Misinformation and the Sins of Memory: False-Belief Formation and Limits on Belief

Revision

Eryn J. Newman,¹ Briony Swire-Thompson,²,³ & Ullrich K. H. Ecker⁴,⁵

¹ Research School of Psychology, The Australian National University, Canberra, Australia
² Network Science Institute, Northeastern University, Boston, USA
³ Institute of Quantitative Social Science, Harvard University, Cambridge, USA
⁴ School of Psychological Science, University of Western Australia, Perth, Australia
⁵ Public Policy Institute, University of Western Australia, Perth, Australia

Author Note

Eryn J. Newman  http://orcid.org/0000-0001-8663-7173.
Briony Swire-Thompson  https://orcid.org/0000-0001-7464-8940
Ullrich K. H. Ecker  https://orcid.org/0000-0003-4743-313X

UKHE is supported by Australian Research Council grant FT190100708 and BST is supported by National Institute of Health Pathway to Independence Award (1K99CA248720-01A). We thank Krissy Kilgallen for research support.

Correspondence concerning this article should be addressed to Eryn Newman, Research School of Psychology, The Australian National University, 39 Science Road, Canberra, 2600, Australia. Email: Eryn.Newman@anu.edu.au
Misinformation and the Sins of Memory: False-Belief Formation and Limits on Belief

Revision
MISINFORMATION AND THE SINS OF MEMORY

Keywords: Misinformation, Memory, False Beliefs, Continued Influence Effect, Truth

Assessment
MISINFORMATION AND THE SINS OF MEMORY

Misinformation and the Sins of Memory: False-Belief Formation and Limits on Belief

Revision

Misinformation has the potential to negatively affect both individual decision making and the common good. For example, belief in misinformation can have a negative impact on public health, environmental behaviors, and democracy (Cook, 2019; Lewandowsky et al., 2017; Loomba et al., 2021; Nisbet et al., 2021; Swire-Thompson & Lazer, 2022). Both in online and offline settings, people’s vulnerability to misinformation arises in part from features of our cognitive system. These cognitive features can lead us to form false beliefs, remember false claims as true, and struggle with memory updating and belief revision when misinformation is corrected. Here, we provide a commentary on the target article Memory Sins in Applied Settings: What Kind of Progress? (Schacter, 2022b), as many of Schacter’s sins of memory are fundamental to understanding the mechanisms behind the cognitive impacts of misinformation. Specifically, the current article discusses these sins in the context of (1) the formation of false beliefs and (2) the continued influence that misinformation can have on cognition after having been corrected.

False-Belief Formation and Inherent Vulnerabilities of Truth Assessment

Whether online or otherwise, we are routinely tasked with assessing the credibility of information we encounter. When sorting true from false, people draw on several criteria to make inferences about veracity (Brashier & Marsh, 2020; Ecker, Lewandowsky, et al., 2022; Schwarz, 2015; Schwarz et al., 2016). For instance, people are cautious of inconsistencies (Unkelbach & Rom, 2017; Winkielman et al., 2012); when information is inconsistent or incompatible with what people already believe, they are less likely to endorse it as factual (Jiang et al., 2021; Oyserman & Dawson, 2020; Pennycook & Rand, 2021b). People are also sensitive to internal
coherence of a message, that is, whether it is logical and forms a rational argument (Pennington & Hastic, 1992; Simon, 2004). Through a more social lens, people draw on consensus in forming attitudes and beliefs (Festinger, 1954; van der Linden et al., 2015); when there is a perception that others agree on a given account of reality, we are inclined to concur (Kerr & van der Linden, 2022). The source of the information matters too; people are generally more inclined to believe credible sources (Brinol & Petty, 2009; Mackie et al., 1990; Nadarevic et al., 2020). While a careful analysis of these criteria may lead to a discerning conclusion regarding truth, some features of truth assessment, such as a reliance on intuition and feelings to infer veracity, can lull us into error, and associated sins of human memory can contribute to false-belief formation.

When people seek and encode information, veracity is not always at the front of their mind, and information can appeal for reasons other than its factual truthfulness (Acerbi, 2019). Moreover, built-in biases can introduce error, which can lead us to prioritize some accounts of truth over others. Schacter identified bias as a key sin of human memory where beliefs and knowledge can shape how we encode and recollect information. For instance, bias in favor of a consistent identity can systematically shape how we recall our pasts (Schacter, 1999; 2022a). A similar bias operates on assessments of truth: Identity can color perceptions of what seems right or feels true, with a tendency for us to endorse ideas and people that align with how we see ourselves and our groups (Ecker et al., 2021; Frenda et al., 2013; Hornsey & Fielding, 2017; Murphy et al., 2019; Pennycook & Rand, 2021a; Swire, Berinsky, et al., 2017; Wang et al., 2022). This identity-oriented bias can make us vulnerable to false beliefs, as identity-aligning information may not be scrutinized sufficiently (Oyserman, 2019).

Another sin of memory that may reduce our tendency to scrutinize information is the sin of absentmindedness, where people make cognitive errors due to inattention on a target task
MISINFORMATION AND THE SINS OF MEMORY

(Schacter, 2022b). Schacter highlights that the sin of absentmindedness may be associated with our digital environment where people are increasingly task-switching across devices and between apps (Schacter, 2022a, 2022b). While Schacter notes that the inattention associated with this sin is a concern for retaining information in educational contexts (Wammes et al., 2019), we suggest this is also a concern for the discernment between high- and low-quality information. Indeed, a lack of analytical, deliberative assessment of truth can contribute to false-belief formation (e.g., Bago et al., 2020). Media multitasking and ongoing absentmindedness may contribute to false belief formation through at least two key routes, by increasing the chances that people’s attention is drawn away from considering the veracity of information, or by increasing the chances that people take a ‘resource-light’ route in their assessment, relying on intuition or gut feelings to infer veracity.

While intuitive impressions allow a rapid and less cognitively demanding assessment of truth, reliance on intuition may lead to false beliefs through the sin of misattribution—drawing on metacognitive cues that can arise out of prior exposure to infer truth. Schacter (2022b) outlines one well-documented example that is highly relevant for misinformation: People judge repeated claims as more likely to be true, a phenomenon called the illusory truth effect (ITE; Hasher et al., 1977). The ITE is thought to occur because people use feelings such as information familiarity, cognitive fluency, and coherence—all of which can emerge as the result of repetition—as evidence that a claim is valid (Unkelbach et al., 2019; Wang et al., 2016). While feelings of familiarity, cognitive fluency, and coherence can serve as rational cues to truth in some social contexts, these feelings can also arise through repetition driven by disinformation campaigns or algorithms and bots on online platforms and can thus be a source of distortion (e.g., Unkelbach & Koch, 2019). The reliance on such repetition cues is insidious and as
MISINFORMATION AND THE SINS OF MEMORY

Schacter alludes, continues to persist even when people have more diagnostic information that they could draw on to assess truth. Indeed, people are inclined to believe repeated claims relative to new claims even when they have general knowledge contradicting the claim (Fazio et al., 2015; Fazio, 2020b), and even when the claim is shared by an untrustworthy source or declared false by a highly accurate source (Henkel & Mattson, 2011; Unkelbach & Greifeneder, 2018). While Schacter’s analysis highlights the distorting effects of repetition, in the absence of repetition, other factors entirely unrelated to truth can also produce feelings of ease of processing, which people routinely draw on to infer truth. For example, sources that are easy to pronounce are rated as more trustworthy and likable than those difficult to pronounce (Laham et al., 2012; Newman et al., 2014; Silva et al., 2017). Familiar faces are rated more favorably and their statements as more convincing (Brown et al., 2002; Weisbuch & Mackie, 2009). Speakers who have clear audio are rated as more credible, reliable and trustworthy (Bild et al., 2021; Newman & Schwarz, 2018). While people are sensitive to the experience of easy processing, they are less sensitive to the origin of the fluency (Jacoby et al., 1989; Schwarz 2015). Thus, people often use fluency as an informative cue to truth even when the real source of fluent processing (e.g., pronunciation ease) is tangential and non-diagnostic.

A lack of skepticism may further contribute to errors in assessments of veracity. Indeed, people have a default to assume incoming information is correct (Gilbert et al., 1993; Grice, 1975; Schwarz et al., 2007). However this assumption of truth can be disrupted, which can benefit decision making in this domain. Explicit warnings about the presence of false information or prompts to think about accuracy can help people engage in more critical analysis, increase truth discernment, and reducing the impact of fluency on judgment (Bago et al., 2020; Brashier et al., 2020; Jalbert et al., 2020). Schacter (2022b) highlights how this might work in
MISINFORMATION AND THE SINS OF MEMORY

subtle ways—by simply presenting claims as questions rather than statements (see Cavillo & Harris, 2022). When people are engaged in more analytical processing or skeptical analysis, they are more likely to identify logical flaws in a claim and are less inclined to apply the default of believing incoming information (Lee et al., 2015; Mayo, 2019). While more analytical, deliberative processing can enhance our ability to discern facts from fiction, the online environment where people often seek information can present significant hurdles that make us vulnerable to false beliefs.

As Schacter (2022b) points out, one can encounter misinformation that has all the features we tend to associate with true information. Even if engaged in optimally skeptical, analytical processing, our internal fact-checks are only helpful to the degree that false information violates the criteria we apply. Deepfakes and fake news may make typical cues to truth less diagnostic, having the potential to lull us into error (Murphy & Flynn, 2022; Nightingale & Farid, 2022). People can also miss or underutilize cues to falsity in digital environments that should suggest cause for hesitation (Dias et al., 2020; Nightingale et al., 2017). Moreover, the contemporary information environment facilitates repeated exposure to false ideas, which can not only lead to the abovementioned illusory truth effects but also give rise to misperceptions of social consensus—for example, minority opinions that are prominent on social media can lead people to assume that their own (majority) opinion is less common, a case of pluralistic ignorance (Lewandowsky et al., 2021; Shamir & Shamir, 1997; see also Weaver et al., 2007). Perhaps even more challenging for discernment is that content is customized for us online, exploiting detailed knowledge about individuals to create appealing messages via microtargeting (Acerbi, 2019; Kozyreva et al., 2020).
MISINFORMATION AND THE SINS OF MEMORY

Tasked with evaluating a large volume of content in our day-to-day environments, false beliefs can be acquired for many of the reasons outlined above. Once a false belief is acquired, to what extent can it be revised? The sins of memory also bear directly on the extent to which corrections hold and belief revision occurs.

The Continued Influence of Misinformation and Implications for Corrections

Once people have formed false beliefs and consider misinformation to be true, revision becomes a challenging task. The sins of memory may help explain the underlying difficulty. One of the main issues with misinformation is that people often continue to rely on misinformation in their reasoning even after it has been corrected. This is known as the continued influence effect (Ecker, Lewandowsky, et al., 2022; Johnson & Seifert, 1994) and can be observed even when people understand, believe, and remember a correction. For example, the Australian government, segments of the media, and social-media groups repeatedly blamed arson for the catastrophic Australian “Black Summer” wildfires in 2019/2020; although these claims were subsequently debunked, some continued to falsely believe that arson played a significant causal role in those fires, contributing to increased polarization about climate change (Mocatta & Hawley, 2020; Weber et al., 2020).

An information-deficit view of communication would argue that the solution to the problem of a false belief is simple: provide a clear, coherent account of the truth, allowing people to update. However, people do not always adjust their knowledge and beliefs even when a correction is compelling—at least not to the extent that the influence of the misinformation is fully mitigated. It seems tempting to attribute this phenomenon to the sin of bias. Indeed, people’s identity and worldview (e.g., their political attitudes) can shape their post-correction reliance on identity- or worldview-related misinformation (Hornsey & Fielding, 2017; Trevors,
MISINFORMATION AND THE SINS OF MEMORY

2022; Trevors & Duffy, 2020). However, the existing evidence is not as strong as one may think. On the one hand, there have been claims that corrections of worldview-congruent misinformation can be ineffective (Ecker & Ang, 2019) or can even backfire (Nyhan & Reifler, 2010). Likewise, a recent meta-analysis suggested that corrections are more successful when the corrected misinformation is worldview-discordant (Walter & Tukachinsky, 2020). On the other hand, findings of ineffective and backfiring corrections have proven difficult to replicate (Ecker et al., 2021; Guess & Coppock, 2020; Wood & Porter, 2019), and accumulating evidence suggests that people by-and-large adjust their beliefs at least somewhat when confronted with counterevidence, even if the misinformation is worldview-congruent (Aird et al., 2018; Ecker, Sanderson, et al., 2022; Nyhan et al., 2020; Swire, Berinsky, et al., 2017; Weeks, 2015).

Moreover, the continued influence of false information also occurs with information that is worldview-neutral. As such, it has been linked to memory lapses, namely failures to retrieve corrective information (i.e., selective retrieval of misinformation) or failures to integrate corrective information into one's mental model or causal account of a situation (Ecker et al., 2022). From this perspective, continued influence is most immediately related to the sin of persistence. Although Schacter (2022b) conceptualizes persistence as “unwanted and emotionally arousing intrusive memories, typically resulting from traumatic events”, we suggest broadening this definition such that false but non-disturbing information and non-intrusive recall are included. As a “treatment” for persistence, Schacter suggests reconsolidation. This refers to a theory that retrieval of a consolidated memory can transfer the memory representation into a labile state that allows it to be updated (Lee et al., 2017). Although Schacter applies this to intrusive memories, “labilization” may also help overcome continued influence. Prior research has found that repeating misinformation directly prior to or within a correction can render the
correction more effective (Ecker et al., 2017; Kowalski & Taylor, 2017; Wahlheim et al., 2020). However, it should be noted that this does not require the assumption of reconsolidation (also see Allanson & Ecker, 2017; Howe et al., 2020). For example, an alternative model suggests that co-activation of a misconception and related corrective information facilitates conflict detection and integration of the corrective information into the relevant mental model, thus boosting knowledge revision (Kendeou & O’Brien, 2014).

Even if people update their belief successfully after misinformation has been corrected, this change is rarely sustained over time (Carey et al., 2022). In other words, belief frequently returns towards pre-correction levels (Kowalski & Taylor, 2017; Paynter et al., 2019; Swire, Ecker, et al., 2017); this is called belief regression. Swire-Thompson et al. (2022) found memory for the correction explained 66% of the variance in belief regression after correcting for measurement reliability. We can conceptualize this into two distinct memory sins: The sin of transience would result in people forgetting that the misinformation has been corrected, leaving them unsure of whether the claim is true or false. By contrast, the sin of misattribution would be involved when people misremember misinformation as having been presented as true. Swire-Thompson et al. (2022) found evidence for both transience and misattribution occurring within the context of belief regression. The proportion of people who believed in the misinformation but were unsure of whether it was presented as true or false (committing the sin of transience) increased from 0.6% immediately after corrections were presented to 5.6% one month later. The proportion of people who believed in the misinformation and thought it had been presented as true (committing the sin of misattribution) increased from 4.1 to 15.6%.

Again, it seems that the role of the sin of bias is limited in belief regression, with beliefs regressing at a similar rate whether a correction is pro- or counter-attitudinal. Swire, Berinsky, et
al. (2017) found that Democrats, Republican supporters of Donald Trump, and Republican non-Trump-supporters all forgot that corrected Trump misinformation was false at a similar rate. In other words, *motivated forgetting*, with a person’s worldview influencing the rate at which a person re-endorse a claim, did not occur. However, this needs to be replicated and extended.

It is tempting to suggest that continued belief in or reliance on corrected misinformation is an irrational expression of a flawed cognitive system. However, in line with Schacter’s (2022b) perspective, one can also entertain a functional account of continued influence. Viewed through a broader lens, effective belief updating is a cognitive challenge arising in part from a core conundrum of memory—namely, balancing the need for stable representations and the need to allow for some representational flexibility to cope with an ever-changing world (Ecker et al., 2014). In the context of misinformation, maintaining a stable representation of corrected misinformation can be functional to the extent that the claim may end up being true after all. Evidence is seldom absolute, and source credibility can rarely be assessed with full confidence; in some situations it can be entirely rational and adaptive to not fully suppress misinformation belief after a correction (Connor Desai et al., 2020). It can also be useful to retain a mental model of an event or causal relation that is known to be false, for example to aid with counterarguing false claims in a debate.

**Future Research Directions**

In considering the formation of false beliefs, the saturated information environment poses a challenge for allocation of attentional resources and analytical assessment of truth. But technology can be leveraged—at least online—to alleviate some of the sins of memory that make us vulnerable to false belief formation (Kozyreva et al., 2020; Lewandowsky et al., 2017). For example, online prompts to consider truth or accuracy may work to reduce absentmindedness
and increase analytical assessment (Fazio, 2020a; Pennycook et al., 2021). While promising, more thorough investigation is warranted to better understand how long such prompts are effective and for which users. Further, emerging research on intellectual humility—one’s willingness to engage in self-reflection and be open to considering disconfirming evidence—may improve our understanding of the sin of bias and may lead to innovations that encourage less identity-oriented assessment of truth (Bowes et al., 2020). Another emerging direction is that digital-literacy and media-literacy education, as well as pre-emptive inoculation techniques specifically designed to boost resilience against misinformation, may equip people with a toolset for more skeptical analysis of content (Cook et al., 2022; Guess et al., 2020; Lewandowsky & van der Linden, 2021; see Modirrousta-Galian et al., 2022).

Considering the continued influence effect through the lens of the sins of memory may also foster new lines of research. For instance, Schacter (2022b) suggests that persistence may be more prevalent in people with lower levels of cognitive control. This meshes well with research suggesting that people’s susceptibility to continued influence relates to their working memory function (Brydges et al., 2018). However, more recent research suggests that people’s episodic memory may be even more important when it comes to continued influence (Sanderson et al., 2021), which highlights the need for future research to further investigate what memory does and does not explain with regards to the continued influence effect. For example, another important factor of the continued influence effect is whether or not a person believes the correction to be valid. O’Rear & Radvansky (2020) found that a large proportion of participants who remembered the correction did not think that it was accurate or genuine, and consequently relied on the misinformation in their reasoning at a similar rate as individuals who never received a correction at all.
Conclusion

There is a clear association between Schacter’s (2022b) sins of memory and (1) the formation of false beliefs and (2) the continued influence effect after misinformation is corrected. This highlights the importance of considering memory in future research, in particular the development of interventions and the measurement of efficacy over time. For instance, interventions that address misinformation may benefit from making memory sins a key focus—reducing bias, absentmindedness, and possible misattributions. One avenue that has been explored to counteract these sins in the continued-influence domain is correction repetition and ensuring a correction can be processed with undivided attention (Ecker et al., 2011; Sanderson et al., 2022). Future research should further explore additional avenues, such as making corrections more distinctive (e.g., via mental imagery). Beyond corrections, interventions will need to be memorable to maximize their long-term efficacy (such as inoculation Maertens et al. 2021; also see Schwarz et al., 2016). In sum, memory is fundamental to belief formation and the correction of information, and the conceptualization of Schacter’s (2022b) sins can be a helpful paradigm for misinformation researchers going forward.
References


Cook, J. (2019). Understanding and countering misinformation about climate change. In Handbook of research on deception, fake news, and misinformation online (pp. 281–
MISINFORMATION AND THE SINS OF MEMORY


Fazio, L. K. (2020a). Pausing to consider why a headline is true or false can help reduce the sharing of false news. *Harvard Kennedy School Misinformation Review, 1*(2). https://doi.org/10.37016/mr-2020-009


MISINFORMATION AND THE SINS OF MEMORY


MISINFORMATION AND THE SINS OF MEMORY

Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences (pp. 353–377). The MIT Press.


MISINFORMATION AND THE SINS OF MEMORY


MISINFORMATION AND THE SINS OF MEMORY


https://doi.org/10.1016/j.cognition.2022.105276


https://doi.org/10.1080/0163853X.2021.1917950


Walter, N., & Tukachinsky, R. (2020). A Meta-Analytic Examination of the continued influence of misinformation in the face of correction: How powerful is it, Why does it happen,
and how to stop it. Communication Research, 47(2), 155–177.
https://doi.org/10.1177/0093650219854600

https://doi.org/10.1016/j.compedu.2018.12.007

https://doi.org/10.1002/ejsp.2909


MISINFORMATION AND THE SINS OF MEMORY


