credibility of the partner, and subjective confidence in one's memory. Some research confirms these assumptions (Wright et al., 2000; Gabbert et al., 2007; French et al., 2011; Allan et al., 2012; Wright and Villalba, 2012; Williamson et al., 2013; Goodwin et al., 2017; Thorley and Kumar, 2017; Sousa and Jaeger, 2022; Kękuś et al., in press).

The third potential mechanism of memory conformity may be connected with memory distortion: The false information provided by other persons may influence the actual memories of the witness. Wright et al. (2009) differentiate between two possibilities: (1) just *believing* that the information provided by the partner is true or (2) the new information becomes part of episodic memory (as defined by Tulving, 1983). Believing that false information is true may also be related to source-monitoring errors; that is, the participant erroneously assumes that details mentioned by the partner were present in the original material (Oeberst and Seidemann, 2014).

The major aim of our research was to examine the effectiveness of a modified version of the MORI-v technique (Cadavid and Luna, 2021) in inducing the memory conformity effect. The original MORI technique is a useful tool for examining the memory conformity effect as—unlike other experimental procedures—it reduces the risk that participants will suspect manipulation, especially when they disagree on certain details during the discussion. In addition, the MORI technique is a good approximation of laboratory conditions to real-life situations where people witness a given crime, talk about it, and are then individually interviewed by enforcement agencies. However, this method requires special apparatus (projectors with polarizing filters, a translucent ground-glass screen, and special polarizing glasses), but not all researchers can afford such equipment. The MORI-v technique is an economical and readily available alternative to the original method. In addition, the MORI-v technique makes it possible to investigate the effects of virtual misinformation (Cadavid and Luna, 2021), which seems important given the increasing number of Internet users each year (Internet World Stats, 2023) and the increasing sharing of unverified information on social networks (Marsh and Rajaram, 2019).

Our basic idea was to make the MORI-v technique suitable for testing participants in full web-based settings instead of laboratories. In this method, participants are presented with original material on their own computers. Next, a pair of participants connected to the experimenter via an instant messenger with cameras and microphones on. The experimenter provides access to their computer screen so the participants can see a PowerPoint presentation with questions about the original material. As in the MORI and MORI-v techniques, the pair provides answers to 12 questions together, four of which refer to dissimilar details, thus leading to mutual misinformation. In contrast to the MORI-v, the participants talked to each other rather than writing their answers. In the end, participants undergo an individual recognition test on their own computers and are asked not to contact each other.

Apart from this, it is worth noting that both MORI-v and our modified online version of it may have some advantages over the original version. Namely, the original MORI procedure requires by design that the participants are told that the glasses will diminish their visual acuity. This could lessen their confidence in what they see. However, confidence in one's memory is an important predictor of memory conformity (e.g., Wright et al., 2000; Wright and Villalba, 2012; Thorley and Kumar, 2017; Yue et al., 2021; Sousa and Jaeger, 2022; Kękuś et al., in press). Therefore, the original procedure may not be as "ecological" as in real life (and research using other methods), where witnesses' confidence is not challenged by design.

Using the online MORI-v technique, we expected the usual memory conformity effect to arise; that is, the proportion of correct answers should be lower in cases in which the partner has mentioned details that are incongruent with what the participant actually saw. Apart from this, our second aim was to replicate the effect on memory conformity of disputing with the partner. In order to explain this, the distinctions between *discussed* and *non-discussed* details and between *disputed* and *non-disputed* ones should be presented.

The main factor in the MORI technique is that of *discussion* between the participants. This term refers simply to situations in which one participant gives answers inconsistent with what their partner saw. For example, one of the participants may have seen a black cap, while the second one sees a red one. If the first participant answers 'black', such a situation is classified as "discussed". This means that the second participant has been misinformed [or misdirected, as Cadavid and Luna (2021) put it]. Now, such discussed (misinforming, misdirecting) details may be divided into disputed and non-disputed ones (see Cadavid and Luna, 2021, Table 1; Kękuś et al., 2023, Figure 1). A given detail is classified as disputed if the participants in a pair disagree with each other during the discussion about it and give different answers (but consistent with their own original information). On the other hand, if, when discussing a dissimilar detail, one participant gives an answer that is consistent with their own original material and the other participant does not dispute the answer, the detail is classified as non-disputed. Thus, disputed details are equated with *mutual* misinformation, which means a pair of participants provide misleading information to each other, while non-disputed details are equated with *unilateral* misinformation, which means one participant misinforms the other (Ito et al., 2019). In existing research using the classic MORI technique, it has been found that when participants dispute a given detail, that is, they dispute what their partner said, they are less likely to be misled (French et al., 2008; Garry et al., 2008; Ito et al., 2019). The same was expected in the present research, not only because it has already been found in existing research but also because it makes logical sense. If the participant does not question their partner's answer, this might be an indication that they are convinced that the 'information' from the partner is the correct answer to a given question. This may be the case even if their own memory is different. This would mean that informational influence is at play.

Lower correctness in the case of non-disputed compared with disputed details may also arise because of the third of the mechanisms mentioned by Wright et al. (2009), namely memory distortions. Failing to question what the partner said may be an indication that a

participant indeed "remembers" the misinformation provided by the partner as if it were their own.

As for the normative impact, it was not considered important in the present study as the final recognition test was performed by the participants individually; therefore, there was no risk of confrontation in this situation.

The third aim of our research was to analyze the relationship between succumbing to misinformation and three individual characteristics: susceptibility to social influence, need for closure, and self-esteem. Susceptibility to social influence was measured by means of a self-description questionnaire called *Measure of Susceptibility to Social Influence* (MSSI; Bobier, 2002), described in detail below. In this tool, three dimensions of susceptibility to social influence were assumed: *Principled Autonomy*, reflecting independence of judgement and beliefs; *Social Adaptability*, referring to compliance with others, that is, allowing oneself to be influenced in order to avoid confrontation; and *Social Friction*, which can be defined as "anticonformity". It was hypothesized that all these three dimensions would be related to memory conformity: Non-conformist people who are independent in their judgments and opinions and are ready to engage in confrontation should rely on their own recollections rather than on information from their partner. In the existing literature, at least one result has been shown to be promising when searching for correlations between memory conformity and other types of suggestibility, namely the positive relationship that has been found between memory conformity and interrogative suggestibility (Thorley, 2013). Interrogative suggestibility (Gudjonsson, 1997) refers to the tendency to answer in accordance with suggestions contained in misleading questions (Yield) and the tendency to change one's answers under the influence of negative feedback (Shift). Memory conformity has been shown to be correlated with Yield but not Shift (Thorley, 2013). In other research, compliance, as measured by the Gudjonsson Compliance Scale (GCS, Gudjonsson, 1997), was positively related to memory conformity (Merckelbach et al., 2007).

Need for closure (NFC) is a construct described by Webster and Kruglanski (1994). It refers to an individual's desire for firm answers and their aversion to ambiguity. People with high NFC have a strong need to reduce the feeling of discomfort experienced in the face of cognitive uncertainty through the quick formulation and validation of a hypothesis. NFC includes five facets: Preference for order, Predictability of future, Decisiveness, Discomfort with ambiguity, and Closed-mindedness. We expected all these to be related to memory conformity as they imply a kind of difficulty with ambiguities, and ambiguities are a natural element of the memory conformity paradigm—when what the partner said is inconsistent with what the participant saw. People with high NFC may be tempted to resolve ambiguities by simply assuming that they are wrong and that their partner is correct. Therefore, our hypothesis was that it would correlate positively with memory conformity.

Finally, we expected self-esteem (SE) to be negatively related to memory conformity because low self-confidence is related to it (e.g., Wright et al., 2000; Wright and Villalba, 2012; Thorley and Kumar, 2017; Yue et al., 2021; Sousa and Jaeger, 2022; Kękuś et al., in press), and self-confidence is, in turn, related to self-esteem (e.g., Campbell, 1990; Coudevylle et al., 2011).

To sum up, the aims of the present study were as follows: (1) to replicate the memory conformity effect by means of the online version of the MORI-v technique, (2) to analyze the difference in correctness between disputed and non-disputed items, and (3) to analyze the correlations between memory conformity and susceptibility to social influence, need for closure, and self-esteem.

2 Method

2.1 Power analysis

The power analysis was performed using the G*Power software (Faul et al., 2009), assuming a desired power of 80%. A repeated-measures ANOVA was used for the analyses (for the main effect of memory conformity, as well as for the differences between disputed and non-disputed items). Also, power was calculated for Pearson's r correlations for the hypotheses concerning the relationship between memory conformity and individual traits.

As for the repeated-measures ANOVA, for the three effect sizes typically considered as small, medium, and large, that is, $f = 0.1$, 0.25, and 0.40, respectively (Cohen, 1988), the required sample sizes were 787, 128, and 52. In fact, in existing research across 10 countries, f was much higher: about 0.95 (translated from Hedges' $g = 1.92$, reported by Ito et al., 2019). As for correlations, the required sample sizes for small, medium, and large effects ($r = 0.10$, 0.30, and 0.50) are 783, 85, and 29, respectively.

Given the existing resources, a sample size of about 60 participants was assumed. This assured satisfactory power in the case of medium–to-large effect sizes but not for small ones.

2.2 Participants

A total of 72 participants (36 pairs, 44 women and 28 men) recruited via social media took part in the research. The mean age of participants was 25.0 (SD = 3.87). The youngest person studied was 18 years old, and the oldest was 54 years old. The participants signed a written informed consent. The Research Ethics Committee at the Faculty of Philosophy of the Jagiellonian University had no objections to the research.

2.3 Materials and design

2.3.1 Modified MORI-v technique

In general, the MORI and MORI-v techniques consist of the following main elements: (1) original material, that is, two versions of the film; (2) the collaborative recognition test, that is, a discussion about the original material, which leads to mutual misinformation), and (3) an individual recognition test that checks the effect of the discussion on the memory account of the original material. These elements are fully derived from the MORI technique and are described below.

2.3.1.1 Original material

The original material (used in previous studies in terms of the MORI technique, for example, Garry et al., 2008; Ito et al., 2019) was a silent movie lasting 6 min and 34 s. The recording depicts an electrician ("Eric") who steals some objects while repairing various household appliances. The movie was created in two versions by

Takarangi et al. (2006). The versions are identical except for eight critical details; for example, in one version, Eric puts on a red cap, but in the other version, it is black.

2.3.1.2 The collaborative recognition test

This consisted of a series of questions, thanks to which the participants discussed half of the discrepant details included in the original material, thus leading to mutual misinformation. The test comprised 12 questions, four of which refer to critical details (e.g., "what color was Eric's cap?"), with the other eight being control questions. Each question had five possible responses and was displayed in a PowerPoint presentation for 60 s while participants discussed the answer. If the participants were unable to reach an agreement, both responses were documented. The discussion was audio recorded.

2.3.1.3 Individual recognition test

The test included 20 questions, with eight of them referring to all critical details included in the original material. The test was used to compare the proportion of correct answers to the questions concerning the discussed details (related to misinformation) against the proportion of correct answers to the questions concerning the non-discussed details.

2.3.2 Individual differences questionnaires

- *The Measure of Susceptibility to Social Influence* (MSSI; Bobier, 2002; Polish adaptation: Polczyk, 2007) consists of 34 statements that evaluate the tendency to succumb to social influence. The tool consists of three subscales: Autonomy, Social Adaptability, and Social Resistance. The answer to each question is ranked on a scale of 1–5.
- *The Need for Closure Scale* (NFCS; Webster and Kruglanski, 1994; Polish adaptation: Kossowska et al., 2012). The shortened version of this tool consists of 15 statements to which the subjects respond on a 6-point scale. The tool includes the following five subscales: Preference for order, Predictability of future, Decisiveness, Discomfort with ambiguity, and Closed-mindedness.
- *Self-Liking/Self-Competence Scale-Revised* (SLCS-R; Tafarodi and Swann, 2001; Polish adaptation: Szpitalak and Polczyk, 2015) consists of 16 statements measuring two dimensions of self-esteem: self-competence and self-liking. The answer to each question is ranked on a scale of 1–5.

2.4 Procedure

Participants were examined in pairs during a single experimental session conducted via the Internet. They were informed that the research concerns the social sharing of information. In addition, they were instructed to use a computer, laptop or tablet rather than a smartphone during the study.

First, participants received a link to the movie, which was the original material. For each person in each pair, the movie was identical except for eight critical details. After watching the movie, participants were asked to complete The Need for Closure Scale and the Self-Liking/Self-Competence Scale, which provided an additional interval between the presentation of the original material and the introduction of the misinformation. Next, the subjects were requested to launch the Skype application. This part began with an instruction given by the experimenter, which was taken from research by Garry et al. (2008) and modified as needed:

"Thank you for participating in the first part of the experiment. I will audio-record our conversation with your permission.

In a moment, I will show you a series of questions about the movie you watched. You will see each question for 60 s. Ten seconds before that time expires, I will ask you for your answer.

You can also talk to each other during this time. If you have no idea what the answer is, talk to each other about the part of the movie the question is about or what was also happening in the movie at that time. This will help you remind each other of the right answer."

Afterward, a series of 12 questions about the original material was shown to the participants using the screen-sharing function. During the discussion of the four critical details, participants inadvertently introduced misinformation. If the pair disagreed with each other on the answer to a given question, the experimenter noted both answers. During the discussion, the experimenter and the participants could see each other via webcams.

In order to counterbalance the discussed details, 50% of the pairs discussed different critical details than the other 50% of the pairs. For example, half the pairs talked about Eric's cap color without discussing his company logo (version A of the discussion), while the other pairs discussed the logo but not the cap color (version B of the discussion). After the discussion, participants were asked to fill out The Measure of Susceptibility to Social Influence. Eventually, participants underwent the individual recognition test and were asked not to contact each other. In addition, the experimenter temporarily deactivated the link to the movie. After the experiment, participants were debriefed.

3 Results

The memory conformity effect was manipulated as a within-subject variable. The proportion of correct answers in the individual recognition test to questions about non-discussed details was compared with that of answers to questions about discussed (misinformed) details. In the case of non-discussed items, it was calculated as the proportion of correct answers to all non-discussed questions. In the case of discussed items, the score was calculated as the proportion of correct answers to questions about details for which a given participant received misinformation from the partner. Participants who were not exposed to any misinformation were excluded from the analysis, which left a sample of 66 subjects.

The results were clear-cut: the proportion of correct answers to questions about non-discussed details was much higher ($M = 0.74$, $SD = 0.23$) than for the questions about discussed and misinformed details ($M = 0.38$, $SD = 0.36$; $F(1, 65) = 54.06$, $p < 0.001$, $\eta^2 = 0.45$, Hedges' $g = 1.17$, 95% CI [0.68, 1.69]). Thus, the effectiveness of the modified version of the MORI-v technique (Mori, 2003; Cadavid and Luna, 2021) in inducing the memory conformity effect was confirmed. Moreover, the size of the effect was large.

In the second analysis, the proportion of correct answers in the individual recognition test for details that were disputed and non-disputed during the discussion was computed. This analysis was performed for participants who disputed an item at least once ($n = 36$).

The proportion of correct answers to questions about disputed details was much higher ($M = 0.84$, SD = 0.33) than for non-disputed ones ($M = 0.17$, SD = 0.38; $F(1, 35) = 51.72$, $p < 0.001$, $\eta^2 = 0.60$, Hedges' $g = 1.84$, 95% CI [1.28, 2.47]). Thus, it was demonstrated that if one participant introduced misinformation, and the other agreed with it during the discussion about a particular item, the latter's answer about that item during the individual recognition test was most often incorrect. However, if one participant introduced misinformation and the other protested during the discussion about an item, the latter's answer about that item during the individual recognition test was most often correct.

In the last set of analyses, individual variables (susceptibility to social influence, need for closure, and self-esteem) were correlated with resistance to the memory conformity effect. The latter was calculated as the difference between the proportions of correct answers to questions about non-discussed details minus the proportion of correct answers to discussed ones. This means that the higher the mean proportion, the better the accuracy when non-discussed and discussed items are compared; in other words, the better the result, the higher the resistance to misinformation. The results are presented in Table 1.

The results indicate that the higher the Principled autonomy and the lower the Social adaptability, the higher the resistance to memory conformity. In other words, participants who were autonomous were resistant to misinformation, while those who preferred to adapt to social rules were less resistant to misinformation. In addition, higher discomfort with ambiguity was connected with lower resistance to misinformation. The remaining correlations were not statistically significant.

4 Discussion

The major aim of our research was to examine the effectiveness of a modified version of the MORI-v technique (Mori, 2003; Cadavid and Luna, 2021) in inducing the memory conformity effect. Obviously, the details of our technique differ from the MORI and MORI-v techniques, but the main idea was preserved Two participants thought that they were watching the same movie while, in fact, it differed in some details. As a result of this, while discussing the movie, the participants mutually misinform each other.

TABLE 1 Correlations between individual variables and resistance to memory conformity.

Questionnaire	Subscale	Resistance to memory conformity
Measure of susceptibility to social influence	Principled autonomy	0.25*
	Social adaptability	−0.35**
	Social friction	0.05
Need for closure	Preference for order	−0.16
	Predictability of future	−0.01
	Decisiveness	0.01
	Discomfort with ambiguity	−0.27*
	Closed-mindedness	−0.05
Self-liking—self-confidence	Self-liking	−0.17
	Self-confidence	−0.11

*: $p < 0.05$. **: $p < 0.01$.

Our assumption was confirmed: The proportion of correct answers concerning discussed and misinformed details was much lower than for non-discussed ones. The size of the effect (Hedges' $g = 1.17$) was large and comparable to other studies. For example, Ito et al. (2019) reported effect sizes obtained using the classic MORI technique from 11 countries. They ranged from 1.01 (Colombia) to 2.97 (Japan). As for Poland, the effect size reported by Ito et al. (2019) was 1.92, with 95% CIs: 1.28, 2.56. It may conclude that the online version of the MORI-v technique generates memory conformity of comparable size to the classic version of this technique.

Thus, the online version of the MORI-v technique may be a useful and handy method for studying memory conformity. Online research may be, for example, the only option during a lockdown caused by a pandemic. In addition, web-based procedures make it possible to obtain a much larger and more diverse sample than most offline studies (Reips and Musch, 2002; Birnbaum, 2004) because Internet studies provide an opportunity to reach participants living in small towns and villages far from university laboratories, which are mostly located in large cities. Moreover, our method can be used to examine individuals who may have difficulty getting to the laboratory, for example, computer-literate disabled persons or seniors. What is more, as virtual contact occurs under relatively 'safe' conditions, it gives the potential to reveal behaviors that would likely be inhibited in face-to-face contact. Online research reduces symptoms of shyness and minimizes fear of evaluation by the experimenter, which is a major problem in laboratory research (Rosenberg, 1965; Dzwonkowska, 2003; Grzyb, 2017).

The hypothesis referred to the distinction between disputed versus non-disputed details, that is, the manner in which participants discuss the original material. It was assumed that the proportion of correct answers would be higher for questions about disputed details than for non-disputed ones. This hypothesis was confirmed. This result is also congruent with existing research using the classic MORI technique in which the proportion of correct answers to questions concerning disputed details was higher than that for non-disputed details (French et al., 2008; Garry et al., 2008; Ito et al., 2019). Thus, this effect was also replicated in the present research using the modified MORI-v technique.

The fact that disputing an item, that is, raising doubts about given misinformation, results in a higher proportion of correct answers might be caused by different factors. It is possible that a given participant had their own correct memories about a given item, and this fosters a correct answer. However, failing to dispute (i.e., to question) certain misinformation does not mean that the participant did not remember anything about a given item (although this is possible). It could be that relevant and correct memories were present, but the participant failed to disclose that they remembered something different than their partner as they were certain the interlocutor was right. In research by Kękuś et al. (in press), up to 58.3% of participants who had different memories from their partners yielded to misinformation and reported afterward that the reason for this was

lack of confidence in their own recollections. In another 41.7% of cases, the participants declared that they preferred to trust information provided by their partners, thinking that their partners "knew better." In sum, informational influence might also have occurred in the present research.

Our next hypothesis concerned the expected relationships between memory conformity and susceptibility to social influence, the need for closure, and self-esteem. These traits were measured by means of self-descriptive questionnaires. As for susceptibility to social influence, two of its dimensions proved to be related to succumbing to misinformation: Principled autonomy was related to higher resistance to misinformation, while Social adaptability was related to lower resistance. Social friction was not significantly statistically related to yielding to misinformation received from the other person. The hypothesis that relationships would be found between general susceptibility to social influence and yielding to misinformation was confirmed, at least in two out of the three facets of susceptibility to social influence. This could mean that the memory conformity effect might be just a kind of susceptibility to social influence. This would be logical as memory conformity is a kind of social influence, after all. However, this result should be taken with caution as the correlations between memory conformity and the same three dimensions of social influenceability have proved nonsignificant in other similar analyses (Kękuś et al., in press).

As for the need for closure, the results were not very compelling: Out of its five dimensions, only higher discomfort with ambiguity was related to lower resistance to misinformation. As for self-esteem, no significant correlations were obtained. Thus, the hypothesis relating to these traits should be treated as not confirmed.

5 Limitations and future directions

The present research was conducted online. Such research is convenient for researchers and has advantages, for example, access to a research sample with a wide range of psychodemographic parameters (Mason and Suri, 2012). However, the modified MORI-v technique is not without disadvantages, including difficulties with controlling various variables, including the main ones. It is possible that the motivation of some participants was not high; this could result in, for example, not paying sufficient attention to the original material, which in turn might artificially augment the memory conformity effect as a result of a lack of memories of the original details.

In future research on memory conformity, it might be useful to apply open-ended questions in the individual test. Such a form of questions might be more ecologically valid as questions in the form of closed alternatives should be avoided in real interrogations. Questions of this kind have already proved promising in research on memory conformity (Kękuś et al., in press).

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by Research Ethics Committee at the Faculty of Philosophy of the Jagiellonian University, Kraków, Poland. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

MK was supported by a grant from the National Science Centre, Poland (2021/41/N/HS6/03810). This research was supported by a grant funded by the Strategic Program Excellence Initiative at the Jagiellonian University to MK.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1239139/full#supplementary-material

References

Allan, K., and Gabbert, F. (2008). I still think it was a banana: memorable 'lies' and forgettable 'truths'. *Acta Psychol.* 127, 299–308. doi: 10.1016/j.actpsy.2007.06.001

Allan, K., Midjord, J. P., Martin, D., and Gabbert, F. (2012). Memory conformity and the perceived accuracy of self versus other. *Mem. Cogn.* 40, 280–286. doi: 10.3758/s13421-011-0141-9

Asch, S. E. (1956). Studies of independence and conformity: a minority of one against a unanimous majority. *Psychol. Monogr. Gen. Appl.* 70, 1–70. doi: 10.1037/h0093718

Baron, R. S., Vandello, J. A., and Brunsman, B. (1996). The forgotten variable in conformity research: impact of task importance on social influence. *J. Pers. Soc. Psychol.* 71, 915–927. doi: 10.1037/0022-3514.71.5.915

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

Birnbaum, M. H. (2004). Human research and data collection via the internet. *Ann. Rev. Psychol.* 55, 803–832. doi: 10.1146/annurev.psych.55.090902.141601

Bobier, D. M. (2002). *A measure of susceptibility to social influence: Scale development and validation*. Unpublished doctoral dissertation. Iowa: University of Iowa.

Cadavid, S., and Luna, K. (2021). Online co-witness discussions also lead to eyewitness memory distortion: the MORI-v technique. *Appl. Cogn. Psychol.* 35, 621–631. doi: 10.1002/acp.3785

Calado, B., Otgaar, H., and Muris, P. (2018). Are children better witnesses than adolescents? Developmental trends in different false memory paradigms. *J. Child Custody* 15, 330–348. doi: 10.1080/15379418.2019.1568948

Campbell, J. D. (1990). Self-esteem and clarity of the self-concept. *J. Pers. Soc. Psychol.* 59, 538–549. doi: 10.1037/0022-3514.59.3.538

Carlucci, M. E., Kieckhaefer, J. M., Schwartz, S. L., Villalba, D. K., and Wright, D. B. (2011). The south beach study: bystanders' memories are more malleable. *Appl. Cogn. Psychol.* 25, 562–566. doi: 10.1002/acp.1720

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Second Edition. Erlbaum Associates.

Coudevylle, G. R., Gernigon, C., and Martin Ginis, K. A. (2011). Self-esteem, self-confidence, anxiety and claimed self-handicapping: a mediational analysis. *Psychol. Sport Exerc.* 12, 670–675. doi: 10.1016/j.psychsport.2011.05.008

Deutsch, M., and Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *J. Abnorm. Soc. Psychol.* 51, 629–636. doi: 10.1037/h0046408

Doughty, N., Paterson, H. M., MacCann, C., and Monds, L. A. (2017). Personality and memory conformity. *J. Individ. Differ.* 38, 12–20. doi: 10.1027/1614-0001/a000217

Dzwonkowska, I. (2003). Nieśmiałość i jej korelaty [shyness and its correlates]. *Przegląd Psychologiczny* 46, 307–322.

Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. (2009). Statistical power analyses using G*power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160. doi: 10.3758/BRM.41.4.1149

French, L., Garry, M., and Mori, K. (2008). You say tomato? Collaborative remembering leads to more false memories for intimate couples than for strangers. *Memory* 16, 262–273. doi: 10.1080/09658210701801491

French, L., Garry, M., and Mori, K. (2011). Relative-not absolute-judgments of credibility affect susceptibility to misinformation conveyed during discussion. *Acta Psychol.* 136, 119–128. doi: 10.1016/j.actpsy.2010.10.009

Gabbert, F., Memon, A., and Allan, K. (2003). Memory conformity: can eyewitnesses influence each other's memories for an event? *Appl. Cogn. Psychol.* 17, 533–543. doi: 10.1002/acp.885

Gabbert, F., Memon, A., and Wright, D. B. (2007). I saw it for longer than you: the relationship between perceived encoding duration and memory conformity. *Acta Psychol.* 124, 319–331. doi: 10.1016/j.actpsy.2006.03.009

Garry, M., French, L., Kinzett, T., and Mori, K. (2008). Eyewitness memory following discussion: using the MORI technique with a Western sample. *Appl. Cogn. Psychol.* 22, 431–439. doi: 10.1002/acp.1376

Goodwin, K. A., Hannah, P. J., Nicholl, M. C., and Ferri, J. M. (2017). The confident co-witness: the effects of misinformation on memory after collaborative discussion. *Appl. Cogn. Psychol.* 31, 225–235. doi: 10.1002/acp.3320

Greene, E., and Loftus, E. (1984). Solving the eyewitness problem. *Behav. Sci. Law* 2, 395–406. doi: 10.1002/bsl.2370020406

Grzyb, T. (2017). *Eksperyment terenowy w psychologii społecznej.* [Field experiment in social psychology]. Warszawa: PWN.

Gudjonsson, G. H. (1997). *The Gudjonsson Suggestibility Scales manual*. Hove: Psychology Press.

Internet World Stats. (2023). World Internet Usage and Population Statistics 2023 Year Estimates. Available at: https://www.internetworldstats.com/stats.htm

Ito, H., Barzykowski, K., Grzesik, M., Gulgoz, S., Gurdere, C., Janssen, S. M. J., et al. (2019). Eyewitness memory distortion following co-witness discussion: a replication of Garry, French, Kinzett, and Mori (2008) in ten countries. *J. Appl. Res. Mem. Cogn.* 8, 68–77. doi: 10.1016/j.jarmac.2018.09.004

Kękuś, M., Chylińska, K., Szpitalak, M., Polczyk, R., Ito, H., Mori, K., et al. (2023). Reinforced self-affirmation as a method for reducing eyewitness memory conformity: an experimental examination using a modified MORI technique. *Appl. Cogn. Psychol.* 37, 660–674. doi: 10.1002/acp.4065

Kękuś, M., Polczyk, R., Ito, H., Mori, K., and Barzykowski, K. (in press). 'Is your memory better than mine?' Investigating the mechanisms and determinants of the memory conformity effect using a modified MORI technique. Registered report stage 2. *Appl. Cogn. Psychol.*

Kossowska, M., Hanusz, K., and Trejtowicz, M. (2012). Skrócona wersja Skali Potrzeby Poznawczego Domknięcia. Dobór pozycji i walidacja skali. [short version of the need for cognitive closure scale. Item selection and scale validation.]. *Psychologia Społeczna* 7, 89–99.

Marsh, E. J., and Rajaram, S. (2019). The digital expansion of the mind: implications of internet usage for memory and cognition. *J. Appl. Res. Mem. Cogn.* 8, 1–14. doi: 10.1016/j.jarmac.2018.11.001

Mason, W., and Suri, S. (2012). Conducting behavioral research on Amazon's mechanical Turk. *Behav. Res. Methods* 44, 1–23. doi: 10.3758/s13428-011-0124-6

Meade, M. L., and Roediger, H. L. (2002). Explorations in the social contagion of memory. *Mem. Cogn.* 30, 995–1009. doi: 10.3758/BF03194318

Merckelbach, H., Van Roermund, H., and Candel, A. (2007). Effects of collaborative recall: denying true information is as powerful as suggesting misinformation. *Psychol. Crime Law* 13:573581, 573–581. doi: 10.1080/10683160601160679

Mori, K. (2003). Surreptitiously projecting different movies to two subsets of viewers. *Behav. Res. Methods Instrum. Comput.* 35, 599–604. doi: 10.3758/BF03195539

Mori, K. (2007). A revised method for projecting two different movies to two groups of viewers without their noticing the duality. *Behav. Res. Methods* 39, 574–578. doi: 10.3758/BF03193028

Oeberst, A., and Seidemann, J. (2014). Will your words become mine? Underlying processes and cowitness intimacy in the memory conformity paradigm. *Can. J. Exp. Psychol.* 68, 84–96. doi: 10.1037/cep0000014

Polczyk, R. (2007). "Mechanizmy efektu dezinformacji w kontekście zeznań świadka naocznego" in *Mechanisms of the disinformation effect in the context of eyewitness testimony*. Ed. L. Sadko (Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego).

Reips, U. D., and Musch, J. (2002). Special issue: internet-based psychological experimenting. *Exp. Psychol.* 49, 241–242. doi: 10.1026//1618-3169.49.4.241

Roediger, H. L., Meade, M. L., and Bergman, E. T. (2001). Social contagion of memory. *Psychon. Bull. Rev.* 8, 365–371. doi: 10.3758/BF03196174

Rosenberg, M. J. (1965). When dissonance fails: on eliminating evaluation apprehension from attitude measurement. *J. Pers. Soc. Psychol.* 1, 28–42. doi: 10.1037/h0021647

Schneider, D. M., and Watkins, M. J. (1996). Response conformity in recognition testing. *Psychon. Bull. Rev.* 3, 481–485. doi: 10.3758/BF03214550

Skagerberg, E. M., and Wright, D. B. (2008). Manipulating power can affect memory conformity. *Appl. Cogn. Psychol.* 22, 207–216. doi: 10.1002/acp.1353

Skagerberg, E. M., and Wright, D. B. (2009). Sibling differentials in power and memory conformity. *Scand. J. Psychol.* 50, 101–107. doi: 10.1111/j.1467-9450.2008.00693.x

Smith, A. M., and Cutler, B. L. (2013). "Introduction: identification procedures and conviction of the innocent" in *Reform of eyewitness identification procedures*. ed. B. L. Cutler (Washington, DC: American Psychological Association), 3–21.

Sousa, W. S., and Jaeger, A. (2022). Memory conformity for high-confidence recognition of faces. *Mem. Cogn.* 50, 1147–1156. doi: 10.3758/s13421-022-01325-y

Szpitalak, M., Polak, M., Polczyk, R., and i Dukała, K. (2015). The influence of social, Para-social, and nonsocial misleading post-event sources on memory performance. *Eur. J. Soc. Psychol.* 46, 185–197. doi: 10.1002/ejsp.2136

Szpitalak, M., and Polczyk, R. (2015). *Samoocena: Geneza, struktura, funkcje i metody pomiaru. [Self-esteem: Origins, structure, functions and measurement methods.]* Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.

Tafarodi, R. W., and Swann, W. B. (2001). Two-dimensional self-esteem: theory and measurement. *Personal. Individ. Differ.* 31, 653–673. doi: 10.1016/S0191-8869(00)00169-0

Takarangi, M. K. T., Parker, S. L., and Garry, M. (2006). Modernizing the misinformation effect: the development of a new stimulus set. *Appl. Cogn. Psychol.* 20, 583–590. doi: 10.1002/acp.1209

Thorley, C. (2013). Memory conformity and suggestibility. *Psychol. Crime Law* 19, 565–575. doi: 10.1080/1068316X.2011.648637

Thorley, C., and Dewhurst, S. A. (2009). False and veridical collaborative recognition. *Memory* 17, 17–25. doi: 10.1080/09658210802484817

Thorley, C., and Kumar, D. (2017). Eyewitness susceptibility to co-witness misinformation is influenced by co-witness confidence and own self-confidence. *Psychol. Crime Law* 23, 342–360. doi: 10.1080/1068316X.2016.1258471

Tulving, E. (1983). *Elements of episodic memory*. Oxford: Clarendon Press.

Webster, D. M., and Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *J. Pers. Soc. Psychol.* 67, 1049–1062. doi: 10.1037/0022-3514.67.6.1049

Williamson, P., Weber, N., and Robertson, M. T. (2013). The effect of expertise on memory conformity: a test of informational influence. *Behav. Sci. Law* 31, 607–623. doi: 10.1002/bsl.2094

Wright, D. B., London, K., and Waechter, M. (2010). Social anxiety moderates memory conformity in adolescents. *Appl. Cogn. Psychol.* 24, 1034–1045. doi: 10.1002/acp.1604

Wright, D. B., Mathews, S. A., and Skagerberg, E. M. (2005). Social recognition memory: the effect of other People's responses for previously seen and unseen items. *J. Exp. Psychol. Appl.* 11, 200–209. doi: 10.1037/1076-898X.11.3.200

Wright, D. B., Memon, A., Skagerberg, E. M., and Gabbert, F. (2009). When eyewitnesses talk. *Curr. Dir. Psychol. Sci.* 18, 174–178. doi: 10.1111/j.1467-8721.2009.01631.x

Wright, D. B., Self, G., and Justice, C. (2000). Memory conformity: exploring misinformation effects when presented by another person. *Br. J. Psychol.* 91, 189–202. doi: 10.1348/000712600161781

Wright, D. B., and Villalba, D. K. (2012). Memory conformity affects inaccurate memories more than accurate memories. *Memory* 20, 254–265. doi: 10.1080/09658211.2012.654798

Yue, W. W. Y., Miyoshi, K., and Yue, W. W. S. (2021). Side- and similarity-biases during confidence conformity. *PLoS One* 16:e0253577. doi: 10.1371/journal.pone.0253577

Zajac, R., and Henderson, N. (2009). Don't it make my brown eyes blue: co-witness misinformation about a target's appearance can impair target-absent line-up performance. *Memory* 17, 266–278. doi: 10.1080/09658210802623950

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 26 January 2024
DOI 10.3389/fpsyg.2024.1232228

OPEN ACCESS

EDITED BY
Ivan Mangiulli,
University of Bari Aldo Moro, Italy

REVIEWED BY
Federica Alfeo,
University of Bari Aldo Moro, Italy
Jacqueline Katzman,
John Jay College of Criminal Justice, United States
Amanda Bergold,
Marist College, United States

*CORRESPONDENCE
Hayley J. Cullen
✉ hayley.cullen@mq.edu.au

RECEIVED 02 June 2023
ACCEPTED 10 January 2024
PUBLISHED 26 January 2024

CITATION
Cullen HJ, Dilevski N, Nitschke FT, Ribeiro G, Brind S and Woolley N (2024) The impact of misinformation presented during jury deliberation on juror memory and decision-making.
Front. Psychol. 15:1232228.
doi: 10.3389/fpsyg.2024.1232228

The impact of misinformation presented during jury deliberation on juror memory and decision-making

Hayley J. Cullen[1,2,3]*, Natali Dilevski[3,4], Faye T. Nitschke[1,5], Gianni Ribeiro[5,6], Shobanah Brind[1] and Nikita Woolley[1]

[1]School of Psychological Sciences, University of Newcastle, Callaghan, NSW, Australia, [2]School of Psychological Sciences, Macquarie University, Sydney, NSW, Australia, [3]School of Psychology, University of Sydney, Camperdown, NSW, Australia, [4]Centre of Investigative Interviewing, Griffith Criminology Institute, Griffith University, Mt Gravatt, QLD, Australia, [5]School of Psychology, The University of Queensland, St Lucia, QLD, Australia, [6]School of Law and Justice, University of Southern Queensland, Ipswich, QLD, Australia

When deliberating, jurors may introduce misinformation that may influence other jurors' memory and decision-making. In two studies, we explored the impact of misinformation exposure during jury deliberation. Participants in both studies read a transcript of an alleged sexual assault. In Study 1 (*N*=275), participants encountered either consistent pro-prosecution misinformation, consistent pro-defense misinformation, or contradictory misinformation (pro-prosecution and pro-defense). In Study 2 (*N*=339), prior to encountering either pro-prosecution or pro-defense misinformation while reading a jury deliberation transcript, participants either received or did not receive a judicial instruction about misinformation exposure during deliberation. Participants in both studies completed legal decision-making variables (e.g., defendant guilt rating) before and after deliberation, and their memory was assessed for misinformation acceptance via recall and source memory tasks. In Study 1, misinformation type did not influence legal decision-making, but pro-prosecution misinformation was more likely to be misattributed as trial evidence than pro-defense or contradictory misinformation. In Study 2, pro-defense misinformation was more likely to be misattributed to the trial than pro-prosecution misinformation, and rape myths moderated this. Furthermore, exposure to pro-defense misinformation skewed legal decision-making towards the defense's case. However, the judicial instruction about misinformation exposure did not influence memory or decision-making. Together, these findings suggest that misinformation in jury deliberations may distort memory for trial evidence and bias decision-making, highlighting the need to develop effective safeguards for reducing the impact of misinformation in trial contexts.

KEYWORDS

juries, legal decision-making, memory, misinformation, jury deliberation

1 Introduction

Jurors have the important task of deciding whether or not a defendant is guilty of a crime. Despite the consequences of their decisions, jurors make mistakes; innocent people can be convicted of crimes they did not commit (Huff et al., 1986; Innocence Project, 2021). For jurors to decide on cases accurately and impartially, they need to correctly remember the

complex, lengthy evidence presented at trial (Ruva et al., 2007; Thorley, 2016; Hirst and Stone, 2017; Ruva and Guenther, 2017). Jurors' memory of the evidence is a predictor of the decisions made in legal trials (Costabile and Klein, 2005). However, discussing the evidence with other jurors during deliberation may taint jurors' memories of the evidence, allowing for inaccuracy in the jury decision-making process.

The deliberation stage of the trial is considered to be a vital component of accurate jury decision-making. Legal systems appear to hold the assumption that discussion among jurors during deliberation enhances their memory of the key details relating to the case, which leads juries to collectively reach more accurate verdicts (Pritchard and Keenan, 1999, 2002; Hirst and Stone, 2017; Jay et al., 2019). However, these assumptions are not well supported by research, as deliberation does not always assist the jury as expected (Devine, 2012). Research suggests that up to 31% of juries engage in a verdict-driven deliberation style, in which they focus on reaching a verdict rather than thoroughly evaluating the evidence (Sandys and Dillehay, 1995; Devine et al., 2004). This deliberation style may reduce the likelihood that jurors uncover mistaken interpretations or recollections of the trial evidence (Devine, 2012). Other research suggests that in group decision-making contexts, individuals are less likely to share uniquely held pieces of information (Stasser and Titus, 2003). This may mean that if jurors misremember key details of a trial and report these inaccurate details during deliberation, other jurors may not correct them or notice the mistakes.

The extensive research on erroneous eyewitness testimony sheds further light on how memory errors may occur in jury deliberations, thus distorting other jurors' memories and interpretations. Eyewitness memory research has shown that eyewitnesses' memories can be distorted through exposure to misinformation—incorrect information that witnesses encounter after an event (see Loftus, 2005, for a review). One way in which misinformation can be introduced is during discussion with other witnesses to the same event (Wright et al., 2000; Frenda et al., 2011). Co-witness discussion can result in memory conformity, where rather than having independent recollections of the witnessed event, witnesses' memory reports start to influence one another's recollections (Gabbert et al., 2003; Hope and Gabbert, 2018; Ito et al., 2019). Witnesses are more likely to produce errors in their testimony if a co-witness has introduced misinformation during a discussion, compared to if witnesses are exposed to the same misinformation through other non-social sources (Gabbert et al., 2004; Paterson and Kemp, 2006a). However, more recent research highlights the potential benefits of collaborative discussion among co-witnesses through correcting one another's errors (see Vredeveldt et al., 2016; Vredeveldt and van Koppen, 2018). Despite these more recent findings, police officers are encouraged to prevent co-witnesses discussing an event with one another wherever practically and ethically possible, given the possible deleterious effects discussion can have on memory (Paterson and Kemp, 2005).

While discussion between witnesses to crimes has been actively discouraged in the past due to well documented issues with discussion on eyewitness memory, discussion between jurors through deliberation is instead *encouraged* (Pritchard and Keenan, 2002; Heydon, 2013). These different approaches across the criminal investigation and trial stages may exist because while eyewitnesses recall experienced events, jurors are required to make decisions about second-hand information learned during the trial. However, it is possible that akin to how witnesses may misremember key details about an eyewitness event, jurors may misremember key details from a legal trial (Pritchard and Keenan, 1999) and introduce misinformation during deliberation (Thorley et al., 2020). This misremembered information may then affect decisions made about the case. Such a possibility has received very little empirical attention (Hirst and Stone, 2017).

Misinformation presented during deliberation may lead to errors in source monitoring, thus altering jurors' recollection of trial information. The *source monitoring framework* (Johnson et al., 1993) has been used to explain the acceptance of misinformation in an eyewitness context. This framework proposes that source monitoring errors occur because people tend to encode the content of memories without any label identifying the source. As such, it is during memory retrieval that one must decide on the source of not only the original memory, but also the misinformation. Thus, misinformation might be remembered as occurring during the original event (e.g., criminal trial) because the source of the misinformation is mistakenly believed to be the original event (Zaragoza and Lane, 1994). Furthermore, if the original event and misinformation share common characteristics, there is an increased likelihood that an individual will misattribute the source of the misinformation (Johnson et al., 1993). Because jurors actively discuss the evidence from trial, it is very likely that any case-related misinformation would share characteristics with the evidence at trial. Thus, jurors might incorrectly remember misinformation items as appearing in the trial, thereby committing a source monitoring error, because the misleading details seemingly fit with the narrative presented during the trial.

Recent research has explored whether misinformation introduced by fellow jurors during jury deliberation may lead to source monitoring errors, thereby affecting juror memory and decision-making. In Thorley et al. (2020), participants viewed a video trial, after which they read a transcript of a deliberation regarding the case. For half of the participants, the deliberation only contained correct information regarding evidence within the trial. For the other half of participants, six items of misinformation favoring the prosecution's case were introduced into the transcript of the deliberation. Participants then provided an individual verdict, and completed a source memory test to determine whether the misinformation within the deliberation was attributed to the trial. The findings showed that those who read the deliberation containing misinformation were more likely to attribute this misinformation as evidence presented during the trial than participants who never received the misinformation. Additionally, acceptance of the misinformation impacted upon decision-making; jurors who misremembered the misinformation as real trial evidence were more likely to deliver a guilty verdict.

Thorley et al.'s (2020) findings provide preliminary evidence that not only can misinformation introduced during deliberation distort jurors' memory for trial evidence, but that this memory distortion may impact the final verdict individual jurors decide on. However, the misinformation presented to participants during deliberation in Thorley et al.'s study only focused on the prosecution's case, causing mock-jurors to evaluate the trial evidence in favor of the prosecution (i.e., rendering them more likely to deliver a guilty verdict). It is plausible that jurors may also mention misinformation that is consistent with the defense's case. This may be a particular issue in sexual assault trials, where the defense case often plays to inaccurate beliefs about how sexual violence is perpetrated (sometimes called

rape myths; e.g., Gray and Horvath, 2018). These inaccurate beliefs are an extra-legal factor, in that jurors are not legally permitted to consider them to make decisions in sexual assault cases (Heydon, 2013). Information which aligns with these inaccurate beliefs about sexual violence, and the defense case, may be more readily accepted by jurors (Süssenbach et al., 2012). Pro-defense misinformation may shift decision making towards the defense's case (i.e., more acquittals than guilty verdicts). Thus, it is important to explore whether different types of misinformation would have different effects on juror memory and decision-making.

The primary aim of the current research was to explore the effect of different types of misinformation—pro-prosecution versus pro-defense—presented during jury deliberation on juror memory and decision-making. We conducted two studies to address this aim. Below, we discuss the procedure and hypotheses for Study 1.

In Study 1, participants read a fictitious trial transcript depicting an alleged sexual assault. Following this, they provided a pre-deliberation verdict, and rated the perceived credibility of the complainant's testimony. Participants then engaged in a simulated deliberation with two other jurors, who provided misinformation on key aspects of the trial. Specifically, in the *consistent pro-prosecution* condition, both jurors consistently mentioned the same misinformation that favored the prosecution's case; in the *consistent pro-defense condition*, both jurors consistently mentioned the same misinformation that favored the defense's case. Given that the group size during deliberation is typically larger than groups of co-witnesses discussing a witnessed event (Paterson and Kemp, 2006b), it is also plausible that different jurors may mention misinformation for the same key detail that is contradictory (Thorley and Dewhurst, 2007; Hirst and Stone, 2017). To account for this possibility, we also included a *contradictory* condition where both jurors mentioned a different misinformation item for the same key detail; one juror's response was consistent with the prosecution's case, while the other juror's response was consistent with the defense's case. Following the deliberation phase, participants re-completed the verdict and credibility measures, as well as a free-recall and source memory test.

We hypothesized that misinformation acceptance would be greatest in both the consistent pro-prosecution and pro-defense misinformation conditions compared to the contradictory misinformation condition. There are several reasons why consistently hearing the same misinformation might lead to an increased likelihood of misinformation acceptance than hearing contradictory information. First, greater acceptance of consistent misinformation may occur because the credibility of the misinformation is heightened when it is consistently recalled by multiple sources (Mojtahedi et al., 2018; Blank et al., 2021). Second, consistent misinformation may be more likely to be accepted than contradictory misinformation because it is remembered better. Research from memory for repeated events (e.g., repeated sexual abuse) suggests that memory is stronger for details that occur in the same way across instances (e.g., the same perpetrator) and weaker memories for details that vary across instances (e.g., different forms of abuse; MacLean et al., 2018; Dilevski et al., 2020a,b; Rubínová et al., 2020a,b; Deck and Paterson, 2021a,b). Therefore, jurors presented with consistent misinformation during deliberation would be more likely to attribute that misinformation as appearing in the trial than those presented with contradictory misinformation, because memory is stronger for the former than the latter. No differences in misinformation acceptance were expected between the consistent pro-prosecution and pro-defense conditions.

As we expected the same patterns of findings for verdicts and ratings of defendant guilt, for brevity, we just report our expectations for guilt ratings below. We hypothesized that post-deliberation guilt ratings and complainant credibility ratings would be significantly higher in the consistent pro-prosecution condition than the other two conditions. However, it was unclear whether there would be a difference in post-deliberation guilt ratings and complainant credibility ratings between the consistent pro-defense and contradictory conditions, because contradictory misinformation may decrease the perceived strength of the evidence (Mojtahedi et al., 2018; Blank et al., 2021), which may reduce guilt and credibility ratings similarly to receiving misinformation that favors the defense. Finally, we hypothesized that the relationship between misinformation condition and post-deliberation ratings of guilt and credibility would be mediated by misinformation acceptance, as per the findings of Thorley et al. (2020).

2 Study 1 method

2.1 Participants

Two-hundred and ninety-eight participants took part in the study. The data from 23 participants was excluded for the following reasons: failing at least one attention check (Cullen and Monds, 2020) ($n = 10$), experiencing technical issues ($n = 8$), not completing the study ($n = 4$), or completing the study more than once ($n = 1$). This left a valid sample of 275 participants. Based on *a priori* power calculations conducting using G*Power 3.1 (Faul et al., 2007), a sample of 267 participants was required to achieve 90% power, given the design, main planned analyses (one-way ANOVAs) and assumed effect size (moderate; Lovakov and Agadullina, 2021). Participants were recruited through undergraduate research participant pools ($n = 231$), or through the paid research recruitment system of the University of Sydney ($n = 44$). Participants were required to be Australian citizens and over 18 years of age to participate in the study, to meet the basic jury eligibility requirements across all Australian states. However, Australian states have other exclusion criteria for jury service that we did not screen for (e.g., criminal history), so it should be noted that some participants may not be jury eligible depending on the jurisdiction. We also required participants to be fluent in English in order to follow and understand all study instructions. The undergraduate and paid research participation pools were based at the same institution and both used the SONA research participation platform; thus, the participants recruited through the two SONA platforms were demographically similar (gender, jury experience, English acquisition, culture). The only difference was that student participants were younger on average than paid participants, $F(1,273) = 7.354$, $p = 0.007$. Given that the samples were equivalent in all other respects, the samples were combined in all analyses. See Table 1 for the breakdown of demographic characteristics based on recruitment strategy.

Overall, participants had a Mean Age of 22.21 years ($SD = 6.99$), and most participants were female (77.5%). Participants identified as the following cultural/ethnic backgrounds: European/White (49.8%), East Asian (23.3%), Southeast Asian (6.9%), mixed (6.9%), South

TABLE 1 Demographic characteristics of Study 1 participants based on recruitment strategy.

Demographics	Student participants (n = 231)	Paid participants (n = 44)
Mean age	21.71 (6.30)	24.80 (9.52)
Gender (%)		
Female	77.5	77.3
Male	22.1	22.7
Non-binary	0.4	-
Previous jury experience (%)		
Yes	0.9	2.3
No	99.1	97.7
English as first language (%)		
Yes	81.4	79.5
No	18.6	20.5
Cultural background (%)		
European/White	52.4	36.4
East Asian	19.9	40.9
Other	27.7	22.7

"Other" for cultural background includes cultural backgrounds with low cell counts: African, Hispanic, Middle Eastern, Mixed, Pacific Islander, South Asian, Southeast Asian. Standard deviations for mean age in parentheses.

Asian (5.5%), other (7.6%). Few participants (n = 3) had previously served on a jury. For most participants (81.1%), English was their first language.

2.2 Design

The current study employed a one-way between-subjects design with three conditions, investigating the effects of misinformation exposure (pro-prosecution vs. pro-defense vs. contradictory) on juror memory and decision-making. We made the decision not to include a pure control group that received no misinformation for two reasons. First, we wanted to ensure that we had sufficient power to detect any effects for our key research questions given practical constraints (e.g., funding, time). Second, decades of research has highlighted that exposure to incorrect information distorts memory across a variety of settings and sources (e.g., Wright et al., 2000; Loftus, 2005). The extant literature suggests that a misinformation effect would occur in a jury setting (Thorley et al., 2020); therefore, our research questions were instead centered around the factors that enhance or reduce this misinformation effect in a jury deliberation context. Participants were randomly allocated to misinformation conditions (consistent pro-prosecution n = 92; consistent pro-defense n = 89; contradictory n = 94). Measurements of juror memory and decision-making are described below.

2.3 Materials

2.3.1 Trial transcript

All participants were required to read a shortened trial transcript depicting an alcohol-involved acquaintance sexual assault. More than half of sexual assaults are alcohol involved, meaning the victim and/or perpetrator have consumed alcohol (Abbey et al., 2004; Cox, 2015). As this is an early investigation of the effects of misinformation in jury deliberation, we opted to use a common type of case that jurors might be asked to consider in a criminal trial. The transcript was modified from that used in Nitschke et al. (2021). As it is common in sexual assault trials to only hear evidence from the complainant (e.g., New South Wales Law Reform Commission, 2020), our transcript features the examination-in-chief of the alleged victim, Chloe Miller, who testifies about the events leading to the sexual assault. Specifically, Chloe testifies that she was out for drinks at a bar with some work colleagues to celebrate a colleagues' promotion. She had been casually dating the defendant, Peter Stanton, who she had met on a dating app 4 weeks prior. Peter had sent her a message to see if she was out, and the two had agreed to meet at the bar. After Peter had arrived at the bar, Chloe had a drink spilled over her, and Peter suggested that they go back to his place down the road so Chloe could clean up. Once at Peter's apartment, Chloe and Peter had two more drinks each, and started to kiss on the sofa. Chloe started to feel uncomfortable when Peter began moving his hand up her thigh. Peter took Chloe's clothes off and pushed Chloe down. Chloe tried to push Peter off but was unsuccessful. Peter then penetrated Chloe with his fingers and penis. The transcript was 1,534 words in length and took participants approximately 6 min, 30 s to read. Pilot testing (n = 15) revealed a conviction rate of 80%, and a mean guilt rating of 5.17 (out of 7) using this trial transcript.

2.3.2 E-deliberation

Participants engaged in a simulated deliberation (approximately 12 min) hosted via an online, text-only chatroom. A similar method of simulated deliberation has been used in previous research (Salerno et al., 2019). Participants were led to believe that they would discuss the case with two other participants taking part in the study. However, the two other "jurors" and their associated text responses were simulated. To simulate what occurs in real legal cases, all participants were assigned a juror number prior to beginning the e-deliberation. Participants were referred to by this juror number throughout the e-deliberation. The two other "jurors" were also referred to by a number. The e-deliberation began with the "moderator" of the chatroom (also simulated) welcoming the other jurors and outlining that the purpose of the deliberation was to answer questions relating to the case. The moderator then asked eight questions about the case that all jurors answered. These questions were asked in a fixed order. The actual participant was always the first person prompted to respond to each question. Our decision to have the actual participant respond first to the question was so that their response was given prior to being exposed to misinformation. Participants could therefore not interact with or respond to the subsequent simulated responses. For four of the questions (questions 1, 3, 5, and 7), both the simulated jurors provided correct answers. For the other four questions (questions 2, 4, 6, and 8), both the simulated jurors provided incorrect answers (i.e., provided misinformation). However, the answers they provided differed depending on the experimental condition to which participants had been assigned.

Research has indicated that certain types of information influence how rape cases are perceived (e.g., Monson et al., 2000). Additionally, rape myths are often expressed throughout jury deliberations in sexual assault cases (Leverick, 2020). Common rape myths include beliefs

that intoxicated victims are somewhat responsible for their rape, that a lack of resistance provides evidence against rape, and that rape cannot occur in intimate relationships (Leverick, 2020). During deliberation, if a juror misremembers the case facts in line with irrelevant rape myths (e.g., the complainant was intoxicated, did not resist, and was in an intimate relationship with the defendant), this could discredit the prosecution's case and add credibility to the defense's case (Dinos et al., 2015). Alternatively, if a juror misremembered the case facts in a way that opposes these rape myths (e.g., the complainant was sober, resisted, and was not in an intimate relationship with the defendant), this could have the opposite effect of adding credibility to the prosecution's case and decreasing credibility of the defense's case. To this end, our different misinformation conditions capture the different types of misinformation that might arise during jury deliberations for sexual assault cases, and the unique effects these types of misinformation will have on credibility and verdict.

2.3.2.1 Misinformation conditions

Table 2 presents questions and responses provided by the simulated jurors based on experimental condition, and correct details for questions where misinformation was provided. Pilot testing was conducted to generate pro-prosecution and pro-defense misinformation items that were equal in similarity to the facts in the trial, to avoid any confounds across conditions.

2.3.2.1.1 Consistent pro-prosecution

In the consistent pro-prosecution condition, both simulated jurors provided the same misinformation item, and this included information that was favorable for the prosecution case. For example, for the question: *"How long had Peter and Chloe known each other for before the night of the alleged rape?,"* both jurors in the consistent pro-prosecution condition answered 2 weeks (as opposed to the correct answer of 4 weeks).

2.3.2.1.2 Consistent pro-defense

In the consistent pro-defense condition, again both simulated jurors provided the same misinformation item, but in this case the misinformation item provided information that was favorable for and served to enhance credibility of the defense's case. For example, for the question: *"How long had Peter and Chloe known each other for before the night of the alleged rape?,"* both jurors in the consistent pro-defense condition answered that the two had known each other for 6 weeks.

2.3.2.1.3 Contradictory

In the contradictory condition, for each question where misinformation was provided, one of the jurors answered with the pro-prosecution misinformation item (e.g., *"2 weeks"*), while the other answered with the pro-defense misinformation item (e.g., *"6 weeks"*).

2.3.2.2 Source credibility

To determine whether perceived source credibility played a role in misinformation acceptance, participants were also asked at the end of the study to rate how accurate they believed both the jurors they deliberated with to be, on a scale from 1 (*Not at all*) to 7 (*Completely accurate*). Analyses relating to these items are presented in the Supplementary Data Sheet S1 as they are not the main focus of the study.

2.4 Measures

2.4.1 Verdict and guilt ratings

Participants were asked to render a verdict of guilty or not guilty for the defendant with a justification. While dichotomous ratings of guilt are reflective of real jury verdicts, these measures can be less sensitive than measures of continuous guilt (Glaser et al., 2015). Therefore, participants were also asked to rate the likelihood that the defendant was guilty on a scale from 1 (*Not at all*) to 7 (*Very likely*) (Matsuo and Itoh, 2016). Participants provided their verdict and completed the guilt rating both prior to and after deliberation (i.e., before and after misinformation exposure).

2.4.2 Credibility

Participants answered four questions regarding their perception of the complainant's honesty, believability, credibility, and accuracy, on a scale from 1 (*Not at all*) to 7 (*Completely*). These questions were adapted from previous research (Connolly et al., 2008). These questions were also completed by all participants at two time points: pre- and post-deliberation. Given that we expected ratings for all four complainant questions to be similar, we checked the internal consistency of the pre- and post-ratings. These ratings revealed high

TABLE 2 Misinformation and correct items for each question in Study 1 based on misinformation condition.

Item	Pro-prosecution	Pro-defense	Correct
Misinformation			
Status of the relationship	Just friends	Intimate relationship	Casually dating*
Length of relationship	2 weeks	6 weeks	4 weeks*
Number of drinks consumed	1 drink	3 drinks	2 drinks*
Chloe's reaction to touching	Left hand on thigh	Moved hand up thigh	Moved hand off thigh*
Correct			
How they met			Dating app
Plan to meet on the night			Peter texted Chloe
Where they were in apartment			On the sofa
Chloe's reaction to taking off underwear			Pushed Peter away

*Indicates correct details from the trial that were not featured in any version of the deliberation. These details are provided in the table to compare to the misinformation items.

internal consistency (pre-deliberation: *Cronbach's* $\alpha = 0.892$; Post-deliberation: *Cronbach's* $\alpha = 0.926$). Therefore, these ratings were aggregated to form a single pre-deliberation and post-deliberation complainant credibility score.

2.4.3 Recall memory

Following the post-deliberation verdict and credibility ratings, participants' memory was measured using a single free recall question. The free recall question asked participants to recall the key details that they regarded as most important to remember about the case. Participants were given a three-minute time limit for the free-recall task, to facilitate focusing only on the most relevant and important details about the case (including those discussed in the deliberation). The free recall task was included to examine the extent to which the misinformation and correct items presented during the deliberation would be spontaneously reported by participants as appearing during the trial.

2.4.3.1 Recall coding

Participants' free recall reports about the key details from the trial were coded to determine whether participants spontaneously mentioned the incorrect (i.e., misinformation) and correct information they encountered during the deliberation. For the current study, only misinformation and correct items presented in the deliberation were coded. For details where misinformation was provided, participants could either accept or reject misinformation items. Therefore, participants were coded as having accepted the misinformation item (i.e., reported the inaccurate misinformation they were exposed to during their deliberation), or correctly rejecting the misinformation item (i.e., reported the correct information instead of the misinformation they were exposed to during the deliberation). For example, if a participant in the pro-prosecution misinformation condition reported that the complainant and defendant had known each other for two weeks, this would be coded as "misinformation accepted." If that same participant had reported that the complainant and defendant had known each other for 4 weeks (i.e., the correct answer), this would be coded as "misinformation correctly rejected." Coders could not be blind to experimental conditions as misinformation acceptance depended on misinformation condition. For items where correct information was provided, participants were coded as having accepted the correct item if they reported the correct item. For example, if participants reported that Chloe and Peter had met on a dating app, this would be coded as "correct accepted".

Two independent scorers completed the coding. Scorer 1 (HC) coded 100% of participant responses. To check for inter-rater reliability, Scorer 2 (FN) coded 50% of participant responses ($n = 159$) in line with APA publishing standards. The Intraclass Correlation Coefficients revealed moderate ($ICC = 0.776$), good ($ICC = 0.848$), and excellent ($ICC = 0.919$) reliability for misinformation accepted, misinformation correctly rejected, and correct accepted, respectively (Koo and Li, 2016). Given the acceptable reliability, the coding from Scorer 1 was used in the analyses.

2.4.4 Source memory

Following the free recall report, participants completed a source memory test. Following previous jury misinformation research (Ruva et al., 2007; Thorley et al., 2020), participants read a series of statements and they were instructed that the information in the statement may have come from different sources (trial only, only deliberation, both trial and deliberation, or neither). Participants were asked to identify the source of the information, and to rely on their own memory for the source of the information. The response options were *Trial Only, Deliberation Only, Both Trial and Deliberation, and Neither Trial nor Deliberation*. These instructions were based on those provided in Mitchell and Zaragoza (2001).

The source memory test consisted of 24 items (or statements). There were 4 types of items: 8 *misinformation* (deliberation only), 4 *correct* (both trial and deliberation), 4 *correct* (trial only), and 8 *new* items. The *misinformation* items restated the misinformation presented during the e-deliberation (correct answer: *Deliberation Only*). However, for participants in the consistent misinformation conditions (pro-prosecution and pro-defense), 4 of the 8 misinformation items were technically filler items, as those items had not been presented during the e-deliberation for these conditions. These four filler items were not included in the scoring for these conditions. The misinformation items were included to provide a measure of participants' proclivity to "misinformation acceptance" (i.e., a critical source monitoring error) after being misled about trial details during the deliberation.

The *correct information* (both trial and deliberation) items restated the correct information presented during the trial and the deliberation (e.g., *Chloe and Peter met on a dating app*) (correct answer: *Both Trial and Deliberation*). The *correct* (trial only) items restated the information presented during the "trial only" (e.g., *As Chloe and Peter kissed, Chloe moved Peter's hand away from her thigh*). The *new* items stated information that appeared in "neither trial nor deliberation" (e.g., *Chloe and Peter first met each other through a co-worker*). These items were based on what appeared during the trial, but they suggested alternative information about what occurred. We included the correct information from deliberation, correct (trial only), and new information items to provide a measure of whether the participants across the three conditions remembered the trial equally well (Thorley et al., 2020). Analyses relating to these items are presented in the Supplementary Data Sheet S1 as they are not the main focus of the study. Overall, the analyses revealed that performance on these items did not differ across misinformation conditions.

Participants' responses to each item in the source memory test were scored to determine whether they had misremembered/remembered the information as appearing during the trial. Specifically, participants received one point each time they had responded to a test item with 'Trial Only' or 'Trial and Deliberation', as both responses indicate that a participant remembered that the information appeared during the trial. After scoring was complete, we summed together participants' 'Trial Only' and 'Trial and Deliberation' scores for each item type separately. For data analysis purposes, proportion scores for each information type were calculated by dividing participants' 'Trial Only' and 'Trial and Deliberation' scores by the number of items for that information type. For example, if a participant in the pro-prosecution condition misremembered that two out of four items of misinformation appeared during the trial, their proportion score would be 0.5 (2/4 = 0.5).

Since participants in the contradictory condition received both pro-prosecution and pro-defense misinformation items, three proportion scores were computed pertaining to performance for these items. We calculated a proportion score for pro-prosecution and pro-defense misinformation items, separately. Then, for the main

analysis relating to misinformation items, we calculated an average proportion score for misinformation acceptance between the pro-prosecution and pro-defense items (i.e., [proportion pro-prosecution score + proportion pro-defense score]/2).

2.4.5 Attention checks and suspicion

At different stages of the study, participants were asked three instructional attention check questions to ensure that they followed the instructions (Oppenheimer et al., 2009). Participants who answered any of these questions wrong were removed from the analyses (Cullen and Monds, 2020). Participants were also asked questions at the end of the study to determine whether they were suspicious about the aims of the study. They were asked if they noticed anything strange about the study, and if so, to report what was strange (Salerno et al., 2019). This was not used as a basis for exclusion, but instead to determine whether participants were suspicious about the simulated deliberation and whether this suspicion mattered. Analyses were conducted with and without participants who were suspicious about the deliberation, to determine whether suspicion impacted upon the study results.

2.5 Procedure

Participants signed up for the study advertised as "Jury decision-making." The study took place in 2020 and 2021. Thus, due to COVID-19 social distancing requirements, the study was conducted online. Once a participant had signed up to the study, an experimenter made contact with that participant via email to arrange a day and time to complete the session. At the time of each participant's appointment, an experimenter emailed the participant the link to the online experiment. The online session began with participants providing informed consent. They were then presented with general instructions about the study. Specifically, they were informed that the study was being conducted by researchers from two different universities, and that they would read about a criminal trial, engage in a deliberation with other participants from the other institution (to increase the realism of the study and the simulated deliberation), and answer some questions about the trial.

Following the general instructions, participants read the trial transcript about a sexual assault case. To ensure that participants attended to the trial transcript, they were given a minimum of three minutes to read the transcript and could not proceed until the time had elapsed. The minimum time limit was determined through pilot testing. After reading the trial transcript, participants completed the pre-deliberation measures of verdict, guilt rating, and complainant credibility. Therefore, this pre-deliberation decision-making occurred prior to any misinformation exposure. Participants were directed via email to log into the chat room where they would engage in the live online deliberation with other participants. Participants then engaged in the 12-min e-deliberation where they either received pro-prosecution, pro-defense, or contradictory misinformation by the simulated jurors.

After the deliberation, participants completed the post-deliberation measures of verdict, guilt rating, and complainant credibility (i.e., after misinformation exposure). Participants then completed the free recall and source memory measures, and then rated how accurate they believed the two "jurors" were during the e-deliberation. Finally, participants completed the suspicion check questions and several demographic questions. Upon completion of the study, participants were fully debriefed about the study. The majority of the study was hosted using Qualtrics survey software. However, the simulated e-deliberation was hosted on AJAX chat. It took approximately 45 min to complete the study. All aspects of the study were approved by the University of Sydney Human Research Ethics Committee (protocol number: 2019/947).

2.6 Transparency statement

We reported how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in this study. The hypotheses, design, measures, and analysis plan were pre-registered on Open Science Framework (OSF). See here for the original registration https://osf.io/kdbma/ and here for the amended version https://osf.io/2rxye/. Any deviations from the pre-registration are reported transparently below. All experimental materials (including the e-deliberation script) and data (dataset, output, and code) are available on OSF.[1]

2.7 Transparent deviations

First, we collected data during periods of lockdown in Australia through the COVID-19 pandemic. This meant we had to switch from lab-based participant recruitment to transform the study to be fully online. As a result, we had to switch to participants having a juror number instead of their name in the online deliberation and we also were unable to record participants' memory responses during the online deliberation to determine their original memory prior to misinformation exposure.

Second, we planned to run mediation analyses to determine whether: (1) misinformation acceptance mediated the relationship between misinformation condition and post-deliberation measures, and (2) perceived credibility of the other jurors mediated the relationship between misinformation condition and misinformation acceptance. We did not find direct effects of misinformation condition on post-deliberation measures or perceived credibility of other jurors. Therefore, we did not run the planned mediations.

3 Study 1 results

3.1 Overview and analysis plan

First, we reported the descriptive statistics relating to suspicion about the simulated deliberation. We then moved on to decision-making. We conducted preliminary analyses (one way ANOVAs and a chi-square test) to ensure that there were no differences in pre-deliberation measures across misinformation conditions (i.e., that there were no pre-existing differences in attitudes and beliefs about the case before the misinformation was introduced). Then, one way

1 https://osf.io/wqgsm/

ANOVAs were conducted using the post-deliberation measures as dependent variables, to determine whether misinformation exposure influenced juror perceptions and decision-making. Next, we focused on memory. We conducted a series of one way ANOVAs with planned contrasts to determine whether misinformation condition influenced participants' memory and misinformation acceptance. The free recall data violated the assumption of normality, so we also conducted robust ANOVAs using 10% trimmed means (Wilcox, 2012). Results of both approaches were the same, so we report the original ANOVA results here for ease of interpretation. Finally, we used a one way ANOVA with planned contrasts to determine whether the misinformation condition affected the perceived credibility of the jurors (see Supplementary Data Sheet S1 for these analyses).

To corroborate non-significant findings, we conducted exploratory Bayesian analyses via the Bayes Factor package (Morey and Rouder, 2018) in R (version 4.0.3; R Core Team, 2020). We implemented default priors to conduct these analyses as they make few assumptions about the data and offer a conservative test of the null hypothesis (Rouder et al., 2012). Bayes Factors quantify the evidence in favor of either the null or alternate hypotheses (Rouder et al., 2012). When reporting Bayes Factors, we use the interpretations provided by Jeffreys (1961) to indicate the strength of evidence for the null or alternate hypothesis. Bayes Factors of 1–3, 3–10, 10–30, 30–100, or > 100 reflect anecdotal, moderate, strong, very strong, and extreme evidence in favor of one hypothesis over the other, respectively.

3.2 Suspicion about deliberation

Participants were asked whether they noticed anything strange about the study, and to elaborate if they had, to determine whether they were suspicious about the nature of the deliberation. As seen in Table 3, 14.5% of participants believed that the deliberation chatroom was simulated and thus the other jurors in the chatroom were not real people (e.g., *"The other participants in the chat were bots"*—Participant 7). A further 24% of participants thought that the other jurors in the deliberation chatroom were either confederates (e.g., *"I do not believe the other jurors were real participants and were actually confederates"*—Participant 267) or that they were real participants, but were given different scenarios (e.g., *"it appears we were given different stories, perhaps to mimic real jurors different interpretations and memories?"*—Participant 15). Additionally, 28.4% of participants thought that the jurors in the deliberation chatroom had an incorrect recollection of the scenario (e.g., *"yes the discussion did not seem accurate and it made me question my own interpretation of the trial"*—Participant 238). Finally, about a third of participants (33.1%) were not suspicious of the simulated deliberation. A chi-square test revealed a significant association between type of suspicion and misinformation condition, χ^2 (N = 275) = 18.008, p = 0.021, φ_c = 0.181, such that participants in the contradictory group reported believing the deliberation was simulated above expected counts.

For each of the analyses reported below, we conducted the same analyses retaining just the participants who were not suspicious about the deliberation (N = 91). We will report when the analyses differed after accounting for suspicious participants.

3.3 Decision-making

Participants delivered a verdict, rated the defendant's guilt, and rated the complainant both before and after deliberation. Before the deliberation, 87.6% of participants delivered a verdict of "guilty," while 12.4% of participants delivered a verdict of "not guilty." After the deliberation, 86.9% of participants delivered a verdict of "guilty," while 13.1% of participants delivered a verdict of "not guilty." The descriptives for the pre- and post-deliberation measures based on misinformation condition are reported in Table 4.

At pre-deliberation, a chi-square analysis revealed no significant relation between misinformation condition and verdict, χ^2 (N = 275) = 4.066, p = 0.131, φ_c = 0.122, BF_{01} = 5.323. The ANOVAs revealed no significant differences in pre-deliberation guilt ratings (F(2,272) = 1.855, p = 0.158, η_p^2 = 0.013, BF_{01} = 4.834) and complainant credibility ratings (F(2,272) = 0.786, p = 0.457, η_p^2 = 0.006, BF_{01} = 12.636) based on misinformation condition. Overall, these analyses suggest that the randomization to misinformation condition was effective.

At post-deliberation, a chi-square analysis revealed no significant relation between misinformation condition and verdict, χ^2 (N = 275) = 2.928, p = 0.231, φ_c = 0.103, BF_{01} = 3.149. Additionally, one-way ANOVAs were conducted to explore whether differences in guilt and credibility ratings from pre- to post-deliberation differed as a function of misinformation condition. For both guilt and complainant credibility difference scores, the effect of misinformation condition was not significant (guilt: F(2,272) = 0.855, p = 0.426, η_p^2 = 0.006, BF_{01} = 11.879; credibility: F(2,272) = 0.842, p = 0.432, η_p^2 = 0.006, BF_{01} = 12.029).

3.4 Memory

3.4.1 Free recall

A one-way ANOVA was conducted to determine whether misinformation acceptance (i.e., reporting the misinformation from the e-deliberation) differed by misinformation condition. For misinformation acceptance, participants could receive a score ranging from 0 to 4 (as it would be implausible for participants in the

TABLE 3 Study 1—type of suspicion about the simulated deliberation, total and across misinformation conditions.

Type of suspicion	Total *n*	Pro-prosecution	Pro-defense	Contradictory	%
Simulated deliberation	40	10	8	22	14.5
Confederate used/ Scenarios manipulated	66	16	23	27	24.0
Jurors incorrect	78	33	24	21	28.4
No suspicion	91	33	34	24	33.1

TABLE 4 Study 1—pre- and post-deliberation verdict, defendant guilt rating, and complainant credibility ratings across misinformation condition.

	Misinformation condition		
	Pro-prosecution (n = 92)	Pro-defense (n = 89)	Contradictory (n = 94)
Pre-deliberation			
% Guilt	89.1	82	91.5
Defendant guilt rating	5.57 (1.32)	5.26 (1.34)	5.59 (1.18)
Complainant credibility	5.49 (0.99)	5.33 (0.92)	5.35 (1.00)
Post-deliberation			
% Guilt	85.9	83.1	91.5
Defendant guilt rating	5.66 (1.48)	5.46 (1.42)	5.63 (1.12)
Complainant credibility	5.47 (1.09)	5.21 (1.12)	5.31 (1.10)
Pre-post deliberation			
Defendant guilt rating	−0.09 (0.81)	−0.20 (0.98)	−0.04 (0.70)
Complainant credibility	0.02 (0.48)	0.12 (0.63)	0.04 (0.57)

Values for defendant guilt rating and complainant credibility represent mean rating (standard deviation in parentheses).

contradictory condition to recall both items of misinformation for the same detail). There was a significant effect of misinformation condition on misinformation acceptance, $F(2,272) = 5.954$, $p = 0.003$, $\eta_p^2 = 0.042$. As shown in Figure 1A, planned contrasts using Tukey's HSD revealed that participants who were exposed to consistent pro-prosecution misinformation accepted more misinformation than participants who were exposed to contradictory misinformation, $t(184) = 3.48$, $p = 0.002$, $d = 0.56$, 95% CI[0.10, 0.52]. No other contrasts were significant. It should be noted that when excluding all participants who were suspicious about the deliberation ($N = 91$), misinformation condition no longer had a significant effect on misinformation acceptance.

3.4.2 Source memory

Overall, 18.6% ($M = 0.186$, SD = 0.23) of misinformation items were misremembered (i.e., misinformation acceptance) as appearing during the trial. A one-way ANOVA was conducted to examine differences in misinformation acceptance between misinformation conditions. For the contradictory condition, the score entered into the ANOVA was the average misinformation acceptance score for pro-prosecution and pro-defense items. The ANOVA was significant, $F(2,272) = 11.30$, $p < 0.001$, $\eta_p^2 = 0.08$. As shown in Figure 1B, planned contrasts using Tukey's HSD procedure revealed that participants in the consistent pro-prosecution condition ($M = 0.27$, SD = 0.27) were significantly more likely to misattribute the misinformation as appearing in the trial than the consistent pro-defense group ($M = 0.13$, SD = 0.22) and contradictory group ($M = 0.16$, SD = 0.17), $t(179) = 4.47$, $p < 0.001$, $d = 0.57$, 95% CI[0.07, 0.23]; $t(184) = 3.64$, $p < 0.001$, $d = 0.49$, 95% CI[0.04, 0.19], respectively. No other contrasts were significant. However, the contrast comparing the pro-prosecution and contradictory groups was no longer significant when excluding participants who were suspicious of the deliberation ($N = 91$).

Given that there were different misinformation items (pro-prosecution and pro-defense), we conducted follow-up analyses to compare misinformation acceptance between conditions for each type of misinformation item separately. A repeated measures ANOVA revealed that participants in the contradictory condition were significantly more likely to misattribute pro-prosecution misinformation items ($M = 0.19$, SD = 0.22) as appearing in the trial than pro-defense items ($M = 0.12$, SD = 0.19), $F(1,93) = 8.32$, $p = 0.005$, $\eta_p^2 = 0.08$. Furthermore, a between-subjects ANOVA revealed that participants consistently exposed to pro-prosecution items were significantly more likely to misattribute this type of misinformation as appearing in the trial than the contradictory condition ($F[1,184] = 4.95$, $p = 0.027$, $\eta_p^2 = 0.03$), while participants consistently exposed to pro-defense items were no more likely to misattribute this type of information to the trial than the contradictory condition, $F(1,181) = 0.09$, $p = 0.765$, $\eta_p^2 = 0$, $BF_{01} = 5.882$.

4 Study 1 discussion

Study 1 evaluated whether different forms of misinformation introduced through jury deliberation influenced juror memory and decision-making about a common type of sexual assault case. Specifically, we compared exposure to misinformation items that consistently favored the prosecution's case, consistently favored the defense's case, or were contradictory (where participants received both pro-prosecution and pro-defense items). A key finding was that while exposure to misinformation during deliberation did not influence post-deliberation decision-making (e.g., verdict), it did influence participants' memory for the trial evidence. Mock-jurors were more susceptible to accepting misinformation that aligned with the prosecution's case than the defense's case, particularly when multiple jurors were in agreement (i.e., consistent condition) rather than disagreement (i.e., contradictory condition) about the misinformation.

The nature of the case and the pre-deliberation decision-making of participants is likely to explain why pro-prosecution misinformation distorted memory more than pro-defense misinformation or contradictory misinformation. The current study contained an excerpt of a sexual assault trial, that only featured the complainant's testimony of the event. We used such a case as this is reflective of many sexual assault trials, where only the complainant provides evidence-in-chief (e.g., New South Wales Law Reform Commission, 2020). The pre-deliberation decision-making measures showed a ceiling effect such that most participants (87.6%) provided a guilty verdict even before deliberation, and pre-deliberation ratings of defendant guilt

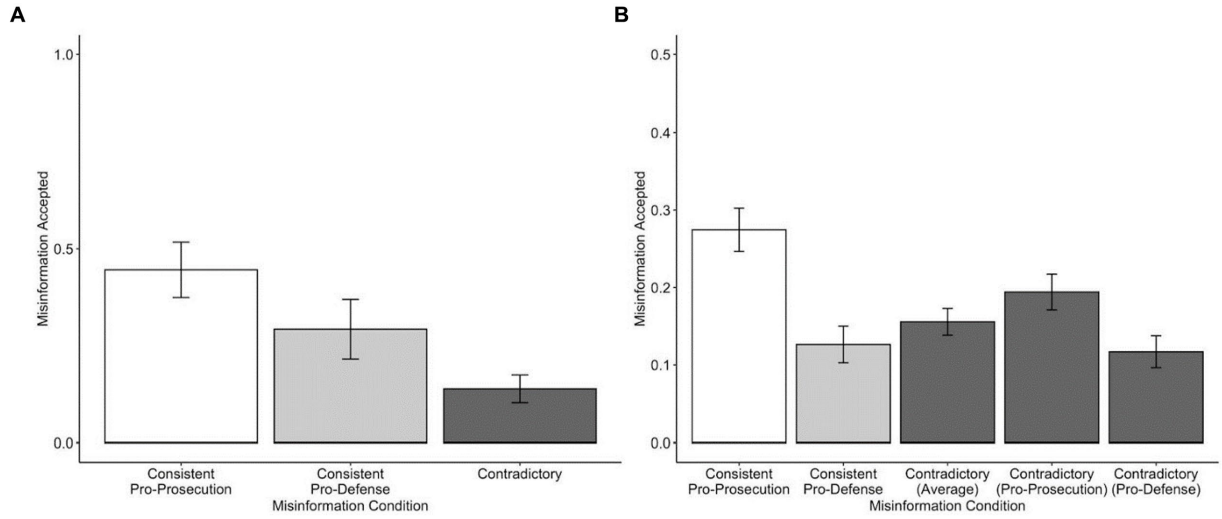


FIGURE 1
Study 1—misinformation acceptance as a function of misinformation condition in the recall **(A)** and source memory test **(B)**. **(A)** shows the total mean number of misinformation items that were misremembered as appearing during the trial during free recall across misinformation conditions. **(B)** shows the proportion of misinformation items that were misremembered as appearing during the trial during the source memory test across misinformation conditions. The "Contradictory (Average)" bar in Panel b represents the average proportion of misinformation items accepted across pro-prosecution and pro-defense items. Error bars represent standard error of the mean.

and complainant credibility were high (5.5 and 5.4 out of 7, respectively). Post-deliberation verdicts and ratings were very similar to their pre-deliberation counterparts. Therefore, the greater endorsement of pro-prosecution misinformation could be explained by the fact that this misinformation most closely reflected participants' beliefs about the case prior to deliberation, thus participants were more likely to misremember this type of information as occurring in the trial. Indeed, jury research has revealed that jurors may engage in *predecisional distortion*, where their evaluation of later case information is unconsciously influenced by the verdict that is leading in their mind (Carlson and Russo, 2001; Hope et al., 2004). The fact that most participants were leaning towards the prosecution's case pre-deliberation might also explain why misinformation exposure did not influence decision-making post-deliberation.

However, in addition to participants' prior beliefs about the case, social factors at the time of deliberation – such as conformity—may partially explain why pro-prosecution misinformation was accepted more in the consistent condition than the contradictory condition (Asch, 1956; Kaplan, 1984; Waters and Hans, 2009). Perhaps when participants saw that multiple jurors agreed that the misinformation was present during the trial, they felt pressure to conform with the group position. In contrast, when the jurors disagreed, participants may have felt less pressure to conform and therefore were more likely to reject the misinformation (Asch, 1956). Together, Study 1 findings suggest that cognitive, and possibly social, factors may influence misinformation acceptance during juror deliberations. More research is required, however, before solid conclusions can be made.

5 Study 2

Since participants' evaluation of the trial information in Study 1 was skewed towards the prosecution's case prior to the deliberation phase, it is difficult to determine to what extent, if any, the misinformation effect found in the pro-prosecution condition was influenced by the misinformation presented during the deliberation phase. To correct for this potential ceiling effect, in Study 2 we re-examined the effect of pro-prosecution and pro-defense misinformation on juror memory and decision-making, but with a more ambiguous sexual assault case (i.e., approximate even split of guilty and not-guilty pre-deliberation verdicts). The contradictory misinformation condition was not included in Study 2.

Another factor that might have impacted the validity of our findings in Study 1 was the high level of suspicion participants reported about the e-deliberation procedure. While our e-deliberation method did allow our participants to actively discuss the case with other 'jurors', just over a third of participants were suspicious about deliberation, citing that they believed it was fully simulated, that confederates were used, or that other participants were provided with alternate versions of the transcript which resulted in them receiving different information. For some analyses, results differed when the sample included versus excluded suspicious participants (e.g., free recall). Therefore, in Study 2 we used a methodology less likely to arouse suspicion in participants. Like Thorley et al. (2020), participants in Study 2 read a transcript of a deliberation, which contained misinformation about the trial evidence.

Finally, given that Study 1 revealed that the misinformation jurors are exposed to during deliberation can alter their memory for the trial, a secondary aim of Study 2 was to explore techniques to inoculate jurors from accepting misinformation mentioned during deliberation. Judicial instructions to the jury are one such technique. In a criminal trial context, jurors can be provided with instructions from the judge at the conclusion of a trial, but prior to the deliberation phase, to assist them in their decision-making. These instructions can include a range of topics, such as instructions to disregard inadmissible evidence (Steblay et al., 2006), instructions to help the jury understand legal concepts such as beyond reasonable doubt (Trimboli, 2008), and Henderson instructions to help them evaluate eyewitness testimony

(Dillon et al., 2017), to name a few. Despite popular support for judicial instructions, there is mixed empirical support for their effectiveness (e.g., Alvarez et al., 2016). However, most relevant to our study, eyewitness memory studies have found that participants who received a warning about having potentially encountered incorrect post-event information about an event showed a reduced misinformation effect compared to those that received no such warning (e.g., Echterhoff et al., 2005; Blank and Launay, 2014; Bulevich et al., 2022). While most of these studies have involved post-warnings where participants received the warning after misinformation exposure, a recent study found that providing a pre-warning (warning before misinformation exposure) was also effective in reducing the misinformation effect (Karanian et al., 2020). Based on this research, we expected that participants who received a judicial instruction about the harmful effects of misinformation would be less likely to accept misinformation mentioned during deliberation than those that received no instruction. From here onwards, we use the term "judicial instruction" to refer to this warning, as this is the language used to describe such warnings given by judges in jury research.

6 Study 2 method

6.1 Participants

Four-hundred and twenty-three participants initially took part in the study. The same eligibility requirements as Study 1 were applied (over 18 years of age, Australian citizen, fluent in English). The data from 84 participants were excluded for the following reasons: failing more than one attention check ($n = 3$), not completing the study ($n = 54$), invalid data entry ($n = 1$), or spending insufficient time reading the trial transcript (as indicated by reading times that were one standard deviation below the mean reading time [$M = 641.04$ s, SD = 397.72 s], $n = 26$). After applying exclusions, 339 participants were retained in the final analyses. An *a priori* power calculation using G*Power 3.1 (Faul et al., 2007) revealed that 265 participants were needed to detect a small to medium effect ($f = 0.20$) with 90% power for a 2 × 2 between-subjects design.

Participants had a Mean age of 29.40 years (SD = 11.72) and were predominantly female (66.1%, Male = 31.3%, Non-binary/Genderqueer/Gender fluid = 2.1%, Prefer not to say = 0.6%). Most participants were of European descent (77.9%), followed by Asian (12.7%), mixed ethnicity (3.8%), and Aboriginal or Torres Strait Islander (2.4%) (Other = 3.3%). Most participants (96.2%) had never served on a jury before.

Participants consisted of undergraduate psychology students ($n = 168$) and members of the community recruited via Prolific ($n = 171$). See Table 5 for a breakdown of demographic characteristics based on recruitment strategy. Prolific participants were significantly older than psychology students, $F(1, 337) = 145.399$, $p < 0.001$, $\eta_p^2 = 0.301$. There were also significant differences in participant gender between Prolific participants and psychology students, χ^2 ($N = 339$) = 45.856, $p < 0.001$, $\varphi_c = 0.368$, with Prolific participants having a more even split between male and female participants than psychology students. There were also differences in cultural background between the samples, χ^2 ($N = 339$) = 47.237, $p < 0.001$, $\varphi_c = 0.373$. While these demographic differences emerged, there were no differences in the frequency of psychology students and Prolific participants across the warning and misinformation conditions (both $ps > 0.485$). Additionally, there were no differences in gender or age distribution among the conditions (all $ps > 0.250$). Therefore, we will not conduct any further analyses between the two participant samples.

TABLE 5 Demographic characteristics of Study 2 participants based on recruitment strategy.

Demographics	Student participants ($n = 168$)	Prolific participants ($n = 171$)
Mean age	22.92 (6.78)	35.77 (12.07)
Gender (%)		
Female	83.3	49.1
Male	14.3	48.0
Non-binary/Genderqueer/Gender fluid	1.8	2.3
Prefer not to say	0.6	0.6
Previous jury experience (%)		
Yes	0.6	7.0
No	99.4	93.0
Cultural background (%)		
European/White	86.3	69.6
East Asian	0.6	9.9
Southeast Asian	0.6	11.1
Other	12.5	9.4

"Other" for cultural background includes cultural backgrounds with low cell counts: Aboriginal and Torres Strait Islander, African, Hispanic, Middle Eastern, Mixed, Pacific Islander, South Asian, Other. Standard deviations for mean age in parentheses.

6.2 Design

The current study employed a 2 × 2 between-subjects design, looking at the effects of a judicial instruction about misinformation (instruction vs. no instruction) and misinformation type (pro-prosecution vs. pro-defense) on juror memory and decision-making. Thus, participants were randomly assigned to one of four conditions: instruction/pro-prosecution ($n = 78$), instruction/pro-defense ($n = 92$), no instruction/pro-prosecution ($n = 83$), and no instruction/pro-defense ($n = 86$).

6.3 Materials

6.3.1 Trial transcript

Similar to Study 1, participants read a shortened trial transcript depicting an alcohol-involved sexual assault. We modified the case from Study 1 to create greater ambiguity, with the goal of achieving a more even split in pre-deliberation verdicts. The transcript in Study 2 featured an opening statement from the judge and both legal parties, with the issue of consent being disputed between the parties. The alleged victim, Daphne Livingstone, was then questioned by both the prosecution and defense. Daphne's testimony detailed that she attended her work Christmas party on the day of the alleged assault. After the Christmas party, she went to a bar with her colleague, Katie, who invited the accused, Alexander Smith, to join them. Daphne and

Alexander knew each other, but had only met once before in passing. A member of their group was removed from the premises by security, and so the group went back to Katie's house. Katie and the others in their group left to get food and drinks, leaving Daphne and Alexander alone. Daphne and Alexander kissed, and Daphne stated that she consented to this. When Alexander tried to take things further, Daphne verbally expressed that she did not want to go further as the others would be back soon. According to Daphne's testimony, Alexander ignored this and penetrated Daphne with his penis. Daphne was shocked and did not know what to do, as she had planned to stay at Katie's overnight and had no way of getting home. One month after the alleged assault, Daphne reported the alleged assault to police and was asked to undertake a medical examination.

At the conclusion of Daphne's testimony, the judge then gave a closing statement. The closing statement reminded jurors of their responsibilities and the burden of proof, and provided instructions about what the jurors should consider when reaching their decision. Pertinently, we manipulated whether the judge provided a specific instruction about the possibility of encountering misinformation from other jurors during the deliberation. Specifically, participants in the instruction condition were given the following information embedded in the judge's instructions:

> "You must be reminded that during your deliberations, it is possible that other jurors will remember the facts of the case differently to you, through no fault of their own. You should be aware of the possibility that your memory of the trial may be tainted or distorted by what other jurors say during the deliberation. You should try to correct these errors during your deliberations as much as possible, so that the decision that you collectively reach is derived from the correct version of events."

The no instruction condition did not receive the judicial instruction.

6.3.2 Deliberation transcript

Many participants were suspicious that the deliberation was fake in Study 1, and this suspicion had to be considered when interpreting the results. To mitigate suspicion in Study 2, we presented participants with a transcript of a fictional deliberation between four jurors, similar to Thorley et al. (2020). Participants were either provided with pro-prosecution or pro-defense misinformation for four of the details in the deliberation. Each juror in the transcript provided one misinformation item and one correct item from the trial during the deliberation. As in Study 1, we selected central misinformation items that, if remembered, would be likely to impact decisions on the case. The misinformation items targeted were those that related to common misconceptions and stereotypes about sexual assault, including the relationship between the complainant and defendant, the actions of the complainant during the assault, the time to report the assault, and the presence of physical injuries (see Carr et al., 2014; Australian Institute of Health and Welfare (AIHW), 2020). Two of the misinformation items were contradictory (i.e., contradicted the correct information from the trial) and two were additive (i.e., referred to details that were not mentioned in the trial). Table 6 presents misinformation and correct information provided by the simulated jurors based on experimental condition, and correct details for questions where misinformation was provided.

Participants were asked two questions about their perceptions of the deliberation. First, they were asked to rate the extent they believed the deliberation would be similar to the discussions that real jury members would have in a real deliberation of a sexual assault case (from 1 to 7, where 1 = *not at all similar* and 7 = *extremely similar*). Second, they were asked to rate how accurate they believed the jurors in the deliberation were in their memory of the information from the trial. An error in the formatting of the response options for this question emerged, therefore responses to the accuracy question will not be considered in the analyses.

6.4 Measures

6.4.1 Verdict and guilt ratings

Like Study 1, participants were asked to render a verdict of *guilty* or *not guilty* for the defendant with a justification, and rate the defendant's guilt, both pre- and post-deliberation. Participants also

TABLE 6 Misinformation and correct items in Study 2 based on misinformation condition.

Item	Pro-prosecution	Pro-defense	Correct
Misinformation			
Status of the relationship	Strangers—had never met before	Friends—had met several times before	Acquaintances—had met once before*
Behavior during assault	Attempted to push Alex off	Did not push Alex off	N/A*
Reporting of assault	One day after	Two months later	One month later*
Presence of injuries	Bruises consistent with being held down	No bruises consistent with being held down	N/A*
Correct			
Doing the day of assault			Attending a work Christmas party
Why they left the bar			Friend removed from premises for spilling a drink
Why the complainant stayed			She had no way of getting home
Information requested by police			Undertake a medical examination

*Indicates correct details from the trial that were not featured in any version of the deliberation. These details are provided in the table to compare to the misinformation items. N/A refers to additive misinformation items (i.e., there was no reference to these details in the trial transcript).

rated the perceived strength of the prosecution and defense cases at pre- and post-deliberation (on a scale from 1 to 7, where 1 = *weak*, 4 = *uncertain*, and 7 = *strong*). Pre- and post-deliberation decision-making occurred before and after misinformation exposure, respectively.

6.4.2 Credibility

Like Study 1, participants were asked to rate their perception of the complainant's honesty, believability, credibility, and accuracy both pre- and post-deliberation. There was high internal consistency in pre- and post-deliberation ratings for all four items (pre-deliberation: *Cronbach's* $\alpha = 0.946$; Post-deliberation: *Cronbach's* $\alpha = 0.961$). Therefore, like in Study 1, these ratings were aggregated to form a single pre-deliberation and post-deliberation complainant credibility score.

6.4.3 Recall memory

Participants completed both a free recall and a cued recall memory questionnaire. For the free recall memory task, participants were given an open-ended prompt and asked to recall what they remember about the alleged sexual assault case described in the trial transcript. They were specifically instructed to only report facts about the case (i.e., what the complainant alleged happened), as opposed to any of the instructions they were provided by the judge. They were also instructed to base their responses off their own memory of the events, and encouraged to report everything they could remember while being as accurate and detailed as possible. Participants were given unlimited time to complete free recall.

The cued recall questionnaire was added to Study 2, to more precisely measure misinformation acceptance. The questionnaire consisted of 12 specific questions about the case, presented to participants in a randomized order. The cued recall questions were selected so that four questions focused on case facts where misinformation was provided during the deliberation (the relationship, time to report, events during alleged assault, information provided to police), four questions focused on case facts where correct information was provided during the deliberation (events of the day, why they left the bar, what the complainant did after alleged assault, information sought by police), and four questions focused on case facts that were not mentioned during the deliberation (discussion at the bar, why the parties were left alone, information being disputed, what the parties did at the bar). Participants were encouraged to rely only on their own memory of the trial when answering the cued recall questions, and to be as accurate and detailed as possible.

6.4.3.1 Recall coding

The same coding system from Study 1 was employed, whereby independent scorers coded participant responses using the categories of misinformation accepted, misinformation rejected, and correct accepted. Using these categories, Scorer 1 (HC) coded 100% of participant responses in both free and cued recall. Scorer 2 (SB) coded 51% of free recall responses, and Scorer 3 (GR) coded 51% of cued recall responses. For both free and cued recall, the Intraclass Correlation Coefficients revealed good to excellent reliability for all coding categories (all *ICC*s > 0.836). Given the acceptable reliability, the coding from Scorer 1 was used in the analyses.

6.4.4 Source memory

Like Study 1, Study 2 included a source memory test. Participants were given the same instructions as they were given in Study 1. The source memory test in Study 2 consisted of 36 items. Eight of the items related to the misinformation from the deliberation; the correct answer to these items depended on the misinformation condition participants were assigned, like Study 1 (*Deliberation only* if a misinformation item was relevant to their experimental condition, or *Neither* if a misinformation item was not relevant to their experimental condition). There were 4 items related to the correct information from the deliberation (correct answer: *Both trial and deliberation*), 8 items related to correct information not covered in the deliberation (correct answer: *Trial only*), and 16 new filler items (correct answer = *Neither trial nor deliberation*). The total number of critical source memory errors (i.e., misinformation acceptance) was the key dependent variable which was calculated in the same way as the consistent conditions in Study 1 (i.e., "Trial Only" and 'Trial and Deliberation' scores for the misinformation items were summed together). Since participants were exposed to the same number of misinformation items in Study 2, we did not calculate proportion scores. Analyses for whether there were differences across misinformation and instruction condition with regard to correct (trial only), correct (trial and deliberation), and new source memory items (16 filler statements +4 statements relating to misinformation that they were not exposed to) are provided in the Supplementary Data Sheet S1.

6.4.5 Rape myth acceptance

As the misinformation items in Study 2 reflected common misconceptions about sexual assault that may most likely be introduced during deliberations in such cases, it is possible that participants with greater rape myth acceptance may be most susceptible to reporting misinformation. Therefore, in Study 2, participants completed the adapted version of the Illinois Rape Myth Acceptance Scale—Subtle Version (IRMA-S; Thelan and Meadows, 2022) to assess rape myth acceptance. We included the 22-items from the IRMA-S that assessed rape myth acceptance, but did not include the filler items. Participants indicated their agreement with each statement on a 5-point Likert scale (where 1 = "*strongly disagree*" and 5 = "*strongly agree*"). Scores on each item were summed to form a total score of rape myth acceptance. Reverse scoring was applied for three of the items. Possible scores ranged from 22 to 110, with higher scores indicating greater rape myth acceptance. The IRMA-S has high internal consistency ($\alpha = 0.93$) and good validity when evaluated with diverse participant samples. The adapted version we used in the current study also had high internal consistency ($\alpha = 0.88$).

6.4.6 Attention and manipulation checks

Like Study 1, we included several attention and manipulation checks. There were three instructional manipulation checks spread throughout the study; participants were required to answer at least two of these questions correctly for their data to be retained in the data analysis. Additionally, we were interested in determining whether participants who received the judicial instruction about being exposed to misinformation during the deliberation remembered receiving this instruction. Memory for the judicial instruction was measured in two ways. First, participants were asked to summarize the judicial instruction in their own words. We coded participants' responses based on whether they mentioned being warned about potential for

misinformation to occur in the deliberation or not. Second, participants were asked three yes/no questions about whether the judge had provided a warning about three different topics. Two of these questions related to distractor topics (unreliability of physical evidence and burden of proof), whereas the other question asked whether participants were warned about memory being tainted by other jurors. Collectively, these manipulation checks provided useful information on the effectiveness of the judicial instruction and assessed whether participants understood the judicial instruction. Participants were also asked what they believed the purpose of the study was to probe suspicion about the aims of the study.

6.5 Procedure

Participants took part in Study 2 online. Given the sensitive nature of the case, participants were provided the contact details of support services before being asked to read the trial transcript. Like Study 1, participants were told to read the transcript in full and not to make any notes while reading the transcript. The transcript was split into separate pages on the online survey host, and we recorded the time participants spent on each page. After reading the transcript, participants completed the pre-deliberation decision-making measures (verdict, guilt rating, strength of case ratings, complainant credibility ratings). Participants were then required to read a transcript of a fictitious deliberation about the case and to imagine that they are forming part of the jury on this case and are involved in the discussion. As with the trial transcript, participants were told to read the deliberation transcript in full and to not make any notes. The deliberation transcript contained either pro-prosecution or pro-defense misinformation, depending on the condition participants had been randomly assigned to. Participants completed the same measures of decision-making post-deliberation. Then, they completed the free recall, cued recall, and source memory tasks. Next, participants answered two questions about their perceptions of the deliberation. They then completed the adapted version of the IRMA-S, following which they completed the manipulation checks. Finally, participants provided demographic information and were debriefed about the study. All aspects of Study 2 were approved by the University of Newcastle Human Research Ethics Committee (protocol number: H-2022-0079).

6.6 Transparency statement

We reported how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in this study. The hypotheses, design, measures, and analysis plan were pre-registered on OSF. The registration can be found at https://osf.io/vcwkj/. Below we note deviations from our pre-registration transparently. All experimental materials and data (dataset, output, and code) are available on the OSF: https://osf.io/wdse5/.

6.7 Transparent deviations

First, after commencing data collection for this study we realized that the survey software we used to host this study was not randomizing participants to all experimental conditions. This meant we had data in some cells and not in others. We recruited more participants to even up the number of participants in each condition (to ensure we could conduct our planned analyses without violating assumptions). For this reason, we exceeded our pre-registered sample size.

Second, we did not pre-register any specific data exclusion criteria focused on checking whether participants had spent sufficient time on the trial transcript pages to ensure that they had read the materials. As the experimental manipulations were contained in the trial materials, it is critical that participants read the materials properly. After data was collected, three authors who had not had contact with the data (ND, FN, GR) decided that it would be reasonable to exclude participants who had a total average reading time of more than one standard deviation below the overall sample total average reading time as participants who viewed the trial materials for this period of time were unlikely to have properly read the materials.

Third, we also planned to look at whether perceptions of the accuracy of the jurors in the deliberation predicted misinformation acceptance, but an error with the programming of the scale anchors for this question meant that the question likely did not make sense to participants. For this reason, we have not analyzed this data as planned.

7 Study 2 results

7.1 Overview and analysis plan

First, we reported on the results of the manipulation check relating to the judicial instruction manipulation using descriptive statistics (paraphrase test) and chi-square analyses (forced choice). Then, we conducted a series of mixed-methods ANOVAs to determine whether there was any effect of misinformation type and judicial instruction on changes in decision-making (from pre- to post-deliberation), as well as to ensure that there were no existing differences in decision-making prior to the deliberation. We conducted two logistic regressions (for pre- and post-deliberation verdict) to determine whether misinformation type and judicial instruction influenced verdicts. Next, we conducted several two way ANOVAs to determine whether misinformation and judicial instruction conditions influenced participants' memory and misinformation acceptance in all three memory tasks (free recall, cued recall, source memory). We conducted moderation analyses to determine whether rape myth acceptance moderated the strength of the relationship between misinformation condition and misinformation acceptance. We also conducted mediation analyses to determine whether the relationship between misinformation condition and post-deliberation decision-making was mediated by misinformation acceptance. Finally, we conducted a regression to determine whether perceptions of the realism of the deliberation predicted misinformation acceptance (see Supplementary Data Sheet S1 for results of this analysis). Like Study 1, we reported Bayes Factors alongside the frequentist analyses.

7.2 Manipulation checks

To check the memorability of the judicial instruction about encountering misinformation during deliberation, participants

completed a paraphrase test (summarizing the judge's instructions in their own words) and answered a yes/no question indicating whether the judge provided such a warning. The paraphrase test revealed that only 11 participants (6.5%) in the instruction condition reported the misinformation instruction. No participants in the no instruction condition spontaneously reported the misinformation instruction. When participants were asked to state whether the judge had warned them about potentially encountering misinformation during the deliberation, 74.7% of participants in the instruction condition correctly responded "yes," compared to 36.1% of participants in the no instruction condition incorrectly responding "yes." The chi-square analysis revealed a significant relation between judicial instruction condition and responses to this manipulation check, with "yes" responses above expected counts for the instruction condition, and below expected counts for the no instruction condition, χ^2 (1, N = 339) = 51.148, $p < 0.001$, $\varphi_c = 0.388$, $BF_{10} > 100$.

7.3 Decision-making

At pre-deliberation, 66.7% of participants delivered a verdict of "guilty" and 33.3% of participants delivered a verdict of "not guilty." After deliberation, 61.4% of participants delivered a verdict of "guilty" and 38.6% of participants delivered a verdict of "not guilty".

Two hierarchical logistic regressions were conducted—one for pre-deliberation and one for post-deliberation—to determine whether misinformation and instruction conditions predicted verdicts. In block 1 of each model, we added the main effects of misinformation type and instruction, and in block 2, we added the misinformation × instruction interaction. As Table 7 demonstrates, misinformation type, instruction, and the interaction did not significantly predict pre-deliberation verdicts, but misinformation type did predict verdicts post-deliberation. Specifically, the odds of delivering a guilty verdict following the deliberation were 2.982 times higher in the pro-prosecution condition compared to the pro-defense condition. A Bayesian chi-square also revealed extreme evidence for a relationship between misinformation condition and post-deliberation guilt ratings, $BF_{10} > 100$.

Mixed methods ANOVAs were conducted to determine whether misinformation and instruction conditions impacted decision-making (guilt ratings, complainant credibility, strength of case) both pre- and post-deliberation. Table 8 provides the descriptive statistics accompanying these analyses. For all decision-making measures, there was a significant time × misinformation interaction. Specifically, there were no differences in guilt ratings, complainant credibility ratings, and the perceived strength of the prosecution and defense cases between the pro-prosecution and pro-defense misinformation conditions before the deliberation (all ps > 0.312, all BF_{01}s > 5.131). After the deliberation, participants in the pro-prosecution misinformation condition gave significantly higher ratings of defendant guilt (using Tukey's LSD: $p = 0.016$, $BF_{10} = 2.033$). None of the other pairwise comparisons between pro-prosecution and pro-defense misinformation at post-deliberation were significant after applying Tukey's LSD (all ps > 0.088, all BF_{01}s > 1.933). Participants in the pro-defense misinformation condition showed a significant decrease in ratings of guilt and complainant credibility from pre- to post-deliberation (ps < 0.001, BF_{10}s > 100), and a significant increase in ratings of the strength of the defense's case ($p < 0.001$, $BF_{10} = 7.980$), but there was no significant change in ratings of the strength of the prosecution's case from pre- to post-deliberation ($p = 0.087$, $BF_{01} = 2.327$). For participants in the pro-prosecution condition, ratings of the strength of the prosecution's case significantly increased from pre- to post-deliberation ($p < 0.001$, $BF_{10} = 73.944$), with no other significant differences emerging after applying Tukey's LSD (all ps > 0.083, BF_{01}s > 1.367). Taken together, these findings suggest that exposure to different forms of misinformation may alter decision-making from pre- to post-deliberation in the direction of the legal party the misinformation favors. This appears to be particularly the case for decision-making related to guilt (verdicts and guilt ratings).

TABLE 7 Study 2—hierarchical logistic regressions for pre- and post-deliberation verdicts with misinformation and judicial instruction conditions as predictors.

	B	S.E.	Sig.	OR	95% CI
Pre-deliberation					
Block 1					
Misinformation	0.082	0.232	0.722	1.086	0.690, 1.710
Instruction	−0.340	0.232	0.142	0.712	0.452, 1.121
Block 2					
Misinformation × Instruction	−0.370	0.465	0.426	0.691	0.278, 1.717
Post-deliberation					
Block 1					
Misinformation	1.092	0.235	< 0.001*	2.982	1.880, 4.728
Instruction	0.193	0.231	0.403	1.213	0.771, 1.908
Block 2					
Misinformation × Instruction	0.296	0.471	0.529	1.345	0.535, 3.383

Verdict was coded as 0 (not guilty) and 1 (guilty). Pro-prosecution misinformation was coded as 0 and pro-defense misinformation was coded as 1. No warning was coded as 0 and warning as 1. OR = odds ratio. 95% confidence intervals are for odds ratio scale.
*$p < 0.05$.

TABLE 8 Study 2—pre- and post-deliberation verdict, defendant guilt rating, complainant credibility rating, and strength of evidence ratings across misinformation conditions.

	Pre-deliberation		Post-deliberation		Pre-post deliberation		Misinformation × time interaction			
	Pro-prosecution (N = 161)	Pro-defense (N = 178)	Pro-prosecution (N = 161)	Pro-defense (N = 178)	Pro-prosecution (N = 161)	Pro-defense (N = 178)	F	p	η^2	BF_{10}
% Guilt	65.84	67.42	74.53	49.44	–	–	–	–	–	
Defendant guilt rating	5.30 (1.64)	5.41 (1.48)	5.44 (1.58)	5.00 (1.69)	−0.13 (0.85)	0.41 (1.04)	27.20	0.001	0.007	>100
Complainant credibility	5.39 (1.24)	5.52 (1.11)	5.43 (1.22)	5.25 (1.26)	−0.04 (0.51)	0.27 (0.57)	27.46	0.001	0.004	>100
Strength of prosecution's case	4.77 (1.71)	4.86 (1.64)	5.05 (1.72)	4.74 (1.59)	−0.27 (0.93)	0.12 (0.99)	14.20	0.001	0.004	15.140
Strength of defense's case	2.90 (1.53)	2.84 (1.52)	2.82 (1.52)	3.10 (1.52)	0.08 (0.77)	−0.26 (1.15)	10.25	0.002	0.003	3.188

Values for defendant guilt rating, complainant credibility, strength of prosecution's case, and strength of defense's case represent mean rating (standard deviation in parentheses).

7.4 Memory

For misinformation acceptance in free and cued recall, there was no significant effect of misinformation type (free recall: $F[1,335] = 0.196$, $p = 0.658$, $\eta^2 < 0.001$, $BF_{01} = 11.212$; cued recall: $F[1,335] = 0.196$, $p = 0.658$, $\eta^2 < 0.001$, $BF_{01} = 9.872$), instruction (free recall: $F[1,335] = 0.157$, $p = 0.692$, $\eta^2 < 0.001$, $BF_{01} = 11.535$; cued recall: $F[1,335] = 2.461$, $p = 0.118$, $\eta^2 = 0.007$, $BF_{01} = 4.008$), and no significant misinformation type × instruction interaction (free recall: $F[1,335] = 0.230$, $p = 0.632$, $\eta^2 < 0.001$, $BF_{01} = 101.248$; cued recall: $F[1,335] = 2.614$, $p = 0.107$, $\eta^2 = 0.008$, $BF_{01} = 15.589$). However, for critical source memory errors, there was a significant main effect of misinformation type ($F[1,335] = 4.513$, $p = 0.034$, $\eta^2 = 0.013$), with participants in the pro-defense condition ($M = 1.39$, SD = 0.95) misattributing the misinformation to the trial more than participants in the pro-prosecution condition ($M = 1.17$, SD = 1.07). However, the Bayes Factor indicated ambiguous evidence for a lack of difference, $BF_{01} = 1.467$. There was no significant effect of instruction ($F[1,335] = 3.857$, $p = 0.050$, $\eta^2 = 0.011$, $BF_{01} = 1.936$) and no misinformation type × instruction interaction ($F[1,335] = 0.016$, $p = 0.898$, $\eta p^2 < 0.001$, $BF_{01} = 6.543$) for critical source memory errors. Mediation analyses conducted in JASP also revealed that misinformation acceptance in any of the memory tasks did not mediate the relationship between misinformation type (pro-prosecution vs. pro-defense) and any of the post-deliberation decision-making outcomes (all $ps > 0.362$).

We also explored whether source misattributions were more likely to occur when exposed to the misinformation (as opposed to spontaneous misattributions of misinformation from the other condition). This was the case. These analyses are reported in the Supplementary Data Sheet S1.

7.5 Rape myth acceptance

The average score on the adapted version the IRMA-S was 40.85 (SD = 11.80), with a minimum score of 22 and a maximum score of 84 (higher scores indicating greater rape myth acceptance). We used a series of linear regression analyses to determine whether rape myth acceptance was a significant moderator in the relationship between misinformation type and misinformation acceptance. As rape myth acceptance was continuous, scores were mean centered using a z-transformation. For both free and cued-recall, rape myth acceptance was not a significant moderator (both $ps > 0.856$). However, for source memory, rape myth acceptance was a significant moderator of the effect of misinformation type on critical source memory errors, $\beta = -0.172$, $t(3,335) = -2.213$, $p = 0.028$. Simple effects analyses were conducted by looking at the relationship between rape myth acceptance scores and critical source memory errors for pro-prosecution and pro-defense participants separately. Rape myth acceptance significantly predicted critical source memory errors in the pro-prosecution condition, $\beta = 0.233$, $t(1,159) = 3.015$, $p = 0.003$, but did not predict critical source memory errors in the pro-defense condition, $\beta = 0.008$, $t(1,176) = 0.107$, $p = 0.915$.

8 General discussion

Across two studies, we evaluated whether different forms of misinformation introduced during jury deliberation influenced mock-juror memory and decision-making in sexual assault trials. In addition, Study 2 explored whether being warned about the harmful effects of misinformation during judicial instructions would inoculate mock-jurors from accepting misinformation presented during deliberation. In general, Study 1 revealed that pro-prosecution misinformation was more likely to be accepted as appearing in the trial than pro-defense misinformation, particularly when jurors were in agreement (i.e., consistent condition) rather than in disagreement (i.e., contradictory condition). Misinformation type did not influence decision-making. Using a more ambiguous sexual assault trial transcript to address the unequal split of pre-deliberation verdicts in Study 1, Study 2 found a limited effect of misinformation on memory, where participants in the pro-defense condition were more likely to misattribute the misinformation to the trial than participants in the

pro-prosecution condition (source memory only). The relationship between misinformation condition and critical source memory errors was also moderated by rape myth acceptance. In contrast to Study 1, Study 2 did find an effect of misinformation type on decision-making. That is, after deliberation, being exposed to pro-defense misinformation led to a decrease in ratings of defendant guilt, complainant credibility, and an increase in the strength of the defense's case. However, there was no effect of the judicial instruction about misinformation on participants' decision-making.

8.1 The effect of misinformation on juror memory

Decades of research into the misinformation effect in an eyewitness context has found that exposure to misinformation can distort later memory (see Loftus, 2005; Frenda et al., 2011, for reviews). Building on this literature, our findings showed that jurors too may be vulnerable to the misinformation effect; the findings from both studies revealed that if jurors encounter misinformation in jury deliberations, they may make source monitoring errors and come to misattribute the misinformation as evidence presented during the trial (Johnson et al., 1993). Our findings align with Thorley et al. (2020), who also found that when participants were exposed to misinformation through a written deliberation, they misattributed the misinformation as trial evidence, compared to participants who were not exposed to misinformation.

The current research expanded on Thorley et al.'s (2020) study by exploring the effects of both pro-prosecution and pro-defense misinformation on juror memory. Our findings revealed that the type of misinformation that our mock-jurors misremembered as trial evidence varied across studies. Participants in Study 1 were more likely to misremember pro-prosecution misinformation as forming part of the trial than pro-defense misinformation, while the opposite was true for Study 2. The differences in findings between studies may be partly attributable to differences in methodology. As we noted above, the greater endorsement of pro-prosecution over pro-defense misinformation in Study 1 could be because participants' beliefs about the case were skewed towards the prosecution's case prior to deliberation. In Study 2, we developed a case vignette with a more even split of pre-deliberation verdicts, finding that participants were more likely to misremember pro-defense than pro-prosecution misinformation as trial evidence (source memory only). Participants' tendency to accept pro-defense misinformation over pro-prosecution misinformation might be explained by laypeople's general mistrust of rape allegations (Webster et al., 2018; Minter et al., 2021). The pro-defense misinformation items were designed to discredit the prosecution case. Thus, perhaps participants more readily endorsed the pro-defense misinformation as evidence from the trial because these items aligned with people's attitudes towards rape allegations, whereas the pro-prosecution items did not.

In Study 2, we found that higher endorsement of rape myths was positively associated with source memory errors for participants exposed to pro-prosecution misinformation. This could be explained by the fact that the pro-prosecution misinformation items in Study 2 aligned with several pervasive rape myths. Take, for example, the following items from the IRMA-S (Thelan and Meadows, 2022) used in our study: *"If a woman does not physically fight back, she cannot really say she was raped"* and *"Sexual assault probably did not happen if the woman has no bruises or marks."* In Study 2, participants were not presented with any information in the trial about whether the victim fought back or had bruises/marks. However, the pro-prosecution misinformation items for these facts were worded in the same direction as the rape myths (*"Attempted to push off"*/*"Bruises consistent with being held down"*), whereas the pro-defense misinformation items were worded in the opposite direction to the rape myths ("*Did not push off"*/*"No bruises consistent with being held down"*). Therefore, it is logical that participants who endorsed these rape myths (i.e., had higher rape myth acceptance scores) were more likely to misattribute the misinformation as having appeared in the trial transcript. This suggests that rape myths, which some scholars argue function as a schema for how sexual violence occurs (Süssenbach et al., 2012), are influencing participants' memories and views of the case. Problematically, most sexual assaults do not occur in ways that are consistent with rape myths (Dinos et al., 2015).

Other research on the misinformation effect shows similar findings. For instance, research that has explored the phenomenon of fake news has shown that people are more likely to misremember fake news when it aligns with their existing beliefs on the topic (e.g., Abortion: Murphy et al., 2019; Feminism: Murphy et al., 2021)—a finding referred to as *ideological congruency.* This idea of accepting misinformation that is more in line with one's pre-existing beliefs, similar to confirmation bias, may also explain why pro-prosecution misinformation was more likely to be accepted in Study 1; most participants already believed in the defendant's guilt before exposure to misinformation, and the pro-prosecution misinformation (serving to give credibility to the complainant's case) may have strengthened these beliefs.

The fact that we only found misinformation effects for source memory and not recall memory in Study 2 might be explained by the type of memory evoked by the different memory tests. The source memory test simply required participants to recognize and then endorse/not endorse statements. When participants were presented with the pro-defense misinformation statements, it might have activated their attitudes about rape, leading them to 'recognize' these details as forming a part of the trial evidence. The recall tests required participants to actively retrieve information about the trial. The cues provided by these questions might not have been enough to elicit information pertaining to the misinformation. Additionally, recall and recognition-based memory tasks place different demands on the individual's ability to monitor and control the information provided (Koriat and Goldsmith, 1996). Specifically, individuals are often more accurate in recall compared to recognition memory tasks, often at the expense of providing fewer details. Confidence plays an important role in how individuals respond, as recall memory tasks allow individuals to withhold low-confidence responses, while recognition memory tasks do not. Measuring memory confidence in future studies will allow us to determine whether differences in misinformation acceptance across tasks can be explained through strategic regulation processes. Notwithstanding, considering misinformation acceptance using both recall and recognition-based memory tasks is a strength of our study, as previous research studies looking at misinformation in

jury contexts have mostly only employed source memory tasks (e.g., Ruva and McEvoy, 2008; Ruva and Guenther, 2015; Thorley et al., 2020), and looking at only source memory does not give a full picture of how misinformation exposure in discussion settings leads to memory conformity.

8.2 The effect of misinformation on juror decision-making

In Thorley et al. (2020), participants were presented with pro-prosecution misinformation during the deliberation phase. They found that the more pro-prosecution misinformation participants attributed to the trial, the more likely they were to give a guilty verdict. We were interested in determining whether misinformation that favors the defense would have the opposite effect, through increased acquittals and unfavorable perceptions of the complainant's case. While we found no evidence of misinformation effects on decision-making in Study 1 (likely due to ceiling effects in decision-making at pre-deliberation), Study 2 revealed a pattern consistent with our hypotheses. From pre- to post-deliberation, participants' decision-making tendencies (e.g., guilty vs. not guilty) shifted based on the type of misinformation they were exposed to during deliberation. Specifically, misinformation that discredited (i.e., pro-defense misinformation) rather than strengthened the prosecution's case had a more pronounced effect on decision-making—exposure to pro-defense misinformation led to less favorable perceptions of the prosecution's case (e.g., decreases in conviction rates, ratings of defendant guilt, complainant credibility).

We expected that misinformation acceptance would be the mechanism by which exposure to different types of misinformation would affect legal decisions. In other words, we expected that the effect of the type of misinformation on post-deliberation decision-making would be mediated by participants' distorted memory about the trial evidence. We did not find any evidence that misinformation acceptance mediated this relationship. Interestingly, this means that the mechanism by which exposure to incorrect trial facts during the deliberation affects post-deliberation decision-making is not one of memory. This is in contrast to previous research looking at the effects of pre-trial publicity on mock-juror memory and decision-making, where critical source memory errors mediated the relationship between exposure to pre-trial publicity and guilt ratings (Ruva et al., 2007; Ruva and Guenther, 2015). However, an alternative mediator that we did not explore in the current study was evidence interpretation (Ruva and Guenther, 2015), such that exposure to different forms of pre-trial publicity appears to alter the way in which trial facts are interpreted (i.e., which legal party they favor), which in turn influences decision-making. It would be useful in future research to also consider the role of evidence interpretation in the relationship between misinformation exposure during deliberation and post-deliberation decision-making.

8.3 Judicial instructions

Although there is popular support for judicial instructions to reduce jurors' reliance on extra-legal factors, there is mixed evidence for their effectiveness (e.g., Alvarez et al., 2016). In Study 2, we found that the judicial instruction had no effect on mock jurors' memory errors or decisions made about the trial. This is consistent with research which indicates that many judicial instructions do not reduce jurors' reliance on extra-legal factors like curative instructions for pre-trial publicity (e.g., Steblay et al., 2006) or Henderson instructions about evaluating eyewitness testimony (e.g., Dillon et al., 2017).

There are several barriers which can prevent judicial instructions from assisting jurors in their decision-making. First, if jurors do not understand what they are being told to do in the instruction, then they were not able to apply the instruction to their decision-making (Baguley et al., 2020). In Study 2, we checked (mock) jurors' comprehension of the judicial instructions about misinformation using a paraphrase test in which participants were asked to summarize the judge's instructions in their own words. We found that only 6.5% of participants mentioned the instruction about misinformation in their responses. When we explicitly asked participants whether they remembered receiving an instruction from the judge about misinformation, 25% of participants said they did not remember this. Collectively, these results suggest that the judge's instruction about the effect of misinformation, while understood, was perhaps not salient. This may explain why the instruction did not affect either memory outcomes or decision-making in Study 2.

A second barrier to judicial instructions helping jurors in their decision-making is if the strategy offered to assist the jury in the instruction is ineffective (Baguley et al., 2020). The instruction we used adopted two common strategies used in instructions that target bias — making the jury aware of the potential bias and encouraging them to challenge inaccurate information they heard from other jurors. Given participants could not actively challenge misinformation from other jurors as the deliberation in Study 2 was presented as a transcript, perhaps the instruction may be effective when participants engage in an interactive or live deliberation. That said, in other research on jury deliberations, jurors find it difficult to contradict information presented by other jurors when they think it is wrong (e.g., Stasser and Titus, 2003). An important avenue for future research would be to test the effectiveness of this instruction when participants engage in a more interactive deliberation.

8.4 Limitations and future directions

There are a few limitations of the current research to consider. The methods we used to simulate jury deliberation in our studies lacked ecological validity. While our e-deliberation method in Study 1 did allow our participants to actively discuss the case with other 'jurors', most participants were suspicious about deliberation. In Study 2, we adopted the same methodology as Thorley et al. (2020) where participants read a transcript of a deliberation. While this method allowed us to account for suspicion, it lacks ecological validity as participants did not actively participate in the deliberation. Despite this, many participants in Study 2 still reported the transcript of the deliberation to be similar to the type of discussions they believed real jurors would have (see Supplementary Data Sheet S1). The fact that we found that the type of misinformation did influence mock juror memory even in our less interactive deliberations suggests that it will be important to examine misinformation effects in more interactive jury deliberations. For example, future research might consider

holding discussions in-person or via video conferencing where a confederate introduces misinformation about the trial evidence. The use of confederates to implant misinformation in a consistent fashion has been used extensively in the eyewitness memory research around co-witness discussion (e.g., Gabbert et al., 2004; Paterson et al., 2009; Eisen et al., 2017), and could be effective in creating a realistic jury deliberation that arouses little suspicion. Alternatively, participants could be exposed to two different versions of the trial transcript and naturally introduce misinformation. Such an approach has also been used successfully in the eyewitness memory literature (e.g., Gabbert et al., 2003, 2006, 2007; Paterson et al., 2011).

There are several important avenues for future research arising from this work. One is investigating how participant characteristics, in particular participant gender, might affect how (mock) jurors remember case evidence and make decisions in sexual assault cases. While many individual studies provide evidence of participant gender effects on decisions made in sexual assault cases, a recent review suggests that evidence for participant gender effects is mixed and context dependent (Gravelin et al., 2019). Many investigations of participant gender in decision-making in sexual assault cases are exploratory (e.g., Nitschke et al., 2021) which may mean that studies are underpowered to adequately detect the interaction of participant gender and other factors in the study (e.g., Giner-Sorolla, 2018). An important avenue for future research in this area will be to investigate potential participant gender effects on acceptance of misinformation in jury deliberations in sexual assault cases.

Another question for future research concerns the efficacy of a warning from the judge to prevent misinformation during jury deliberations from affecting memory and decision-making. In Study 2, we found that the warning delivered via judicial instruction did not seem to assist jurors to avoid misinformation during deliberations. Research suggests that judicial instructions can be revised to make them more effective (e.g., Steblay et al., 2006). Future research should consider whether the content of the judicial instruction can be made more effective by drawing on the research literature to support accurate memory. As with warnings to eyewitnesses, the timing of a judicial instruction can also be important to whether it assists the jury (e.g., Alvarez et al., 2016). Future research should investigate when and how many times the judicial instruction needs to be given to help jurors avoid misinformation during deliberation.

Finally, both of our studies explored the impact of different types of misinformation on mock-juror memory and decision-making in a sexual assault trial. We chose sexual assault as the offense because of the diverse attitudes that jurors may hold and thus bring into deliberations, making it particularly important to understand the consequences of misinformation exposure in these cases. Future research should determine whether our findings generalize to other crime types, as well as whether misinformation introduced about other forms of evidence can similarly influence memory and decision-making (e.g., forensic experts, eyewitnesses).

9 Conclusion

Across two studies, our findings provide initial insight into the effect of misinformation exposure during jury deliberations on juror memory. The findings suggest that jurors may misremember trial details when exposed to misinformation provided by fellow jurors, but that this memory distortion may depend on the nature of the misinformation items, the consistent repetition of these misinformation items, and the beliefs and attitudes held about legal cases and how certain offences occur (e.g., rape myths). Future research should continue to consider the important role of memory in jury deliberation contexts and the factors that increase or decrease memory distortion due to misinformation exposure. With a clearer understanding of memory conformity effects in criminal trial settings, we can determine to what extent memory distortion subsequently biases legal decision-making.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: Study 1: https://osf.io/wqgsm/, Study 2: https://osf.io/wdse5/.

Ethics statement

The studies involving humans were approved by the University of Sydney Human Research Ethics Committee (protocol number: 2019/947) and the University of Newcastle Human Research Ethics Committee (protocol number: H-2022-0079). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

HC, ND, FN, and GR contributed to funding acquisition, data curation, methodology, project administration, and validation for the current project, contributed to the conceptualization and investigation of Study 1, contributed to the formal analysis of Study 1, and contributed to the original manuscript preparation and writing. HC, ND, FN, GR, and NW contributed to the conceptualization and investigation of Study 2. HC, GR, and SB contributed to the formal analysis of Study 2. ND contributed to data visualization in Study 1. HC contributed to data visualization in Study 2 and developed software and was involved in supervision for the current project. HC and ND provided resources for the current project. All authors contributed to the article and approved the submitted version.

Funding

We would like to thank the Society of Australasian Social Psychologists, the School of Psychology at the University of Queensland, and the School of Psychological Sciences at the University of Newcastle for providing funding for the current project. We would also like to thank the Macquarie Minds and Intelligence Initiative at Macquarie University for providing funding for publication costs.

Acknowledgments

The authors would like to thank Sarah Deck (Centre of Investigative Interviewing, Griffith Criminology Institute) and Emily Murphy (School of Psychological Sciences, the University of Newcastle) for their assistance on this project.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1232228/full#supplementary-material

References

Abbey, A., Zawacki, T., Buck, P. O., Clinton, A. M., and McAuslan, P. (2004). Sexual assault and alcohol consumption: what do we know about their relationship and what types of research are still needed? *Aggress. Violent Behav.* 9, 271–303. doi: 10.1016/S13591789(03)00011-9

Alvarez, M. J., Miller, M. K., and Bornstein, B. H. (2016). ""It will be your duty…:" the psychology of criminal jury instructions" in *Advances in psychology and law*. eds. M. K. Miller and B. H. Bornstein, vol. *1* (Switzerland: Springer International Publishing), 119–158.

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychol. Monogr.* 70, 1–70. doi: 10.1037/h0093718

Australian Institute of Health and Welfare (AIHW) (2020). Sexual assault in Australia. Available at: https://www.aihw.gov.au/getmedia/0375553f-0395-46cc-9574-d54c74fa601a/aihw-fdv5.pdf.aspx?inline=true

Baguley, C. M., McKimmie, B. M., and Masser, B. M. (2020). Re-evaluating how to measure jurors' comprehension and application of jury instructions. *Psychol. Crime Law* 26, 53–66. doi: 10.1080/1068316X.2019.1634195

Blank, H., and Launay, C. (2014). How to protect eyewitness memory against the misinformation effect: a meta-analysis of post-warning studies. *J. Appl. Res. Mem. Cogn.* 3, 77–88. doi: 10.1016/j.jarmac.2014.03.005

Blank, H., Panday, A., Edwards, R., Skopicz-Radkiewicz, E., Gibson, V., and Reddy, V. (2021). Double misinformation: effects on eyewitness remembering. *J. Appl. Res. Memory Cogn.* 11, 97–105. doi: 10.1016/j.jarmac.2021.08.001

Bulevich, J. B., Gordon, L. T., Hughes, G. I., and Thomas, A. K. (2022). Are witnesses able to avoid highly accessible misinformation? Examining the efficacy of different warnings for high and low accessibility postevent misinformation. *Mem. Cogn.* 50, 45–58. doi: 10.3758/s13421-021-01255-1

Carlson, K. A., and Russo, J. E. (2001). Biased interpretation of evidence by mock jurors. *J. Exp. Psychol. Appl.* 7, 91–103. doi: 10.1037//1076-898X.7.2.91

Carr, M., Thomas, A. J., Atwood, D., Muhar, A., Jarvis, K., and Wewerka, S. S. (2014). Debunking three rape myths. *J. Forensic Nurs.* 10, 217–225. doi: 10.1097/JFN.0000000000000044

Connolly, D. A., Price, H. L., Lavoie, J. A. A., and Gordon, H. M. (2008). Perceptions of predictors of children's credibility of a unique event and an instance of a repeated event. *Law Hum. Behav.* 32, 92–112. doi: 10.1007/s10979-006-9083-3

Costabile, K. A., and Klein, S. B. (2005). Finishing strong: Recency effects in juror judgments. *Basic Appl. Soc. Psychol.* 27, 47–58. doi: 10.1207/s15324834basp2701_5

Cox, P. (2015). Violence against women: additional analysis of the Australian Bureau of Statistics Personal Safety Survey, 2012. Australia's National Research Organisation for Women's safety. Available at: https://www.anrows.org.au/publication/violence-against-women-in-australia-additional-analysis-of-the-australian-bureau-of-statistics-personal-safety-survey-2012/

Cullen, H. J., and Monds, L. A. (2020). Jury simulation studies: to exclude or not to exclude participants based on a lack of comprehension of the case? *Appl. Cogn. Psychol.* 34, 1224–1233. doi: 10.1002/acp.3695

Deck, S. L., and Paterson, H. M. (2021a). Adults' ability to particularise an occurrence of a repeated event. *Appl. Cogn. Psychol.* 35, 685–692. doi: 10.1002/acp.3792

Deck, S. L., and Paterson, H. M. (2021b). Adults also have difficulty recalling one instance of a repeated event. *Appl. Cogn. Psychol.* 35, 286–292. doi: 10.1002/acp.3736

Devine, D. J. (2012). *Jury decision making: the state of science*. New York and London: New York University Press.

Devine, D. J., Olafson, K. M., Jarvis, L. L., Bott, J. P., Clayton, L. D., and Wolfe, J. M. T. (2004). Explaining jury verdicts: is leniency bias for real? *J. Appl. Soc. Psychol.* 34, 2069–2098. doi: 10.1111/j.1559-1816.2004.tb02691.x

Dilevski, N., Paterson, H. M., and van Golde, C. (2020b). Investigating the effect of emotional stress on adult memory for single and repeated events. *Psychol. Public Policy Law* 26, 425–441. doi: 10.1037/law0000248

Dilevski, N., Paterson, H. M., Walker, S. A., and van Golde, C. (2020a). Adult memory for instances of a repeated event: a preliminary review. *Psych. Psychol. Law* 28, 711–732. doi: 10.1080/13218719.2020.1837031

Dillon, M. K., Jones, A. M., Bergold, A. N., Hui, C. Y. T., and Penrod, S. D. (2017). Henderson instructions: do they enhance evidence evaluation? *J. Forensic Psychol. Res. Pract.* 17, 1–24. doi: 10.1080/15228932.2017.1235964

Dinos, S., Burrowes, N., Hammond, K., and Cunliffe, C. (2015). A systematic review of juries' assessment of rape victims: do rape myths impact on juror decision-making? *Int. J. Law Crime Justice* 43, 36–49. doi: 10.1016/j.ijlcj.2014.07.001

Echterhoff, G., Hirst, W., and Hussy, W. (2005). How eyewitnesses resist misinformation: social postwarnings and the monitoring of memory characteristics. *Mem. Cogn.* 33, 770–782. doi: 10.3758/BF03193073

Eisen, M. L., Gabbert, F., Ying, R., and 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. *J. Appl. Res. Mem. Cogn.* 6, 274–282. doi: 10.1016/j.jarmac.2017.01.009

Faul, F., Erdfelder, E., Lang, A., and Buchner, A. (2007). G*power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/BF03193146

Frenda, S. J., Nichols, R. M., and Loftus, E. F. (2011). Current issues and advances in misinformation research. *Curr. Dir. Psychol. Sci.* 20, 20–23. doi: 10.1177/0963721410396620

Gabbert, F., Memon, A., and Allan, K. (2003). Memory conformity: can eyewitnesses influence each other's memories for an event? *Appl. Cogn. Psychol.* 17, 533–543. doi: 10.1002/acp.885

Gabbert, F., Memon, A., Allan, K., and Wright, D. B. (2004). Say it to my face: examining the effects of socially encountered misinformation. *Leg. Criminol. Psychol.* 9, 215–227. doi: 10.1348/1355325041719428

Gabbert, F., Memon, A., and Wright, D. B. (2006). Memory conformity: Disentangling the steps toward influence during a discussion. *Psychon. Bull. Rev.* 13, 480–485. doi: 10.3758/BF03193873

Gabbert, F., Memon, A., and Wright, D. B. (2007). I saw it for longer than you: the relationship between perceived encoding duration and memory conformity. *Acta Psychol.* 124, 319–331. doi: 10.1016/j.actpsy.2006.03.009

Giner-Sorolla, R. (2018). Powering your interaction. Available at: https://approachingblog.wordpress.com/2018/01/24/powering-your-interaction-2/

Glaser, J., Martin, K. D., and Kahn, K. B. (2015). Possibility of death sentence has divergent effect on verdicts for black and white defendants. *Law Hum. Behav.* 39, 539–546. doi: 10.1037/lhb0000146

Gravelin, C. R., Biernat, M., and Bucher, C. E. (2019). Blaming the victim of acquaintance rape: individual, situational, and sociocultural factors. *Front. Psychol.* 9. doi: 10.3389/fpsyg.2018.02422

Gray, J. M., and Horvath, M. A. H. (2018). "Rape myths in the criminal justice system" in *Women and the criminal justice system*. eds. E. Milne, K. Brennan, N. South and J. Turton (Switzerland: Palgrave Macmillan), 15–41.

Heydon, J. D. (2013). *Cross on evidence* (*9th ed.*). Chatswood, New South Wales: Lexis Nexus Butterworths.

Hirst, W., and Stone, C. B. (2017). "The effects of collaborative remembering on trial verdicts" in *The psychology of juries*. ed. M. B. Kovera (Worcester, MA: American Psychological Association), 37–57.

Hope, L., and Gabbert, F. (2018). Memory at the sharp end: the costs of remembering with others in forensic contexts. *Top. Cogn. Sci.* 11, 609–626. doi: 10.1111/tops.12357

Hope, L., Memon, A., and McGeorge, P. (2004). Understanding pretrial publicity: Predecisional distortion of evidence by mock jurors. *J. Exp. Psychol. Appl.* 10, 111–119. doi: 10.1037/1076-898X.10.2.111

Huff, C. R., Rattner, A., and Sagarin, E. (1986). Guilty until proved innocent: wrongful conviction and public policy. *Crime Delinq.* 32, 518–544. doi: 10.1177/0011128786032004007

Innocence Project (2021). The causes of wrongful conviction. Available at: https://innocenceproject.org/causes-wrongful-conviction./

Ito, H., Barzykowski, K., Grzesik, M., Gülgöz, S., Gürdere, C., Janssen, S. M., et al. (2019). Eyewitness memory distortion following co-witness discussion: a replication of Garry, French, Kinzett, and Mori (2008) in ten countries. *J. Appl. Res. Mem. Cogn.* 8, 68–77. doi: 10.1016/j.jarmac.2018.09.004

Jay, A. C. V., Stone, C. B., Meksin, R., Merck, C., Gordon, N. S., and Hirst, W. (2019). The mnemonic consequences of jurors' selective retrieval during deliberation. *Top. Cogn. Sci.* 11, 627–643. doi: 10.1111/tops.12435

Jeffreys, H. (1961). *The theory of probability*. New York, NY: Oxford University Press.

Johnson, M. K., Hashtroudi, S., and Lindsay, D. S. (1993). Source monitoring. *Psychol. Bull.* 114, 3–28. doi: 10.1037/0033-2909.114.1.3

Kaplan, M. F. (1984). How do people influence in jury deliberation? A social psychological view. *Behav. Sci. Law* 2, 407–412. doi: 10.1002/bsl.2370020407

Karanian, J. M., Rabb, N., Wulff, A. N., Torrance, M. G., Thomas, A. K., and Race, E. (2020). Protecting memory from misinformation: warnings modulate cortical reinstatement during memory retrieval. *Proc Natl Acad Sci U S A* 117, 22771–22779. doi: 10.1073/pnas.2008595117

Koo, T. K., and Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* 15, 155–163. doi: 10.1016/j.jcm.2016.02.012

Koriat, A., and Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychol. Rev.* 103, 490–517. doi: 10.1037/0033-295x.103.3.490

Leverick, F. (2020). What do we know about rape myths and juror decision making? *Int. J. Evid. Proof* 24, 255–279. doi: 10.1177/136571270923157

Loftus, E. F. (2005). Planting misinformation in the human mind: a 30-year investigation of the malleability of memory. *Learn. Mem.* 12, 361–366. doi: 10.1101/lm.94705

Lovakov, A., and Agadullina, E. R. (2021). Empirically derived guidelines for effect size interpretation in social psychology. *Eur. J. Soc. Psychol.* 51, 485–504. doi: 10.1002/ejsp.2752

MacLean, C. L., Coburn, P. I., Chong, K., and Connolly, D. L. (2018). Breaking script: Deviations and postevent information in adult memory for a repeated event. *Appl. Cogn. Psychol*, 32, 474–486. doi: 10.1002/acp.3421

Matsuo, K., and Itoh, Y. (2016). Effects of emotional testimony and gruesome photographs on mock jurors' decisions and negative emotions. *Psychiatry Psychol. Law* 23, 85–101. doi: 10.1080/13218719.2015.1032954

Minter, K., Carlisle, E., and Coumarelos, C. (2021). "Chuck her on a lie detector"—investigating Australians' mistrust in women's reports of sexual assault (research report, 04/2021). ANROWS. Available at: https://anrowsdev.wpenginepowered.com/wp-content/uploads/2021/10/ANROWS-NCAS-Mistrust-of-reporting-SV.pdf

Mitchell, K. J., and Zaragoza, M. S. (2001). Contextual overlap and eyewitness suggestibility. *Mem. Cogn.* 29, 616–626. doi: 10.3758/BF03200462

Mojtahedi, D., Ioannou, M., and Hammond, L. (2018). Group size, misinformation and unanimity influences on co-witness judgements. *J. Forensic Psychiatry Psychol.* 29, 844–865. doi: 10.1080/14789949.2018.1439990

Monson, C. M., Langhinrichsen-Rohling, J., and Binderup, T. (2000). Does "no" really mean "no" after you say "yes"?: attributions about date and marital rape. *J. Interpers. Violence* 15, 1156–1174. doi: 10.1177/088626000015011003

Morey, R. D., and Rouder, J. N. (2018). Bayes factor: Computation of Bayes factors for common designs (0.9.12–4.2.) [computer software]. Available at: https://CRAN.R-project.org/package=BayesFactor

Murphy, G., Loftus, E. F., Hofstein-Grady, R. H., Levine, L. J., and Greene, C. M. (2019). False memories for fake news during Ireland's abortion referendum. *Psychol. Sci.* 30, 1449–1459. doi: 10.1177/0956797619864887

Murphy, G., Murray, E., and Gough, D. (2021). Attitudes towards feminism predict susceptibility to feminism-related fake news. *Appl. Cogn. Psychol.* 35, 1182–1192. doi: 10.1002/acp.3851

New South Wales Law Reform Commission. (2020). *Report 148: consent in relation to sexual offences*. New South Wales Law Reform Commission. Available at: https://www.lawreform.justice.nsw.gov.au/Documents/Publications/Reports/Report%20148.pdf

Nitschke, F. T., Masser, B. M., McKimmie, B. M., and Riachi, M. (2021). Intoxicated but not incapacitated: are there effective methods to assist juries in interpreting evidence of voluntary complainant intoxication in cases of rape? *J. Interpers. Violence* 36, 4335–4359. doi: 10.1177/0886260518790601

Oppenheimer, D. M., Meyvis, T., and Davidenko, N. (2009). Instructional manipulation checks: detecting satisficing to increase statistical power. *J. Exp. Soc. Psychol.* 45, 867–872. doi: 10.1016/j.jesp.2009.03.009

Paterson, H. M., and Kemp, R. I. (2005). Co-witness discussion: a survey of police officers' attitudes, knowledge, and behavior. *Psychiatry Psychol. Law* 12, 424–434. doi: 10.1375/pplt.12.2.424

Paterson, H. M., and Kemp, R. I. (2006a). Co-witnesses talk: a survey of eyewitness discussion. *Psychol. Crime Law* 12, 181–191. doi: 10.1080/10683160512331316334

Paterson, H. M., and Kemp, R. I. (2006b). Comparing methods of encountering post-event information: the power of co-witness suggestion. *Appl. Cogn. Psychol.* 20, 1083–1099. doi: 10.1002/acp.1261

Paterson, H. M., Kemp, R. I., and Forgas, J. P. (2009). Co-witnesses, confederates, and conformity: effects of discussion and delay on eyewitness memory. *Psychiatry Psychol. Law* 16, S112–S124. doi: 10.1080/13218710802620380

Paterson, H. M., Kemp, R. I., and Ng, J. R. (2011). Combating co-witness contamination: attempting to decrease the negative effects of discussion on eyewitness memory. *Appl. Cogn. Psychol.* 25, 43–52. doi: 10.1002/acp.1640

Pritchard, M. E., and Keenan, J. M. (1999). Memory monitoring in mock-jurors. *J. Exp. Psychol. Appl.* 5, 152–168. doi: 10.1037/1076-898x.5.2.152

Pritchard, M. E., and Keenan, J. M. (2002). Does jury deliberation really improve jurors' memories? *Appl. Cogn. Psychol.* 16, 589–601. doi: 10.1002/acp.816

R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Available at: https://www.R-project.org/

Rouder, J. N., Morey, R. D., Speckman, P. L., and Province, J. M. (2012). Default Bayes factors for ANOVA designs. *J. Math. Psychol.* 56, 356–374. doi: 10.1016/j.jmp.2012.08.001

Rubínová, E., Blank, H., Koppel, J., and Ost, J. (2020b). Schema and deviation effects in remembering unfamiliar stories. *Br. J. Psychol.* 112, 180–206. doi: 10.1111/bjop.12449

Rubínová, E., Blank, H., Ost, J., and Fitzgerald, R. (2020a). Structured word-lists as a model of basic schemata: deviations from content and order in a repeated event paradigm. *Memory* 28, 309–322. doi: 10.1080/09658211.2020.1712421

Ruva, C. L., and Guenther, C. C. (2015). From the shadows into the light: how pretrial publicity and deliberation affect mock jurors' decisions, impressions, and memory. *Law Hum. Behav.* 39, 294–310. doi: 10.1037/lhb0000117

Ruva, C. L., and Guenther, C. C. (2017). Keep your bias to yourself: how deliberating with differently biased others affects mock-jurors' guilt decisions, perceptions of the defendant, memories, and evidence interpretation. *Law Hum. Behav.* 41, 478–493. doi: 10.1037/lhb0000256

Ruva, C., and McEvoy, C. (2008). Negative and positive pretrial publicity affect juror memory and decision making. *J. Exp. Psychol. Appl.* 14, 226–235. doi: 10.1037/1076-898X.14.3.226

Ruva, C., McEvoy, C., and Bryant, J. B. (2007). Effects of pre-trial publicity and jury deliberation on juror bias and source monitoring errors. *Appl. Cogn. Psychol.* 21, 45–67. doi: 10.1002/acp.1254

Salerno, J. M., Peter-Hagene, L. C., and Jay, A. C. V. (2019). Women and African Americans are less influential when they express anger during group decision making. *Group Process. Intergroup Relat.* 22, 57–79. doi: 10.1177/1368430217702967

Sandys, M., and Dillehay, R. C. (1995). First-ballot votes, predeliberation dispositions, and final verdicts in jury trials. *Law Hum. Behav.* 19, 175–195. doi: 10.1007/bf01499324

Stasser, G., and Titus, W. (2003). Hidden profiles: a brief history. *Psychol. Inq.* 14, 304–313. doi: 10.1080/1047840x.2003.9682897

Steblay, N., Hosch, H. M., Culhane, S. E., and McWethy, A. (2006). The impact on juror verdicts of judicial instruction to disregard inadmissible evidence: a meta-analysis. *Law Hum. Behav.* 30, 469–492. doi: 10.1007/s10979-006-9039-7

Süssenbach, P., Bohner, G., and Eyssel, F. (2012). Schematic influences of rape myth acceptance on visual information processing: an eye-tracking approach. *J. Exp. Soc. Psychol.* 48, 660–668. doi: 10.1016/j.jesp.2011.12.005

Thelan, A. R., and Meadows, E. A. (2022). The Illinois Rape Myth Acceptance Scale - Subtle Version: Using an adapted measure to understand the declining rates of rape myth acceptance. *Journal of Interpersonal Violence*, 37, NP17807-NP17833. doi: 10.1177/08862605211030013

Thorley, C. (2016). Note taking and note reviewing enhance jurors' recall of trial information. *Appl. Cogn. Psychol.* 30, 655–663. doi: 10.1002/acp.3240

Thorley, C., Beaton, L., Deguara, P., Jerome, B., Khan, D., and Schopp, K. (2020). Misinformation encountered during a simulated jury deliberation can distort jurors' memory of a trial and bias their verdicts. *Leg. Criminol. Psychol.* 25, 150–164. doi: 10.1111/lcp.12174

Thorley, C., and Dewhurst, S. A. (2007). Collaborative false recall in the DRM procedure: effects of group size and group pressure. *Eur. J. Cogn. Psychol.* 19, 867–881. doi: 10.1080/09541440600872068

Trimboli, L. (2008). Juror understanding of judicial instructions in criminal trials. *Crime and Justice Bulletin* 119, 1–15.

Vredeveldt, A., Hildebrandt, A., and van Koppen, P. J. (2016). Acknowledge, repeat, rephrase, elaborate: witnesses can help each other remember more. *Memory* 24, 669–682. doi: 10.1080/09658211.2015.1042884

Vredeveldt, A., and van Koppen, P. J. (2018). Recounting a common experience: On the effectiveness of instructing eyewitness pairs. *Front. Psychol.* 9:284. doi: 10.3389/fpsyg.2018.00284

Waters, N. L., and Hans, V. P. (2009). A jury of one: opinion formation, conformity, and dissent on juries. *J. Empir. Leg. Stud.* 6, 513–540. doi: 10.1111/j.1740-1461.2009.01152.x

Webster, K., Diemer, K., Honey, N., Mannix, S., Mickle, J., Morgan, J., et al. (2018). Australians' attitudes to violence against women and gender equality. Findings from the 2017 National Community Attitudes towards violence against women survey (NCAS) (research report, 03/2018). ANROWS. Available at: https://d2rn9gno7zhxqg.cloudfront.net/wp-content/uploads/2019/12/05062144/anr001-NCAS-report-WEB-1019.pdf

Wilcox, R. R. (2012). *Introduction to robust estimation and hypothesis testing. 3rd* Edn. San Diego, CA: Academic Press.

Wright, D. B., Self, G., and Justice, C. (2000). Memory conformity: exploring misinformation effects when presented by another person. *Br. J. Psychol.* 91, 189–202. doi: 10.1348/000712600161781

Zaragoza, M. S., and Lane, S. M. (1994). Source mis-attribution and the suggestibility of eyewitness memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 934–945. doi: 10.1037//0278-7393.20.4.934

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 28 March 2024
DOI 10.3389/fpsyg.2023.1213996

OPEN ACCESS

EDITED BY
Henry Otgaar,
Maastricht University, Netherlands

REVIEWED BY
David Copeland,
University of Nevada, Las Vegas, United States
Charlie Frowd,
University of Central Lancashire,
United Kingdom

*CORRESPONDENCE
Heather D. Flowe
✉ h.flowe@bham.ac.uk

RECEIVED 28 April 2023
ACCEPTED 09 October 2023
PUBLISHED 28 March 2024

CITATION
Deering K, Colloff MF, Bennett TC and Flowe HD (2024) Does presenting perpetrator and innocent suspect faces from different facial angles influence the susceptibility of eyewitness memory? An investigation into the misinformation effect and eyewitness misidentification.
Front. Psychol. 14:1213996.
doi: 10.3389/fpsyg.2023.1213996

Does presenting perpetrator and innocent suspect faces from different facial angles influence the susceptibility of eyewitness memory? An investigation into the misinformation effect and eyewitness misidentification

Kara Deering, Melissa F. Colloff, Tia C. Bennett and Heather D. Flowe*

Applied Memory Lab, School of Psychology, University of Birmingham, Birmingham, United Kingdom

Introduction: This study investigated the effects of face angle congruency across stages of a misinformation paradigm on lineup discrimination accuracy.

Methods: In a between-subjects design, participants viewed a mock crime with the perpetrator's face from the front or profile angle. They then read a news report featuring an innocent suspect's image from the same or different angle as the perpetrator had been shown. A subsequent lineup manipulated perpetrator presence and viewing angle of the lineup members, who were all shown either from the front or in profile.

Results: No significant difference emerged in identification errors based on angle congruency between stages. However, accuracy was higher when faces were shown from the front angle, both during the initial event and the lineup, compared to the profile angle.

Discussion: The results of this research underscore the importance of considering viewing angles in the construction of lineups.

KEYWORDS

misinformation effect, eyewitness misidentification, eyewitness accuracy, eyewitness susceptibility, facial angles

1. Introduction

In June 1984, notorious serial killer Ted Bundy challenged his lineup identification process in the Court of Appeal, arguing that he was innocent, and that the witness identified him in error, solely because she had previously seen his picture in a newspaper story about the crime (Bundy, 1984). The prosecution countered that the newspaper image did not influence the witness' memory because it showed Bundy's face from the front, whereas the witness observed the perpetrator from a different angle, namely in profile view, during the crime. Supporting this, the witness stated that her identification was based on her initial memory of Bundy's face from the profile angle. Further, the image of Bundy's face she identified from the 10-image photographic lineup was also in profile view. The court dismissed Bundy's appeal, implying the prosecution's argument was more

convincing. This study empirically examines the arguments put forward in this case, testing whether memory impairment arising from exposure to a suspect's face depends on the congruence between the angle from which the perpetrator and the suspect are viewed by a witness.

From the defense's perspective, whether the viewing angles of the perpetrator and newspaper suspect corresponded was immaterial. Rather, the witness's post-event encounter with Bundy's newspaper image, regardless of angle, altered her original memory of the perpetrator, leading her to misidentify Bundy, exemplifying a phenomenon known as the misinformation (MI) effect. The MI effect refers to a memory impairment that arises from exposure to misleading information about an earlier witnessed event that individuals subsequently integrate or substitute into their memory of the original event (Ayers and Reder, 1998).

Research has shown that eyewitness identification accuracy can be influenced by misleading post event information, including misleading face descriptions (Loftus and Greene, 1980) and composites (Topp-Manriquez et al., 2014; Sporer et al., 2020). The mechanism behind the incorporation of MI into the witness's memory for original event has been widely researched and the subject of numerous debates. Some argue that misinformation overwrites or weakens the original memory traces (e.g., Loftus et al., 1978). Others have proposed that memory traces for the original and misleading information coexist, with interference (Bekerian and Bowers, 1983; Chandler, 1991) or source monitoring difficulties (Johnson et al., 1993) hindering accurate memory retrieval. Researchers have also extensively studied the boundary conditions of the effect, such as the whether the source of the misinformation is authoritative (Zaragoza et al., 2007).

Poorly encoded event details have been reported to be particularly susceptible to the influence of misleading information (Loftus and Greene, 1980). This susceptibility may be especially notable when the encoding involves a profile view of a perpetrator's face. Key facial features such as the eyes, nose, and mouth, critical for accurate facial identification, are less visible from a profile angle (McKelvie, 1976; Fraser et al., 1990). This observation, combined with the holistic nature of face processing (Taubert et al., 2011), may lead to incomplete face encoding from profile views. Recent studies support this claim, indicating a decrease in lineup discrimination accuracy when witnesses encode a perpetrator's face in profile rather than from a frontal view (Colloff et al., 2021). Consequently, a witness may be more susceptible to misleading post-event information when the perpetrator's face is encoded solely from a profile angle, a hypothesis that we will refer to hereafter as the *encoding strength hypothesis*.

The impact of the angle of face presentation extends beyond the encoding phase to post-event information processing. The similarity between the original event and misleading information significantly contributes to the misinformation effect (Loftus, 1977). For instance, witnesses are more likely to incorporate post-event information into their memories when it is similar in nature to the original event, as demonstrated by the impact of shared contextual information on false memory formation (Carpenter et al., 2022). In line with the prosecution's argument, these results lead to the hypothesis that the misinformation effect is more likely when the intervening innocent suspect's face is presented from the same angle as the perpetrator, a hypothesis that we will refer to hereafter as the *facial angle congruency hypothesis*.

In testing our hypotheses, it is important to control for the angle of the lineup faces at test, even though police lineups typically show the faces from the front. The encoding specificity principle posits that the overlap between the cues at learning and test impacts memory performance (Tulving and Thomson, 1973). Consistent with this, discrimination accuracy, defined as the witness's ability to distinguish between guilty and innocent suspects, is higher when the angle of the lineup faces aligns with the encoding angle (Colloff et al., 2021). This alignment of cues across encoding and the lineup might reduce the size of the misinformation effect, particularly if the angle of the test faces matches the angle of the perpetrator's face during the crime, as was the case for the witness who identified Bundy.

The misinformation stage itself is an integral part of the encoding process and therefore necessitates consideration of face angle. Both Campbell et al. (2007) and Yamashita (1996) have argued that a recognition test presented in a format like the misinformation leads to an increased misinformation effect. Therefore, this study also explores the impact of face angle congruence between the misinformation face and the lineup members on discrimination accuracy. Specifically, we explored the possibility that witnesses are more easily misled when the angle of the faces shown during the misinformation and test stages matches.

2. Method

Full ethical approval for the current research was granted by the University of Birmingham Ethics Committee.

2.1. Design

The current hypotheses and analysis plan were pre-registered on the Open Science Framework before data were collected. A factorial between-subjects design was used, where participants were randomly assigned to one of eight conditions: 2 (encoding and test view: front, profile) x 2 (misinformation suspect view: front, profile) x 2 (lineup type: target-absent, target-present). Target-absent (TA) lineups contained the misinformation suspect presented among five fillers. The target-present (TP) lineups contained the guilty culprit among five fillers. A between-subjects design was used to avoid learning effects. The misinformation suspect and guilty culprit were never presented in the same lineup, and suspect position in the lineup was randomized for each participant. The facial angle shown during the lineup (i.e., at test) always matched the facial position shown at encoding. Therefore, it was also possible to collapse across conditions such that participants either received congruent facial angles (front encoding, front misinformation suspect, front lineup (FFF); profile encoding, profile misinformation suspect, profile lineup (PPP)) or incongruent facial angles (front encoding, profile misinformation suspect, front lineup (FPF); profile encoding, front misinformation suspect, profile lineup (PFP)) information. Table 1 summarizes each condition and the attendant abbreviation.

2.2. Participants

Our preregistered data collection stopping rule was 2,000 participants.[1] The sample size was based on collapsing across conditions to answer the research questions. Using mean differences and standard deviations observed in Mickes et al. (2012) as a guide, a

1 www.osf.io/5pr8n

TABLE 1 Table to show Front (F), Profile (P), Target-Present (TP) and Target-Absent (TA) experimental conditions.

Encoding facial position	Facial position of misinformation suspect	Test: lineup condition and facial position	Condition summary	Facial angle congruency	Total per condition
Front	Front	Front; Target-Present	FFF-TP	Congruent	258
Front	Front	Front; Target-Absent	FFF-TA	Congruent	269
Front	Profile	Front; Target-Present	FPF-TP	Incongruent	268
Front	Profile	Front; Target-Absent	FPF-TA	Incongruent	252
Profile	Profile	Profile; Target-Present	PPP-TP	Congruent	251
Profile	Profile	Profile; Target-Absent	PPP-TA	Congruent	251
Profile	Front	Profile; Target-Present	PFP-TP	Incongruent	251
Profile	Front	Profile; Target-Absent	PFP-TA	Incongruent	251

power analysis indicated that, with a minimum of 250 participants per between-subjects condition, power would exceed 80%. We determined the sample size needed for >80% power to detect significant misinformation effect within each lineup condition. A bespoke power calculation tool developed for eyewitness lineup procedures was used.[2] The misinformation effect size was based on effect sizes from the literature (Longmore et al., 2008; Bülthoff et al., 2019; Colloff et al., 2021), and it was reframed in terms of possible condition pAUC ratios, and used a Bonferroni-corrected alpha level based on the number of comparisons to be made (i.e., alpha = 0.05/2). An initial 2,947 participants were recruited using Amazon Mechanical Turk; all of whom were in the United Kingdom or America and aged 18 years or older. Individuals who had previously taken part in studies using the same crime video or lineup photographs were prevented from taking part in this study. Participants were paid 35 cents for taking part in the study, which took approximately 5 minutes. Participants were excluded from the final analysis if they incorrectly answered the attention check question or stated they had experienced significant technical issues that prevented them from witnessing either video (total *N* excluded = 896).

The final sample was 2,051 participants (55% female, 44% male, 1% preferred not to say or stated "other"; 18–89 years old, *M* age = 38.63, *SD* age = 12.74; 71% White Caucasian, 9% Black or African American, 6% Hispanic or Latino or Spanish, 5% East Asian, 2% South Asian, <1% American Indian or Alaska Native, <1% Native Hawaiian or Other Pacific islander, 3% said other and 3% preferred not to say).

2.3. Materials

A traditional misinformation paradigm was used in this study. The traditional paradigm involves three stages: encoding or experiencing an event, being presented with misinformation about the event, and then being asked to recall information about the event (Loftus, 2005). The misinformation paradigm allows researchers to test how an individual takes an external suggestion and misattributes this to their own personal memory of an event (Zhu et al., 2013).

2.3.1. Mock crime videos

The video stimuli presented at the encoding stage was a mock crime video from Colloff et al. (2021), lasting approximately 17 s, depicting a Caucasian male perpetrator, approximately 30 years old, stealing a handbag from a female victim. There were two videos: one video presented the perpetrator from frontal view and the other presented the perpetrator from profile view.

The video stimuli presented at the misinformation stage was a news report video containing a photograph of the misinformation suspect. The video lasted approximately 36 s and contained an auditory narrative and subtitles explaining that a suspect had been arrested in connection with a recent handbag theft in the area. Specifically, the news report explained that the suspect was apprehended after police reviewed CCTV footage of the crime and found that the culprit looked like a local resident. A picture of an innocent suspect's face was then shown on screen, either from a front facing or profile view. The misinformation suspect was male, aged approximately 30 years, and was similar in appearance to the perpetrator in the encoding video. The misinformation suspect was chosen based on pre-existing data from Colloff et al. (2021). These data showed that amongst the six filler faces used in the target-absent condition in the study, the misinformation suspect chosen was considered the most similar in appearance to the perpetrator. Faces shown in the encoding stage and the misinformation stage were both displayed for a duration of 7 seconds.

2.3.2. Lineups

For the final stage of the misinformation paradigm, participant memories were tested using a six-person simultaneous photo lineup procedure – this method is not used by policing in the United Kingdom (which instead uses nine-person sequential video lineups) (Police and Criminal Evidence Act 1984, Code D, 2017), but it is used in many countries worldwide, including the United States (Fitzgerald et al., 2021). The photos showed the lineup members from the shoulder upwards, and the materials have been successfully used in prior research (Colloff et al., 2021). In the target-present lineup conditions, the guilty suspect (i.e., the perpetrator presented in the mock crime video) was shown amongst five fillers. In the target-absent lineup conditions, the misinformation suspect (i.e., the innocent suspect presented in the news report) was shown amongst five fillers. In line with police guidelines, Colloff et al. (2021) selected fillers who had similar facial attributes to the perpetrator in the mock crime video

2 https://github.com/E-Y-M/poweROC

such that the suspect did not stand out (Police and Criminal Evidence Act 1984, Code D, 2017; Technical Working Group for Eyewitness Evidence, 1999). Colloff et al. (2021) established through mock witness-testing that the lineups were fair.

Lineups were presented with either right profile view or frontal facing lineup members (see Figure 1) that always matched the facial position presented to the participant at encoding. At present, there is a dearth of literature examining the effects of the different sides of the face on facial recognition performance. For example, some research has suggested that the right side of the human face has greater saliency as it bears more resemblance to the face as a whole (Gilbert and Bakan, 1973). On the other hand, Butler et al. (2005) found that when chimeric faces are used (where the left and right side of the face are combined from two different people), participants were more likely to bias their responses towards information on the left-hand side of the face. The current research did not use chimeric faces, it used photographs and videos of sole individuals. Therefore, the right profile faces were used in the "profile" conditions.

2.4. Procedure

Participants were initially provided with an on-screen participant information sheet that included information about the study and the participant's right to withdraw. Participants were required to select "continue" on-screen to consent before they could take part. When they began the study, participants were asked several demographic questions (i.e., age, sex, and ethnicity/race).

All participants completed the three primary stages of the misinformation procedure: the encoding stage, the misinformation stage, and lineup test stage. First, in the encoding stage, participants were randomly assigned to watch one of two versions of the video: 1) the perpetrator's face was shown in right profile view for the duration of the video, or 2) the perpetrator's face was shown from the front, head on, for the duration of the video. After watching the video, participants completed a one-minute filler task consisting of anagram puzzles.

Next, the misinformation stage began. Participants watched the news report video and were randomly assigned to view the misinformation suspect either in the same pose as the mock crime video (front encoding, front MI; profile encoding, profile MI) or different pose (front encoding, profile MI; profile encoding, front MI). After viewing the news report video, participants then completed a further one-minute anagram filler task.

Finally, participants were presented with a simultaneous lineup test displayed in 2 rows of 3 photos. Participants were randomly assigned to view either a target-present or target-absent lineup. Before the lineup, participants were told that they needed to identify the person who they saw in the mock crime video. They were also informed that the guilty suspect may or may not be present in the lineup. Participants were asked to identify whether the guilty suspect was present, or to indicate "not present" if they believed the perpetrator was not present in the lineup. If a suspect was selected, participants were asked to indicate how confident they were in their identification response on a scale ranging from "guessing that he is the culprit" (50%) to "completely certain this is the culprit" (100%). If "not present" was selected, participants were presented with a forced choice lineup, comprising the same lineup members in the same position in the lineup as they had seen before, and asked to guess which suspect was the one they had seen in the crime video. They were then asked to indicate how confident they were that the person selected was *not* the person seen in the original crime video, on scale from "completely certain he is not the culprit" (−100%) to "guessing this is not the culprit" (−50%). This allowed for generating a "fullest possible" ROC curve that includes suspect IDs for the full range of the confidence scale (i.e., −100 to 100%). On completion of the confidence scale, participants were asked an attention check question ("How many people were in the first video you watched?") and a technical check question ["Did you experience any technical issues when watching the mock-crime video (the first video) or the news report video (the second video)]. If "yes" was selected for the technical check question, participants were then asked to briefly explain the technical issue they had experienced. Participants who answered the attention check incorrectly, or who described experiencing significant technical issues (that prevented them from watching the videos), had their data excluded from final analysis. Upon completing these checks, participants were shown an on-screen debrief form which reiterated the details of the study, withdrawal procedures, and provided contact details for the researchers. Participants completed the study by closing the study tab on their computer.

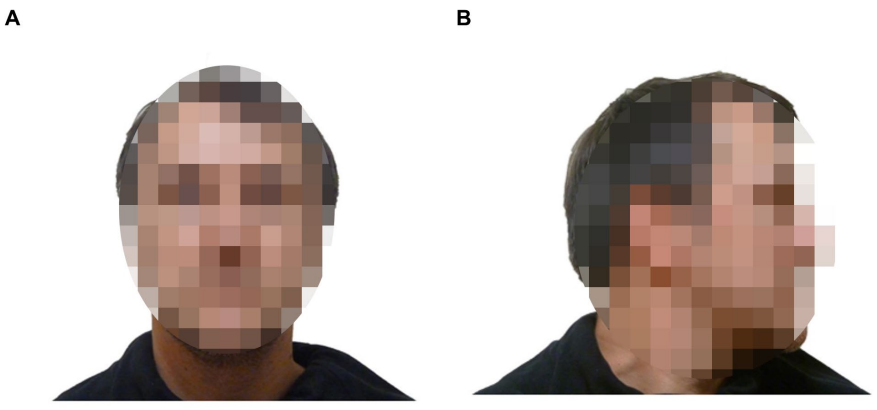


FIGURE 1
Guilty suspect lineup faces from the front **(A)** and right-profile **(B)**.

10.3389/fpsyg.2023.1213996

3. Results

The number of subjects in each of the eight conditions is displayed in Table 1. Recall that when presented with the lineup at test, participants either selected a suspect from the six photographs presented (first lineup selection), or selected "Not Present," which subsequently led to a second forced choice lineup. Response frequencies for the perpetrator, misinformation suspect, filler, and rejection (i.e., not present) decisions at each level of confidence for each condition are shown in Tables 2, 3 for first lineup selection and second forced choice lineup selection, respectively. The overall incorrect ID rate of the misinformation suspect (displayed in the proportion row in Table 2) is equal to the total number of misinformation suspect IDs from the target-absent lineups divided by the total number of target-absent lineups for each facial angle condition. Similarly, the overall correct ID rate of the guilty suspect (also displayed in the proportion row in Table 2) is equal to the total number of perpetrator IDs from target-present lineups divided by the total number of target-present lineups for each facial angle condition.

The overall ID rates of the suspect (TA lineups = misinformation suspect selection, TP lineups = guilty suspect selection) when a selection was made during the first lineup (Table 2) were FFF-TA = 0.12, FFF-TP = 0.77, PPP-TA = 0.50, PPP-TP = 0.61, FPF-TA = 0.14, FPF-TP = 0.79, PFP-TA = 0.49, PFP-TP = 0.65. For the second forced choice lineup (Table 3), the overall ID rates of the suspect were FFF-TA = 0.34, FFF-TP = 0.66, PPP-TA = 0.69, PPP-TP, 0.71 FPF-TA = 0.46, FPF-TP, 0.75, PFP-TA = 0.76. PFP-TP = 0.70. Further analyzes were conducted to explore these results, analyzing discrimination accuracy.

3.1. ROC analysis

Receiver Operating Characteristic (ROC) analysis (see Wixted and Mickes, 2015) was used to explore (1) the facial angle congruency hypothesis – that is, whether discrimination accuracy is higher when facial angles are incongruent across the misinformation paradigm (e.g., frontal encoding, profile misinformation, frontal test), and (2) the encoding strength hypothesis – that is, whether discrimination accuracy is higher when participants view a misinformation suspect from a profile facial angle at the misinformation stage when the encoding and test faces are frontal, compared to those who view the misinformation suspect's face from a frontal angle when the guilty suspect's face at the encoding stage and test faces are shown in profile.

In the current study, the ROC curves were created by plotting the hit rate (HR; the proportion of correct identifications of guilty suspects in TP lineups) against the false alarm rate (FAR; the proportion of incorrect identifications of misinformation suspects in TA lineups). Much previous lineup literature has plotted only positive IDs in ROC curves. Here, because participants in the study were forced to make an identification decision in the second forced choice lineup task, it was possible to extend the curves to contain negative IDs (second forced choice lineup selections). In order to plot the extended ROC curves, we took the six-point confidence scale from the first lineup selections (50%: guessing he is the culprit to 100%: certain he is the culprit) and the six-point confidence scale from the second, forced-choice lineup selections (−50%: guessing he is not the culprit to −100%: certain he is not the culprit) and combined them to create a

TABLE 2 Frequencies of perpetrator, misinformation suspect, and filler identification decisions by pose condition for first lineup respondents.

Confidence rating	FFF						PPP						FPF						PFP					
	Target-Present			Target-Absent			Target-Present			Target-Absent			Target-Present			Target-Absent			Target-Present			Target-Absent		
	Perp	Filler	Reject	MI	Filler	Reject	Perp	Filler	Reject	MI	Filler	Reject	Perp	Filler	Reject	MI	Filler	Reject	Perp	Filler	Reject	MI	Filler	Reject
100	82	2	–	5	2	–	36	2	–	45	0	–	94	0	–	3	1	–	46	0	–	31	0	–
90	63	1	–	8	5	–	51	3	–	27	1	–	58	1	–	6	5	–	46	1	–	32	3	–
80	28	2	–	12	4	–	32	1	–	18	3	–	27	1	–	8	5	–	33	6	–	24	1	–
70	12	4	–	2	1	–	17	1	–	17	0	–	19	2	–	8	4	–	17	0	–	21	1	–
60	10	0	–	3	5	–	13	3	–	14	3	–	10	3	–	9	1	–	19	2	–	11	1	–
50	3	1	–	2	0	–	5	1	–	4	0	–	3	2	–	1	1	–	3	1	–	5	2	–
Total	198	10	50	32	17	220	154	11	86	125	7	119	211	9	48	35	17	200	164	10	77	124	8	119
Proportion	0.77	0.04	0.19	0.12	0.06	0.82	0.61	0.04	0.34	0.50	0.03	0.47	0.79	0.03	0.18	0.14	0.07	0.79	0.65	0.04	0.31	0.49	0.03	0.47

TABLE 3 Frequencies of perpetrator, misinformation suspect, and filler identification decisions by pose condition for second forced choice lineup respondents.

Confidence rating	FFF				PPP				FPF				PFP			
	Target-Present		Target-Absent		Target-Present		Target-Absent		Target-Present		Target-Absent		Target-Present		Target-Absent	
	Perp	Filler	MI	Filler	Perp	Filler	MI	Filler	Perp	Filler	MI	Filler	Perp	Filler	MI	Filler
−50	10	3	4	8	18	2	3	4	12	2	7	11	12	5	16	4
−60	5	2	3	10	8	2	2	3	10	2	9	10	14	2	12	3
−70	4	2	10	11	12	3	14	3	1	4	13	10	9	1	10	2
−80	6	3	9	18	5	3	11	3	2	0	18	15	8	6	9	4
−90	3	3	13	28	8	4	22	9	7	1	20	17	7	3	13	8
−100	5	4	35	71	10	11	30	15	4	3	24	46	4	6	30	8
Total	33	17	74	146	61	25	82	37	36	12	91	109	54	23	90	29
Proportion	0.66	0.34	0.34	0.66	0.71	0.29	0.69	0.31	0.75	0.25	0.46	0.55	0.70	0.30	0.76	0.24

Each proportion is the proportion within each of the second forced choice lineups only. MI, misinformation suspect.

single twelve-point scale (−100 to 100%). This followed a similar analysis procedure used by Colloff and Wixted (2020), where both partial and full ROCs were plotted. In both partial and full ROC analysis, the procedure with the ROC curve that falls furthest from the dashed line is the best at enhancing empirical discriminability (Colloff and Wixted, 2020).

To statistically compare ROC curves, pairwise comparisons between two conditions were made. To complete this pairwise comparison, the partial area under the curve (*p*AUC) was computed using the statistical package *p*ROC (Robin et al., 2011). The difference between the two *p*AUCs was then calculated and divided by the standard deviation of the difference estimated by bootstrapping, and therefore D is the measure of effect size. D is defined as $\frac{(\mathrm{AUC1}-\mathrm{AUC2})}{\mathrm{S}}$, where s is the standard error of the difference between the two AUCs estimated by the bootstrap method, with the number of bootstraps set to 10,000 (Mickes et al., 2012). In a *p*AUC analysis, the specificity cut-off must be set in the analysis. In each set of analyzes, a cut-off that was applied at the most liberal ROC point on the most conservative procedure.

As noted above, to increase the power of our analysis, "extended" ROCs were constructed that included both first lineup decisions (positive IDs where a face was selected) and second forced choice decisions (made after a negative "not present" decision), and the plan was to calculate the *p*AUC for the extended ROCs. However, when the extended ROCs were plotted, it was evident that the portion of the ROCs for the second forced choice lineup decisions were noisy. Previous research has found different results for positive and negative portions of ROCs (see Colloff et al., 2018; Colloff and Wixted, 2020). Therefore, for each research question, we plotted the extended ROCs (as we had initially planned), and also plotted the ROCs for the first lineup decisions only (i.e., the positive IDs, in the way that has typically been done in the lineup literature). For each research question, we present the *p*AUC results for extended ROCs that contain the positive and negative IDs (following our preregistered plan) and then the *p*AUC results for the positive IDs in the first lineups.

3.1.1. Testing the facial angle congruency hypothesis

First, we investigated if discrimination accuracy differed depending on the congruency of facial angles. For the full ROC analyzes, the incongruent facial angle condition (FPF and PFP, $n=1{,}022$) yielded a slightly higher pAUC (0.377, 95% CI [0.358–0.402]) than the congruent condition (FFF and PPP, $n=1{,}029$) which was 0.362 (95% CI [0.336–0.387]). However, this difference was not statistically significant ($D=0.78$, $p=0.44$; specificity cut-off of 0.60, Figure 2A). Considering only the initial identification decisions, the incongruent condition yielded a slightly higher pAUC (0.131, 95% CI [0.114–0.148]) than the congruent condition (0.114, 95% CI [0.096–0.131]), yet this difference was also not statistically significant ($D=1.33$, $p=0.19$; specificity cut-off of 0.30, Figure 2B). Together, the results indicate that discrimination accuracy is similar regardless of facial angle congruency.

3.1.2. Testing the encoding strength hypothesis

Second, we investigated if encoding strength was stronger for frontal-view faces compared to profile-view faces. That is, whether participants are more likely to accept the misinformation (i.e., identify

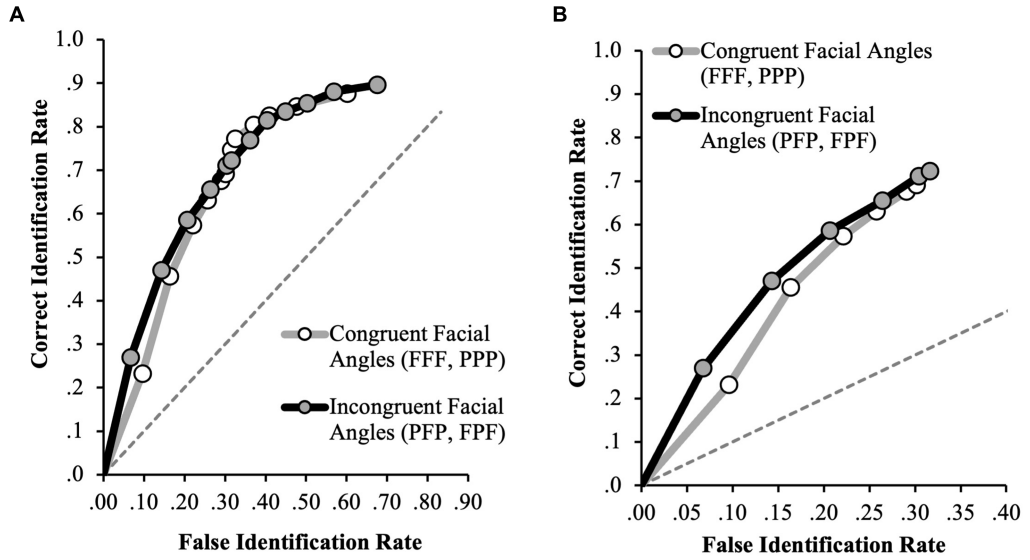


FIGURE 2
ROC data in the congruent facial angle (FFF, PPP) and incongruent facial angle (FPF, PFP) conditions for **(A)** positive IDs and negative ID decisions (extended ROCs) and **(B)** positive ID decisions only. The circular icons represent the empirical data. The dashed line indicates chance-level performance.

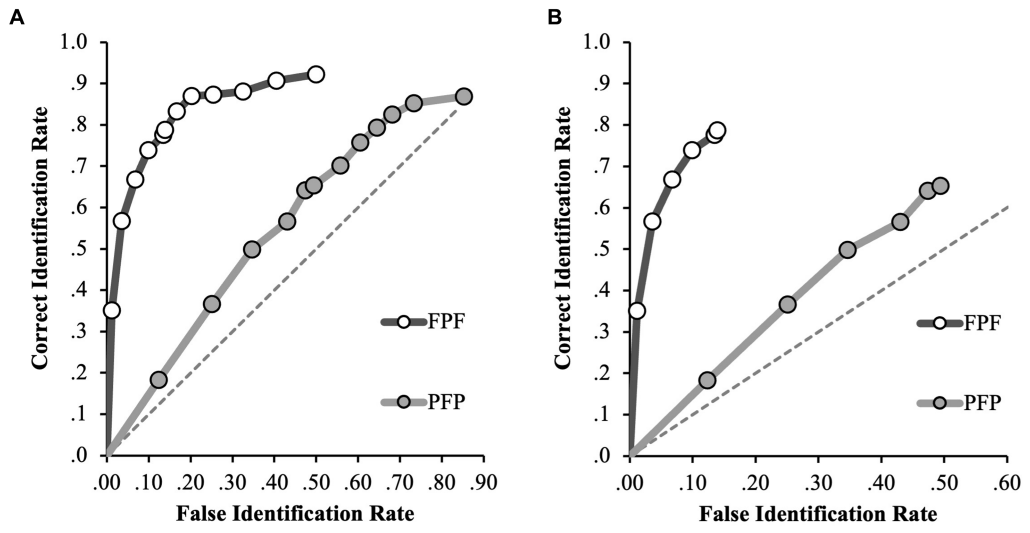


FIGURE 3
ROC data in the FPF and PFP conditions for **(A)** positive IDs and negative ID decisions (extended ROCs) and **(B)** positive ID decisions only. The circular icons represent the empirical data. The dashed line indicates chance-level performance.

the misinformation suspect) when the perpetrator is presented from the profile view and misinformation presented from the front, compared to when the perpetrator is presented from the front and the misinformation is presented from the profile. To answer that question, we compared the ROC curves for the incongruent facial angle conditions – FPF and PFP (see Figure 3).

For the full ROC analyzes in the incongruent facial angle conditions (FPF and PFP), the FPF condition yielded a significantly higher pAUC (0.404, 95% CI [0.381–0.426]) than the PFP condition (0.176, 95% CI [0.143–0.209]), $D = 10.97$, $p < 0.001$ (specificity cut-off of 0.50, Figure 3A). This difference was also found considering only the initial identification decisions, where the pAUC for the FPF condition (0.101, 95% CI [0.088–0.112]) was significantly greater than the pAUC for the PFP condition (0.015, 95% CI [0.010–0.021]); $D = 11.97$, $p < 0.001$ (specificity cut-off of 0.14, Figure 3B). Therefore, for any false identification rate, the correct identification rate was increased by 129% in the FPF compared to the PFP condition when all identification decisions are considered and by 14.8% when only initial decisions are considered.

To further explore the differences in discrimination accuracy between the incongruent conditions (i.e., FPF and PFP), ROC curves for every condition (FFF, PPP, FPF, PFP) were plotted on a single plot. Figure 4 shows the ROC curves for the FFF, PPP, FPF and PFP conditions.

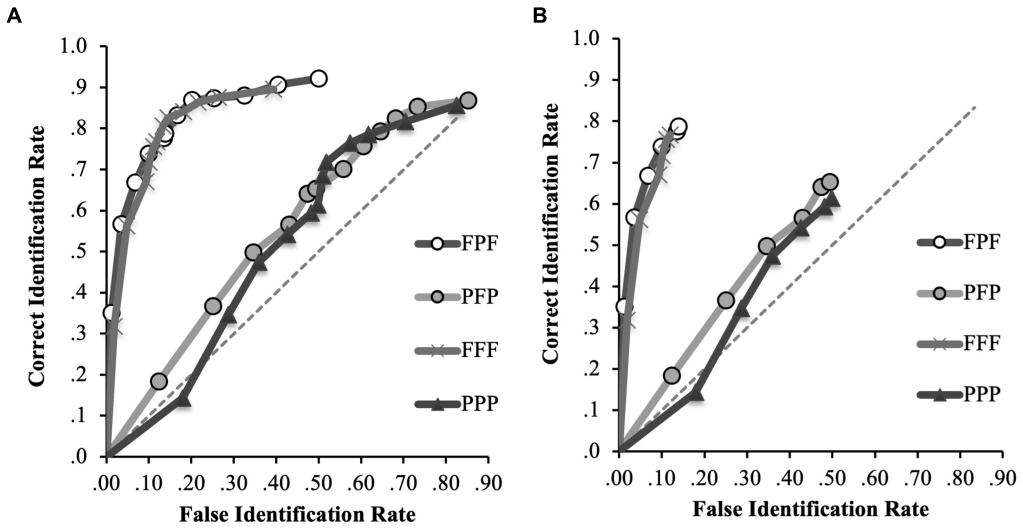


FIGURE 4
ROC data in the FFF, PPP, FPF and PFP conditions for **(A)** positive IDs and negative ID decisions (extended ROCs) and **(B)** positive ID decisions only. The circular icons represent the empirical data. The dashed line indicates chance-level performance.

In our evaluation of the full ROC curves across all conditions, several noteworthy patterns emerged (specificity cut-off of 0.39; Figure 4A). In the FFF condition, participants exhibited a pAUC of 0.297 (95% CI [0.275–0.317]), significantly outperforming those in the PPP condition (0.084, 95% CI [0.06–0.109]), $D = 12.73$, $p < 0.001$, and the PFP condition (0.110, 95% CI [0.086–0.138]), $D = 11.39$, $p < 0.001$. Therefore, for any false identification rate, the correct identification rate in the FFF condition increased by 253% compared to the PPP condition and by 170% compared to the PFP condition when first and second identification decisions are considered. The FPF condition (0.303, 95% CI [0.282–0.323]) also significantly surpassed the PPP condition, $D = 13.48$, $p < 0.001$, indicating that correct identifications for any possible false alarm rate increased by 175% in the FPF compared to the PPP condition. When comparing the FPF and PFP conditions directly, we found that the FPF condition had a significantly higher pAUC, $D = 11.97$, $p < 0.001$, indicating that correct identifications for any possible false alarm rate increased by 175% in the FPF compared to the PPP condition. However, we found no significant differences between the FFF and FPF conditions, $D = 0.43$, $p = 0.67$, or between the PPP and PFP conditions, $D = 1.45$, $p = 0.15$.

For completeness, the *p*AUC for the FPF and PFP conditions were calculated again for this analysis using the new specificity cut-off. Again, the *p*AUC for the FPF condition (0.303) was significantly higher than that for the PFP condition (0.110), $D = 11.97$, $p < 0.001$. This indicates that discrimination accuracy was significantly higher when participants were exposed to a frontal face at encoding and test compared to when they were exposed to a profile face at encoding and test. This suggests that the difference between the FPF and PFP in the previous analysis was due to a beneficial effect of viewing frontal faces at encoding and test, rather than a detrimental effect of viewing frontal faces at the misinformation stage.

Turning to the analysis of the initial identification decisions (specificity cut-off of 0.12; Figure 4B), we noted the following. The FPF condition (0.083, 95% CI [0.071–0.094]) significantly outperformed the FFF condition (0.063, 95% CI [0.051–0.076], $D = 2.20$, $p = 0.01$), indicating that correct identifications for any possible false alarm rate increased by 31.7% in the FPF compared to the FFF condition. However, the difference between the PFP (0.169, 95% CI [0.136–0.200]) and PPP (0.140, 95% CI [0.110–0.173]) conditions was not significant, $D = 1.29$, $p = 0.20$.

This suggests that discrimination accuracy was significantly higher when participants were exposed to the incongruent frontal encoding conditions (FPF) compared to the congruent frontal encoding conditions (FFF), but only for those who made IDs in the first lineup.

3.2. Confidence-accuracy characteristic (CAC) analysis

The relationship between confidence and accuracy was also explored in the current study. The link between high confidence ratings taken at the time of the identification and accurate lineup IDs has been well documented in recent research (Kebbell et al., 1996; Wixted et al., 2015; Wixted and Wells, 2017; Seale-Carlisle et al., 2019). Yet, there is a dearth of research looking at CACs for misinformation studies.

CAC analysis consists of plotting identification accuracy of suspect IDs (ignoring fillers IDs) for each level of confidence. For a six-person lineup procedure, CAC is given by;

$$CAC = \frac{CIDconf}{CIDconf + FIDconf}$$

CID*conf* is the number of correct guilty suspect IDs made with each level of confidence from target-present lineups. Alternatively, FID*conf* is the number of false IDs of misinformation suspects made with that same level of confidence from the target-absent lineups (Mickes, 2015; Seale-Carlisle et al., 2019). In this study, confidence ratings were binned into four levels of confidence: −100 to −80 and −70 to −50 (for the forced-choice lineup decisions, or negative IDs), and 50–70 and 80–100 (for the first lineup decisions, or positive IDs). Unlike ROC analysis, the goal of CAC is to measure the

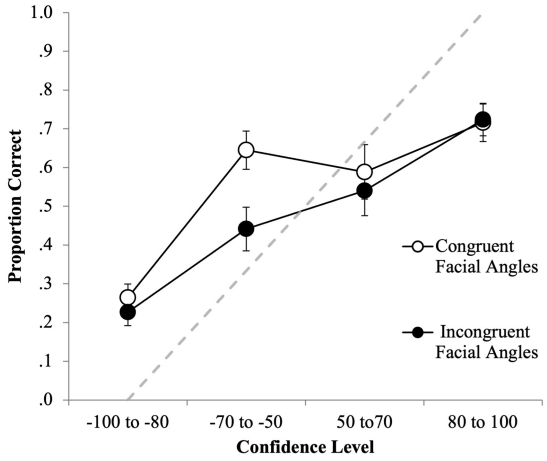


FIGURE 5
CAC data for the facial angle congruent (FFF, PPP) and incongruent (PFP, FPF) conditions for first- and second-line up decisions. The circular icons represent the empirical data. The dashed line indicates chance-level performance. The error bars also represent the standard error.

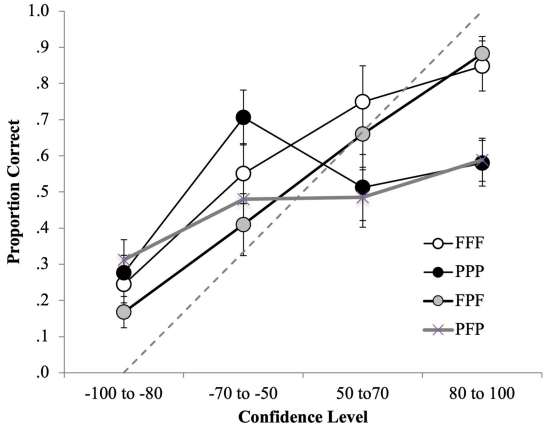


FIGURE 6
CAC data for the all four facial angle (FFF, PPP, FPF, PFP) conditions for first- and second-line up decisions. The circular icons represent the empirical data. The dashed line indicates chance-level performance.

relationship between confidence and accuracy (Mickes, 2015). As such, accuracy is plotted on the *y*-axis and confidence is plotted on the *x*-axis. This is useful from a practical standpoint, whereby the legal system is most interested in knowing the probability that a suspect who has been identified is actually guilty (Wilson et al., 2018).

First, CAC curves were plotted for the congruent facial angle (FFF, PPP) and incongruent facial angle (FPF, PFP) conditions. Figure 5 shows that there appeared to be a relationship between confidence and ID accuracy in both conditions, because, generally speaking, as accuracy increased, so did confidence. However, the relationship was stronger in the incongruent facial angle conditions. In the congruent facial angle conditions, there was a relationship within the negative IDs (i.e., −70 to −50 yielded a higher proportion correct than −100 to −80) and within the positive IDs (i.e., 80 to 100 yielded a higher proportion correct than 50 to 70) but, for some reason, IDs made with a confidence rating of 50 to 70 were less accurate than those made with −70 to −50. For both conditions, it is important to note that high confidence did not indicate high accuracy, as participant were overconfident at high confidence. Participants who made 80–100% confidence judgments where only approximately 70% accurate in their suspect IDs. This is likely due to the deleterious effect of misinformation.

To further explore the relationship between confidence and accuracy in the frontal and profile facial angle encoding conditions, all four conditions were plotted for the CAC analysis. Figure 6 shows the CAC analysis for the FFF, FPF, PPP and PFP conditions.

For the frontal encoding and test conditions (FFF and FPF) there was a relationship between confidence and accuracy, because as confidence increased, so did accuracy. IDs made with high confidence (i.e., 80–100% confidence rating) were also higher in accuracy (around 80% accurate) in the frontal facial angle encoding conditions compared to the profile facial angle encoding conditions. For the profile encoding conditions (PPP and PFP), there appeared to be a weaker relationship between confidence and accuracy. Moreover, participants were overconfident that they had identified the perpetrator when they provided high confidence ratings; they were only approximately 55% accurate when they were 80 to 100 confident.

4. Discussion

The current research explored the impact of facial angle on misinformation susceptibility. It was hypothesized that facial angle congruence between encoding/test and misinformation (e.g., FFF and PPP) would decrease discrimination accuracy compared to facial angle incongruence (e.g., FPF and PFP). The facial angle congruence hypothesis was not supported, as there was no significant difference between congruent and incongruent facial angle conditions. This suggests that participants were no more likely to be misinformed if the misinformation was more like encoding and test compared to when the misinformation was more different to encoding and test. One explanation for this could be that because the facial angle at encoding and test were always congruent, this may have had a stronger impact compared to congruent facial angles at misinformation and test stages. That is, matching the context at misinformation and test is less problematic for discrimination accuracy, so long as the test context remains the same as that experienced at encoding. This supports previous research by Bruce (1982), who found that when individuals learned a frontal face and were subsequently tested with a frontal face, they were able to recognize faces more accurately and quickly compared to when they were tested with faces posed a 45° angle (profile). Although the prediction was not met, a dearth of previous research has fully explored congruent and incongruent facial angles at different stages of the misinformation paradigm. Therefore, this finding has contributed to the growing understanding of facial angle manipulations in the misinformation paradigm.

Based on previous research regarding the strength of frontal face encoding, it was also hypothesized that front-view encoding would enable better discrimination accuracy compared to profile-view encoding. This hypothesis was supported because performance was generally better when the encoded face was front facing compared to profile. This suggests that frontal face encoding and test is superior in memory to profile face encoding and test.

An additional encoding strength hypothesis was considered, proposing that discrimination accuracy would be higher when participants were presented with a profile facing misinformation suspect when the encoding and test faces are frontal (FPF), compared to when

participants view a frontal misinformation face when the encoding and test faces are profile (PFP). This hypothesis was supported, as discrimination accuracy was better in the FPF condition than the PFP condition. To explain this result, we initially proposed that discrimination accuracy may have been better in the FPF condition than the PFP condition due to the strength of the facial angle at the misinformation stage. Put another way, discrimination accuracy in the PFP condition may have been lower than the FPF condition due to the stronger encoding of the front facing misinformation, opposed to the profile facing encoding and test stages. Likewise, higher discrimination accuracy found in the FPF condition may be because profile misinformation would not have had the same encoding strength as the original front facing perpetrator, making it easier for participants to discriminate between faces. This would support previous research, whereby frontal faces have been considered to provide more information than a profile face (McKelvie, 1976), thus leaving a stronger memory trace (Fraser et al., 1990; Meltzer and Bartlett, 2019).

However, our further analyzes suggest this is not the case. When we compared all four facial angle conditions (FFF, PPP, FPF, PFP), further support for a front face encoding benefit was evident. That is, a frontal encoding benefit over profile encoding was observed in the FFF and FPF conditions compared to the PPP and PFP conditions. This difference cannot be explained by differences of facial angle at the misinformation stage, and instead must be explained by difference of facial angle at encoding (and test). Together, the findings support the encoding strength hypothesis and previous face memory literature, where frontal face encoding is argued to be superior to other poses (Colloff et al., 2021). This also supports the holistic encoding hypothesis, which suggests that instead of processing faces as a collection of separate, distinct, facial features, we instead process the face as a perceptual whole (Taubert et al., 2011). Therefore, seeing a criminal's face from a frontal view at encoding and test means that participants can engage in holistic facial encoding and recognition. We also know that a frontal face provides more perceptual information than a profile face (Meltzer and Bartlett, 2019) and that this perceptual information can be beneficial for facial recognition.

For most of the findings, the ROC analysis of the positive lineup IDs (first lineup decisions) replicated the findings from the extended ROC analysis including negative IDs. However, when results for the partial positive portion of the curve were calculated for the FFF and FPF condition, discrimination accuracy was significantly higher in the FPF condition compared to the FFF condition ($p = 0.03$). This significant difference was not observed in the extended ROC analysis. A possible explanation for the significant finding is that the congruence between encoding, misinformation, and test in the FFF condition may have made it more difficult for participants to discriminate between the guilty suspect and the misinformation suspect than the FPF. This would, in part, support the proposed facial angle congruence hypothesis. But if that were true, it is not clear why the same pattern of results was not observed in the profile encoding conditions (i.e., no significant difference between PPP and PFP), or on the extended ROC. What we do know is that the analysis found significantly better discrimination accuracy in the frontal encoding conditions compared to the profile encoding conditions. One reason this finding may not have been observed in the profile encoding condition is due to the overall poor discrimination accuracy in the PPP and PFP conditions, where discrimination accuracy was only marginally better than chance. Moreover, other research has found the predicted pattern of results only in the positive IDs and not the negative IDs (see Colloff et al., 2018; Colloff and Wixted, 2020), but it is not yet clear why that is the case. Nevertheless, because this result was only found in front encoding conditions (i.e., FFF, FPF), but not profile encoding conditions (i.e., PPP, PFP), and was only observed in the positive ID portion of the ROC and not the extended ROC including negative IDs, the significant result should be interpreted with caution and further research is needed.

4.1. Practical implications

We found that the angle of the misinformation (congruent or incongruent with study and test) was not an important determinant of identification accuracy. Instead, we found that when the encoding face was presented from a profile view discrimination accuracy was significantly poorer than when the encoding face was presented from the front. The witness in the Ted Bundy case *did* encode Ted Bundy from the profile view. Whilst it is highly probable that she correctly identified Bundy (considering the abundance of evidence implicating him), the lower discrimination accuracy results for profile encoding in the current study are noteworthy. This underscores the importance of ensuring that the angle of the lineup faces matches the angle(s) shown during encoding. Previous research has found that discrimination accuracy for faces encoded in profile view is higher when the lineup faces are also presented in profile view (Colloff et al., 2021). Interestingly, the lineup the witness in the Bundy case viewed showed the lineup members also in profile view, providing cues that likely matched the encoding context and supported her memory retrieval.

Moreover, the results suggest that witnesses who have encoded perpetrators from profile view may be less reliable because they were found to have lower accuracy at high-confidence and have a poorer confidence-accuracy relationship than witnesses who have encoded perpetrators from the front. One explanation for this is that because the discrimination performance was so low in the PPP and PFP conditions (only marginally higher than chance), this impacted participant's ability to assign appropriate confidence ratings. The poor confidence-accuracy relationship in the PPP and PFP conditions are consistent with findings from previous research that has found a poor confidence-accuracy relationship when memory accuracy is below chance (see Weber and Brewer, 2003; Nguyen et al., 2017). Theoretically, participants who are guessing should not be more confident in their guess that resulted in a correct identification than a guess that resulted in an incorrect identification (Nguyen et al., 2017). Furthermore, participants who are guessing (i.e., whose memory signal is weak) would have more relaxed criterion for identifying faces. Therefore, they are predicted to be less confident in their responses than participants who make recognition judgments based on more information in memory (i.e., stronger feelings of familiarity with a face). This suggests that accuracy is more likely to fluctuate around chance levels at lower levels of confidence.

Court systems may not always consider confidence when evaluating eyewitness IDs (Juslin et al., 1996). It can be argued that the reason for this is because confidence ratings are susceptible to influence. For example, other research has found that a poor correspondence between confidence and accuracy has also been associated with conformity to misinformation, whereby participants are misled but still provide high confidence ratings (Mudd and Govern, 2004; Foster et al., 2012; Spearing and Wade, 2021).

4.2. Limitations and future directions

In considering these findings, it is important to note a methodological limitation of the current research. Given that participants were always

exposed to the same facial position at encoding and test, this research has not considered the potential influence that incongruent facial angles between encoding and test in the misinformation paradigm may have on misinformation susceptibility. Previous research suggests that people will be slower to recognize a face and less accurate in their recognition if the viewing angle of a face is changed (for example, between front facing and ¾ facing) between initial presentation and test compared to when it remains unchanged (Bruce, 1982). However, it is noted that this finding has not been explicitly explored in the misinformation paradigm. Likewise, the full impact of facial viewing angle manipulations across the three stages of the misinformation paradigm have not been explored in this single study. It will be important for future research to explore how further facial manipulations at test could impact misinformation susceptibility.

Similarly, the present study only included one suspect and one misinformation face, however, to counter any mediating factors that may be involved in eyewitness discrimination accuracy (for a discussion about these factors, see Ryder et al., 2015), it would be useful for future research to investigate the misinformation effect using a variety of perpetrator and misinformation faces. For example, future research could explore own-race bias in the context of misinformation and facial angles.

Like many other studies that adopt a lineup paradigm, a limitation of this research is the length of the distractor task – one minute. In real cases, the median average delay between witnessing a crime and being presented with a lineup is around 11 days in the United States (Flowe et al., 2018), and 31 days in the United Kingdom (Horry et al., 2012). Whilst this might seem concerning at face value, some studies have demonstrated that length of delay between encoding and test does not necessarily harm identification accuracy (Valentine et al., 2012; Wetmore et al., 2015). Nevertheless, other research finds that longer retention intervals are associated with decreased face recognition performance (Deffenbacher et al., 2008), and therefore, it would be valuable to investigate whether delay mediates the misinformation effect.

It might also be fruitful if future research considers whether a combined lineup procedure would have implications for these findings. That is, the lineup procedure at test could contain both the guilty suspect and misinformation suspect amongst fillers in a single lineup. A similar procedure has been used by some police departments, whereby everyone in the lineup is suspected of being the person (all-suspect design) who committed the offense (Wells and Luus, 1990). Whilst this lineup design has been used in forensic contexts, it is certainly not the norm and it would be unusual to have multiple suspects (i.e., one guilty and one innocent) in a single lineup. Nevertheless, it may be interesting for future research to explore this different method.

5. Conclusion

The impact of facial angle on recognition and discrimination accuracy was explored using a traditional misinformation paradigm (encoding, misinformation, test). Participants were not differentially likely to be misled by misinformation (i.e., an innocent suspect) depending on facial angle congruency across encoding, the misinformation, and lineup phases. This suggests that participants are no less likely to be misled if the innocent suspect's face is presented in the same as opposed to different angle across encoding, misinformation, and test. Discrimination accuracy was significantly higher overall when the participants encoded the perpetrator from the front compared to the profile angle, suggesting that memory is stronger for faces that are originally encoded in frontal view. ROC analysis for all four conditions (FFF, PPP, FPF, PFP) also supported the encoding benefit of encoding a face from the front compared to the profile. Moreover, CAC analysis revealed a weak relationship between confidence and accuracy in the profile encoding (PPP and PFP) conditions compared to a stronger relationship in the frontal encoding (FFF and FPF) conditions. Given that legal decision makers rely on eyewitness confidence in court (Mickes, 2015; Garrett et al., 2020), they should be particularly aware that the reliability of eyewitness identifications could be impaired when a witness has encoded a perpetrator from a profile posed face (and discrimination accuracy is poor) compared to when the face is encoded from the front.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://osf.io/fsmr9/ or https://osf.io/vdq63/.

Ethics statement

The studies involving humans were approved by University of Birmingham STEM Ethics Committee. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

KD, MC, and HF conceived of the idea and designed the experiments. KD collected the data. KD and MC analyzed the data. KD, TB, and HF wrote the paper with input from all authors. All authors contributed to the article and approved the submitted version.

Funding

We are grateful to the School of Psychology at the University of Birmingham for an internal grant award that supported participant recruitment.

Acknowledgments

We would like to thank the Editor and reviewers for helpful comments and advice on the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Ayers, M. S., and Reder, L. M. (1998). A theoretical review of the misinformation effect: predictions from an activation-based memory model. *Psychon. Bull. Rev.* 5, 1–21. doi: 10.3758/BF03209454

Bekerian, D. A., and Bowers, J. M. (1983). Eyewitness testimony: were we misled? *J. Exp. Psychol. Learn. Mem. Cogn.* 9, 139–145. doi: 10.1037/0278-7393.9.1.139

Bruce, V. (1982). Changing faces: visual and non-visual coding processes in face recognition. *Br. J. Psychol.* 73, 105–116. doi: 10.1111/j.2044-8295.1982.tb01795.x

Bülthoff, I., Mohler, B. J., and Thornton, I. M. (2019). Face recognition of full-bodied avatars by active observers in a virtual environment. *Vis. Res.* 157, 242–251. doi: 10.1016/j.visres.2017.12.001

Bundy, V. State, 455 So. 2d 330. (1984). Available at: https://law.justia.com/cases/florida/supreme-court/1984/57772-0.html

Butler, S., Gilchrist, I. D., Burt, D. M., Perrett, D. I., Jones, E., and Harvey, M. (2005). Are the perceptual biases found in chimeric face processing reflected in eye-movement patterns? *Neuropsychologia* 43, 52–59. doi: 10.1016/j.neuropsychologia.2004.06.005

Campbell, J. M., Edwards, M. S., Horswill, M. S., and Helman, S. (2007). Effects of contextual cues in recall and recognition memory: the misinformation effect reconsidered. *Br. J. Psychol.* 98, 485–498. doi: 10.1348/000712606X160768

Carpenter, C., Webb, C., Overman, A., and Dennis, N. (2022). Within-category similarity negatively affects associative memory performance in both younger and older adults. *Memory* 31, 77–91. doi: 10.1080/09658211.2022.2123524

Chandler, C. C. (1991). How memory for an event is influenced by related events: interference in modified recognition tests. *J. Exp. Psychol. Learn. Mem. Cogn.* 17, 115–125. doi: 10.1037/02787393.17.1.115

Colloff, M., Seale-Carlisle, T., Karoğlu, N., Rockey, J., Smith, H. M., Smith, L., et al. (2021). Perpetrator pose reinstatement during a lineup test increases discrimination accuracy. *Sci. Rep.* 11:13830. doi: 10.1038/s41598-021-92509-0

Colloff, M. F., Wade, K. A., Strange, D., and Wixted, J. T. (2018). Filler-siphoning theory does not predict the effect of lineup fairness on the ability to discriminate innocent from guilty suspects: reply to Smith, Wells, Smalarz, and Lampinen (2018). *Psychol. Sci.* 29, 1552–1557. doi: 10.1177/0956797618786459

Colloff, M. F., and Wixted, J. T. (2020). Why are lineups better than showups? A test of the filler siphoning and enhanced discriminability accounts. *J. Exp. Psychol. Appl.* 26, 124–143. doi: 10.1037/xap0000218

Deffenbacher, K. A., Bornstein, B. H., McGorty, E. K., and Penrod, S. D. (2008). Forgetting the once-seen face: estimating the strength of an eyewitness's memory representation. *J. Exp. Psychol. Appl.* 14, 139–150. doi: 10.1037/1076-898X.14.2.139

Fitzgerald, R. J., Rubínová, E., and Juncu, S. (2021). "Eyewitness identification around the world" in *Methods, measures, and theories in eyewitness identification tasks*. eds. D. M. Bialer, V. F. Reyna and C. J. Brainerd (England: Routledge), 294–322.

Flowe, H. D., Carline, A., and Karoğlu, N. (2018). Testing the reflection assumption: a comparison of eyewitness ecology in the laboratory and criminal cases. *Int. J. Evid. Proof.* 22, 239–261. doi: 10.1177/1365712718782996

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

Fraser, I. H., Craig, G. L., and Parker, D. M. (1990). Reaction time measures of features saliency in schematic faces. *Perception* 19, 661–673. doi: 10.1068/p190661

Garrett, B. L., Liu, A., Kafadar, K., Yaffe, J., and Dodson, C. S. (2020). Factoring the role of eyewitness evidence in the courtroom. *J. Empir. Leg. Stud.* 17, 556–579. doi: 10.1111/jels.12259

Gilbert, C., and Bakan, P. (1973). Visual asymmetry in perception of faces. *Neuropsychologia* 11, 355–362. doi: 10.1016/0028-3932(73)90049-3

Horry, R., Memon, A., Wright, D. B., and Milne, R. (2012). Predictors of eyewitness identification decisions from video lineups in England: a field study. *Law Hum. Behav.* 36, 257–265. doi: 10.1037/h0093959

Johnson, M. K., Hashtroudi, S., and Lindsay, D. S. (1993). Source monitoring. *Psychol. Bull.* 114, 3–28. doi: 10.1037/0033-2909.114.1.3

Juslin, P., Olsson, N., and Winman, A. (1996). Calibration and diagnosticity of confidence in eyewitness identification: comments on what can be inferred from the low confidence–accuracy correlation. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 1304–1316. doi: 10.1037/0278-7393.22.5.1304

Kebbell, M. R., Wagstaff, G. F., and Covey, J. A. (1996). The influence of item difficulty on the relationship between eyewitness confidence and accuracy. *Br. J. Psychol.* 87, 653–662. doi: 10.1111/j.2044-8295.1996.tb02614.x

Loftus, E. F. (1977). Shifting human color memory. *Mem. Cogn.* 5, 696–699. doi: 10.3758/BF03197418

Loftus, E. F. (2005). Planting misinformation in the human mind: a 30-year investigation of the malleability of memory. *Learn. Mem.* 12, 361–366. doi: 10.1101/lm.94705

Loftus, E. F., and Greene, E. (1980). Warning: even memory for faces may be contagious. *Law Hum. Behav.* 4, 323–334. doi: 10.1007/BF01040624

Loftus, E. F., Miller, D. G., and Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *J. Exp. Psychol. Hum. Learn. Mem.* 4, 19–31.

Longmore, C. A., Liu, C. H., and Young, A. W. (2008). Learning faces from photographs. *J. Exp. Psychol. Hum. Percept. Perform.* 34:77. doi: 10.1037/0096-1523.34.1.77

McKelvie, S. J. (1976). The role of eyes and mouth in the memory of a face. *Am. J. Psychol.* 89, 311–323. doi: 10.2307/1421414

Meltzer, M. A., and Bartlett, J. C. (2019). Holistic processing and unitization in face recognition memory. *J. Exp. Psychol. Gen.* 148, 1386–1406. doi: 10.1037/xge0000640

Mickes, L. (2015). Receiver operating characteristic analysis and confidence–accuracy characteristic analysis in investigations of system variables and estimator variables that affect eyewitness memory. *J. Appl. Res. Mem. Cogn.* 4, 93–102. doi: 10.1016/j.jarmac.2015.01.003

Mickes, L., Flowe, H. D., and Wixted, J. T. (2012). Receiver operating characteristic analysis of eyewitness memory: comparing the diagnostic accuracy of simultaneous versus sequential lineups. *J. Exp. Psychol. Appl.* 18, 361–376. doi: 10.1037/a0030609

Mudd, K., and Govern, J. M. (2004). Conformity to misinformation and time delay negatively affect eyewitness confidence and accuracy. *N. Am. J. Psychol.* 6, 227–238.

Nguyen, T. B., Pezdek, K., and Wixted, J. T. (2017). Evidence for a confidence–accuracy relationship in memory for same-and cross-race faces. *Q. J. Exp. Psychol.* 70, 2518–2534. doi: 10.1080/17470218.2016.1246578

Police and Criminal Evidence Act 1984, Code D. (2017) Available at: https://www.gov.uk/government/publications/pace-code-d-2017

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., et al. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* 12, 1–8. doi: 10.1186/1471-2105-12-77

Ryder, H., Smith, H. M. J., and Flowe, H. D. (2015). "Estimator variables and memory for faces" in *Forensic facial identification: Theory and practice of identification from eyewitnesses, composites and CCTV*. eds. T. Valentine and J. P. Davis (Hoboken, NJ: Wiley Blackwell), 159–183.

Seale-Carlisle, T. M., Colloff, M. F., Flowe, H. D., Wells, W., Wixted, J. T., and Mickes, L. (2019). Confidence and response time as indicators of eyewitness identification accuracy in the lab and in the real world. *J. Appl. Res. Mem. Cogn.* 8, 420–428. doi: 10.1016/j.jarmac.2019.09.003

Spearing, E. R., and Wade, K. A. (2021). Providing eyewitness confidence judgements during versus after eyewitness interviews does not affect the confidence-accuracy relationship. *J. Appl. Res. Mem. Cogn.* 11, 54–65. doi: 10.1037/h0101868

Sporer, S. L., Tredoux, C. G., Vredeveldt, A., Kempen, K., and Nortje, A. (2020). Does exposure to facial composites damage eyewitness memory? A comprehensive review. *Appl. Cogn. Psychol.* 34, 1166–1179. doi: 10.1002/acp.3705

Taubert, J., Apthorp, D., Aagten-Murphy, D., and Alais, D. (2011). The role of holistic processing in face perception: evidence from the face inversion effect. *Vis. Res.* 51, 1273–1278. doi: 10.1016/j.visres.2011.04.002

Technical Working Group for Eyewitness Evidence. (1999). Eyewitness evidence: A guide for law enforcement. United States Department of Justice, Office of Justice Programs, National Institute of Justice. Available at: https://www.ojp.gov/pdffiles1/nij/178240.pdf

Topp-Manriquez, L. D., McQuiston, D., and Malpass, R. S. (2014). Facial composites and the misinformation effect: how composites distort memory. *Leg. Criminol. Psychol.* 21, 372–389. doi: 10.1111/lcrp.12054

Tulving, E., and Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychol. Rev.* 80, 352–373. doi: 10.1037/h0020071

Valentine, T., Davis, J. P., Memon, A., and Roberts, A. (2012). Live showups and their influence on a subsequent video line-up. *Appl. Cogn. Psychol.* 26, 1–23. doi: 10.1002/acp.1796

Weber, N., and Brewer, N. (2003). The effect of judgment type and confidence scale on confidence-accuracy calibration in face recognition. *J. Appl. Psychol.* 88, 490–499. doi: 10.1037/0021-9010.88.3.490

Wells, G. L., and Luus, C. A. (1990). Police lineups as experiments: social methodology as a framework for properly conducted lineups. *Personal. Soc. Psychol. Bull.* 16, 106–117. doi: 10.1177/0146167290161008

Wetmore, S. A., Neuschatz, J. S., Gronlund, S. D., Wooten, A., Goodsell, C. A., and Carlson, C. A. (2015). Effect of retention interval on showup and lineup performance. *J. Appl. Res. Mem. Cogn.* 4, 8–14. doi: 10.1016/j.jarmac.2014.07.003

Wilson, B. M., Seale-Carlisle, T. M., and Mickes, L. (2018). The effects of verbal descriptions on performance in lineups and showups. *J. Exp. Psychol. Gen.* 147:113. doi: 10.1037/xge0000354

Wixted, J. T., and Mickes, L. (2015). Evaluating eyewitness identification procedures: ROC analysis and its misconceptions. *J. Appl. Res. Mem. Cogn.* 4, 318–323. doi: 10.1016/j.jarmac.2015.08.009

Wixted, J. T., Mickes, L., Clark, S. E., Gronlund, S. D., and Roediger, H. L. (2015). Initia leyewitness confidence reliably predicts eyewitness identification accuracy. *Am. Psychol.* 70, 515–526. doi: 10.1037/a0039510

Wixted, J. T., and Wells, G. L. (2017). The relationship between eyewitness confidence and identification accuracy: a new synthesis. *Psychol. Sci. Public Interest* 18, 10–65. doi: 10.1177/1529100616686966

Yamashita, M. (1996). A re-examination of the misinformation effect by means of visual and verbal recognition tests. *Jpn. Psychol. Res.* 38, 47–52. doi: 10.1111/j.1468-5884.1996.tb00007.x

Zaragoza, M. S., Belli, R. F., and Payment, K. E. (2007). "Misinformation effects and the suggestibility of eyewitness memory" in *Do justice and let the sky fall: Elizabeth Loftus and her contributions to science, law, and academic freedom*. eds. M. Garry and H. Hayne (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 35–63.

Zhu, B., Chen, C., Loftus, E. F., Lin, C., and Dong, Q. (2013). The relationship between DRM and misinformation false memories. *Mem. Cogn.* 41, 832–838. doi: 10.3758/s13421-013-0300-2

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 12 April 2024
DOI 10.3389/fpsyg.2024.1233782

OPEN ACCESS

EDITED BY
Fabiana Battista,
University of Bari Aldo Moro, Italy

REVIEWED BY
Charlie Frowd,
University of Central Lancashire, United Kingdom
Jamal K. Mansour,
University of Lethbridge, Canada

*CORRESPONDENCE
Margaret Bull Kovera
✉ mkovera@jjay.cuny.edu

RECEIVED 02 June 2023
ACCEPTED 18 March 2024
PUBLISHED 12 April 2024

CITATION
Jones JM, Katzman J and Kovera MB (2024) Phenotypic mismatch between suspects and fillers but not phenotypic bias increases eyewitness identifications of Black suspects. *Front. Psychol.* 15:1233782. doi: 10.3389/fpsyg.2024.1233782

Phenotypic mismatch between suspects and fillers but not phenotypic bias increases eyewitness identifications of Black suspects

Jennifer M. Jones[1,2], Jacqueline Katzman[1,2] and Margaret Bull Kovera[1,2]*

[1]Department of Psychology, The Graduate Center, City University of New York, New York, NY, United States, [2]Department of Psychology, John Jay College of Criminal Justice, City University of New York, New York, NY, United States

Introduction: Despite converging evidence that people more closely associate the construct of criminality with Black people who exhibit a more African facial phenotype than Black people who express a more European phenotype, eyewitness researchers have largely ignored phenotypic bias as a potential contributor to the racial disparities in the criminal legal system. If this form of phenotypic bias extends to eyewitness identification tasks, eyewitnesses may be more likely to identify Black suspects with an African rather than European phenotype, regardless of their guilt status. Further, in cases where the witness's description of the perpetrator does not contain phenotypic information, phenotypic mismatch between the suspect and the other lineup members may bias identification decisions toward or against the suspect. If witnesses can use elements of the lineup construction to guide their identification decisions rather than relying on their recognition memory, then the lineup should be deemed unfair due to suspect bias. The current study also investigated lineup presentation method as a procedural safeguard, predicting that that when lineups were presented simultaneously, there would be a significant two-way interaction of phenotypic bias and lineup composition, with a larger simple main effect of phenotypic bias when lineups were suspect-biased (i.e., the fillers were a phenotypic mismatch to the suspect) than when all lineup members shared the same phenotype. We expected that this interaction would be significantly smaller or non-significant for sequential lineups.

Methods: Participants watched a mock crime video that contained a Black culprit with either a more African phenotype or a less African phenotype before attempting identifications from a photo array that contained a suspect whose phenotype always matched the culprit viewed in the video, but varied in culprit-presence, phenotypic match of the suspect and fillers, and presentation method.

Results: Participants did not identify Black suspects with Afrocentric features more often than Black suspects with Eurocentric features. However, witnesses made more identifications of suspects when the fillers did not match the suspect's phenotype compared to when all lineup members possessed similar phenotypic features.

Discussion: In sum, phenotypic bias did not influence our participant-witnesses' identification decisions, nor interact with lineup composition and lineup presentation type to affect identifications of suspects, suggesting that phenotypic bias may be less influential in match-to-memory tasks than other types of legal decision-making (e.g., determining guilt and sentencing). However, the suggestiveness created by failing to match fillers' phenotypes to the suspect's phenotype can be avoided with proper attention to fair lineup construction.

KEYWORDS

phenotypic bias, own-race bias, cross-race, eyewitness identification, lineup construction, suspect bias, lineup fairness

1 Introduction

Mistaken eyewitness identification has been identified as the leading contributing factor to wrongful convictions, as evidenced by 69% of DNA exoneration cases involving a mistaken identification (Innocence Project, 2023). Moreover, there are great racial disparities in these data as 65% of misidentified exonerees were Black defendants (The National Registry of Exonerations, 2022). Scholars who have explored the contribution of race to mistaken identifications have almost exclusively examined the role of an own-race bias in identification accuracy (Katzman and Kovera, 2023). Also known as the cross-race effect, people are more accurate when identifying people from their own racial group than they are identifying people of other racial groups (Meissner and Brigham, 2001a; Lee and Penrod, 2022; Katzman and Kovera, 2023). Yet the own-race bias does not provide a sufficient explanation for these disparities in wrongful convictions because (a) five times as much crime occurs with victims and perpetrators of the same race versus different races (National Archive of Criminal Justice Data, 2016), (b) the size of the own-race bias in identification accuracy is small relative to the large racial disparities in the exoneration data, and (c) a meta-analysis of the own-race bias literature demonstrated that both White and Black participants were better able to discriminate among previously seen White than Black faces (Katzman and Kovera, 2023). Thus, to reduce wrongful convictions, it is important to explore additional psychological mechanisms that may explain how race contributes to eyewitness identification accuracy.

1.1 Beyond own-race bias: external influences on witnesses' mistaken identifications of Black faces

Moreover, it is not a new proposition that memories for perpetrators might be altered through external events like the process of providing a description of a face (Schooler and Engstler-Schooler, 1990; Meissner and Brigham, 2001b; Alogna et al., 2014), exposure to mugshots (Deffenbacher et al., 2006), or the act of engaging in repeated identification procedures (Wells et al., 2020; Wixted et al., 2021). Witnesses' decision-making processes are susceptible to influence from external cues, particularly when they do not experience an immediate sense of recognition (Bradfield et al., 2002). For example, both the selection of dissimilar fillers and differences between photos of fillers and the suspect (e.g., background and clothing) in a photographic lineup cause witnesses to be more likely to identify the suspect, irrespective of the suspect's guilt (for a review, see Wells et al., 1998; Wells et al., 2020 for a review). Rather than distorting or degrading witnesses' memory for the perpetrator, these factors affect decision-making through creating suspect bias in the lineup procedure (Smalarz, 2021). If a lineup procedure biases a witness toward identifying the suspect through non-memorial cues, the identification procedure could be deemed unfair and inadmissible in court, as the legal system requires that an eyewitness identification must be based on the independent memory of the witness, free from suggestive influences (*Perry v. New Hampshire, 2012*). A suspect-biased lineup not only puts innocent suspects at greater risk of being misidentified, but also fails to provide guilty culprits with due process. Eyewitness researchers have made a distinction between two types of factually correct identifications: *legitimate hits*, where the witness makes a correct identification based on their memory of the culprit, and *illegitimate hits*, where suggestive procedures ultimately produce a correct identification (Wells et al., 2012).

Internal cognitive structures of witnesses, like racial stereotypes, may also alter witnesses' identification decisions. One potential internal phenomenon that may explain racial disparities in misidentifications that has received relatively little attention among eyewitness researchers is phenotypic bias. *Phenotypic bias* refers to stereotypes, prejudice, and discrimination based on race-related facial characteristics. Whereas racial bias involves comparisons between different racial groups (e.g., White people vs. Black people), phenotypic bias involves comparisons between people of the same racial group who possess varying phenotypic characteristics (e.g., light-skinned Black people vs. dark-skinned Black people). This form of subgroup prejudice has manifested in several phenomena, including *colorism* (Okazawa-Rey et al., 1987; Russel et al., 1992), *Afrocentric bias* (Blair et al., 2002), and *bleaching syndrome* (Hall, 1994, 1995). Previous research investigating the effect of skin tone bias on perceptions of Black people in the United States has revealed that individuals

who possess features that are more typical of their racial group are perceived and treated more negatively (Maddox, 2004).

Phenotypic bias is rooted in the fact that criminality is a central feature of the stereotype that people have about Black individuals, irrespective of whether people are highly prejudiced toward that group (Devine, 1989) or whether Black people are members of their in-group (i.e., other Black people) or out-group (e.g., White people; Maddox and Gray, 2002). Black stereotypes are automatically activated with exposure to Black people, but may be more strongly activated by some members of that group than others as there is variation in the extent to which Black people exhibit a facial phenotype that is associated with being African. Black people who have more phenotypically African facial features (e.g., darker skin, fuller lips, wider nose, afro-textured hair, and prominent brow) are viewed as more representative of their race (Knuycky et al., 2013) and thus activate racial stereotypes more strongly than do those whose features are more phenotypically European (Blair et al., 2002, 2004b; Maddox, 2004). People with more Afrocentric facial features are more likely to prime stereotypes about Black aggressiveness (Blair et al., 2005) and criminality (Eberhardt et al., 2004; Knuycky et al., 2013), and be seen as more intimidating than people with less Afrocentric features (Kleider-Offutt et al., 2018). It is difficult to consciously inhibit criminal inferences activated by Afrocentric features, even when aware of the consequences associated with phenotypic bias (Blair et al., 2004a).

Phenotypic bias has detrimental consequences for the treatment of Black people with Afrocentric features within the criminal justice system. For example, laypeople and police officers infer that Black people with Afrocentric features are more likely to engage in criminal activities (Eberhardt et al., 2004; Kahn and Davies, 2011). In studies employing shooting simulations to study shooter bias—the tendency to mistakenly shoot unarmed Black men and fail to shoot armed White men—both Black and White participants were more likely to show shooter bias when the Black men in the stimulus materials had more Afrocentric features compared to Eurocentric features (Kahn and Davies, 2011). Among Black people found guilty of a crime, those with more Afrocentric features receive harsher punishments, including the death penalty, than those with less Afrocentric features (Blair et al., 2004b; Eberhardt et al., 2006; Peterson, 2016). Moreover, in an archival study of use of force cases, independent raters coded the phenotypic stereotypicality of each suspect's booking photograph and found that the more phenotypically White an individual was perceived to be, the less police force was used during the interaction. In other words, police used less force with highly stereotypical White people compared to less stereotypically White people, resulting in a pro-White protective bias (Kahn et al., 2016). Taken together, there is converging evidence that people make inferences about others' culpability and deservingness of punishment based on their phenotypic features.

Less is known about whether this form of phenotypic bias extends to the recognition memory and decision-making processes involved in eyewitness identification. In one study, participants viewed a series of slides depicting a Black man leaving a building and were told that he was accused of committing either a stereotypically White or stereotypically Black crime (Osborne and Davies, 2013). Participants were then asked to identify the man from a series of 100 pictures that were created by morphing the target face with a face that was more phenotypically African and with a face that was less phenotypically African than the target face. The 50th picture in the series was the target face. Participants who were told that the target was committing a stereotypically Black crime chose a morphed picture that was more phenotypically African and significantly different from the target face.

In another series of experiments evaluating misidentifications in lineup identification procedures, witnesses perceived Black faces with more African phenotypes as more familiar than Black faces with a less African phenotype (Knuycky et al., 2013, Experiment 2), and made more identifications from culprit-absent lineups when they contained more, rather than fewer, lineup members with phenotypically African features (Knuycky et al., 2013, Experiment 3). This finding could be explained by the tendency for people to mistakenly report novel faces as previously seen more often when the novel faces are "typical" rather than "distinctive" (Vokey and Read, 1992; Dewhurst et al., 2005), combined with the finding that witnesses perceived Black faces with more African phenotypes as more "prototypical" of Black faces (Knuycky et al., 2013, Experiment 1). Although these studies are provocative, none of them were conducted within an eyewitness identification paradigm, with participants who watched a crime and then were asked to identify the culprit from a properly conducted identification procedure. Perhaps there are identification procedures that safeguard against phenotypic bias from biasing identification decisions and contaminating eyewitness accuracy.

1.2 Suspect bias versus general impairment in making eyewitness identification decisions

Scholars have posed two categories of factors that influence witness accuracy: general impairment and suspect-bias factors (Brewer and Wells, 2011; Smalarz, 2021). General impairment factors increase the likelihood that an eyewitness will make an identification error, but do not increase the rate of mistaken identifications of the suspect relative to identifications of known innocent lineup fillers. For example, errors caused by variables that increase the likelihood that the witness will make an identification (versus reject the lineup), such as failing to instruct the witness that the perpetrator may not be present in the lineup or presenting the lineup simultaneously (all photos at once) instead of sequentially (one photo at a time, no second lap), are expected to be distributed equally among lineup members, rather than being disproportionately directed to the suspect. The signal detection framework conceptualizes this willingness to make an identification of anyone as response bias (Wixted and Mickes, 2014).

In contrast, suspect bias factors encourage the witness to choose the suspect rather than any of the known-innocent fillers, meaning suspect identifications increase while filler identifications decrease (Kovera and Evelo, 2017; Smalarz, 2021). Factors that create suspect bias make suspects more vulnerable to identification (i.e., increased suspect identification rates) than they would be without that feature—regardless of whether the suspect is guilty or innocent. For example, when the lineup is biased such that the known-innocent fillers do not match the suspect's appearance, the likelihood the eyewitness will choose the suspect increases relative to the choice of

fillers (Fitzgerald et al., 2013). If a witness's identification decision is driven by something other than their own memory for the culprit, the lineup should be deemed unfair, even in cases when the culprit is present in the lineup and the witness accurately identifies them, because the identifications are induced through suggestiveness and therefore are legally illegitimate hits (Wells et al., 2012). If this increase of both correct and incorrect identifications of suspects occurs in the absence of an increase in overall choosing (rejection rates remain unchanged), neither discriminability nor response criterion have changed. This pattern of effects was documented in the non-blind lineup administration literature (Kovera and Evelo, 2017), in which there is a reliable filler-to-suspect shift in identifications that reflects neither changes in discriminability nor changes in response criterion. When a lineup identification procedure is tainted by suspect bias, the signal detection framework does not provide appropriate analysis for identifying its prejudicial nature, which affects the due process of both innocent and guilty suspects.

However, little to no research directly tests whether general impairment factors will moderate the influence of suspect bias factors but not other general impairment factors (Brewer and Wells, 2011). Phenotypic bias functions as a general impairment factor in eyewitness identifications: more stereotypically Black faces facilitated erroneous feelings of familiarity and recognition errors, with stereotypically Black faces being *more* likely and less phenotypically African faces being *less* likely to be mistakenly identified as previously seen (Knuycky et al., 2013). Moreover, phenotypic bias (a general impairment factor) has the potential to interact with biased lineup composition (a suspect bias factor) in a unique way. As discussed above, a lineup procedure can be suspect-biased if the lineup composition does not adequately protect the suspect, such as in cases when the fillers do not match the suspect's appearance (perhaps because the fillers do not share the same phenotypic facial features as the suspect). Suspect bias factors bias the witness toward choosing the suspect. General impairment factors may magnify this bias toward the suspect through a shift toward leniency in witnesses' criteria to make a positive identification from the lineup, decreasing rejections of the lineup overall (Brewer and Wells, 2011).

However, consider a case in which a Black suspect is less phenotypically African than his surrounding lineup members. The phenotypic differences between the suspect's and the fillers' appearance should steer witnesses toward picking that suspect over other fillers, regardless of his guilt status. Yet, previous research on phenotypic bias suggests that witnesses may implicitly associate people with more African phenotypes with criminality (Eberhardt et al., 2004, 2006), causing them to identify the fillers with more African phenotypes over the suspect with a less Afrocentric phenotype. Rather than biasing witnesses toward identifying the suspect, a phenotypic mismatch between suspect and fillers may steer witnesses away from identifying a suspect with less phenotypically African features (despite matching the phenotypic expression of the witnessed culprit) in favor of one of the more Afrocentric looking fillers, due to phenotypic bias. In this case, biased lineup composition—traditionally considered a suspect bias factor (Smalarz, 2021)—may protect more Eurocentric suspects and bias witnesses toward choosing a known-innocent filler.

1.3 Current study: goals and research questions

To investigate whether there are procedural safeguards that can be used to reduce the potential effects of phenotypic bias on eyewitness identifications, we evaluated whether phenotypic bias interacts with lineup presentation and composition to affect suspect identifications and accuracy. When witnesses are tasked with making an identification from a lineup, they might identify the police's suspect, which represents a correct *hit* in cases when the guilty culprit is present, and a problematic *false alarm* when the suspect is innocent (Cutler and Kovera, 2010). Alternatively, they might identify a known-innocent filler, which is a relatively harmless misidentification as the fillers will not be prosecuted, but is consistent with the witness using a more liberal response criterion (Cutler and Kovera, 2010; Lee and Penrod, 2019). Finally, witnesses might reject the lineup, which represents a *correct rejection* when the lineup contains an innocent suspect instead of the guilty culprit and a *miss* when the lineup contains the guilty culprit (Cutler and Kovera, 2010). The current study seeks to investigate racial bias and lineup fairness, rather than memory optimization, by observing whether a suspect's phenotype in relation to the phenotype of the fillers creates suspect bias by disproportionately increasing the likelihood that a witness will identify them. To investigate suspect bias (Smalarz, 2021), we will report correct hits and false alarms collapsed into an encompassing suspect identification variable. To observe potential effects on accuracy, we will compare identification decisions across culprit-present and culprit-absent lineups. If an independent variable had a main effect on affects suspect identifications, yet failed to on its own but does not interact with the culprit-presence variable, then the independent variable increased correct it influences both accurate identifications of the culprit and mistaken inaccurate identifications of the innocent suspect at the same rate. If equal rates. Obtaining this pattern of effects were to obtain, then the results would indicate that discriminability of the culprit from the innocent suspects would be was unaffected, despite an increased likelihood of choosing the designated suspect (suspect bias; Smalarz, 2021). However, if the culprit-presence variable were to interact with an independent variable to influence suspect identification decisions, the independent variable would have affected discriminability as it was increasing or decreasing identifications of the culprit from culprit-present (guilty suspect) and culprit-absent (innocent suspect) at different rates (Wixted and Mickes, 2014).

Photographic lineups are typically presented in one of two ways: simultaneously, meaning all photos are visible to the witness at once, or sequentially, meaning each photo is presented one at a time, and the witness is asked to render a judgment on each photo before moving onto the next. Presenting a lineup sequentially reduces the rate at which witnesses mistakenly identify innocent suspects (Steblay et al., 2011). One explanation for this reduction in misidentifications is that sequential lineups prompt witnesses to make absolute judgments about whether a lineup member matches their memory for the perpetrator, whereas simultaneous lineups prompt witnesses to make relative judgments about which lineup member looks the most like the perpetrator, much like choosing the "best" answer on a multiple-choice test (Wells, 1984). In lineups that do not contain the perpetrator of the crime, simultaneous

lineups pose a danger to innocent suspects who look the most like the perpetrator, relative to the other lineup members. Other scholars argue that the reduction in misidentifications seen in sequential lineups is caused by a higher degree of certainty required for witnesses to make an identification from a sequential lineup compared to a simultaneous lineup (Meissner et al., 2005; Flowe and Ebbesen, 2007; Goodsell et al., 2010). This type of criterion shift should affect the rate of mistaken choosing for both suspects and fillers by increasing or decreasing the criterion the witness sets for how well a face needs to match their memory for the culprit before choosing to identify them.

These interpretations of the sequential superiority effect have been challenged by proponents of diagnostic feature detection theory (DFDT), which predicts a memory advantage for simultaneous lineups compared to sequential lineups, because witnesses are better able to compare features of different lineup members when presented simultaneously (Wixted and Mickes, 2014). However, although responses tend to be more conservative for sequential lineups, some studies find little or no difference between the two procedures in underlying discriminability (Palmer and Brewer, 2012; Kaesler et al., 2020). Further, any differences in underlying discriminability between simultaneous and sequential lineups may simply reflect methodological choices that lack ecological validity (e.g., the first-yes-counts instruction; Horry et al., 2021; Winter et al., 2023). Because the current study's focus is investigating suspect bias (e.g., the likelihood of witnesses identifying the designated suspect due to non-memorial factors), this study will not focus on discriminability.

1.3.1 Hypotheses

If it is the case that general impairment factors will not interact to influence eyewitness accuracy, and both phenotypic bias and simultaneous presentation are general impairment factors, they should not interact to affect eyewitness decisions. Both factors should increase the likelihood that a witness chooses to identify someone from a lineup (reduce lineup rejections), but the increase should occur for both fillers and suspects if the lineup is fairly constructed. If so, sequential lineups will not serve as a safeguard against the general impairment caused by phenotypic bias. However, when the suspect has phenotypically African features that are not shared by the other lineup members, relative judgments made during simultaneous lineups may shift choosing toward that suspect. In this way, a general impairment factor (phenotypic bias) could act as a suspect biasing factor (unfair lineup composition). Because sequential lineups do not allow for this type of relative comparison, sequential lineup presentation should show less evidence of phenotypic bias when the lineup composition is unfair due to a mismatch between the phenotype of the suspect and the fillers. This prediction is consistent with findings that the superiority of the sequential procedure is maximized when the suspect is distinctive (Carlson et al., 2008) and that sequential presentation reduces other forms of bias (Lindsay et al., 1991).

To test these propositions, we conducted an experiment in which we manipulated the phenotype of the culprit, whether the fillers in a photo-array matched the suspect's phenotype, whether the photo-array was conducted simultaneously or sequentially, and whether the photo-array was culprit-absent or -present. The phenotype of innocent suspects in culprit-absent lineups always matched the phenotype of the witnessed culprit. We predicted that suspect phenotype would produce a main effect, with witnesses more likely to identify the suspect when the suspect had features that were more, rather than less, Afrocentric (hypothesis 1). We also predicted that presentation method would produce a main effect, with witnesses more likely to identify the suspect when the lineup was presented simultaneously compared to sequentially (hypothesis 2). Finally, we predicted a three-way interaction between phenotypic bias, lineup construction, and lineup presentation style (hypothesis 3). Specifically, we hypothesized that when lineups were presented simultaneously, there would be a significant two-way interaction of phenotypic bias and lineup composition, such that the simple main effect of phenotypic bias will be greater when lineups were suspect-biased (i.e., the fillers are a phenotypic mismatch to the suspect) than when all lineup members shared features with the same phenotype (hypothesis 3a). We expected that this interaction would be significantly smaller or non-significant for sequential lineups (hypothesis 3b).

2 Materials and methods

To test these predictions, we conducted an experiment with a 2 (Culprit/Innocent Suspect Phenotype: More African vs. Less African) × 2 (Lineup Fillers: Phenotypic Match to Culprit/Innocent Suspect vs. Phenotypic Mismatch) × 2 (Photo Presentation: Simultaneous vs. Sequential) × 2 (Array Type: Culprit-present vs. Culprit-absent) between-subjects factorial design. Our dependent variables of interest were the witness's identification decision and their confidence in that decision.

2.1 Participants

Because there are no statistical packages for reliably determining needed sample sizes to power for a three-way interaction effect within a logistic regression analysis (Aberson, 2019), we approached estimating the sample size we needed to sufficiently power the test of our predicted effects in two ways. First, we conducted a power analysis for the test of the predicted three-way interaction using an ANOVA with GPower 3.1.9.1 (Faul et al., 2009). That analyses suggested that a sample of 560 participants would be sufficient to detect relatively small effects (partial $\eta^2 = 0.02$) with alpha = 0.05 and power = 0.80, even after taking into account that power is multiplicative across our three predicted effects (Maxwell, 2004; Schimmack, 2012). Second, we compared this suggested sample size with crude estimates for sufficient sample size in logistic regression models, including having at least 10 cases per predictor with a sample size of 500 typically producing adequate power (Long, 1997). In our analyses, there were at most 11 predictors (4 main effects, 6 two-way interactions, and 1 three-way interaction) in the model.

Thus, we recruited 600 White adults (50% women, 3% White Hispanic) through Qualtrics Panels, which was the maximum number of participants we could recruit given our funding constraints. Participants ranged in age from 18 to 87 ($M = 43.94$, SD = 14.25). Participants were compensated $1.50 for their time.

2.2 Materials

Materials included the mock theft video, the photo arrays, and a Qualtrics questionnaire.

2.2.1 Mock theft video

Participants viewed a video of a mock theft in which a Black man entered an office and stole an iPhone from a backpack located in that office. The culprit's face was visible for 8 seconds. For purposes of stimulus sampling to increase the generalizability of our results (Wells and Windschitl, 1999), we video recorded six different versions of the theft, with a different actor portraying the culprit in each version. The videos are available to view online at https://osf.io/am5qh/?view_only=7cf96a56de9946348e001ef2885df2d7.

2.2.2 Culprit phenotype

To ensure that the culprits were good representations of the categories they were to represent (e.g., Black men with facial features that were representative of a more or less African facial phenotype), graduate students rated headshots of 86 Black men who responded to a solicitation on Craigslist.org to play a perpetrator in a mock crime video. Students rated the headshots on the actors' attractiveness and distinctiveness on 1 (*not at all*) to 7 (*extremely*) Likert-type scales. They also rated the extent to which the actors looked stereotypically Black on a 1 (*not at all*) to 9 (*extremely*) Likert-type scale. They rated age on a 10-point scale, with each point representing a 5-year increment, starting with 15–20 and ending with 60+. Finally, they also indicated the perceived race of the actors. We selected six actors, three of whom received low stereotypicality scores (mean scores ranging from 3.5 to 4.4) for use in our Less African condition) and three of whom received high stereotypicality scores (mean scores ranging from 6.2 to 6.8) for use in our More African condition. All six actors were correctly categorized as Black by over 90% of the pilot participants and were perceived as being between 20 and 40 years old by over 85% of participants. The actors were rated similarly on attractiveness, distinctiveness, and age. When we filmed the stimulus materials, we took photographs of the perpetrators displayed against a white background for later use in the culprit-present photo arrays.

2.2.3 Photo array

We created six-person photo arrays for each culprit, each containing one suspect (either the culprit or a designated innocent suspect) and five fillers. The photo arrays orthogonally varied whether the culprit was present and whether the fillers matched the suspect's phenotype. Fillers and innocent suspects were generated from a pool of 127 photos of Black men in the Chicago Face Database (Ma et al., 2015) or used in previous research (Eberhardt et al., 2004). All photographs were edited to depict the fillers' and innocent suspects' faces in front of a white background. Faces did not have facial hair or other distinctive features (i.e., unique hair styles, piercings, and tattoos). We also ensured that the fillers matched the characteristics most frequently mentioned in descriptions of culprits obtained from MTurk workers (240 workers provided descriptions, with 12 workers providing descriptions of each face in our facial database).

We selected one photograph from the facial databases to serve as the innocent suspect for each of the culprit-absent lineups (three innocent suspects in total). Similarly, we selected fillers for both the culprit-present and -absent lineups from these two databases. MTurk workers rated the faces distinctiveness, attractiveness, age, and stereotypicality and categorized the faces based on perceived race, with 15–22 participants rating each face. Participants correctly categorized all faces chosen to serve as innocent suspects and fillers as Black. Faces chosen to serve as innocent suspects matched the culprit on phenotype, attractiveness, distinctiveness, and age in that the 95% confidence interval (CI) of the mean difference between the rating of the culprit and of the innocent suspect included zero for attractiveness, distinctiveness, and age. The selected fillers had ratings of average distinctiveness (defined as ±2 SD of filler distinctiveness mean, M = 3.94, SD = 0.43), average attractiveness (±2 SD of filler attractiveness mean, M = 3.37, SD = 0.54), and an age similar to the culprits (M = 3.08, SD = 1.10). Fillers selected to serve as more African phenotypic fillers had stereotypicality ratings of 5.36–7.45 (M = 6.47, SD = 0.62). Fillers selected to serve as less African phenotypic fillers had stereotypicality ratings of 2.41–6.06 (M = 4.37, SD = 0.77). All 24 photo-arrays are available at https://osf.io/am5qh/?view_only=7cf96a56de9946348e001ef2885df2d7.

2.3 Procedure

This study was approved by the Institutional Review Board at John Jay College of Criminal Justice (City University of New York) and run online using Qualtrics. After consenting to participate and providing their demographic information, participants were given the following instructions, "You are going to be shown a short film. Pay close attention because you will be asked some questions afterwards. There will be no audio in this film. When you are ready to view the film, please press the NEXT button." On the next page of the survey, participants were provided further instructions, which were displayed above the simulated crime video and read, "Please watch the video. There is no sound. Do not click or pause the video. You can advance this page once the timer is finished. The page will automatically advance after 4 minutes." Each sentence in both instructions was displayed as a separate bullet point. Participants watched one version of the mock crime video, in which they viewed one of our Black male culprits who appeared either more phenotypically African or less phenotypically African. After viewing the crime video, participants were instructed, "What you just witnessed was a crime recorded on video surveillance. In a few minutes, you will be asked to identify the individual who stole the phone in the video. Before that occurs, we want to measure your attention". Participants were then instructed on and completed a 3-min word puzzle as a filler task. Following the filler task, participants were provided lineup instructions. All participants were instructed, "Time: You may spend as long as you want on this page. You must spend at least 1 minute (see timer) to ensure comprehension."

Participants in the simultaneous lineup conditions were instructed, "On the next page you will view six photos. Your task is to identify which photo, if any, depicts the culprit from the video who stole the phone. There is no time limit for this task. You may take as long as you like to make a decision and you may change your selection as many times as you like before you submit. Once you submit your answer (by clicking NEXT) you will not be able

to return to the lineup. You should know: (1) The culprit might not be in the lineup at all, so the correct answer might be 'not present,' (2) If you feel unable to make a decision, you have the option of responding 'don't know,' (3) After making a decision you will be asked to state how confident you are in that decision, (4) The investigation will continue if no identification is made." On the next page displayed above the lineup, participants in the simultaneous lineup conditions were further instructed, "Below are photos of six faces. Please identify which photo depicts the individual who stole the phone in the video. If you do not think any of these photos depicts the person who stole the phone, please select 'The culprit is not present.' If you are unsure if the culprit is present or not, please select 'I don't know if the culprit is present.' You will be asked to give your confidence in this decision on the next page."

Participants in the sequential lineup conditions were instructed: "On the following pages you will view photos of several individuals. Each photo will be shown one at a time. Your task is to decide if the photo is of the same individual who stole the smart phone in the video you watched. For each photo, select YES if you do think that individual in the photo is the person who stole the smart phone. Select NO if you do not think that the individual in the photo is the person who stole the smart phone. There is no time limit for this task. You may take as long as you want on each page and you may change your selection (YES/NO) as many times as you like before you submit. Once you submit your answer (by clicking NEXT), you will not be able to go back. You will only be able to see each photo once. You should know: (1) The culprit might not be in the lineup at all, so the correct answer might be to choose NO for all photos, (2) After making a decision you will be asked to state how confident you are in that decision, (3) The investigation will continue even if no identification is made." On the next pages displayed above each photo, participants in the simultaneous lineup conditions were further instructed, "Please decide if the photo below depicts the culprit from the video who stole the phone."

Participants were presented one six-person lineup that contained either the culprit they saw in the video (culprit-present) or an innocent suspect (culprit-absent) who always matched the culprit's phenotype, and had been rated as equally attractive, equally distinct, and around the same age (rating and selection procedure described above). The suspect in the lineup always matched the phenotype of the culprit witnesses viewed in the video (i.e., if a witness viewed a less Afrocentric perpetrator in the video, they were presented with a less Afrocentric suspect in the lineup, and if a witness viewed a more Afrocentric perpetrator in the video, they were presented with a more Afrocentric suspect in the lineup). The lineups also contained fillers who either matched the culprit/suspect's phenotype (Phenotypic Match: less Afrocentric suspects/culprits surrounded by less Afrocentric fillers; more Afrocentric suspects/culprits surrounded by more Afrocentric fillers) or did not match the culprit/suspect's phenotype (Phenotypic Mismatch: less Afrocentric suspects/culprits surrounded by more Afrocentric fillers; more Afrocentric suspects/culprits surrounded by less Afrocentric fillers). The suspect was always in position number five.

Lineups were either presented simultaneously (all photos at once) or sequentially (one at a time). Sequential lineups presented each photo in order one at a time and were concluded once the witness made an identification (first identification stopping rule). For each photo, participants in the sequential conditions answered "Yes" or "No" and were not given a "Don't Know" response option. If participants went through all the photograph options without identifying anyone, their identification decision was recorded as a lineup rejection. Lineup photos were only presented once. Participant-witnesses in the simultaneous lineup conditions were instructed to either identify one of the six photos as the culprit, report that the culprit is not present, or indicate they do not know if the culprit is present. Not present and do not know responses were categorized as lineup rejections (i.e., non-identifications). Finally, participants reported their confidence in their identification decision (0%–100%) and completed a demographics questionnaire that asked for their age, gender, race/ethnicity, and education level.

3 Results

Participants' identification decisions were re-categorized into three dichotomous variables: (1) suspect identification versus other (other = filler identifications and rejections), (2) filler identifications versus other (other = suspect identifications and rejections), and (3) rejections versus other (other = filler and suspect identifications). These variables allow us to investigate suspect bias, our main dependent variable of interest, while using culprit presence to detect any differences between accurate and inaccurate witnesses (where suspect identifications are correct *hits* in culprit-present lineups, but *false alarms* in culprit-absent lineups, and rejections are *correct rejections* in culprit-absent lineups but *misses* in culprit-present lineups). Identification decisions across experimental conditions are presented in **Table 1**. Measures of discriminability (d') and response bias (c) for each experimental condition are available in the supplemental materials (**Supplementary Table 1**).

We used IBM's SPSS Version 28 to conduct forced-entry, hierarchal binary logistic regressions to test our predicted effects on our three dependent variables: suspect identifications, filler identifications, and photo array rejections. We entered suspect/culprit phenotype, lineup presentation style, filler composition, and culprit-presence as predictors in the first block, all two-way interactions into the second block, and our predicted three-way interaction between phenotypic bias, lineup construction, and lineup presentation into the third block. We ran analyses twice: once including all participants, both White Hispanics (n = 18, 3%) and White Europeans (n = 582, 97%), and once excluding White Hispanic participants and only including White European participants. Excluding White Hispanics did not produce significant differences in the pattern of the results, so we have reported the results based on the data provided by both White Europeans and White Hispanics.

This study was designed and powered to evaluate identification decision rates, but participant-witnesses also reported their numeric confidence in their identification decision. Our confidence analysis is exploratory, as we did not make prior predictions regarding differences in confidence nor collect enough observations to appropriately power for confidence-accuracy analysis. For our exploratory analysis, we ran an analysis of variance using IBM's SPSS Version 28 general linear model univariate procedure to compare confidence means across groups. Confidence-accuracy characteristic (CAC) analysis comparing the accuracy of highly

TABLE 1 Frequency of identification decisions in each condition.

Independent variables										
Culprit/suspect features	Lineup composition	Presentation	Array type	Suspect		Filler		Rejections		Total
				n	%	*n*	%	*n*	%	*n*
More phenotypic	Mismatch	Simultaneous	Present	16	41	10	26	13	33	39
			Absent	6	19	8	26	17	55	31
			Total	22	31	18	26	30	43	70
		Sequential	Present	10	28	14	39	12	33	36
			Absent	4	9	18	41	22	50	44
			Total	14	18	32	40	34	43	80
		Total		36	24	50	33	64	43	150
	Match	Simultaneous	Present	10	29	14	40	11	31	35
			Absent	1	3	14	39	21	58	36
			Total	11	15	28	39	32	45	71
		Sequential	Present	4	10	25	63	11	28	40
			Absent	1	3	19	49	19	49	39
			Total	5	6	44	56	30	38	79
		Total		16	11	72	48	62	41	150
	Total			52	17	122	41	126	42	300
Less phenotypic	Mismatch	Simultaneous	Present	11	29	10	26	17	45	38
			Absent	5	13	11	29	22	58	38
			Total	16	21	21	28	39	51	76
		Sequential	Present	7	19	11	30	19	51	37
			Absent	8	22	12	32	17	46	37
			Total	15	20	23	31	36	49	74
		Total		31	21	44	29	75	50	150
	Match	Simultaneous	Present	11	24	15	33	20	43	46
			Absent	4	11	16	43	17	46	37
			Total	15	18	31	37	37	45	83
		Sequential	Present	6	21	13	45	10	35	29
			Absent	1	3	18	47	19	50	38
			Total	7	10	31	46	29	43	67
		Total		22	15	62	41	66	44	150
	Total			53	18	106	35	141	47	300
Total				105	18	228	38	267	45	600

Values are given in raw counts and percentages within conditions. Suspect identifications are correct hits in culprit-present lineups and false alarms in culprit-absent lineups. Rejections are correct rejections in culprit-absent lineups and misses in culprit-present lineups. Rejections include both "not present" and "don't know" responses. The percentages in some rows do not total 100 because of rounding.

10.3389/fpsyg.2024.1233782

confident suspect identifications in the phenotypic match versus phenotypic mismatch conditions was calculated using Microsoft Excel and plotted using R Version 4.3.1 ggplot package. Results of our confidence analyses are available in **Supplementary Tables 2–5** and **Supplementary Figure 1**.

3.1 Suspect identification rates

Using suspect identifications as our dependent variable, we entered suspect/culprit phenotype, lineup presentation style, filler composition, and culprit-presence as predictors in the first block, all two-way interactions into the second block, and our predicted three-way interaction between suspect/culprit phenotype, lineup presentation style, and filler composition into the third block. The second block of the analyses did not show significant improvement in the model, Wald's χ^2 (6, N = 600) = 9.36, p = 0.155, nor did the third block, Wald's χ^2 (1, N = 600) = 0.01, p = 0.946. Because our interaction terms did not improve model fit, we report results from the first block below in-text (Field, 2013). The full-factorial model is presented in **Supplementary Table 6**. All two-way interactions and our predicted three-way interaction between suspect/culprit phenotype, lineup presentation style, and filler composition were not statistically significant, and the main effects for lineup composition and lineup presentation reported below lose significance when interaction terms are included.

Contrary to our first hypothesis, there was no main effect of phenotypic bias. Witnesses were no more likely to identify the suspect when the suspect had more Afrocentric features (17%) than when he had less Afrocentric features (18%), B = 0.005, SE = 0.22, Wald's χ^2 (1, N = 600) = 0.001, p = 0.982, OR = 1.01, 95% CI [0.65, 1.56]. However, the phenotypic match of the lineup fillers to the suspect produced a significant main effect, such that witnesses were more likely to identify the suspect when the other lineup fillers did not match the suspect's phenotype (**Figure 1**). Witnesses who were presented with a suspect with relatively more Afrocentric features surrounded by fillers with relatively less Afrocentric (more Eurocentric) features or a suspect with relatively less Afrocentric features surrounded by fillers with more Afrocentric features were more likely to identify the mismatched suspect (22%) than were witnesses who were presented with a lineup composed of faces that all matched in phenotype (13%), B = 0.73, SE = 0.23, Wald's χ^2 (1, N = 600) = 10.35, p = 0.001, OR = 2.08, 95% CI [1.33, 3.26], irrespective of whether that phenotype was more or less African.

We found support for our second hypothesis, as there was a significant main effect for lineup presentation, such that witnesses were more likely to identify the suspect when the lineup was presented simultaneously (21%) compared to sequentially (14%), B = −0.53, SE = 0.23, Wald's χ^2 (1, N = 600) = 5.53, p = 0.019, 1/OR = 1.70, 95% CI [1.09, 2.65]. This analysis also reproduced the well-established finding that participant-witnesses make more identifications of guilty culprits from culprit-present lineups (25%) than identifications of innocent suspects from culprit-absent lineups (10%), B = 1.10, SE = 0.24, Wald's χ^2 (1, N = 600) = 21.5, p < 0.001, OR = 3.00, 95% CI [1.82, 4.78]. Despite no significant interaction emerging between lineup presentation and culprit presence, a closer look at the data reveals that the increase in suspect identifications in simultaneous lineups was driven by choosing in culprit-present lineups: participant-witnesses were more likely to accurately identify the guilty culprit in simultaneous lineups compared to sequential lineups (30% vs. 19%), but nearly equally likely to inaccurately misidentify the innocent suspect (11% vs. 9%).

3.2 Filler identifications

Using filler identifications as our dependent variable, we entered suspect/culprit phenotype, lineup presentation style, filler composition, and culprit-presence as predictors in the first block, all two-way interactions into the second block, and our predicted three-way interaction between suspect/culprit phenotype, lineup presentation style, and filler composition into the third block. The second block of the analyses did not show significant improvement in the model, Wald's χ^2 (6, N = 600) = 2.96, p = 0.814, nor did the third block, Wald's χ^2 (1, N = 600) = 0.07, p = 0.791. Because our interaction terms did not improve model fit, we report results from the first block below in-text (Field, 2013). The full-factorial model is presented in **Supplementary Table 7**. All two-way interactions and our predicted three-way interaction between suspect/culprit phenotype, lineup presentation style, and filler composition were not statistically significant, and the main effects for lineup composition and lineup presentation reported below lose significance when interaction terms are included.

Again, there was no main effect of phenotypic bias. Witnesses were not significantly more likely to identify a known-innocent filler when the suspect had more stereotypically African features (41%) than less stereotypically African features (35%), B = 0.21, SE = 0.17, Wald's χ^2 (1, N = 600) = 1.43, p = 0.232, OR = 1.23, 95% CI [0.88, 1.72]. The manipulation of the phenotypic match of the fillers to the suspects produced a significant main effect, B = −0.59, SE = 0.17, Wald's χ^2 (1, N = 600) = 11.86, p < 0.001, 1/OR = 1.81, 95% CI [1.29, 2.53]. Witnesses were more likely to identify a known-innocent filler when the lineup fillers matched the suspect's phenotype (**Figure 1**). Witnesses who were presented with a suspect with more Afrocentric features surrounded by fillers with less Afrocentric (more Eurocentric) features or a suspect with less Afrocentric features surrounded by fillers with more Afrocentric features were less likely to identify a known-innocent filler (31%) than were witnesses who were presented with a lineup composed of faces that all matched in phenotype (45%).

There was a significant effect of lineup presentation such that witnesses were more likely to identify a known-innocent filler when the lineup was presented sequentially (43%) as opposed to simultaneously (33%), B = 0.47, SE = 0.17, Wald's χ^2 (1, N = 600) = 7.36, p = 0.007, OR = 1.6, 95% CI [1.14, 2.24]. Finally, participant-witnesses were no more likely to make filler identifications from culprit-absent lineups (39%) than they were from culprit-present lineups (37%), B = −0.03, SE = 0.17, Wald's χ^2 (1, N = 600) = 0.04, p = 0.846, 1/OR = 1.03, 95% CI [0.74, 1.45].

3.3 Lineup rejections

Using lineup rejections as our dependent variable, we entered suspect/culprit phenotype, lineup presentation style, filler

FIGURE 1
Identification decisions by lineup compositions.

composition, and culprit-presence as predictors in the first block, all two-way interactions into the second block, and our predicted three-way interaction between suspect/culprit phenotype, lineup presentation style, and filler composition into the third block. As mentioned above, our "lineup rejections" variable includes both "not present" and "don't know" responses. The second block of the analyses did not show significant improvement in the model, Wald's χ^2 (6, N = 600) = 4.86, p = 0.563, nor did the third block, Wald's χ^2 (1, N = 600) = 0.09, p = 0.760. Because our interaction terms did not improve model fit, we report results from the first block below (Field, 2013), along with one significant non-predicted two-way interaction that emerged in the second block and remained significant in the third block. The full-factorial model is presented in **Supplementary Table 8**. All other two-ways and our predicted three-way interaction are not statistically significant, and the main effect for culprit-presence reported below loses significance when interaction terms are included.

Once again, there was no main effect of suspect phenotype. Witnesses were not significantly more likely to reject the lineup when the suspect had more Afrocentric features (42%) than when he had less Afrocentric (more Eurocentric) features (47%), B = −0.20, SE = 0.17, Wald's χ^2 (1, N = 600) = 1.42, p = 0.234, 1/OR = 1.22, 95% CI [0.88, 1.69].

In contrast to its effects on suspect and filler identifications, whether the phenotype of the fillers matched that of the suspect did not affect the rate of lineup rejections. Witnesses were no more likely to reject the lineup when the suspect's phenotype mismatched the fillers' phenotype (46%) than when the suspect's phenotype matched the fillers' phenotype (43%), B = 0.16, SE = 0.17, Wald's χ^2 (1, N = 600) = 0.88, p = 0.348, OR = 1.20, 95% CI [0.85, 1.62] (**Figure 1**). Similarly, lineup presentation did not affect the rate of lineup rejections; witnesses were equally likely to reject the lineup when the lineup was presented sequentially (43%) as opposed to simultaneously (46%), B = −0.15, SE = 0.17, Wald's χ^2 (1, N = 600) = 0.78, p = 0.377, OR = 0.86, 95% CI [0.62, 1.20].

Participant-witnesses were more likely to reject the lineup when the culprit was absent (51%) than when the culprit was present (37%), B = −0.57, SE = 0.17, Wald's χ^2 (1, N = 600) = 11.60, p < 0.001, 1/OR = 1.76, 95% CI [1.27, 2.44]. This main effect was qualified by a significant two-way interaction that emerged between suspect phenotype and culprit presence, B = −0.66, SE = 0.34, Wald's χ^2 (1, N = 600) = 3.83, p = 0.050, 1/OR = 1.93, 95% CI [1.00, 3.73]. Although correct rejection rates were about equal when the lineup contained an innocent suspect (more Afrocentric = 53%, less Afrocentric = 50%), participant-witnesses were more likely to inaccurately reject a culprit-present lineup when the guilty suspect had more Eurocentric features (44% incorrect rejection) than when the guilty suspect had more Afrocentric features (31% incorrect rejection).

4 Discussion

There are large racial disparities in the number of wrongful convictions based on eyewitness misidentifications of Black versus White defendants. The size of the own-race bias effects is not sufficient to explain these disparities (Katzman and Kovera, 2023). Although some of these disparities may be the result of disparate policing practices that lead more innocent Black than White men to be subjected to the risk of misidentification (Katzman and Kovera, 2023), there may be other racial biases that contribute to them. Phenotypic bias, a bias against individuals who have more Afrocentric facial features (Knuycky et al., 2013), may help to explain this disparity. Because phenotypic bias operates on Black rather than White target lineups, it is possible that it may put innocent Black suspects at greater risk of misidentification. This experiment had three goals: (a) to examine whether phenotypic bias affects eyewitness identification decisions, (b) to investigate whether a phenotypic mismatch between fillers and a suspect may bias the lineup against the suspect, and (c) to explore whether sequential lineup presentation might guard against the harmful effects of phenotypic bias and phenotypic mismatch.

4.1 Suspect phenotype effects

Phenotypic bias did not influence our participant-witnesses' identification decisions in the way that we hypothesized. Indeed, we found little evidence that witnesses were more likely to identify suspects if they had more Afrocentric features rather than more Eurocentric features, whether they were the culprit or an innocent suspect. The only evidence to support the supposition that a more African phenotype promotes mistaken identifications comes from our finding that participants were more likely to incorrectly reject a culprit-present lineup when the guilty suspect had less stereotypically African features than when the guilty suspect had more stereotypically African features. This finding suggests that our participants used a higher criterion for identifying the less Afrocentric culprit than the more Afrocentric culprit. However, the increased choosing rates for more Afrocentric suspects were distributed evenly across both suspects and fillers, and thus did not differentially increase suspect identifications in this study. Further, this finding was unexpected, obtained from a logistic regression model that did not improve model fit, and was significant at $p = 0.050$, all of which suggest that this finding should be interpreted with caution. Other than this one effect, phenotype—on its own—had little influence on witnesses' decisions.

There are several possibilities for why phenotypic bias failed to appear in this eyewitness context. First, the type of phenotypic bias observed in other research may simply not extend to eyewitness recognition tasks. In previous studies, phenotypic bias affected inferences about criminality (Eberhardt et al., 2004; Kahn and Davies, 2011) and deservingness of punishment (Blair et al., 2004b; Eberhardt et al., 2006). These types of inferences may be more susceptible to bias than a facial recognition task, in which witnesses' judgments are at least somewhat constrained by their memory for the culprit and whether any of the photos before them provide a good match to their memory (Clark, 2003). However, we also may have simply failed to produce phenotypic bias because of our study design. In all our lineups, the suspect matched the culprit's phenotype. Future research should manipulate phenotypic match between the culprit and the suspect to investigate whether witnesses are more likely to misidentify an innocent suspect with more Afrocentric features when the culprit had more Eurocentric features than when the reverse is true (i.e., an innocent person with more Eurocentric features is suspected of being a culprit who had more Afrocentric features).

It is also possible that phenotypic bias would have extended to eyewitness recognition tasks in the past, but increased societal attention to implicit racial bias provided our participants with the self-awareness and motivation to avoid acting on these biases. With the massive boom in the Black Lives Matter movement after the murder of George Floyd in 2020, experimental work on racial bias has diverged from real-world field data which consistently demonstrates racial bias, in part due to social desirability effects (Salerno et al., 2023; Smalarz et al., 2023). This explanation is somewhat less likely as the data were collected before the COVID-19 pandemic began, thus before the racial unrest prevalent in the aftermath of Floyd's murder. However, the possibility that social desirability concerns were present prior to these events remains.

4.2 Lineup composition effects

Even though our participants did not exhibit phenotypic bias in their identifications of suspects, they were sensitive to variations of phenotype among the people depicted in the photo arrays. Participant identification decisions were affected by whether the suspect had a different phenotypic expression than the known-innocent fillers. When there was a phenotypic mismatch between the suspect and the lineup fillers, witnesses were more likely to identify the suspect, regardless of whether the suspect was guilty. That is, witnesses were (a) more likely to identify the suspect from a biased rather than an unbiased photo array (b) less likely to identify fillers from a suspect biased photo array, and (c) no more likely to state that a culprit is not present from a biased than unbiased photo array (**Figure 1**). This pattern of results resembles findings from the double-blind administration literature, known as the "filler-to-suspect shift" (Kovera and Evelo, 2017). The filler-to-suspect shift represents the phenomenon that when a lineup administration is single-blind (i.e., when the lineup administrator knows the identity of the suspect), the witness is more likely to identify the suspect and less likely to identify a filler than when the lineup administrator is double-blind (i.e., when the lineup administrator does *not* know the identity of the suspect). However, administrator knowledge does not affect the likelihood that the witness rejects the lineup. Because administrator knowledge of the suspect does not increase the proportion of witnesses who make an identification, administrator knowledge does not affect witnesses' criterion to make an identification (Kovera and Evelo, 2017). Thus, just as the non-blind administrator communicates the identity of the suspect to the witness, our mismatched phenotypic lineups communicated to witnesses which photo depicted the suspect, especially to witnesses who were willing to identify someone from the photo array but may not have a strong match between their memory of the culprit and any particular member of the lineup. We observed no shift in decision criterion to make an identification, as rejection rates were the same across phenotypically matched lineups and phenotypically mismatched lineups. However, their ability to discern which lineup member is the suspect among the fillers increased in mismatched lineups (suspect bias). This increased discernment of the suspect did not translate into the ability to discriminate guilty suspects from innocent suspects; instead, it simply created suspect bias, rather than improving signal discriminability versus noise. The phenotypic mismatch of the suspect with the fillers leads those willing witnesses to choose the mismatched suspect rather than a filler. Thus, it may be particularly important to ensure that fillers match suspects on phenotypic expression when witness memories are weaker or when their criterion for choosing may be low.

Our findings underscore the theoretical importance of examining the variety of ways that suspect bias manifests (Smalarz, 2021) and the methodological importance of designating an innocent suspect in culprit-absent lineups, as these analyses would not have been possible otherwise. Put simply, mismatched phenotypes can make the suspect "stand out" in the lineup and put innocent suspects at greater risk of misidentification. A large body of research has investigated biasing factors of lineups, including mismatched backgrounds, clothing, and lighting (Lindsay et al., 1987; Harvard et al., 2023). Indeed, witnesses in real cases are more

likely to identify suspects when the lineup is demonstrably biased toward the suspect according to mock witness studies of lineup fairness (Steblay and Wells, 2020). Yet phenotypic mismatching has escaped empirical notice.

Overall, to reduce disproportionate identifications of Black suspects, lineup fillers should always match the suspect's phenotype. However, an archival study of 250 offender descriptions by witnesses of armed bank robberies revealed that when describing the offender, witnesses reported few identifying details and information related to phenotype was not among the frequent descriptors used (Fahsing et al., 2004). One method that police officers use to construct lineups is to find known-innocent fillers who match the witness's description of the culprit (e.g., build, hair, and race). However, if phenotype expression is not included in these descriptions, fillers who match the suspect on every other descriptor will still not provide adequate protection for the suspect. When investigators rely on witnesses' imprecise descriptions to construct their photo arrays, known-innocent fillers that match the general description of the culprit provided by the witness may possess disqualifying features that ultimately reduce the lineup's functional size and the protections provided to the suspect. Researchers are developing interview strategies to elicit bountiful and accurate offender descriptions from witnesses. For example, the person description interview (PDI), which includes a general-to-specific instruction (GSI) and a down-to-up instruction (DUI) tested both in the laboratory and in the field, meaningfully increased the amount and accuracy of facial descriptors (Demarchi and Py, 2009). If a witness's memory for the perpetrator is not strong enough to provide a detailed description, the reliability of any positive identification they make should be questioned.

4.3 Lineup presentation effects

Participants' identification decisions were influenced by how the lineup was presented. When the lineup was presented simultaneously rather than sequentially, participants were more likely to identify the suspect, less likely to identify a known-innocent filler, and equally likely to reject the lineup. However, lineup presentation was included in our study design for its potential to mitigate the problematic effects of both phenotypic bias and phenotypic mismatch in lineup composition. Although we did not find an effect for phenotypic bias (and thus no intervention is required to address it), sequential presentation failed to protect innocent suspects: a closer look at our data revealed that the increase in suspect identifications we observed in simultaneous lineups was driven by choosing in culprit-present lineups, such that participant-witnesses were much more likely to accurately identify the guilty culprit in simultaneous lineups compared to sequential lineups (30% vs. 19%), but nearly equally likely to inaccurately misidentify the innocent suspect (11% vs. 9%).

This pattern of results partially mirrors findings from other studies in which witnesses were more likely to positively identify perpetrators from culprit-present lineups presented simultaneously rather than sequentially (Steblay et al., 2001, 2011; Steblay and Wells, 2020). However, meta-analyses also find that witnesses are more likely to correctly reject lineups from culprit-absent sequential than culprit-absent simultaneous lineups (Steblay et al., 2001, 2011). The current study did not reproduce this effect: participant-witnesses correctly rejected the culprit-absent lineup 54% of the time when it was presented simultaneously, and 49% of the time when presented sequentially. Instead, our participant-witnesses were more likely to make a filler identification from sequential lineups (culprit-present = 44%, culprit-absent = 42%), than from simultaneous lineups (culprit-present = 31%, culprit-absent = 35%). Thus, although suspect identification rates were higher overall in simultaneous lineups, sequential lineup presentation did not provide protections for innocent suspects, and only acted to reduce accurate culprit identifications.

Although scholars argue and there is empirical evidence that sequential presentation can reduce mistaken identifications resulting from suspect bias in photo arrays, perhaps by diminishing eyewitnesses' reliance on relative judgment processes (Lindsay and Wells, 1985), we found that sequential lineup presentation was an inadequate safeguard for suspect bias based on phenotypic mismatch. Perhaps the strength of our manipulation of phenotypic mismatch was strong enough and noticeable enough to allow witnesses to hold that information in mind when making their decisions about sequentially presented photos. Whatever the reason, given that phenotypic match seems to operate differently than other types of suspect bias, it is ripe for continued empirical examination.

5 Future research and conclusion

This study was conducted entirely online. Although we took care to maximize the study's ecological validity by filming a realistic mock crime video and including a filler task, the social context in which identifications are made can influence the identifications made by witnesses (Kovera and Evelo, 2021). Future researchers could benefit from exploring these questions using in-person paradigms. Additionally, the filler task in this study only provided a 3-min retention interval between viewing the perpetrator and being asked to make an identification. Ecological validity would be heightened if future researchers use a retention interval that more accurately matches the average interval witnesses experience in the field. In addition, to provide better recommendations to law enforcement, future research could tease apart which of these prototypically African features witnesses rely on most by isolating and manipulating each feature. For example, perhaps fillers need only match the suspect on skin tone and hair texture, but not nose shape.

Future research should also examine the generalizability of these effects to contexts in which the encoding conditions are more favorable to witness memory. To explore whether phenotypic bias affects eyewitness identification decisions, we intentionally created encoding conditions (e.g., an 8 s exposure duration) that were likely to produce weak memory traces that would allow for bias to operate. As a result, we obtained more lineup rejections and fewer suspect identifications than are typically seen in actual eyewitness identification decisions (see Wells et al., 2020 for a review of estimates of the types of eyewitness decisions made by witnesses in actual cases). Scholars should explore the extent to which these findings hold under better encoding conditions.

Moreover, this study was not designed to investigate the role phenotypic bias plays on the *own-race bias* (ORB), as we only investigated White participant-witnesses attempting identifications of Black perpetrators from lineups composed entirely of Black men. Although empirical studies consistently produce the ORB, there is substantial (and currently inexplicable) variation in the size of this effect (Lee and Penrod, 2022). Within-race differences in appearance could be a meaningful contributor to this variation (Chiroro et al., 2008). Future researchers interested in evaluating how phenotypic bias may affect the ORB should fully cross the design by collecting data from both Black and White participant-witnesses making identifications from both Black and White lineups. Future research should examine whether both Black and White witnesses are similarly affected by phenotypic mismatch.

Finally, facial recognition scholars have spent decades investigating the causes of the own-race bias, but virtually no research has examined why Black suspects are misidentified at higher rates than White suspects. Although scholars suggest that racial disparities in exonerations based on eyewitness misidentifications may be largely explained by an officer's decision to place Black suspects in lineups when there is little evidence connecting them to the crime (Katzman and Kovera, 2023), lineup construction issues may also contribute. For instance, recent meta-analytic findings suggest that both Black and White witnesses may perform worse on Black than White target lineups (Katzman and Kovera, 2023). Additionally, in a study examining lineup fairness, Black suspects were more likely to be identified from lineups by *both* Black and White mock witnesses (Brigham et al., 1990). Thus, including phenotypic bias as a factor in future investigations could provide (a) greater understanding of the psychological mechanisms responsible for the variations in the size of the ORB, and (b) an explanation for the finding that under certain conditions, Black suspects are at uniquely high risk of being misidentified as the perpetrator of a crime. In the meantime, the findings from this study strongly support that police take care to match the facial phenotype of the suspect when choosing fillers to appear in photo arrays and lineups to eliminate one form of suspect bias.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://osf.io/am5qh/?view_only=7cf96a56de9946348e001ef2885df2d7.

Ethics statement

The studies involving humans were approved by John Jay College – City University of New York. The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional review board waived the requirement of written informed consent for participation from the participants or the participants' legal guardians/next of kin because data were collected on the internet so written consent was not possible and would have been the only thing linking participants to their data.

Author contributions

JMJ conducted most of the data analyses, wrote much of the initial draft of the manuscript, and revised and edited the manuscript. JK provided data analytic support, wrote sections of the initial draft and revision of the manuscript, and edited the manuscript. MBK conceived of the research, secured the funding, oversaw the stimulus creation and data collection, provided the data analytic support, and edited the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Science Foundation under Grant No. 1655265 awarded to MBK.

Acknowledgments

The authors thank Melanie J. Close, now at the City University of New York Institute for State and Local Governance, and Andrew J. Evelo, now at the University of Waikato, for their assistance in creating the stimulus materials used in the study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Author disclaimer

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1233782/full#supplementary-material

References

Aberson, C. L. (2019). *Applied Power Analysis for the Behavioral Sciences*, 2nd Edn. London: Routledge.

Alogna, V. K., Attaya, M. K., Aucoin, P., Bahník, Š, Birch, S., Birt, A. R., et al. (2014). Registered replication report: Schooler and Engstler-Schooler (1990). *Perspect. Psychol. Sci.* 9, 556–578. doi: 10.1177/1745691614545653

Blair, I. V., Chapleau, K. M., and Judd, C. M. (2005). The use of Afrocentric features as cues for judgment in the presence of diagnostic information. *Eur. J. Soc. Psychol.* 35, 59–68. doi: 10.1002/ejsp.232

Blair, I. V., Judd, C. M., and Fallman, J. L. (2004a). The automaticity of race and Afrocentric facial features in social judgments. *J. Pers. Soc. Psychol.* 87, 763–778. doi: 10.1037/0022-3514.87.6.763

Blair, I. V., Judd, C. M., and Chapleau, K. M. (2004b). The influence of Afrocentric facial features in criminal sentencing. *Psychol. Sci.* 15, 674–679. doi: 10.1111/j.0956-7976.2004.00739.x

Blair, I. V., Judd, C. M., Sadler, M. S., and Jenkins, C. (2002). The role of Afrocentric features in person perception: Judging by features and categories. *J. Pers. Soc. Psychol.* 83, 5–25. doi: 10.1037/0022-3514.83.1.5

Bradfield, A. L., Wells, G. L., and Olson, E. A. (2002). The damaging effect of confirming feedback on the relation between eyewitness certainty and identification accuracy. *J. Appl. Psychol.* 87, 112–120. doi: 10.1037//0021-9010.87.1.112

Brewer, N., and Wells, G. L. (2011). Eyewitness identification. *Curr. Dir. Psychol. Sci.* 20, 24–27. doi: 10.1177/0963721410389169

Brigham, J. C., Ready, D. J., and Spier, S. A. (1990). Standards for evaluating the fairness of photograph lineups. *Basic Appl. Soc. Psychol.* 11, 149–163. doi: 10.1207/s15324834basp1102_3

Carlson, C. A., Gronlund, S. D., and Clark, S. E. (2008). Lineup composition, suspect position, and the sequential lineup advantage. *J. Exp. Psychol.* 14, 118–128. doi: 10.1037/1076-898X.14.2.118

Chiroro, P. M., Tredoux, C. G., Radaelli, S., and Meissner, C. A. (2008). Recognizing faces across continents: The effect of within-race variations on the own-race bias in face recognition. *Psychon. Bull. Rev.* 15, 1089–1092. doi: 10.3758/PBR.15.6.1089

Clark, S. E. (2003). A memory and decision model for eyewitness identification. *Appl. Cogn. Psychol.* 17, 629–654. doi: 10.1002/acp.891

Cutler, B. L., and Kovera, M. B. (2010). *Evaluating Eyewitness Identification*. Oxford: Oxford University Press.

Deffenbacher, K. A., Bornstein, B. H., and Penrod, S. D. (2006). Mugshot exposure effects: Retroactive interference, mugshot commitment, source confusion, and unconscious transference. *Law Hum. Behav.* 30, 287–307. doi: 10.1007/s10979-006-9008-1

Demarchi, S., and Py, J. (2009). "A method to enhance person description: A field study," in *Handbook of Psychology of Investigative Interviewing: Current Developments and Future Directions*, eds R. Bull, T. Valentine, and T. Williamson (Hoboken, NJ: John Wiley & Sons Ltd), 241–256. doi: 10.1002/9780470747599.ch14

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *J. Pers. Soc. Psychol.* 56, 5–18. doi: 10.1037/0022-3514.56.1.5

Dewhurst, S. A., Hay, D. C., and Wickham, L. H. V. (2005). Distinctiveness, typicality, and recollective experience in face recognition: A principal components analysis. *Psychon. Bull. Rev.* 12, 1032–1037. doi: 10.3758/BF03206439

Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., and Johnson, S. (2006). Looking deathworthy: Perceived stereotypicality of Black defendants predicts capital-sentencing outcomes. *Psychol. Sci.* 17, 383–386. doi: 10.1111/j.1467-9280.2006.01716.x

Eberhardt, J. L., Goff, P., Purdie, V. J., and Davies, P. G. (2004). Seeing Black: Race, crime, and visual processing. *J. Pers. Soc. Psychol.* 87, 876–893. doi: 10.1037/0022-3514.87.6.876

Fahsing, I. A., Ask, K., and Granhag, P. A. (2004). The man behind the mask: Accuracy and predictors of eyewitness offender descriptions. *J. Appl. Psychol.* 89, 722–729. doi: 10.1037/0021-9010.89.4.722

Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160.

Field, A. (2013). *Discovering Statistics using IBM SPSS Statistics*, 4th Edn. Thousand Oaks, CA: Sage, 760–790.

Fitzgerald, R. J., Price, H. L., Oriet, C., and Charman, S. D. (2013). The effect of suspect-filler similarity on eyewitness identification decisions: A meta-analysis. *Psychol. Public Policy Law* 19, 151–164. doi: 10.1037/a0030618

Flowe, H. D., and Ebbesen, E. B. (2007). The effect of lineup similarity on recognition accuracy in simultaneous and sequential lineups. *Law Hum. Behav.* 31, 33–52. doi: 10.1007/s10979-006-9045-9

Goodsell, C. A., Gronlund, S. D., and Carlson, C. A. (2010). Exploring the sequential lineup advantage using WITNESS. *Law Hum. Behav.* 34, 445–459. doi: 10.1007/s10979-009-9215-7

Hall, R. E. (1994). The "bleaching syndrome": Implications of light skin for Hispanic American assimilation. *Hispanic J. Behav. Sci.* 16, 307–314. doi: 10.1177/07399863940163008

Hall, R. E. (1995). The bleaching syndrome: African American's response to cultural domination vis-à-vis skin color. *J. Black Stud.* 26, 172–184. doi: 10.1177/002193479502600205

Harvard, C., Breese, E., Thirkettle, M., Kask, K., Leol, K. L., and Madamurk, K. (2023). A background of bias: Subtle changes in lineup backgrounds increase the own race bias. *J. Police Crim. Psychol.* doi: 10.1007/s11896-023-09578-2

Horry, R., Fitzgerald, R. J., and Mansour, J. K. (2021). "Only your first yes will count": The impact of prelineup instructions on sequential lineup decisions. *J. Exp. Psychol. Appl.* 27, 170–186. doi: 10.1037/xap0000337

Innocence Project (2023). *DNA Exonerations in the United States (1989-2020)*. Available online at: https://innocenceproject.org/dna-exonerations-in-the-united-states/ (accessed July 9, 2018).

Kaesler, M., Dunn, J. C., Ransom, K., and Semmler, C. (2020). Do sequential lineups impair underlying discriminability? *Cogn. Res. Princ. Implic.* 5:35. doi: 10.1186/s41235-020-00234-5

Kahn, K. B., and Davies, P. G. (2011). Differentially dangerous? Phenotypic racial stereotypicality increases implicit bias among ingroup and outgroup members. *Group Process. Intergr. Relat.* 14, 569–580. doi: 10.1177/1368430210374609

Kahn, K. B., Goff, P. A., Lee, J. K., and Motamed, D. (2016). Protecting Whiteness: White phenotypic racial stereotypicality reduces police use of force. *Soc. Psychol. Pers. Sci.* 7, 403–411. doi: 10.1177/1948550616633505

Katzman, J., and Kovera, M. B. (2023). Potential causes of racial disparities in wrongful convictions based on mistaken identifications: Own-race bias and differences in evidence-based suspicion. *Law Hum. Behav.* 47, 23–35. doi: 10.1037/lhb0000503

Kleider-Offutt, H. M., Bond, A. D., Williams, S. E., and Bohil, C. J. (2018). When a face type is perceived as threatening: Using general recognition theory to understand biased categorization of Afrocentric faces. *Mem. Cogn.* 46, 716–728. doi: 10.3758/s13421-018-0801-0

Knuycky, L. R., Kleider, H. M., and Cavrak, S. E. (2013). Lineup misidentifications: When being prototypically Black is perceived as criminal. *Appl. Cogn. Psychol.* 28, 39–46. doi: 10.1002/acp.2954

Kovera, M. B. and Evelo, A. J. (2017). The case for double-blind lineup administration. *Psychol. Public Policy Law* 23, 421–437. doi: 10.1037/law0000139

Kovera, M. B., and Evelo, A. J. (2021). Eyewitness identification in its social context. *J. Appl. Res. Mem. Cogn.* 10, 313–327. doi: 10.1016/j.jarmac.2021.04.003

Lee, J., and Penrod, S. (2022). Three-level meta-analysis of the other-race bias in facial identification. *Appl. Cogn. Psychol.* 36, 1106–1130. doi: 10.1002/acp.3997

Lee, J., and Penrod, S. D. (2019). New signal detection theory-based framework for eyewitness performance in lineups. *Law Hum. Behav.* 43, 436–454. doi: 10.1037/lhb0000343

Lindsay, R. C., Wallbridge, H., and Drennan, D. (1987). Do the clothes make the man? An exploration of the effect of lineup attire on eyewitness identification accuracy. *Can. J. Behav. Sci.* 19, 463–478.

Lindsay, R. C. and Wells, G. L. (1985). Improving eyewitness identifications from lineups: Simultaneous versus sequential lineup presentation. *J. Appl. Psychol.* 70, 556–564. doi: 10.1037/0021-9010.70.3.556

Lindsay, R. C. L., Lea, J. A., Nosworthy, G. J., Fulford, J. A., Hector, J., LeVan, V., et al. (1991). Biased lineups: Sequential presentation reduces the problem. *J. Appl. Psychol.* 76, 796–802. doi: 10.1037/0021-9010.76.6.796

Long, J. S. (1997). *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage Publications, Inc.

Ma, D. S., Correll, J., and Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behav. Res. Methods* 47. doi: 10.3758/s13428-014-0532-5

Maddox, K. B. (2004). Perspectives on racial phenotypicality bias. *Pers. Soc. Psychol. Rev.* 8, 383–401. doi: 10.1207/s15327957pspr0804_4

Maddox, K. B., and Gray, S. A. (2002). Cognitive representations of Black Americans: Reexploring the role of skin tone. *Pers. Soc. Psychol. Bull.* 28, 250–259. doi: 10.1177/0146167202282010

Maxwell, S. E. (2004). The persistence of underpowered studies in psychological research: Causes, consequences, and remedies. *Psychol. Methods* 9, 147–163. doi: 10.1037/1082-989X.9.2.147

Meissner, C. A., and Brigham, J. C. (2001a). Thirty years of investigating the own-race bias in memory for: A meta-analytic review. *Psychol. Public Policy Law* 7, 3–35.

Meissner, C. A., and Brigham, J. C. (2001b). A meta-analysis of the verbal overshadowing effect in face identification. *Appl. Cogn. Psychol.* 15, 603–616. doi: 10.1002/acp.728

Meissner, C. A., Tredoux, C. G., Parker, J. F., and MacLin, O. H. (2005). Eyewitness decisions in simultaneous and sequential lineups: A dual-process signal detection theory analysis. *Mem. Cogn.* 33, 783–792. doi: 10.3758/BF03193074

National Archive of Criminal Justice Data (2016). *National Incident-Based Reporting System, 2016: Extract files [Data set]*. Ann Arbor, MI: Inter-University Consortium for Political and Social Research, doi: 10.3886/ICPSR37066

Okazawa-Rey, M., Robinson, T., and Ward, J. V. (1987). Black women and the politics of skin color and hair. *Women Therapy* 6, 89–102. doi: 10.1300/J015V06N01_07

Osborne, D., and Davies, P. G. (2013). Eyewitness identifications are affected by stereotypes about a suspect's level of perceived stereotypicality. *Group Process. Intergr. Relat.* 16, 488–504. doi: 10.1177/1368430212454927

Palmer, M. A., and Brewer, N. (2012). Sequential lineup presentation promotes less-biased criterion setting but does not improve discriminability. *Law Hum. Behav.* 36, 247–255. doi: 10.1037/h0093923

Perry v. New Hampshire (2012). *565 U.S. 228*. Available online at: https://case-law.vlex.com/vid/perry-v-new-hampshire-890705750 (accessed January 11, 2012).

Peterson, A. M. (2016). Complicating race: Afrocentric facial feature bias and prison sentencing in Oregon. *Race Justice* 7, 59–86. doi: 10.1177/2153368716663607

Russel, K. Y., Wilson, M., and Hall, R. E. (1992). *The Color Complex: The Politics of Skin Color Among African Americans*. San Diego, CA: Harcourt Brace Jovanovich.

Salerno, J. M., Kulak, K., Smalarz, L., Eerdmans, R. E., Lawrence, M. L., and Dao, T. (2023). The role of social desirability and establishing nonracist credentials on mock juror decisions about Black defendants. *Law Hum. Behav.* 47, 100–118. doi: 10.1037/lhb0000496

Schimmack, U. (2012). The ironic effect of significant results on the credibility of multi-study articles. *Psychol. Methods* 17, 551–566. doi: 10.1037/a0029487

Schooler, J. W., and Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cogn. Psychol.* 22, 36–71. doi: 10.1016/0010-0285(90)90003-M

Smalarz, L. (2021). Suspect bias: A neglected threat to the reliability of eyewitness identification evidence. *J. Appl. Res. Mem. Cogn.* 10, 356–362. doi: 10.1016/j.jarmac.2021.06.005

Smalarz, L., Eerdmans, R. E., Lawrence, M. L., Kulak, K., and Salerno, J. M. (2023). Counterintuitive race effects in legal and nonlegal contexts. *Law Hum. Behav.* 47, 119–136. doi: 10.1037/lhb0000515

Steblay, N., Dysart, J. E., Fulero, S., and Lindsay, R. C. L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. *Law Hum. Behav.* 25, 459–473. doi: 10.1023/A:1012888715007

Steblay, N., Dysart, J. E., and Wells, G. L. (2011). Seventy-two tests of the sequential lineup superiority effect: A meta-analysis and policy discussion. *Psychol. Public Policy Law* 17, 99–139. doi: 10.1037/a0021650

Steblay, N. K., and Wells, G. L. (2020). Assessment of bias in police lineups. *Psychol. Public Policy Law* 26, 393–412. doi: 10.1037/law0000287

The National Registry of Exonerations (2022). *Exoneration Detail List*. https://www.law.umich.edu/special/exoneration/Pages/detaillist.aspx (accessed January, 2023).

Vokey, J. R., and Read, J. D. (1992). Familiarity, memorability, and the effect of typicality on the recognition of faces. *Mem. Cogn.* 20, 291–302. doi: 10.3758/BF03199666

Wells, G. L. (1984). The psychology of lineup identifications. *J. Appl. Soc. Psychol.* 14, 89–103. doi: 10.1111/j.1559-1816.1984.tb02223.x

Wells, G. L., Kovera, M. B., Douglass, A. B., Brewer, N., Meissner, C. A., and Wixted, J. (2020). Policy and procedure recommendations for the collection and preservation of eyewitness identification evidence. *Law Hum. Behav.* 44, 3–36. doi: 10.1037/lhb0000359

Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., and Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law Hum. Behav.* 22, 603–647. doi: 10.1023/A:1025750605807

Wells, G. L., Steblay, N. K., and Dysart, J. E. (2012). Eyewitness identification reforms: Are suggestiveness-induced hits and guesses true hits? *Perspect. Psychol. Sci.* 7, 264–271. doi: 10.1177/1745691612443368

Wells, G. L., and Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Pers. Soc. Psychol. Bull.* 25, 1115–1125. doi: 10.1177/01461672992512005

Winter, K., Menne, N. M., Bell, R., and Buchner, A. (2023). Evaluating the impact of first-yes-counts instructions on eyewitness performance using the two-high threshold eyewitness identification model. *Sci. Rep.* 13, 1–10. doi: 10.1038/s41598-023-33424-4

Wixted, J. T., and Mickes, L. (2014). A signal-detection-based diagnostic-feature-detection model of eyewitness identification. *Psychol. Rev.* 121, 262–276. doi: 10.1037/a0035940

Wixted, J. T., Wells, G. L., Loftus, E. F., and Garrett, B. L. (2021). Test a witness's memory of a suspect only once. *Psychol. Sci. Public Interest* 22, 1S–18S. doi: 10.1177/15291006211026259

frontiers | Frontiers in Psychology

TYPE Review
PUBLISHED 22 July 2024
DOI 10.3389/fpsyg.2024.1294730

OPEN ACCESS

EDITED BY
Ivan Mangiulli,
University of Bari Aldo Moro, Italy

REVIEWED BY
Jacques Py,
Université Toulouse - Jean Jaurès, France
Lawrence Patihis,
University of Portsmouth, United Kingdom

*CORRESPONDENCE
Deborah Davis
✉ debdavis@unr.edu

RECEIVED 15 September 2023
ACCEPTED 13 March 2024
PUBLISHED 22 July 2024

CITATION
Davis D, Hogan AA and Hart DJ (2024) Myths of trauma memory: on the oversimplification of effects of attention narrowing under stress. *Front. Psychol.* 15:1294730. doi: 10.3389/fpsyg.2024.1294730

Myths of trauma memory: on the oversimplification of effects of attention narrowing under stress

Deborah Davis*, Alexis A. Hogan and Demi J. Hart

Department of Psychology, University of Nevada, Reno, NV, United States

The present article addresses claims commonly made by prosecution witnesses in sexual assault trials: that attention narrows under stress, and that these attended aspects of the event are encoded in a way that ensures accuracy and resistance to fading and distortion. We provide evidence to contradict such claims. Given that what is encoded is largely the gist of one's interpretation of experience, we discuss the way in which attention and emotion can bias the interpretation of experience. We illustrate with issues of memory reports in cases of acquaintance rape, where the primary issue is the presence or absence of consent. We provide some specific illustrations concerning effects of emotion on interpretation of sexual consent. Finally, based on what is known regarding priming effects on memory retrieval and judgment, we conclude with discussion of the potential of some "trauma-informed" interviewing strategies to promote false memories (such as FETI: Forensic Experiential Trauma Interview).

KEYWORDS

trauma, memory, sexual assault, memory distortion, tunnel memory, trauma-informed

1 Introduction

Many trials of child or adult sexual assault feature two kinds of experts addressing memory: often specifically trauma memory. On the plaintiff or prosecution side will be what is often referred to as a "counterintuitive behavior" expert (Henceforth, CBE). This expert will testify concerning behaviors that seem intuitively inconsistent with having been sexually assaulted, typically suggesting that these are actually common among victims. These include such things as continuing contact with and other seemingly positive reactions toward the perpetrator, long delays in reporting, failure to fight back during the assault, and other behaviors that are subjects of what are claimed to be "rape myths" and other lay misunderstandings of the behavior of victims. Often such experts also testify concerning memory: to explain, for example, why victims' accounts are assertedly often fragmented, incomplete, and inconsistent over time. They also commonly assert that whereas memory for some aspects of a sexual assault might fail or be distorted, memory for the fact of the assault and central details are highly resistant or impervious to failure or distortion. These experts are often clinical or development psychologists, social workers, or law enforcement personnel. On the defense side will typically be an academic expert, often a cognitive or social psychologist, who is called to contest some of the claims of CBEs regarding "rape myths" or "child sex abuse accommodation syndrome" and/or those regarding memory.

The purpose of this paper is to address a specific issue commonly addressed by CBEs: that of the combined effects of attention and emotion on information processing and memory. We focus on the specific claims of many CBEs concerning "tunnel memory," or attention narrowing under stress, and its effects on memory strength and accuracy. We do not purport to review the full range of issues or research literature regarding trauma

and memory. Nor do we attempt to provide full reviews of the specific points we address. We address only the issue of what effects attention exerts on memory in the context of strong emotions. Our goal is to demonstrate that the claims of CBEs regarding combined effects of attention-narrowing and emotion are both oversimplified and inaccurate or unjustifiably extreme.

For example, these claims include the idea that memories formed under stress are accurate regarding the central features of the event: period. Among the experts promoting this claim is Dr. Rebecca Campbell: a psychology professor at Michigan State University. She commonly testifies as a CBE expert, and has offered training to Universities, police departments, and training organizations such as John Reid Associates (one of the premier interview/interrogation training organizations in America) on the "neurobiology of trauma" and its implications for memory of traumatic events such as sexual assault.

Dr. Campbell asserts as follows:

"… memory can be slow and difficult — because the encoding and the consolidation went down in a fragmented way. It went down on little tiny post-it notes and they were put in all different places in the mind. And you have to sort through all of it, and it's not well-organized, because remember I told you to put some of them in folders that had nothing to do with this. But the question everybody wants to know about is the accuracy of that information, okay. And what we know from the research is that the laying down of that memory is *accurate* and the recall of it is *accurate*. So what gets written on the post-it notes — *accurate*. The storage of it is disorganized and fragmented.

So, victims who are assaulted under the influence of alcohol, may not have anything to retrieve. So to speak, their post-it notes are just blank. They may not have it, okay? But for those who are able to remember it, either in pieces or parts, it does go in *accurately,* it does come out *accurately,* but it comes out slow, steady, fragmented, and disorganized.…And again, they interpret this victim's behavior as evasiveness or lying. And again, what it really is, most often, is that the victim is having difficulty accessing the memories. Again, the content of the memory the research tells us very clearly is *accurate*. It's just going to take some time and patience for it to come together" (Campbell, 2012).

> https://nij.ojp.gov/media/video/24056#transcript\protect\T1\textdollar-\protect\T1\textdollar0

Similarly, Dr. Jim Hopper is a clinical psychologist who also often testifies as a CBE, guest-writes for popular magazines, and educates Harvard Medical School students about counterintuitive behavior and memory for trauma. He likewise emphasizes the accuracy of memory for traumatic events. He maintains an elaborate website describing his analysis of the neurobiology of trauma and its implications for the behavior and memory of victims (Jim Hopper, Ph.D.).

"Research also shows that usually people *accurately* recall the 'gist' and 'central details' of highly stressful experiences" (Hopper, 2023).

It is important to note that nothing in the memory literature suggests that any memories, particularly for complex events, can be presumed accurate. The literature documents factors that increase or decrease the likelihood of accuracy. Moreover, in further contradiction to many such claims, we present evidence that attention does not guarantee accuracy, and instead, particularly in combination with strong emotions, can promote inaccuracy. We include a discussion of how some prominent "trauma informed" strategies of interviewing sexual assault victims can exacerbate memory distortion rather than facilitate accurate retrieval.

As we proceed, we discuss the way in which the potential for memory distortion is important for litigation of claims of disputed sexual encounters between acquaintances, where the primary issue is whether there was or was not consent (or at least a reasonable belief in consent on the part of the accused). In such cases, many details, both historical and surrounding the encounter itself, become important for issues of actual and perceived consent. We suggest that it is unrealistic to assume that remembered details of consent-related communications and other consent-related details that are so important in claims of acquaintance rape will be unerringly accurate for either party. As we shortly review, in no way do emotions, whether rising to the level of trauma or not, guarantee such accuracy: nor do any of the well-established functions of cognition or memory.

Considering the context of acquaintance rape, it is important to address the wide range of emotions that are relevant. Such incidents do not always rise to the level of trauma or intense fear. They can also be an issue of annoyance, provoking emotions such as irritation, anger, embarrassment, or disgust. Moreover, the accused's emotions are relevant as well: and can include those such as sexual arousal, or happiness. As we consider in sections to come, different emotions can have different effects on processing and interpretation: and thus emotion-related arousal will function somewhat differently in different emotional contexts.

It is also important to emphasize that regardless of what emotions an accuser might experience during a sexual encounter, no matter how extreme the emotions are, and no matter how firmly she is convinced that she was raped, the event may not qualify legally as rape. This judgment will rest in part on the nature of the interaction itself and features of the context going to the issue of what a reasonable man in the situation would believe regarding consent. The woman can absolutely feel that she was raped, even though her consent-related communications and behaviors could have allowed a reasonable belief that she consented (see Wood et al., 2019 for a review of legal standards of consent). As some studies have indicated, there can be a discrepancy between internal feelings of consent and external expressions of it (e.g., Willis et al., 2019). Likewise, a man can have an unjustified view of consent, believing there was no rape even though legal standards would suggest there was. Given the wide range of consent-related details relevant to claims of acquaintance rape, it is crucial to understand how emotion can cause predictable forms of distortion for their memory.

We begin our discussion with a brief review of the encoding function of memory and the determinants of what is encoded. Next, we turn to a brief review of the concept of tunnel memory and claims regarding the effects of attention narrowing under stress. We then explore the way in which claims regarding emotion, tunnel memory, and accuracy are often oversimplified and misleading. We specifically address (1) difficulties in the prediction

of what attention will narrow to, (2) how attention affects information processing, (3) how these processes are exacerbated and/or modified by emotion, (4) why predictable failures and distortions in memory encoding can result from attention itself and its combination with emotion, and (5) the importance of consideration of post encoding processes promoting distortion. We consider these issues in the context of claims of counterintuitive behavior experts regarding memory for alleged sexual assault. As context for these discussions it is important to emphasize that much of the social interaction that is relevant for judgment of disputed sexual interactions is not necessarily stressful for either participant. Particularly in cases of alleged date rape, the interactions that convey likely or actual sexual consent can take place across substantial time in advance of the specific actions alleged to be assault. Accordingly, the claims of CBEs regarding memory for trauma are irrelevant to much of the information sought from the parties concerning potentially consent-related behaviors and communications leading up to the disputed actions: though the many effects of specific emotions on cognition documented by cognitive scientists can be very relevant.

2 What is encoded?

Errors in memory are not just about the process of "remembering." They begin with the initial processes of perception during the event and how event gist and details are encoded into memory. Though this point is second nature to cognitive psychologists, as context for discussion of why accuracy in encoding is never guaranteed, it is nevertheless worth emphasizing that what is initially encoded includes some verbatim representations, along with the gist of our interpretation of what is attended to about the event. Over time, verbatim images progressively fade and memory favors the gist of our interpretation of what was attended to during the event.

In contrast to the myth endorsed by many among the public, and even by some professionals, our observations and memory do not work like a video camera. Our observations of any given situation are selective, in that we do not attend to all aspects of an event. Moreover, memory follows the focus of attention, such that the more something is attended to, and the greater the processing devoted to it, the more likely it will be remembered. Finally, and most importantly for our discussion to come, encoding is "interpretative:" in that an interpretation is imposed on what is observed, and that interpretation is encoded along with the gist of the surface characteristics of what is interpreted. That interpretation might be simple categorization: such as "bear," "table," "man," or "soldier." Or it might be a characterization of behavior (such as "hostile," "coercive," "consenting," or "resistant") or emotion (such as "happy," "sad," "afraid," or "embarrassed": for reviews of these principles see Davis and Loftus, 2016; Reisberg and Heuer, 2020; Davis et al., 2023). It is these functions of selectivity and interpretation that pose the potential for inaccuracy and so soundly contradict the idea that memory accuracy can be guaranteed (or almost guaranteed) under any circumstances. As we discuss in sections to come, the interplay of emotion with processes of selective processing and interpretation provides a double-edged sword: in some ways enhancing the strength of memory for an experience, but also risking greater error in what interpretation of that experience is recalled.

3 Fundamentals of the tunnel memory hypothesis

The fundamentals of the tunnel memory hypothesis (originally called the "Easterbrook Hypothesis") were first proposed by Easterbrook (1959). Easterbrook suggested that the arousal associated with emotion causes a narrowing of attention to "central" aspects of an event (and therefore better memory for central information), at the expense of attention to "peripheral" aspects (and therefore poorer memory for peripheral details). The phenomenon was later dubbed "tunnel memory," and the phrase is now frequently employed to depict this severe and concentrated focusing of attention on specific facets of a situation (Mackworth, 1965; Safer et al., 1998).

There has been some theoretical debate concerning the specific cause of emotion-related attention narrowing. Whereas, Easterbrook (1959) viewed the arousal associated with emotion as the cause, others have pointed to such possibilities as defensive strategies that direct attention (such as the disassociation that has been suggested to occur in rape victims (Brokke et al., 2022; Lynch et al., 2023); or the fact that "attention magnets," or stimuli that naturally draw attention (such as horrific violence) can both cause attention to narrow to themselves and also cause the emotion, rather than the reverse cause where emotion causes narrowing of attention to the stimulus (see Reisberg and Heuer, 2020 for review). Nevertheless, there is considerable empirical support for the proposition that emotion is associated with the narrowing of attention (Levine and Edelstein, 2009; Mitchell, 2023), and general agreement between memory scientists and both defense and prosecution experts on this point. The disagreement concerns the direction and consequences of that narrowing.

4 Why stop there? Oversimplification of effects of attention and emotion

The simple view of emotion and attention narrowing suggests that information that is attended to will be more successfully encoded into memory. Clear support exists for the idea that the gist of emotional events is more successfully encoded into memory, and there is general agreement concerning the resistance of memories for highly emotional events to forgetting. Indeed, Daniel Schacter has included this resistance, dubbed "persistence," as one of his "seven sins of memory" (Schacter, 1999, 2001; see also Bonsall and Holmes, 2023): though it might be considered a "sin" mostly for negative or traumatic events that are resistant to efforts to forget.

However, as previously established, counterintuitive behavior experts, such as Dr. Rebecca Campbell, Dr. Jim Hopper, and others, further argue that this information will be encoded accurately, and that it will be highly resistant, if not completely impervious, to both fading and memory distortion (Hopper, 2018a).

For example, Dr. Bessel van der Kolk, psychiatrist, author, and frequent CBE expert asserts:

"What is so extraordinary about trauma is that these images or sounds or physical sensations *don't change over time*. So people who have been molested as kids continue to see the wallpaper of the room in which they were molested. Or when they examine all these priest-abuse victims, they keep seeing the silhouette of the priest standing in the door of the bathroom and stuff like that. And so it's these images, these sounds that don't get changed" (Tippett, 2021).

> https://onbeing.org/programs/bessel-van-der-kolk-how-trauma-lodges-in-the-body-revisited/

Such claims as these suggest that the accuracy of the accounts of alleged victims should not be questioned, but rather should be presumed almost certainly accurate. We suggest, in contrast to this simple view, that the combined effects of emotion and attention are more complicated. In the sections to come we consider a number of processes that can instead undermine the accuracy of information encoded in the context of specific emotions.

4.1 What is actually encoded?

Given that attention provides the opportunity for encoding, what is it that actually gets encoded? Memory scientists generally agree that this consists of the gist of one's interpretation of what was attended to (e.g., Reisberg and Heuer, 2020). Whereas, verbatim images can be included during encoding, they are nevertheless interpreted, and it is the gist representations that persist more strongly over time. Therefore, it is important to ask what determines what is attended to and how it is interpreted. As we discuss in the sections to come, what is encoded is not necessarily what is legally most relevant, as many CBEs state or imply. Nor is the encoded interpretation always accurate.

4.2 Where does attention go under stress?

A necessity for predicting which features of an emotional event will be remembered is accurate understanding of where attention will go in the circumstance. This point has been central to understanding why, for example, the presence of a weapon has led to poorer memory for the face of a criminal perpetrator. Attention goes to the weapon instead, and therefore the face is remembered more poorly (Loftus et al., 1987; Steblay, 1992; Pickel, 1999).

Research has indicated that it is more difficult to identify what might be "central" than many contemplate. For example, emotions experienced during a threatening event are assumed to direct attention toward the aspects of the event that we determine to be the most useful for survival in the moment. More generally, Levine and Edelstein (2009) and Kaplan et al. (2012) have noted that specific emotions tend to activate specific goals, and to direct attention toward goal-relevant information (see also Fredrickson, 2000; Levine and Pizarro, 2004; Huntsinger, 2012, 2013; Harmon-Jones et al., 2013). The valence of emotions can also direct attention to emotion-consistent or inconsistent stimuli, depending upon the emotion-provoked goals in the situation (e.g., Clore et al., 2018; Clore and Schnall, 2019; Yu et al., 2021).

In some cases, emotion can direct attention away from, rather than toward, stimuli: such as when disgust directs attention away from the disgust-provoking stimulus; when shame directs attention away from a rapist's face; or when attention is directed away from the sexual activity itself in order to suppress extreme emotions (as is claimed regarding the tendencies of rape victims to disassociate: Kindelan, 2018). Emotion regulation strategies provoked by extreme emotions can have strategy-specific directive effects on attention and interpretation (such as distancing vs. reappraisal: e.g., Schmidt et al., 2010). In others, the goal might be mood maintenance, and therefore selective attention toward mood-consistent, and away from mood-inconsistent, information. Where details of an interaction become crucial to disputed sexual events, it is important to note that attempts to suppress the outward expression of emotions have been shown to impair memory for factual details of an interaction but increase memory for emotional reactions: presumably, because attention is directed toward the emotions and away from the interaction (Richards and Gross, 1999, 2000, 2006; Chang et al., 2018). Such findings suggest that it is no simple matter to identify what aspects of a disputed sexual encounter would have been attended to.

Again, it is important to note that "trauma" is not likely to occur during all time periods relevant to an alleged victim's account. For example, whereas details concerning interactions between accuser and accused leading up to the sexual encounter can include many consent-relevant communications and behaviors, the trauma itself (if any) should begin during that encounter or when it becomes clear that unwanted sex will occur. The timing of the trauma relative to the to-be-remembered information is crucial, in that any of the effects of the purported "neurobiology of trauma" on memory should not include pre-trauma information (see Marr et al., 2021 regarding effects of timing of stress relative to to-be-remembered information). But what has research shown us about where attention goes during a sexual assault? Anecdotal reports point to a fairly large variety of targets of attention during alleged sexual assaults. For example, sexual assault survivors often recount instances of tunnel memory, wherein they might recall only specific details such as the expressions on the perpetrator's face, the smell of his cologne, or the sound of his voice, the wallpaper in the room, upholstery in the car, the weapon they were carrying, sensory perceptions such as sound or smell, and many more details that might seem peripheral: rather than, or, in addition to, the actions of the perpetrator and the rape itself (Steblay, 1992; Percy, 2023).

A common viewpoint expressed by many CBEs is that rape victims tend to disassociate during the event, paying attention to anything but the assault and resulting feelings (van der Kolk, 2014; Hopper, 2015; Kindelan, 2018).

According to Dr. van der Kolk: "Dissociation is a temporary putting aside, not knowing, and not noticing. It's a way to survive. Blocking things out allows many traumatized people to go on. It may be very helpful in order to make it through the crisis, but in the long-range, living your life in a dissociative way only keeps the trauma alive" (Melaragno, 2018) (https://www.dailygood.org/story/1901/trauma-in-the-body-an-interview-with-dr-bessel-van-der-kolk-elissa-melaragno/; https://www.besselvanderkolk.com/resources/the-body-keeps-the-score).

Assuming, however, that disassociation does not occur for a particular accuser, where might attention go? Each party's behaviors

might be neglected by an accuser if disassociation occurred. But either way, where might each party's attention be focused? While there might be commonalities in the focus of attention between people, there are also likely individual differences. Some might focus on their own emotions and sensations. Others might devote more attention to the behaviors and reactions of the other person. Still, others might experience wide-ranging divided or rapidly shifting focus of attention (e.g., Kern et al., 2005). Attention might be focused on the behaviors of the other person leading up to sexual activity, in an effort to read the other's interest and intentions: and become more self-focused when sexual activity commences. If the woman feels that the encounter is unwanted, she may focus on how to escape. The possibilities for allocation of attention are extensive, making generalizations concerning what behaviors will be "central" and most likely remembered as inappropriate. Arguably, the specific behaviors going to a reasonable belief in consent may be less likely remembered than the feelings generated (recall the common advice that while you might not remember everything a person says or does, you will remember how they made you feel).

The issue of what is remembered is, of course, central to litigation of disputed claims of sexual assault. It is unfortunate that many of the most crucial details for litigation might not have drawn attention during the event or later be remembered. This issue is particularly important with respect to the example of alleged acquaintance rape, litigation of which arguably demands consideration of more wide-ranging information than that of stranger rape.

For the sake of any subsequent reports and associated litigation, a crucial set of issues for allegations of acquaintance rape concerns the behaviors of each party going to whether consent did or did not occur. Were there coercive behaviors? What behaviors occurred leading up to and during the event that indicated consent vs. refusal?. Were there alcohol or drugs involved? If so, when and how much did each person ingest? As previously noted, a person can feel that she was raped, even if the situation did not meet the legal standards for the crime: for example, if the accused could have a "reasonable" belief that the woman had consented (People v. Mayberry, 1975). To understand this, it is important to have accurate information concerning what was said and done by each party leading up to and during the disputed encounter.

Clearly, many details that might be considered important for litigation would not be the primary "attention magnets" during the period leading up to and during the event. These details may be encoded only vaguely or not at all: leaving them more susceptible to distortion based on context or suggestion. As shown by research on "fuzzy trace theory," contextual and suggestive influences on memory exert greater impact when the original encoding is more vague, or when the original traces become more vague over time (e.g., Brainerd and Reyna, 2019; Bialer et al., 2021; Brainerd et al., 2021, 2022).

Before leaving this discussion of attention, we note that it is important to consider, as well, that while both academic and legal attention is most often devoted to the emotions and memory of the accuser, those of the accused are important for understanding what he will remember and report (and what might be reasonable beliefs regarding consent for a person in his shoes).

4.3 What does attention do?

At the most basic level, attention provides the opportunity for encoding. This opportunity, however, does not guarantee retention. Information encoded into short term memory does not necessarily reach long term memory: as can occur with alcohol blackout (Lee et al., 2009), with head trauma (Vanderploeg et al., 2014), with sufficiently high levels of stress (Trammell and Clore, 2014), or with superficial attention (Schacter, 1999, 2001).

Retention is most likely to occur with elaborative encoding: where the person thinks about what is observed, forming more links to other information in memory that can later facilitate retrieval (Coane, 2013). Indeed, even early academic discussions of tunnel memory noted the relationship of narrowed attention to elaborative processing:

"Participants comprehend a neutral scene by automatically extending its boundaries and understanding the visual information in a broader external context. However, when participants are negatively aroused by a scene, they process more elaborately those critical details that were the source of the emotional arousal, and they maintain or restrict the scene's boundaries. 'Tunnel memory' results from this greater elaboration of critical details and more focused boundaries. Tunnel memory may explain the superior recognition and recall of central, emotion-arousing details in a traumatic event, as shown in previous research on emotion and memory" (Safer et al., 1998, p. 116).

However, elaborative encoding does not ensure a reliably objective representation of what occurred. When attention is brought to bear on something, the observer's general knowledge and expectations affect what is subjectively perceived. As perception theorists have routinely demonstrated, perception is inherently constructive, adding to what is physically perceived, filling in with what is expected (e.g., Hoffman, 2019). Cognitive psychology has further demonstrated the effects of expectations on perception of physical objects: showing, for example, that the same physical object can be seen as a rabbit vs. a duck, a number vs. a letter, or an old witch vs. a young girl depending upon which concept or expectation is activated. They have likewise argued that without such expectations we would not even know how to label what we observe or how to react to it (Brosch et al., 2013; Hoffman, 2019). Moreover, illustrations within all areas of psychology abound of "cognitive bias" in interpretation due to chronic and situationally activated expectations (Lord and Taylor, 2009).

In these respects attention can be regarded as a double-edged sword, in many cases affording us the opportunity to accurately understand and encode what we experience: but at the same time serving as a potentially biasing machine, leading us to slant the interpretation of what we experience toward consistency with salient expectations.

4.4 What does emotion do?

Beyond the attention narrowing effects of emotion, counterintuitive behavior experts often testify to a number of additional effects of stress on encoding. These include, for example,

claims regarding the physiology and "neurobiology" of responses of victim-survivors to trauma: such as their effects on memory. These claims tend to be a mixture of fact and fiction. Many claims regarding how emotion-provoked physiological/neurological responses potentiate encoding are empirically supported, though a review of these is beyond the scope of this review. In essence, though, research has shown that emotion and "trauma" tend to amplify the strength (as distinct from the accuracy) of encoding of the person's interpretation of the gist of the attended aspects of the experience (Brosch et al., 2013; Schoch et al., 2017), and this basic effect of strong emotions is relatively uncontested[1].

However, CBEs tend to go beyond the potentiation of encoding to claim that what is encoded is accurate, and so strongly encoded as to defy subsequent distortion: Hopper (n.d.-b) as reflected in the quotes of CBEs Rebecca Campbell and Jim Hopper provided earlier (Campbell, 2014; Hopper, 2023).

Such claims belie the basic truth that what is encoded is the gist of the interpretation of the attended information. Moreover, they ignore research concerning the impact of emotion on interpretation, and the implications of this for accuracy. Generally, work on implicit associations has shown that mental associations activated, even outside awareness, shape judgments outside of awareness (e.g., Greenwald and Banaji, 2017). In part, context affects which associations are activated, and thereby inevitably affects interpretation. Emotion is one feature of context, and as such also affects interpretation: in part through associations between emotion and the conditions that tend to produce them (such as when a person interprets an event in a way that explains his or her emotional reactions: Davis et al., 2023). Research has pointed to at least three mechanisms through which emotion affects interpretation.

4.4.1 Affect as information

First, through the "affect as information" mechanism, the person may interpret the event or aspects of the event in a way that makes sense in light of current emotions. The person's own emotional reaction is the primary basis of judgment, without reliance on all potentially relevant additional information in the situation. This mechanism of influence is more likely to occur in circumstances that likely apply in most disputed sexual encounters, as well as in uncontested instances of rape: when emotions are strong, when emotions are produced in the situation that is to be judged, when there is room for interpretation, and when intense processing is applied to the judgment [see reviews by Greifeneder et al. (2011), Clore et al. (2018), and Clore and Schnall (2019)].

4.4.2 Affective priming

The second mechanism is "affect as context" or "affective priming." This mechanism was identified in the Affect Infusion Model of Forgas (2002). In this view, emotion serves to contextually activate emotion-consistent schemas and expectations that direct the processing of information and affect judgments. Consistent with the AIM model, emotions exert more impact on judgments when the person engages in more elaborative processing of the event: such as one might expect of sexual assault victims (at least after the fact, if not during). As Davis et al. (2023) pointed out, emotions that can be provoked in an unpleasant sexual encounter might include those such as irritation, resentment, anger, fear, and disgust, which can be associated in memory with concepts of coercion or rape: thereby potentially biasing interpretation and memory toward consistency with rape.

As with any form of priming, emotions can also prompt constructive memory processes such that individuals can supplement their recollections with what they anticipate ought to exist in the surrounding context, guided by their emotion-related schemas and scripts. As Brainerd et al. (2008) have noted, the activation of a network of emotional connections can lead to intensified conceptual priming and an elevated feeling of familiarity, promoting both accurate and erroneous memories of emotional stimuli.

4.4.3 Emotion and intuitive processing

A third effect of emotion concerns the disengagement of System 2 (More Elaborative, Analytical, Reflective) processing (Davis and Loftus, 2009; Kahneman, 2011). To put this more strongly, emotion can impair frontal lobe functioning: generally leading the person to rely more strongly on ingrained habits and instinctual behaviors. Similarly, some CBEs and defense experts agree that habitual behaviors, such as behaviors informed by instinct, schemas, and scripts, can be rendered more dominant during a traumatic experience due to the impact of stress on the brain regions responsible for regulating our thoughts, emotions, and actions.

For example, Dr. Jim Hopper asserts: "When the fear circuitry kicks in [during the midst of a sexual assault], basically there's a subnucleus in the amygdala (a brain region responsible for emotional processing and connecting emotions to memory) called the central nucleus and it sends a signal to the brain stem that says 'hit the prefrontal cortex with more norepinephrine and dopamine.' So the fear circuitry triggers chemicals that hit the prefrontal cortex

1 A full review of the literature of effects of stress on memory is beyond our scope. However, it is important to note that this literature is not fully consistent in showing either negative effects of negative emotion (or stress) on accuracy (most prominently shown in the eyewitness memory literature) vs. positive effects [most prominently shown in the basic memory literature: see review by Marr et al. (2021)]. Most importantly, however, many studies have created stress that is independent of the to-be-remembered event (e.g., exposure to the cold pressor task before or during the stimulus materials or event. Research that would be relevant to our current discussion would involve strong emotions produced by the event to be remembered, such as is most often the case in the eyewitness memory literature, and is rarely the case in the general memory literature (see Marr et al., 2021). It would also involve events, rather than face memory or memory for pictures or words (which is common in neither literature), and assessment of memory for aspects of the event with potential for emotion-consistent distortion (also common in neither literature). As such, little of the basic literature on stress and witness memory is directly relevant to our discussion. Therefore, we focus largely on literature on emotion and interpretation: which is directly relevant to what is encoded into memory.

(a brain region responsible for regulating rational thoughts, actions and behaviors) and impair it."

> (Hopper: Sexual Assault & the Brain in 6 minutes, 2018; Hopper and Lisak, 2014).

As Dr. Hopper further explains, "When a larger predator is coming at you or has you in its grip, thinking through a response with your rational prefrontal cortex is too slow and could get you killed. But reflexes and habits, which your brain can automatically cue up and execute in fractions of a second, could save your life. So evolution selected brains in which stress and trauma impair the prefrontal cortex, because that allows fast reflexes and habits to take over."

> (Hopper: Sexual Assault & the Brain in 6 minutes, 2018; Hopper and Lisak, 2014).

CBE Lisak also offers an elaborate discussion of the stress-induced disengagement of the frontal lobes and the implications for habitual behaviors (Neurobiology of Trauma—Dr. David Lisak—YouTube) (Lisak, 2013).

Interestingly, however, Dr. Hopper, Dr. Lisak and other CBEs stop short of recognizing the effects of the dominance of habits on cognitive functions. Habits of the mind are equally promoted by the disabling of executive functioning: shunting the person into System 1 modes of intuitive thought driven by habits, schemas, and expectations (Davis and Loftus, 2009). According to Kahneman (2011), System 1 operates automatically and quickly without conscious effort or voluntary control. In this mode, a person's initial impressions of what they are experiencing go uncorrected by the more deliberate, rational analysis carried out in System 2 (Kahneman, 2011). This is because the automaticity of System 1 cannot be turned off at will, so any perceptual errors that occur during this stage are difficult to prevent. The correction of errors is left up to the slow, enhanced monitoring of System 2. However, many errors go unnoticed and uncorrected, because it's impractical to constantly question the accuracy of one's own thinking. Thus, we place trust in the snap decisions made by System 1 (Kahneman, 2011). Since emotion serves as the context triggering the expectations and associations served up by the intuitive system and used for interpretation, these interpretations are likely to be at least somewhat biased by the emotions.

5 What do we know about the role of emotion in interpretation and judgment of sexual consent?

A large literature has accumulated regarding how sexual consent is conceptualized, conveyed, and interpreted (see Wertheimer, 2003; Muehlenhard et al., 2016; Fenner, 2017; Wood et al., 2019; Kabota and Nakazawa, 2022 for reviews). This literature includes an array of studies designed to understand sources of miscommunication of consent. Among these are cultural scripts that promote misunderstanding: such as belief in "token resistance," whereby women may say "no" when really meaning "yes": offering "token" refusals before consenting to sex (e.g., Muehlenhard and Hollabaugh, 1988). Also included are studies of individual differences in beliefs that underlie many misunderstandings: such as traditional sex role beliefs, rape myth acceptance or "rape supportive attitudes", and others: many of which are also predictive of verdicts in trials of sex crimes (see Ryan, 2011; Rerick et al., 2019 for reviews). These comprise an array of cultural, perpetrator, victim and situational variables underlying miscommunication of consent and/or perpetration or victimization (Adams-Curtis and Forbes, 2004).

While many studies have addressed how and why women can fail to communicate resistance effectively, studies of errors in the interpretation of consent behaviors have been almost entirely restricted to males. There is, of course, considerable concern with why men coerce women, and it is assumed that the misperception of consent cues is among the causes. Accordingly, studies have addressed issues such as whether men on average tend to "overperceive" consent relative to the actual intentions of women; which female behaviors are most commonly perceived as indicating consent vs. refusal; which men are most prone to overperceive cues of consent; and what circumstances promote such misperceptions (see Wertheimer, 2003; Muehlenhard et al., 2016; Fenner, 2017; Wood et al., 2019; Kabota and Nakazawa, 2022 for reviews). These sorts of studies are crucial to understanding sources of disagreements between accusers and accused concerning whether consent did or did not occur. However, they also illustrate how vulnerable sexual interactions are to subjective interpretation, and how precarious are assumptions that reports of these interactions will be fully accurate on either side.

Nevertheless, much remains to be addressed. Studies of misperceptions of sexual consent have almost exclusively focused on the misperceptions among males of female consent. We are unable to locate studies focusing on female misinterpretation of coercion or of accuracy in understanding the clarity of their own sexual consent behaviors and communications. Memory reports of these behaviors by accusers are crucial to judgments of claims of sexual assault, and as such, research is needed to address the personal, situational, and cultural forces that might compromise accuracy.

Additionally, almost no research has addressed the manner in which emotions can impact the interpretation and memory of sexual consent interactions. Given that the issue of whether there was or was not consent to sexual activity is central to disputes regarding acquaintance rape, it is important to know how emotions might specifically affect judgments and memory of sexual consent. To date, little research has investigated this topic, particularly as it concerns female interpretation/memory of coercive male behaviors or their own behaviors communicating consent or non-consent.

6 Research on emotion and judgment of sexual intentions

Davis et al. have conducted a series of studies relevant to the effects of emotion on judgments of female sexual willingness. Though we have recently begun to study perceptions of male potentially coercive behaviors, these largely concern emotions likely experienced by the initiator of sex: and perhaps only the accused when there is a disputed sexual assault.

6.1 Sexual arousal

Several of these concerned the impact of male sexual arousal on interpretation of the extent to which specific female behaviors implied sexual willingness. The authors suggested that sexual arousal might lead aroused men to infer that specific behaviors reflect greater sexual interest or willingness compared to unaroused men (see also Murray et al., 2017). Several lines of research support such a prediction.

Steele and Josephs's (1990) theory of "alcohol myopia" proposed that alcohol narrows attention to impulse consistent cues and inhibits attention to impulse inconsistent inhibitory cues. Similarly, in line with the previously cited work on attention to emotion related goals, Loewenstein (1996) argued that strong emotions can lead the person to focus on how to resolve or satisfy the emotion quickly, without full consideration of reasons to avoid the behavior in question; or to focus attention inwardly and compromise concern for others. Similar in its effects to "alcohol myopia" (Steele and Josephs, 1990), sexual arousal could lead the person to rely on promotional cues (favoring sexual activity) more strongly than inhibitory cues (disfavoring it). To the extent that sexual arousal narrows attention to cues consistent with sexual activity, or biases interpretation of all cues in that direction, one would expect arousal to promote stronger perception of sexual consent.

Across three studies, Davis and colleagues asked male participants to either write an arousing sexual (vs. non-sexual) fantasy, or to view relatively arousing (vs. non-arousing) pictures of females. They were then asked to rate the extent to which each of 25 specific behaviors reflected sexual willingness (though the specific questions regarding willingness varied). In all studies, sexually aroused males (particularly single males) rated the female behaviors as reflecting greater sexual willingness (Livingston and Davis, 2020; Rerick et al., 2020). Relatedly, Bouffard and Miller (2014) found that though manipulated sexual arousal did not affect ratings of female sexual willingness in a dating scenario, self-reported sexual arousal did do so. Perhaps likewise reflecting sexual motivation, Rerick and Livingston (2022) found that specific behaviors were perceived as reflecting greater sexual willingness for attractive than unattractive women: and this effect was mediated by sexual arousal.

Miller and Davis (2024) recently found that arousal among both men and women was associated with perceptions that specific female behaviors reflected sexual willingness: as well as that they did so more strongly if toward a man who was attractive and had stronger financial credentials. However, arousal itself (while reading about and reacting to the female behaviors) was strongly predicted by political conservatism and religiosity. As these characteristics are associated with many of the rape-supportive attitudes mentioned earlier (such as rape myth acceptance, belief in token resistance to sex, and others e.g., Bohner et al., 2009; Rerick et al., 2019) it may be that effects of individual differences in the tendency to get sexually aroused when reading about sex-related behaviors (and therefore self-reported arousal) are actually reflecting effects of individual differences in rape supportive attitudes. However, this relationship might well be bi-directional, in that Rerick et al. (2022) found that sexual arousal led to greater agreement with attitudes consistent with greater permissiveness for sexual activity: including those regarding female token resistance to sexual activity, and assertive sexual strategies.

Indirect evidence is also consistent with such biasing effects of sexual arousal. Sexual arousal has been shown to shift motivation away from longer term desires toward satisfaction of more immediate ones (see Kim and Zauberman, 2013 for review). Sexually aroused males find females more attractive (e.g., Stephan et al., 1971; Ditto et al., 2006), find female faces to reflect greater sexual arousal (e.g., Maner et al., 2005), and find sexual material less disgusting (e.g., Stevenson et al., 2011). They also report greater willingness to engage in forms of sex they might otherwise find unacceptable: such as sex with unattractive or older women, sex without protection, or inappropriate coercive behaviors (Blanton and Gerrard, 1997; Ariely and Loewenstein, 2006). Such findings are consistent with the notion that sexual arousal facilitates perception of the social world as consistent with sexual activity: which would include perceiving potential partners as willing. And, not surprisingly, sexual arousal is associated with sexual disinhibition (e.g., Bouffard and Miller, 2014; Imhoff and Schmidt, 2014).

Our lab has recently begun to study determinants of perceptions of the coerciveness of male behaviors. Hogan et al. (2024) asked male and female participants to rate the extent to which a set of potential male behaviors seeking a date or sex put pressure on the female and their appropriateness in the circumstances depicted. Mirroring the Miller and Davis findings, the authors found that among both males and females male behaviors were seen as exerting less pressure and as more appropriate if the male was depicted as physically attractive and as possessing better financial potential. Self-reported sexual arousal predicted these ratings as well. Additionally, political conservatism, religiosity, and endorsement of rape-supportive attitudes predicted both arousal while reading about the behaviors and ratings of pressure and appropriateness. Such strong situational (characteristics of the male involved) and individual attitudinal differences in interpretation of sexual consent -related behaviors undermine the idea that memory for sexual assault incidents will be inevitably accurate.

6.2 Effects of felt power

While sexual arousal is perhaps the most pervasively present emotion in sexual encounters, other feelings can also affect interpretation of sexual willingness: among them, "felt power." Felt power is defined as a person's sense of agency and ability to exert influence over others (Fiske, 1993; Galinsky et al., 2003; Guinote, 2007, 2010; Guinote and Vescio, 2010).

Research has indicated that felt power exerts a number of effects that can become relevant in sexual interactions. Among the most relevant of these are the tendencies to selectively attend to and notice cues consistent with goal pursuit, to fail to notice or ignore goal-inconsistent cues, to interpret social situations as consistent with one's goals, and to feel more confident. Generally, these tendencies, and the power-induced disinhibition they entail, lead persons to more likely and vigorously pursue their goals (Galinsky et al., 2003, 2008, 2016; Keltner et al., 2003; Guinote, 2007, 2010; Lammers et al., 2008; Smith and Bargh, 2008; Guinote and Vescio, 2010; Slabu and Guinote, 2010; Hirsh et al., 2011; Whitson et al., 2013; Pike and Galinsky, 2020).

Though power-induced disinhibition can lead to prosocial behaviors when those are the preferred goals, it has also been shown to lead people to cheat, steal, or violate traffic laws, and generally to disregard social norms (see Lammers et al., 2015 for review): perhaps partly the result of increased tendencies toward social distance from others (see Magee and Smith, 2013 for review). A number of these norm-violating behaviors include sexual behaviors. For example, power increases infidelity among both men and women (e.g., Lammers et al., 2011). Moreover, power is associated with positive reactions (e.g., sexual arousal) to counter-normative sexual behaviors, such as sadistic behaviors among women and masochistic behaviors among men (Lammers and Imhoff, 2016): and is associated with perceptions and expectations of sexual interest from subordinates (Kunstman and Maner, 2011), sexual harassment-consistent cognitions (Pryor and Stoller, 1994; Bargh and Raymond, 1995; Bargh et al., 1995), and sexual aggression (Zurbriggen, 2000).

Based on findings regarding the effects of felt power on selective attention to goal-consistent cues, and biased interpretation of those cues toward consistency with goal pursuit, one might predict that in sexual situations this would include bias toward cues indicating that a potential sexual partner is willing. Consistent with this expectation, Livingston and Davis (2020) showed that males, but not females, regarded specific female behaviors as more indicative of sexual willingness when they were primed to feel more powerful.

6.3 Effects of alcohol

The relationship of alcohol to sexual behavior has been extensively studied. In part, alcohol can promote sexual arousal, sexual motivation, sexual activity, and sexual pleasure (see Davis and Loftus, 2004): but it has also been heavily implicated in sexual coercion and victimization (Villalobos et al., 2016; Caamano-Isorna et al., 2021; Steele et al., 2022). Of greatest interest for the purposes of this review are its effects on attention and interpretation in sexual situations.

In this respect, alcohol shares much with both felt power and sexual arousal. Each fuels alterations in cognition and promotes disinhibition. Moreover, they do so through largely comparable mechanisms: as elaborated by Steele and Josephs (1990) in their theory of "alcohol myopia" (see also Davis and Loftus, 2016), by Hirsh et al.'s (2011) analysis of effects of alcohol and power, and by Imhoff and Schmidt (2014) in their application to sexual arousal.

Each affects attention, narrowing it to cues consistent with motivations that are salient in the situation, reducing the depth of processing of the cues that are attended to, and reducing the ability to access existing knowledge and relate it to incoming information. They thereby reduce the complexity with which an event is processed and the extent to which all relevant information is brought to bear on a particular judgment or decision.

Analyses of the effects of power, strong emotion (including sexual arousal), and alcohol all point to the importance of reduced functionality or use of executive functions associated with each (Steele and Josephs, 1990; Hirsh et al., 2011; Imhoff and Schmidt, 2014; Davis and Loftus, 2016). As Daniel Kahneman put it, when executive functions are impaired, such as occurs with alcohol and strong emotions, judgment is based on the assumption "WYSIATI" ("What you see is all there is!" Kahneman, 2011). As such, mistakes in interpretation become more likely, and biased in a direction consistent with the emotion and motivation in question.

In addition to these processes, Davis and Loftus (2004, 2016) noted the relevance of "alcohol expectancy" effects regarding the tendency for alcohol to promote sexual motivation and activity (see also Villalobos et al., 2016; Wood et al., 2019 for review). Because alcohol is expected to promote interest in sex, many report using alcohol as a tool of seduction. Both sexes perceive intoxicated others as more sexually aroused, easy to seduce, and willing to consent. Women report being more likely to use alcohol when willing to have sex, and men believe the same. Reflecting such assumptions, those accused of sexually assaulting intoxicated alleged victims are less likely to be convicted (see Rerick et al., 2019 for review). These expectancy effects would increase the degree to which sexual motivation is assumed among intoxicated others. Moreover, when executive functions are impaired, as they are with alcohol and strong emotions, expectancies exert more influence on judgment (see Kahneman, 2011 for review).

As Davis and Loftus (2016) noted, during a sexual encounter these processes of selective attention and impaired executive functions can lead the perceiver to ignore contextual cues that should inform interpretation of both persons' behaviors and the degree of consent vs. coercion involved: such as the other's level of intoxication, the relationship history between them, the context of their current encounter, behaviors immediately preceding initiation of sexual activity, historical information about each person's behaviors and preferences while not intoxicated, one's own behaviors and effects of those on the other person and much more. Conflicting reports concerning disputed sexual consent can be based in these processes of selective attention, incomplete consideration of relevant information, and biased interpretation reflecting goals of sexual engagement vs. avoidance of the encounter.

7 But what about negative emotions associated with trauma?

At present we are not aware of studies examining the influence of negative emotions on judgments of sexual consent. Nevertheless, there is reason to expect that emotions experienced during a sexual encounter or later when recalling it will affect judgments of consent. Davis et al. (2023) have recently outlined reasons for such an expectation, as well as some pathways through which negative emotions can lead to false memories and false allegations of sexual assault.

The authors first suggested that several negative emotions can occur, even during subjectively or legally consensual sexual encounters. Surely, they will occur in instances of genuine rape. However, the authors argued that there are two other common circumstances that can generate negative emotions, such as anger, fear, or disgust: (1) when the person voluntarily chooses to engage in sex when they would prefer not to (such as to please a partner, to avoid conflict, pity sex with an unattractive partner, to secure other benefits from the partner, and others); and (2) when the person does not want to have sex and does not subjectively consent, but does

not effectively communicate refusal to the other person in a way a "reasonable person" would understand.

The authors further suggested that negative emotions felt during such encounters can color the interpretation of the interaction such that it is interpreted in a way that makes sense, given the emotions felt. One such interpretation that might explain such emotions could be that the encounter was actually involuntary and/or that the actions of the partner were coercive. Such effects would be consistent with the known mechanisms through which emotions affect judgment discussed earlier. Also as discussed earlier, such emotion priming can affect memory and judgment at the time of the encounter, or later when it is recalled.

Unfortunately, at this time we were unable to locate studies directly addressing the role of negative emotions in judgments of either female consent or male coercion. This is a gap in the literature in need of research. Our lab is in the process of initial efforts to test such effects. Meanwhile, there is some research consistent with the hypothesis that negative emotions such as fear, anger and disgust are likely to affect judgments of consent vs. coercion.

Some such research has come from the domain of moral psychology (e.g., Haidt, 2012). Emotions can be provoked by moral judgments, they can intensify moral judgments, or they can provoke moral judgments for morality neutral behaviors (Haidt, 2001, 2012; Avramova and Inbar, 2013; McAuliffe, 2019). Each of these effects can occur in sexual encounters. Strong emotions can be triggered by the encounter, and the encounter can produce moral judgments for each person's behavior. Such judgments can be made for behaviors that would generally be considered moral violations (such as sexual coercion, infidelity, incest, and others), or be imposed on behaviors that might be considered morally neutral in the absence of the emotion (many objectively non-coercive or permissible sexual invitations, advances, or activities between adults). A number of moral emotions with relevance to sexual situations have been addressed in the literature: including fear, anger, disgust, and guilt.

Davis et al. (2023) argued that a person feeling such emotions as fear, anger or disgust during a sexual encounter might tend to view the behaviors of the other person as wrong in a way consistent with force or coercion: such as dangerous, unwanted, or otherwise consistent with coercion. Or an emotion such as shame or disgust can be associated with the view that the encounter is inappropriate for other reasons.

The emotion of disgust (both as a state or a trait), for example, has enjoyed considerable interest in the moral psychology literature and has been associated with a variety of moral judgments. Moreover, incidental triggers of disgust (unrelated to the issues to be judged) have also been shown to provoke harsher moral judgments, though the size of such effects is sometimes small (see; Haidt, 2012; Landy and Goodwin, 2015; Schnall et al., 2015; van Leeuwen et al., 2017 for reviews).

Sex can provoke disgust in any number of ways that don't involve coercion: such as sex with unattractive partners, specific sexual acts, sex with inappropriate partners, bodily fluids and smells, and more. As such, it has the potential to also provoke moral judgments involving coercion. Relevant to this, Haidt (2012) showed that when confronted with stories with themes of disgust and disrespect, but involving no harm to anyone, 38% of participants nevertheless claimed someone was harmed.

The moral psychology literature supplements the emotion priming literature in suggesting that emotions can infect judgments. It remains for future research to specifically investigate such effects specifically on judgments of consent and coercion.

7.1 Are "central details" of traumatic events really impervious to fading and distortion?

Some counterintuitive behavior experts have made claims regarding the fidelity of memory for trauma over time that could be viewed as controversial or problematic by memory researchers. For example, Dr. Jim Hopper's claim that "...memories of highly stressful or traumatic experiences, at least their most central details, don't tend to fade over time," (Hopper, 2018b; see also van der Kolk, 1998). Researchers do agree that people tend to have stronger, longer lasting memories for emotional events, but also agree that all memories (even those for emotional/traumatic events) are subject to forgetting and will degrade over time (McNally et al., 2004; Laney and Loftus, 2005; Reisberg and Heuer, 2020). In fact, research has shown that people do forget the central and peripheral details of highly stressful/traumatic memories (Wagenaar and Groeneweg, 1990; Hirsh et al., 2011; see also research on vagaries of "flashbulb" memories for highly emotional events): (Rubin and Kozin, 1984; Christianson, 1989, 1992; Talarico and Rubin, 2003, 2007).

CBE Jim Hopper and others have also made claims regarding the imperviousness of traumatic memories to distortion. For example, while Dr. Hopper has conceded that the peripheral details of an event can easily be distorted, he also claimed that "...decades of research have shown that the most central details are not easy to distort, which typically requires repeated leading questions from people in authority or a very strong internal motivation for doing so" (Hopper, 2018b). Research has also soundly contradicted this claim. It is beyond our scope to offer a comprehensive review on this point, but the following are a few examples.

Since laboratory studies cannot create the level of distress experienced by victims of trauma, one way to address this claim is to look at changes in the reports of trauma victims, or those subject to highly stressful events, over time. Southwick et al. (1997), for example, examined the consistency of memories of combat-related traumatic events among veterans of Operation Desert Storm. Participants completed a questionnaire 1 month and at 2 years after returning from the war. Results showed that reports of 88% of participants changed over time. That is, 46% first reported traumatic events that they did not recall 2 years later and 70% recalled traumatic events at the 2-year evaluation that they did not report during the first evaluation. The researchers explained that many of these event memories that changed involved highly traumatic events that were specific and objective. For example, about 27% of participants changed their memory report for the event of "seeing others killed or wounded". The researchers provided several explanations for the inconsistency in memory for these traumatic events, including the possibility that post-event information may have led to distortion. Though the mechanisms of change no doubt varied across persons, the results indicated that the presumed central details of a traumatic event (e.g., seeing others killed) are not indelible. Additionally,

they provide presumptive evidence of the possibility that memory of traumatic events can be distorted in the absence of leading questions or other suggestive interviewing tactics. Similar changes in memory for highly emotional events have been documented in the previously referenced "flashbulb" memory research. People both forgot over time and changed memories for aspects of the events (Rubin and Kozin, 1984; Christianson, 1989, 1992; Talarico and Rubin, 2003, 2007).

Other studies have examined the effects of common sources of memory distortion on memories for highly stressful events. For example, Morgan et al. (2013) assessed the impact of misinformation on memories of military personnel in SERE (Survive, Evade, Resist, Escape) training, which has been shown to result in very high levels of stress and stress hormones. Participants had to survive in the wilderness, try to evade capture, endure capture and placement in a mock prisoner-of-war camp, undergo a stressful interrogation, and try to escape. The experience is meant to be realistic, and therefore, highly stressful. After the experience, participants completed a questionnaire with or without misinformation and leading questions. Results showed a significant influence of misinformation on participants' memory of the event: for example, 27% of participants in the misinformation condition falsely remembered a weapon, compared to only 3% of participants in the no misinformation condition. Participants were also asked to make an eyewitness identification of their interrogator in a target-absent photo array. Results showed that 91% of participants in the misinformation condition made a false-positive eyewitness identification, compared to 53% of participants in the no misinformation condition. While exposure to misleading information led to a significant increase in memory distortions compared to those who were not exposed to misinformation, more than half of the participants who did not receive misinformation still showed memory distortion in the form of false identifications. Memory for the central details of a highly stressful event was readily distorted by exposure to misleading information, even in this group of military personnel who are trained to withstand stress, propaganda, and other exploitation efforts.

These results are consistent with large bodies of research showing that memory for forensically important aspects of highly stressful events is subject to failure and distortion: such as memory for a perpetrator, war trauma experiences, and other traumatic event details (see McNally, 2003; Deffenbacher et al., 2004; Morgan et al., 2004, 2013). Moreover, there is no reason to believe that traumatic memories cannot be altered through the same processes that have been repeatedly shown to distort memories for countless real-life events: or to lead to memories of events that never happened at all (see McNally, 2003; Brainerd and Reyna, 2005; Davis and Loftus, 2007, 2020; Bialer et al., 2021).

7.2 Some foibles of "trauma-informed" interviewing

Our culture is replete with "trauma-informed" strategies of dealing with alleged victims of trauma (e.g., Reicherter et al., 2022). Among these are methods of interviewing alleged trauma victims that are intended to maximize the completeness and accuracy of their reports. Among the most commonly taught of these is "FETI," or the "Forensic Experiential Trauma Interview" (Strand and Heitman, 2017).

Informed by the many researchers and educators on the "neurobiology of trauma," the FETI method relies on the common claims that memories of trauma are fragmented, and not organized in the coherent sequence needed by those in the legal system: as reflected in our earlier quotes from Rebecca Campbell, one of the adopters of Strand's ideas.

Many other CBEs also express this opinion. Dr. David Lisak, a clinical psychologist and forensic consultant who researches the causes and consequences of interpersonal violence, and teaches the neurobiology of trauma to law enforcement, describes what happens when someone experiences a traumatic event as such: "When the amygdala responds to a life-threatening stimulus and reacts, we are no longer able to encode experiences in the same way. When the amygdala is firing due to something traumatizing, experiences get encoded as intense sensory fragments rather than coherent, sequential events" (Lisak, 2013). As a result, Dr. Lisak suggests, one should focus on asking the person what is remembered about these specific fragments, and should not ask for a sequential narrative. He does, at least, acknowledge that respondents will tend to try to cooperate and guess when asked about things they have not successfully encoded, and that any sequential narrative elicited by questions concerning sequence can be inaccurate.

> https://www.youtube.com/watch?v=py0mVt2Z7nc

Still, the trauma-related fragmentation assumption itself seems to be flawed, in that the best research on the topic has revealed no differences in the fragmentation of memory for real life positive, vs. important, vs. traumatic events (see McNally, 2022 for review). If traumatic memories are actually no more fragmented than other memories, is it really necessary to have specialized interviewing procedures for trauma? Or should interviewing regarding all events be conducted using the same special procedures?

The claim of fragmentation of trauma memories, though not new (e.g., van der Kolk and Fisler, 1995), is now much more widespread: and recommended strategies of interviewing alleged trauma victims is based on the idea that memory for sensory fragments of the event will be accurate and should be the focus of questions. Trauma victims are presumed unable to have coherent sequential narratives and these should not be the (at least initial) focus of forensic interviews. Instead, trauma-informed interviewing strategies such as "FETI" suggest asking alleged victims about these kinds of fragments and peripheral details (sounds, smells, feelings, and so on) as a pathway of association that will lead to memory of litigation-relevant details. For example, John Reid Associates (arguably the most prominent interview/interrogation training organization in America), now trainers of FETI, recommends on their website that interviewers focus on seven questions: "What are you able to tell me about your experience?" "Tell me more about…(the room; the person; etc.)." "What was your thought process during this experience?" "What are you able to remember about…5 senses?" "What were your reactions to this experience." "What is the most difficult part of this experience for you?" "What if anything can't you forget about

your experience." While it is true that some details elicited in this manner can lead through associative pathways to other relevant information, they are nevertheless risky in some respects.

First, the account elicited through such questions is highly likely to be fragmented and disorganized, seeming to confirm the testimony of CBEs concerning the nature of trauma memories, and inviting jurors to believe that because the account is fragmented the victim was traumatized (and impliedly the event was indeed rape). Though CBEs are quite right in noting the impact of the way trauma victims are interviewed on the coherence of the narratives elicited, there are other ways of questioning that increase coherence without suggestion. For example, consistent with the principles of cognitive interviewing (Fisher and Geiselman, 1992, 2010), Taylor and colleagues showed that simply asking the person to describe their experience results in a more coherent narrative than use of a series of specific questions regarding details of sequence, persons, context, or events (Taylor et al., 2020). Though FETI does incorporate the relatively open-ended "tell me about" question types of the cognitive interview, it also imposes a different structure by directing attention firmly away from narrative structure to unorganized fragments.

The phenomenon of "retrieval-induced forgetting" (see Bäuml and Kliegl, 2017 for a review), whereby selective retrieval (and particularly repeated retrieval) of some aspects of an event can lead to greater forgetting of others, would also suggest that emphasis on retrieval of non-crucial details might be unwise, particularly when repeated.

More relevant to the focus of this review is the issue of how the procedure of asking the alleged victim to focus on feelings and emotions (and even other sensory details) can serve to prime those emotions and serve as context for recall. As Davis et al. have pointed out elsewhere (Davis and Loftus, 2016, 2019; Davis et al., 2023), emotion can serve similar biasing functions at recall to those at encoding (see also Bower and Forgas, 2001; Forgas, 2008; Gibbons et al., 2018).

First, the emotions felt at retrieval are not necessarily the same as those present at encoding (e.g., Levine et al., 2006; Schmidt et al., 2021). This can happen for at least two reasons. The emotions felt during a sexual encounter might not have been so negative as they later became. Clancy (2011) documented this with many sexual abuse victims who did not begin to feel traumatized by their abuse until they were old enough to understand what it was and why it was so inappropriate. This is also discussed in literature on "unacknowledged rape," regarding those who initially did not regard themselves as victims, but later came to "understand" that they had actually been raped (e.g., Kahn and Mathie, 2000; Wilson and Miller, 2016). Finally, appraisal theorists (e.g., Schmidt et al., 2010) have suggested that biases in memory for emotion can be explained by how much significance a past event is given by the person in the present (e.g., Scherer et al., 2001), and that this might explain some findings that people tend to overestimate the negativity of past emotions (e.g., Schrader et al., 1990; Bryant, 1993; Parkinson et al., 1995; Cutler et al., 1996; Barrett, 1997; Safer et al., 2001; Lench and Levine, 2010; Levine et al., 2021). Moreover, it poses the potential that current negative emotions might provoke a more negative view of a past sexual encounter than warranted by the behaviors at the time.

The emotions a person relies upon for recall can also become different than those experienced at encoding because when the person is asked to describe the emotions they experienced during the event at a later time, they might misremember them. A substantial literature has documented inconsistency in recall of emotions across time. Memory for emotion can be distorted by the same processes as can other memories. Emotion memories can also become distorted to serve current motivations or goals, and to be consistent with current beliefs about oneself and other issues. There is also a tendency toward recalling one's own autobiography and experiences as consistent with current feelings and views of oneself. These and other processes undermine the consistency of emotion between then and now (see Levine et al., 2009; Davis and Loftus, 2019 for reviews). Thus, the emotion relied on for retrieval might both misdirect retrieval of relevant facts, and also bias memory toward consistency with it, as discussed next.

Second, whereas emotion can direct attention during encoding, it can also direct retrieval, leading the person to selectively recall emotion-consistent information as well as to interpret it in an emotion-consistent manner (as incorporated in the previously referenced AIM model: Forgas, 2002). Judgments at recall can then be biased because they are based on incomplete information, the set of which is consistent with the emotion, even though emotion-inconsistent information might have originally predominated.

Third, given that many sexual assaults are not reported immediately, it is important to note that as memories become more vague with time the potential for distortion increases. As suggested by "fuzzy trace" theory (e.g., Brainerd and Reyna, 2005; Brainerd et al., 2022), when verbatim memories are vague, strong emotion memories can lead to constructive memory errors consistent with the emotion and how it would likely be produced. If FETI theorizing is correct regarding the vagueness of trauma memories, this leaves open greater opportunity for memory distortion in the direction of expectations triggered by emotion and other remembered details of context, as well as expectations based on one's self-concept and other general knowledge that tends to infect "fuzzy" memories. As Davis and colleagues (Davis and Loftus, 2016, 2019; Davis et al., 2023) have pointed out, and as discussed earlier regarding effect of emotion on encoding, knowledge and expectations tend to include those regarding circumstances likely to produce the emotions in question: such as when fear or disgust during a sexual encounter will be viewed as consistent with coercion or rape. Given the above considerations, the wisdom of selective priming of emotions and sensory fragments during retrieval, and of relying on the accuracy of resulting reports is questionable.

8 Conclusions

"Extraordinary Claims Require Extraordinary Evidence"
Carl Sagan

What has become known as "the Sagan standard," or "ECREE" seems particularly relevant to the claims of CBEs regarding traumatic memories (Sagan, 1979; Sagan and Druyan, 1997). The claims that trauma causes the brain functions underlying memory to operate in fundamentally different ways than those underlying memories for less stressful events truly are extraordinary: as are those that traumatic memories that are laid down are unerringly accurate, do not fade over time, and are resistant or impervious to distortion. As Sagan and others have pointed out, the burden of proof for such extraordinary claims is on the claimant, not on those

who doubt it. Yet, what extraordinary evidence has been offered for such claims regarding trauma memories? We suggest that there is no extraordinary evidence in support of the claims of CBEs regarding trauma memory, where there is, as we have documented herein, significant evidence to contradict them.

Author contributions

DD: Writing – original draft. AH: Writing – original draft. DH: Writing – original draft.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Adams-Curtis, L. E., and Forbes, G. B. (2004). College women's experiences of sexual coercion: a review of cultural, perpetrator, victim, and situational variables. *Trauma Violence Abuse* 5, 91–122. doi: 10.1177/1524838003262331

Ariely, D., and Loewenstein, G. F. (2006). The heat of the moment: the effect of sexual arousal on sexual decision making. *J. Behav. Decis. Mak.* 19, 87–98. doi: 10.1002/bdm.501

Avramova, Y. R., and Inbar, Y. (2013). Emotion and moral judgment. *Cogn. Sci.* 4, 169–178. doi: 10.1002/wcs.1216

Bargh, J. A., and Raymond, P. (1995). The naive misuse of power: nonconscious sources of sexual harassment. *J. Soc. Issues* 51, 85–96. doi: 10.1111/j.1540-4560.1995.tb01310.x

Bargh, J. A., Raymond, P., Pryor, J. B., and Strack, F. (1995). Attractiveness of the underling: an automatic power → sex association and its consequences for sexual harassment and aggression. *J. Pers. Soc. Psychol.* 68, 768–781. doi: 10.1037/0022-3514.68.5.768

Barrett, L. F. (1997). The relationships among momentary emotion experiences, personality descriptions, and retrospective ratings of emotion. *Pers. Soc. Psychol. Bull.* 23, 1100–1110. doi: 10.1177/01461672972310010

Bäuml, K.-H. T., and Kliegl, O. (2017). "Retrieval-induced remembering and forgetting," in *Learning and Memory: A Comprehensive Reference, 2nd Edn, Vol. 2*, ed J. H. Byrne (San Diego, CA: Academic Press), 27–51.

Bialer, D. M., Reyna, V. F., and Brainerd, C. J. (2021) "False memory: what are the effects, how does fuzzy-trace theory predict them, and how does this matter for eyewitness testimony?," in *Methods, Measures, and Theories in Eyewitness Identification Tasks*, eds A. M. Smith, M. Toglia, and J. M. Lampinen (New York, NY: Routledge/Taylor & Francis Group), 325–352

Blanton, H., and Gerrard, M. (1997). Effect of sexual motivation on men's risk perception for sexually transmitted disease: there must be 50 ways to justify a lover. *Health Psychol.* 16:374. doi: 10.1037/0278-6133.16.4.374

Bohner, G., Eyssel, F., Pina, A., Siebler, F., and Viki, G. T. (2009). "Rape myth acceptance: cognitive, affective and behavioural effects of beliefs that blame the victim and exonerate the perpetrator," in *Rape: Challenging Contemporary Thinking*, 1st Edn, eds. M. Horvath and J. Brown (Uffculme: Willan Publishing, 17–45.

Bonsall, M. B., and Holmes, E. A. (2023). Temporal dynamics of trauma memory persistence. *J. R. Soc. Interface* 20:20230108. doi: 10.1098/rsif.2023.0108

Bouffard, J. A., and Miller, H. A. (2014). The role of sexual arousal and overperception of sexual intent within the decision to engage in sexual coercion. *J. Interpers. Violence* 29, 1967–1986. doi: 10.1177/0886260513515950

Bower, G. H., and Forgas, J. P. (2001). "Mood and social memory," in *Handbook of Affect and Social Cognition*, ed J. P.Forgas (Mahwah, NJ: Lawrence Erlbaum Associates), 95–120.

Brainerd, C. J., Bialer, D. M., and Chang, M. (2022). Fuzzy-trace theory and false memory: meta-analysis of conjoint recognition. *J. Exp. Psychol. Learn. Mem. Cogn.* 48, 1680–1697. doi: 10.1037/xlm0001040

Brainerd, C. J., Chang, M., and Bialer, D. M. (2021). Emotional ambiguity and memory. *J. Exp. Psychol. Gen.* 150, 1476–1499. doi: 10.1037/xge0001011

Brainerd, C. J., and Reyna, V. F. (2005). *The Science of False Memory*. Oxford University Press.

Brainerd, C. J., and Reyna, V. F. (2019). Fuzzy-trace theory, false memory, and the law. *Policy Insights Behav. Brain Sci.* 6, 79–86. doi: 10.1177/2372732218797143

Brainerd, C. J., Stein, L. M., Silveira, R. A., Rohenkohl, G., and Reyna, V. F. (2008). How does negative emotion cause false memories? *Psychol. Sci.* 19, 919–925. doi: 10.1111/j.1467-9280.2008.02177

Brokke, S. S., Bertelsen, T. B., Landrø, N. I., and Haaland, V. Ø. (2022). The effect of sexual abuse and dissociation on suicide attempt. *BMC Psychiatry* 22:29. doi: 10.1186/s12888-021-03662-9

Brosch, T., Scherer, K. R., Grandjean, D., and Sander, D. (2013). The impact of emotion on perception, attention, memory, and decision-making. *Swiss Med. Wkly.* 143, 1–10. doi: 10.4414/smw.2013.13786

Bryant, R. A. (1993). Memory for pain and affect in chronic pain patients. *Pain* 54, 347–351. doi: 10.1016/0304-3959(93)90036-O

Caamano-Isorna, F., Adkins, A., Moure-Rodriguez, L., Conley, A. H., and Dick, D. (2021). Alcohol use and sexual and physical assault victimization among university students: three years of follow-up. *J. Interpers. Viol.* 36, NP3574–NP3595. doi: 10.1177/0886260518780413

Campbell, R. (2012). *The Neurobiology of Sexual Assault: Implications for Law Enforcement, Prosecution and Victim Advocacy*. Research for the Real. Available online at: https://nij.ojp.gov/media/video/24056#transcript$-$0 (accessed June 26, 2024).

Campbell, R. (2014). Sexual assault: A trauma informed to law enforcement first response, part 1 [Video]. *YouTube*. Available online at: https://www.youtube.com/watch?v=CnlXzD2pYSA (accessed June 26, 2024).

Chang, V. T., Overall, N. C., Madden, H., and Low, R. S. T. (2018). Expressive suppression tendencies, projection bias in memory of negative emotions, and well-being. *Emotion* 18, 925–941. doi: 10.1037/emo0000405

Christianson, S. Å. (1992). Emotional stress and eyewitness memory: a critical review. *Psychol. Bull.* 112:284. doi: 10.1037/0033-2909.112.2.284

Christianson, S. Á. (1989). Flashbulb memories: special, but not so special. *Mem. Cogn.* 17, 435–443. doi: 10.3758/BF03202615

Clancy, S. A. (2011). *The Trauma Myth: The Truth About the Sexual Abuse of Children—anD its Aftermath*. New York, NY: Basic Books.

Clore, G. L., Schiller, A. J., and Shaked, A. (2018). Affect and cognition: three principles. *Curr. Opin. Behav. Sci.* 19, 78–82. doi: 10.1016/j.cobeha.2017.11.010

Clore, G. L., and Schnall, S. (2019). "The influence of affect on attitude," in *The Handbook of Attitudes, Volume I: Basic Principles, 2nd Edn*, eds. D. Albarracín and B. T. Johnson (Routledge), 259–290.

Coane, J. H. (2013). Retrieval practice and elaborative encoding benefit memory in younger and older adults. *J. Appl. Res. Mem. Cogn.* 2, 95–100. doi: 10.1016/j.jarmac.2013.04.001

Cutler, S. E., Larson, R. J., and Bunce, S. C. (1996). Repressive coping style and the experience and recall of emotion: a naturalistic study of daily affect. *J. Pers.* 64, 379–405. doi: 10.1111/j.1467-6494.1996.tb00515.x

Davis, D., Cano, J., Miller, G., and Loftus, E. F. (2023). The multiple roles of emotion in memory for disputed sexual encounters. *Top. Cogn. Sci.* 1–17. doi: 10.1111/tops.12691

Davis, D., and Loftus, E. F. (2004). "What's good for the goose cooks the gander: inconsistencies between the law and psychology of voluntary intoxication and sexual assault," in *Handbook of Forensic Psychology: Resource for Mental Health and Legal Professionals*, eds. W. T. O'Donohue and E. R. Levensky (Cambridge, MA: Academic Press), 997–1032.

Davis, D., and Loftus, E. F. (2007). "Internal and external sources of distortion in adult witness memory," in *Handbook of Eyewitness Memory, Vol. 1: Memory for Events*, eds M. P. Toglia, J. D. Read, D. R. Ross, and R. C. L. Lindsay (Mahwah, NJ: Erlbaum), 195–237.

Davis, D., and Loftus, E. F. (2009). "Expectancies, emotion, and memory reports for visual events," in *The Visual World in Memory, 1st Edn*, ed. J. Brockmole (London: Psychology Press), 178–214.

Davis, D., and Loftus, E. F. (2016). Remembering disputed sexual encounters: A new frontier for witness memory research. *J. Crim. L. Criminol.* 105:811.

Davis, D., and Loftus, E. F. (2019). Title IX and "Trauma-Focused" investigations: the good, the bad and the ugly. *J. Appl. Res. Mem. Cogn.* 8, 403–410. doi: 10.1016/j.jarmac.2019.08.001

Davis, D., and Loftus, E. F. (2020). "Recovered memories and false memories," in *New Oxford Textbook of Psychiatry, 3rd Edn*, eds N. Andreasen, J. Geddes, and G. Goodwin (Oxford: Oxford University Press).

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., and McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law Hum. Behav.* 28, 687–706. doi: 10.1007/s10979-004-0565-x

Ditto, P. H., Pizarro, D. A., Epstein, E. B., Jacobson, J. A., and MacDonald, T. K. (2006). Visceral influences on risk taking behavior. *J. Behav. Decis. Mak.* 19, 99–113. doi: 10.1002/bdm.520

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

Fenner, L. (2017). Sexual consent as a scientific subject: a literature review. *Am. J. Sex. Educ.* 12, 451–471. doi: 10.1080/15546128.2017.1393646

Fisher, R. P., and Geiselman, R. E. (1992). *Memory-Enhancing Techniques for Investigative Interviewing: The Cognitive Interview*. Springfield: Charles C Thomas Publisher.

Fisher, R. P., and Geiselman, R. E. (2010). The cognitive interview method of conducting police interviews: eliciting extensive information and promoting therapeutic jurisprudence. *Int. J. Law Psychiatry* 33, 321–328. doi: 10.1016/j.ijlp.2010.09.004

Fiske, S. T. (1993). Controlling other people: the impact of power on stereotyping. *Am. Psychol.* 48, 621–628. doi: 10.1037/0003-066X.48.6.621

Forgas, J. P. (2002). Feeling and doing: affective influences on interpersonal behavior. *Psychol. Inq.* 13, 1–28. doi: 10.1207/S15327965PLI1301_01

Forgas, J. P. (2008). "Affect, cognition, and social behavior: the effects of mood on memory, social judgments, and social interaction," in *Memory and Mind: A Festschrift for Gordon H. Bower*, eds M. A. Gluck, J. R. Anderson, and S. M. Kosslyn (Mahwah, NJ: Lawrence Erlbaum Associates), 261–279.

Fredrickson, B. L. (2000). Extracting meaning from past affective experiences: the importance of peaks, ends, and specific emotions. *Cogn. Emot.* 14, 577–606. doi: 10.1080/026999300402808

Galinsky, A. D., Gruenfeld, D. H., and Magee, J. C. (2003). From power to action. *J. Pers. Soc. Psychol.* 85, 453–466. doi: 10.1037/0022-3514.85.3.453

Galinsky, A. D., Magee, J. C., Gruenfeld, D. H, Whitson, J. A., and Liljenquist, K. A. (2008). Power reduces the press of the situation: implications for creativity, conformity, and dissonance. *J. Pers. Soc. Psychol.* 95, 1450–1466. doi: 10.1037/a0012633

Galinsky, A. D., Rucker, D. D., Magee, J. C. (2016). Power and perspective-taking: a critical examination. *J. Exp. Soc. Psychol.* 67, 91–92. doi: 10.1016/j.jesp.2015.12.002

Gibbons, H., Seib-Pfeifer, L.-E., Koppelheie-Gossel, J., and Schnuerch, R. (2018). Affective priming and cognitive load: event-related potentials suggest and interplay of implicit affect misattribution and strategic inhibition. *Psychophysiology* 55, 1–19. doi: 10.1111/psyp.13009

Greenwald, A. G., and Banaji, M. R. (2017). The implicit revolution: reconceiving the relation between conscious and unconscious. *Am. Psychol.* 72:861. doi: 10.1037/amp0000238

Greifeneder, R., Bless, H., and Pham, M. T. (2011). When do people rely on affective and cognitive feelings in judgment? A review. *Pers. Soc. Psychol. Rev.* 15, 107–141. doi: 10.1177/1088868310367640

Guinote, A. (2007). Power and goal pursuit. *Pers. Soc. Psychol. Bull.* 33, 1076–1087. doi: 10.1177/0146167207301011

Guinote, A. E. (2010). In touch with your feelings: power increases reliance on bodily information. *Soc. Cogn.* 28, 110–121. doi: 10.1521/soco.2010.28.1.110

Guinote, A. E., and Vescio, T. K. (2010). *The Social Psychology of Power*. New York, NY: Guilford Press.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychol. Rev.* 108, 814–834. doi: 10.1037/0033-295X.108.4.814

Haidt, J. (2012). *The Righteous Mind: Why Good People are Divided by Politics and Religion*. New York, NY: Pantheon.

Harmon-Jones, E., Gable, P. A., and Price, T. F. (2013). Does negative affect always narrow and positive affect always broaden the mind? Considering the influence of motivational intensity on cognitive scope. *Curr. Dir. Psychol. Sci.* 22, 301–307. doi: 10.1177/0963721413481353

Hirsh, J. B., Galinsky, A. D., and Zhong, C. B. (2011). Drunk, powerful, and in the dark: how general processes of disinhibition produce both prosocial and antisocial behavior. *Perspect. Psychol. Sci.* 6, 415–427. doi: 10.1177/1745691611416992

Hoffman, D. D. (2019). *The Case Against Reality: Why Evolution Hid the Truth From Our Eyes*. New York, NY: W.W. Norton & Company.

Hogan, A., Hart, D., Miller, G., and Davis, D. (2024). *Perceptions of Sexually Coercive Behaviors* [Poster presentation]. San Francisco, CA: The Association for Psychological Science Annual Convention.

Hopper, J. (2015). *Scientific Research*. Available online at: https://jimhopper.com/topics/child-~abuse/recovered-memories-of-sexual-abuse/scientific-research/ (accessed June 26, 2024).

Hopper, J. (2018a). *How Reliable Are the Memories of Sexual Assault Victims?* Scientific American. Available online at: https://blogs.scientificamerican.com/observations/how-reliable-are-the-memories-of-sexual-assault-victims/ (accessed June 26, 2024).

Hopper, J. (2018b). *Sexual Assault & the Brain in Six Minutes*. Available online at: https://www.youtube.com/watch?v=7G8weF4jsZw&t=1s (accessed June 26, 2024).

Hopper, J. (2023). *Recovered Memories of Sexual Abuse*. Available online at: https://jimhopper.com/topics/child-abuse/recovered-memories-of-sexual-abuse/ (accessed June 26, 2024).

Hopper, J. (n.d.-b). *Potential Overlaps of Dissociation with Freezing, Habit Behaviors, and/or Extreme Survival Reflexes [PowerPoint slide]*. Available online at: https://jimhopper.com/pdf/hopper_dissociation_overlaps.pdf

Hopper, J., and Lisak, D. (2014). Why rape and trauma survivors have fragmented and incomplete memories. *Time Magazine*. Available online at: https://time.com/3625414/rape-trauma-brain-memory/

Huntsinger, J. R. (2012). Does positive affect broaden and negative affect narrow attentional scope? A new answer to an old question. *J. Exp. Psychol. Gen.* 141, 595–600. doi: 10.1037/a0027709

Huntsinger, J. R. (2013). Does emotion directly tune the scope of attention? Curr. *Dir. Psychol. Sci.* 22, 265–270. doi: 10.1177/0963721413480364

Imhoff, R., and Schmidt, A. F. (2014). Sexual disinhibition under sexual arousal: Evidence for domain specificity in men and women. *Arch. Sex. Behav.* 43, 1123–1136. doi: 10.1007/s10508-014-0329-8

Kabota, S., and Nakazawa, E. (2022). Concept and implications of sexual consent for education: a systematic review of empirical studies. *Sex. Relat. Therapy* 1–23. doi: 10.1080/14681994.2022.2039617

Kahn, A. S., and Mathie, V. A. (2000). "Understanding the unacknowledged rape victim," in *Sexuality, Society, and Feminism, 1st Edn*, eds. C. B. Travis and J. W. White (Washington, DC: American Psychological Association), 377–403. doi: 10.1037/10345-015

Kahneman, D. (2011). *Thinking, Fast and Slow*. Macmillan. New York, NY: Farrar, Straus, & Giroux.

Kaplan, R. L., Van Damme, I., and Levine, L. J. (2012). Motivation matters: differing effects of pre-goal and post-goal emotions on attention and memory. *Front. Psychol.* 3:404. doi: 10.3389/fpsyg.2012.00404

Keltner, D., Gruenfeld, D. H., and Anderson, C. (2003). Power, approach, and inhibition. *Psychol. Rev.* 110:265. doi: 10.1037/0033-295X.110.2.265

Kern, R. P., Libkuman, T. M., Otani, H., and Holmes, K. (2005). Emotional stimuli, divided attention, and memory. *Emotion* 5, 408–417. doi: 10.1037/1528-3542.5.4.408

Kim, B. K., and Zauberman, G. (2013). Can Victoria's Secret change the future? A subjective time perception account of sexual-cue effects on impatience. *J. Exp. Psychol. Gen.* 142, 328–335. doi: 10.1037/a0028954

Kindelan, K. (2018). *Rose McGowan's Account of Detaching From Reality During Alleged Rape Is Common: Experts*. Good Morning America. Available online at:https://www.goodmorningamerica.com/wellness/story/rose-mcgowans-account-~detaching-reality-alleged-rape-common-52747906 (accessed June 26, 2024).

Kunstman, J. W., and Maner, J. K. (2011). Sexual overperception: Power, mating motives, and biases in social judgment. *J. Pers. Soc. Psychol.* 100:282. doi: 10.1037/a0021135

Lammers, J., Galinsky, A. D., Dubois, D., and Rucker, D. D. (2015). Power and morality. *Curr. Opin. Psychol.* 6, 15–19. doi: 10.1016/j.copsyc.2015.03.018

Lammers, J., Galinsky, A. D., Gordijn, E. H., and Otten, S. (2008). Illegitimacy moderates the effects of power on approach. *Psychol. Sci.* 19, 558–564. doi: 10.1111/j.1467-9280.2008.02123.x

Lammers, J., and Imhoff, R. (2016). Power and sadomasochism: understanding the antecedents of a knotty relationship. *Soc. Psychol. Personal. Sci.* 7, 142–148. doi: 10.1177/1948550615604452

Lammers, J., Stoker, J. I., Jordan, J., Pollmann, M., and Stapel, D. A. (2011). Power increases infidelity among men and women. *Psychol. Sci.* 22, 1191–1197. doi: 10.1177/0956797611416252

Landy, J. F., and Goodwin, G. P. (2015). Does incidental disgust amplify moral judgment? A meta-analytic review of experimental evidence. *Perspect. Psychol. Sci.* 10, 518–536. doi: 10.1177/17456916155831

Laney, C., and Loftus, E. F. (2005). Traumatic memories are not necessarily accurate memories. *Can. J. Psychiatry* 50, 823–828. doi: 10.1177/070674370505001303

Lee, H., Roh, S., and Kim, D. J. (2009). Alcohol-induced blackout. *Int. J. Environ. Res. Public Health* 6, 2783–2792. doi: 10.3390/ijerph6112783

Lench, H. C., and Levine, L. J. (2010). Motivational biases in memory for emotions. *Cogn. Emot.* 24, 401–418. doi: 10.1080/02699930802650788

Levine, L. J., and Edelstein, R. S. (2009). Emotion and memory narrowing: a review and goal-relevance approach. *Cogn. Emot.* 23, 833–875. doi: 10.1080/02699930902738863

Levine, L. J., Lench, H. C., and Safer, M. A. (2009). Functions of remembering and misremembering emotion. *Appl. Cogn. Psychol.* 23, 1059–1075. doi: 10.1002/acp.1610

Levine, L. J., Murphy, G., Lench, H. C., Greene, C. M., Loftus, E. F., Tinti, C., et al. (2021). Remembering facts versus feelings in the wake of political events. *Cogn. Emot.* 35, 936–955. doi: 10.1080/02699931.2021.1910496

Levine, L. J., and Pizarro, D. A. (2004). Emotion and memory research: a grumpy overview. *Soc. Cognit.* 22(5: Special issue), 530–554. doi: 10.1521/soco.22.5.530.50767

Levine, L. J., Safer, M. A., and Lench, H. C. (2006). "Remembering and misremembering emotions," in *Judgments Over Time: The Interplay of Thoughts, Feelings, and Behaviors*, eds. L. J. Sanna, and E. C. Chang (Oxford: Oxford University Press), 271–290.

Lisak, D. (2013). Neurobiology of trauma - Dr. David Lisak [Video]. [ACASA Arkansas]. *YouTube*. Available online at: https://www.youtube.com/watch?v=py0mVt2Z7nc (accessed June 26, 2024).

Livingston, T. N., and Davis, D. (2020). Power affects perceptions of sexual willingness: Implications for litigating sexual assault allegations. *Viol. Gender* 7, 116–120. doi: 10.1089/vio.2019.0068

Loewenstein, G. (1996). Out of control: visceral influences on behavior. *Org. Behav. Hum. Decis. Process* 65, 272–292. doi: 10.1006/obhd.1996.0028

Loftus, E. F., Loftus, G. R., and Messo, J. (1987). Some facts about "weapon focus." *Law Hum. Behav.* 11, 55–62. doi: 10.1007/BF01044839

Lord, C. G., and Taylor, C. A. (2009). Biased assimilation: effects of assumptions and expectations on the interpretation of new evidence: biased assimilation. *Soc. Personal. Psychol. Compass* 3, 827–841. doi: 10.1111/j.1751-9004.2009.00203.x

Lynch, S., Weber, S., Kaplan, S., and Craun, E. (2023). Childhood and adult sexual violence exposures as predictors of PTSD, dissociation, and substance use in women in jail. *J. Child Sex. Abuse* 1–17. doi: 10.1080/10538712.2023.2226132

Mackworth, N. H. (1965). Visual noise causes tunnel vision. *Psychon. Sci.* 3, 67–68. doi: 10.3758/BF03343023

Magee, J. C., and Smith, P. K. (2013). The social distance theory of power. *Person. Soc. Psychol. Rev.* 17, 158–186. doi: 10.1177/1088868312472732

Maner, J. K., Kendrick, D. T., Neuberg, S. L., Becker, D. V., Robertson, T., Hofer, B., et al. (2005). Functional projection: how fundamental social motives can bias interpersonal perception. *J. Pers. Soc. Psychol.* 88, 63–78. doi: 10.1037/0022-3514.88.1.63

Marr, C., Sauerland, M., Otgaar, H., Quaedflieg, C. W. E. M., and Hope, L. (2021). The effects of acute stress on eyewitness memory: an integrative review for eyewitness researchers. *Memory* 20, 1091–1100. doi: 10.1080/09658211.2021.1955935

McAuliffe, W. H. (2019). Do emotions play an essential role in moral judgments? *Think. Reason.* 25, 207–230. doi: 10.1080/13546783.2018.1499552

McNally, R. J. (2003). *Remembering Trauma*. Cambridge, MA: The Belknap Press of Harvard University Press.

McNally, R. J. (2022). Are memories of sexual assault fragmented? *Memory* 30, 26–30. doi: 10.1080/09658211.2020.1871023

McNally, R. J., Clancy, S. A., Barrett, H. M. (2004). "Forgetting trauma?" in *Memory and Emotion*, eds. D. Reisberg and P. Hertel (Oxford University Press), 129–154. doi: 10.1093/acprof:oso/9780195158564.003.0004

Melaragno, E. (2018). Trauma in the body: An interview with Dr. Bessel van der Kolk. *Still Harbor*. Available online at: https://www.dailygood.org/story/1901/trauma-in-the-body-an-interview-with-dr-bessel-van-der-kolk-elissa-melaragno/ (accessed June 26, 2024).

Miller, G., and Davis, D. (2024). Effects of male physical attractiveness and career potential on ratings of implications of female behaviors for sexual willingness (Unpublished Manuscript).

Mitchell, J. (2023). Emotion and attention. *Philos. Stud.* 180, 73–99. doi: 10.1007/s11098-022-01876-5

Morgan, C. A. III, Hazlett, G., Doran, A., Garrett, S., Hoyt, G., Thomas, P., et al. (2004). Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress. *Int. J. Law Psychiatry* 27, 265–279. doi: 10.1016/j.ijlp.2004.03.004

Morgan, C. A. III, Southwick, S., Steffian, G., Hazlett, G. A., and Loftus, E. F. (2013). Misinformation can influence memory for recently experienced, highly stressful events. *Int. J. Law Psychiatry* 36, 11–17. doi: 10.1016/j.ijlp.2012.11.002

Muehlenhard, C., Humphreys, T., Jozkowski, K., and Peterson, J. (2016). The complexities of sexual consent among college students: a conceptual and empirical review. *J. Sex Res.* 53, 1–31. doi: 10.1080/00224499.2016.1146651

Muehlenhard, C. L., and Hollabaugh, L. C. (1988). Do women sometimes say no when they mean yes? The prevalence and correlates of women's token resistance to sex. *J. Pers. Soc. Psychol.* 54, 872–879. doi: 10.1037/0022-3514.54.5.872

Murray, D. R., Murphy, S. C., von Hippel, W., Trivers, R., and Haselton, M. G. (2017). A preregistered study of competing predictions suggests that men do overestimate women's sexual intent. *Psychol. Sci.* 28, 253–255. doi: 10.1177/0956797616675474

Parkinson, B., Briner, R. B., Reynolds, S., and Totterdell, P. (1995). Time frames for mood: relations between momentary and generalized ratings of affect. *Pers. Soc. Psychol. Bull.* 21, 331–339. doi: 10.1177/0146167295214003

People v. Mayberry (1975). *15 Cal.3d 143. SCOCAL*. Available online at: https://scocal.stanford.edu/opinion/people-v-mayberry-23014 (accessed June 26, 2024).

Percy, J. (2023). *What People Misunderstand About Rape*. The New York Times. Available online at: https://www.nytimes.com/2023/08/22/magazine/immobility-rape-trauma-freeze.html#:$\sim$:text$=$Sexual%20assault%20often%20goes%20unpunished,an%20involuntary%20response%20to%20trauma (accessed June 26, 2024).

Pickel, K. (1999). The influence of context on the "weapon focus" effect. *Law Hum. Behav.* 23, 299–311. doi: 10.1023/A:1022356431375

Pike, B. E., and Galinsky, A. D. (2020). Power leads to action because it releases the psychological brakes on action. *Curr. Opin. Psychol.* 33, 91–94. doi: 10.1016/j.copsyc.2019.06.028

Pryor, J. B., and Stoller, L. M. (1994). Sexual cognition processes in men high in the likelihood to sexually harass. *Pers. Soc. Psychol. Bull.* 20, 163–169. doi: 10.1177/0146167294202003

Reicherter, D. Wang, S., Ohrtman, T. H., Ndukwe, N., Vaatainen, S., Alcalay, S., and Brown, L. M. (2022). Implementation of trauma-informed best practices for international criminal investigations conducted by the United Nations Investigative Team to promote accountability for crimes committed by Da'esh/ISIL (UNITAD). *Psychol. Inj. Law* 15, 319–329. doi: 10.1007/s12207-022-09457-x

Reisberg, D., and Heuer, F. (2020). "Emotion's (varied) impact on memory for sexual misconduct," in *Memory and Sexual Misconduct: Psychological Research for Criminal Justice*, eds. J. Pozzulo, E. Pica, and C. Sheahan (New York, NY: Routledge; Taylor & Francis Group), 7–41. doi: 10.4324/9780429027857-2

Rerick, P.O., Livingston, T. N., and Davis, D. (2020). Does the horny man think women want him too? Effects of male sexual arousal on perceptions of female willingness. *J. Soc. Psychol.* 160, 520–533. doi: 10.1080/00224545.2019.1692330

Rerick, P. O., and Livingston, T. N. (2022). Beauty is in the eye of the beholder, and so is intent: men's interpretations of the sexual intent of attractive versus unattractive women. *Violence Gend.* 9, 193–200. doi: 10.1089/vio.2022.0014

Rerick, P. O., Livingston, T. N., and Davis, D. (2019). "Rape and the jury," in *Handbook of Sexual Assault and Sexual Assault Prevention*, eds. W. O. O'Donohue, C. Cummings, and P. Schewe (Springer), 551–571. doi: 10.1007/978-3-030-23645-8_33

Rerick, P. O., Livingston, T. N., and Davis, D. (2022). Let's just do it: Sexual arousal's effects on attitudes regarding sexual consent. *J. Soc. Psychol.* 164, 456–472. doi: 10.1080/00224545.2022.2106174

Richards, J. M., and Gross, J. J. (1999). Composure at any cost? The cognitive consequences of emotion suppression. *Pers. Soc. Psychol. Bull.* 25, 1033–1044. doi: 10.1177/014616729925110

Richards, J. M., and Gross, J. J. (2000). Emotion regulation and memory: the cognitive costs of keeping one's cool. *J. Pers. Soc. Psychol.* 79:410. doi: 10.1037/0022-3514.79.3.410

Richards, J. M., and Gross, J. J. (2006). Personality and emotional memory: how regulating emotion impairs memory for emotional events. *J. Res. Pers.* 40, 631–651. doi: 10.1016/j.jrp.2005.07.002

Rubin, D. C., and Kozin, M. (1984). Vivid memories. *Cognition* 16, 81–95. doi: 10.1016/0010-0277(84)90037-4

Ryan, K. (2011). The relationship between rape myths and sexual scripts: the social construction of rape. *Sex Roles* 65, 774–782. doi: 10.1007/s11199-011-0033-2

Safer, M. A., Bonanno, G. A., and Field, N. P. (2001). "It was never that bad": biased recall of grief and long-term adjustment to the death of a spouse. *Memory* 9, 195–203. doi: 10.1080/09658210143000065

Safer, M. A., Christianson, S.-Å., Autry, M. W., and Österlund, K. (1998). Tunnel memory for traumatic events. *Appl. Cogn. Psychol.* 12:99. doi: 10.1002/(SICI)1099-0720(199804)12:2<99::AID-ACP509>3.0.CO;2-7

Sagan, C. (1979). *Broca's Brain: The Romance of Science*. New York, NY: Random House, 317–319.

Sagan, C., and Druyan, A. (1997). *The Demon-Haunted World: Science as a Candle in the Dark*. New York, NY: Ballantine Books.

Schacter, D. L. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *Am. Psychol.* 54, 182–203. doi: 10.1037/0003-066X.54.3.182

Schacter, D. L. (2001). *The Seven Sins of Memory: How the Mind Forgets and Remembers*. Houghton, Mifflin and Company.

Scherer, K. R., Schorr, A., and Johnstone, T. (2001). *Appraisal Processes in Emotion: Theory, Methods, Research*. Oxford: Oxford University Press.

Schmidt, S., Muzzulini, B., Levine, L. J., and Tinti, C. (2021). Sources of bias in memory for emotional reactions to Brexit: current feelings mediate the link between appraisals and memories. *Appl. Cogn. Psychol.* 35, 819–827. doi: 10.1002/acp.3806

Schmidt, S., Tinti, C., Levine, L. J., and Testa, S. (2010). Appraisals, emotions and emotion regulation: an integrative approach. *Motiv. Emot.* 34, 63–72. doi: 10.1007/s11031-010-9155-z

Schnall, S., Haidt, J., Clore, G. L., and Jordan, A. H. (2015). Landy and Goodwin (2015) confirmed most of our findings then drew the wrong conclusions. *Perspect. Psychol. Sci.* 10, 537–538. doi: 10.1177/1745691615589078

Schoch, S. F., Cordi, M. J., and Rasch, B. (2017). Modulating influences of memory strength and sensitivity of the retrieval test on detectability of the sleep consolidation effect. *Neurobiol. Learn. Mem.* 145, 181–189. doi: 10.1016/j.nlm.2017.10.009

Schrader, G., Davis, A., Stefanovic, S., and Christie, P. (1990). The recollection of affect. *Psychol. Med.* 20, 105–109. doi: 10.1017/S0033291700013271

Slabu, L., and Guinote, A. (2010). Getting what you want: power increases the accessibility of active goals. *J. Exp. Soc. Psychol.* 46, 344–349. doi: 10.1016/j.jesp.2009.10.013

Smith, P. K., and Bargh, J. A. (2008). Nonconscious effects of power on basic approach and avoidance tendencies. *Soc. Cogn.* 26, 1–24. doi: 10.1521/soco.2008.26.1.1

Southwick, S. M., Morgan, C. A., Nicolaou, A. L., and Charney, D. S. (1997). Consistency of memory for combat-related traumatic events in veterans of Operation Desert Storm. *Am. J. Psychiatry* 154, 173–177. doi: 10.1176/ajp.154.2.173

Steblay, N. (1992). A meta-analytic review of the weapon focus effect. *Law Hum. Behav.* 16, 413–424. doi: 10.1007/BF02352267

Steele, B., Martin, M., Yakubovich, A., Humphreys, D. K., and Nye, E. (2022). Risk and protective factors for men's sexual violence against women at higher education institutions: a systematic and meta-analytic review of the longitudinal evidence. *Trauma Violence Abuse* 23, 716–732. doi: 10.1177/1524838020970900

Steele, C. M., and Josephs, R. A. (1990). Alcohol myopia: its prized and dangerous effects. *Am. Psychol.* 45:921. doi: 10.1037/0003-066X.45.8.921

Stephan, W., Berscheid, E., and Walster, E. (1971). Sexual arousal and heterosexual perception. *J. Pers. Soc. Psychol.* 20, 93–101. doi: 10.1037/h0031943

Stevenson, R. J., Case, T. I., and Oaten, M. J. (2011). Effect of self-reported sexual arousal on responses to sex-related and non-sex-related disgust cues. *Arch. Sex. Behav.* 40, 79–85. doi: 10.1007/s10508-009-9529-z

Strand, R. W., and Heitman, L. D. (2017). *The Forensic Experiential Trauma Interview (FETI)*. Arizona Coalition for Victim Services. Available online at: http://www.azcvs.net/wp-content/uploads/FETI-Public-Description-Jan-2017.pdf

Talarico, J. M., and Rubin, D. C. (2003). Confidence, not consistency, characterizes flashbulb memories. *Psychol. Sci.* 14, 455–461. doi: 10.1111/1467-9280.0245

Talarico, J. M., and Rubin, D. C. (2007). Flashbulb memories are special after all; in phenomenology, not accuracy. *Appl. Cogn. Psychol.* 21, 557–578. doi: 10.1002/acp.1293

Taylor, A., Jordan, K., Zajac, R., Takarangi, M. K. T., and Garry, M. (2020). Judgments of memory coherence depend on the conditions under which a memory is retrieved, regardless of reported PTSD symptoms. *J. Appl. Res. Mem. Cogn.* 9, 396–409. doi: 10.1037/h0101855

Tippett, K. (2021). Bessel van der Kolk: How trauma lodges in the body, revisited. *On Being*. Available online at: https://onbeing.org/programs/bessel-van-der-kolk-how-trauma-lodges-in-the-body-revisited/ (accessed June 26, 2024).

Trammell, J. P., and Clore, G. L. (2014). Does stress enhance or impair memory consolidation? *Cogn. Emot.* 28, 361–374. doi: 10.1080/02699931.2013.822346

van der Kolk, B. (2014). *The Body Keeps the Score: Brain, Mind, and Body in the Healing of Trauma*. London: Penguin Books.

van der Kolk, B. A. (1998). The psychology and psychobiology of developmental trauma. *Prax. Kinderpsychol. Kinderpsychiatr.* 47, 19–35.

van der Kolk, B. A., and Fisler, R. (1995). Dissociation and the fragmentary nature of traumatic memories: overview and exploratory study. *J. Trauma. Stress* 8, 505–525. doi: 10.1002/jts.2490080402

van Leeuwen, F., Dukes, A., Tybur, J. M., and Park, J. H. (2017). Disgust sensitivity relates to moral foundations independent of political ideology. *Evol. Behav. Sci.* 11:92. doi: 10.1037/ebs0000075

Vanderploeg, R. D., Donnell, A. J., Belanger, H. G., and Curtiss, G. (2014). Consolidation deficits in traumatic brain injury: the core and residual verbal memory defect. *J. Clin. Exp. Neuropsychol.* 36, 58–73. doi: 10.1080/13803395.2013.864600

Villalobos, J. G., Davis, D., and Leo, R. A. (2016). *History, Her Story: Sexual Miscommunication, Motivated Remembering, and Intoxication as Pathways to Honest FALSE Testimony Regarding Sexual Consent*. Vilified: Wrongful Allegations of Sexual and Child Abuse. (Oxford: Oxford University Press, 2016, Forthcoming)., Univ. of San Francisco Law Research Paper, (2014-33). Available online at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2480049 (accessed June 26, 2024).

Wagenaar, W. A., and Groeneweg, J. (1990). The memory of concentration camp survivors. *Appl. Cogn. Psychol.* 4, 77–87. doi: 10.1002/acp.2350040202

Wertheimer, A. (2003). *Consent to Sexual Relations*. Cambridge: Cambridge University Press.

Whitson, J. A., Liljenquist, K. A., Galinsky, A. D., Magee, J. C., Gruenfeld, D. H., and Cadena, B. (2013). The blind leading: power reduces awareness of constraints. *J. Exp. Soc. Psychol.* 49, 579–582. doi: 10.1016/j.jesp.2012.10.009

Willis, M., Blunt-Vinti, H. D., and Jozkowski, K. N. (2019). Associations between internal and external sexual consent in a diverse national sample of women. *Pers. Individ. Dif.* 149, 37–45. doi: 10.1016/j.paid.2019.05.029

Wilson, L. C., and Miller, K. E. (2016). Meta-analysis of the prevalence of unacknowledged rape. *Trauma Violence Abuse* 17, 149–159. doi: 10.1177/1524838015576391

Wood, E. F., Rikkonen, K. J., and Davis, D. (2019). "Definition, communication, and interpretation of sexual consent," in *Handbook of Sexual Assault and Sexual Assault Prevention, 1st Edn*, eds. W. T. O'Donohue and P. A. Schewe (New York, NY: Springer), 399–421.

Yu, A. N. C., Iodice, P., Pezzulo, G., and Barca, L. (2021). Bodily information and top-down affective priming jointly affect the processing of fearful faces. *Front. Psychol.* 12:625986. doi: 10.3389/fpsyg.2021.625986

Zurbriggen, E. L. (2000). Social motives and cognitive power-sex associations: predictors of aggressive sexual behavior. *J. Pers. Soc. Psychol.* 78, 559–581. doi: 10.1037/0022-3514.78.3.559

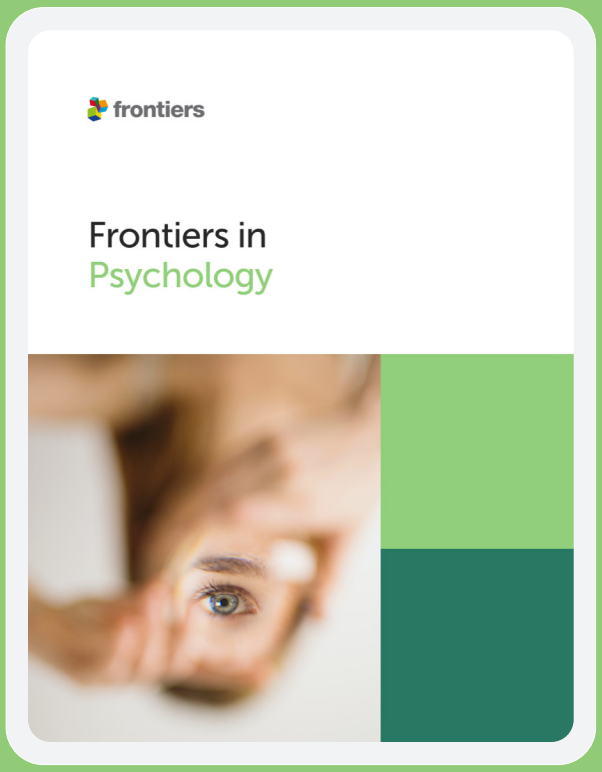
Frontiers in Psychology

Paving the way for a greater understanding of human behavior

The most cited journal in its field, exploring psychological sciences - from clinical research to cognitive science, from imaging studies to human factors, and from animal cognition to social psychology.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact