8

FALSE BELIEFS

Byproducts of an adaptive knowledge base?

Elizabeth J. Marsh and Matthew L. Stanley

Pizzagate. The Bowling Green Massacre. Pope Francis endorsing Donald Trump for president. Such “fake news” stories often go viral, hijacking typical cues for credibility while spinning fictions about familiar people and places. It is almost impossible to completely eliminate exposure to factually incorrect information (misinformation), especially as such stories rarely (if ever) really disappear, living on through the internet (Shin, Jian, Driscoll, & Bar, 2018). Although some stories were never meant to be believed at face value (e.g., satires from *The Onion*), others are concerning because they serve as highly effective propaganda tools (e.g., the stories attributed to Russian Bots on Twitter and Facebook). Oftentimes, they play on emotions and motivations by scaring viewers with, for example, images of immigrants crossing the southern border or false information about the future implementation of bureaucratic “death panels” with the power to terminate an individual’s medical care.

But this “hot cognition” occurs within a system tuned to learn, maintain, and update knowledge and beliefs about the world. The likelihood that we come to believe fake news stories (and our ability to correct false beliefs) is dependent upon everyday cognitive processes underlying knowledge and belief, and not upon unique, special processes or strategies (Isberner & Richter, 2014; Marsh, Cantor, & Brashier, 2016; Rapp & Donovan, 2017; Rapp & Salovich, 2018). In this chapter, we begin by briefly describing the knowledge base, defining its contents and comparing it to other forms of memory. We then describe five properties of the knowledge base that often make it efficient and powerful – but which sometimes yield false beliefs, with implications for why people sometimes believe fake news stories. In the last section of the chapter, we summarize potentially valuable strategies for combatting the problem of fake news.

The knowledge base

Humans know a lot, and this knowledge can take many forms, from discrete facts (e.g., knowing the capital of Peru) to scripts (abstractions of past experience, such as a script for navigating airport security) to understanding complex systems (e.g., how the economy works, and how changing interest rates will have downstream implications for other parts of the system). It is accumulated over a lifetime, sometimes formally (in school, through deliberate instruction) and other times informally (through trial-and-error, exposure to media, interpersonal interactions, etc.).

In psychological science, knowledge is often defined by what it is not: It is *not* thinking back to a specific event that occurred at a particular time and place in the past. The distinction between knowledge and retrieving specific events (episodic memory) is perhaps best captured with examples. An individual might know what a prom is, for example, but might recollect a specific episode that involves receiving a pink corsage for her own senior prom. An individual might know the steps involved in taking a typical airplane flight, or that individual might think back to a particularly bumpy flight from Columbus to St. Louis. An individual might know who Abraham Lincoln is, without remembering having learned about him in a specific lesson in a particular classroom.

We suggest that combatting errors in the knowledge system (false beliefs, illusions of knowledge, misconceptions) may require different strategies than the ones that are sometimes recommended to combat false memories of specific events. Correcting someone's false beliefs about the purported link between vaccines and autism, for example, differs from helping a person remember a specific event accurately and completely. Some of the current suggestions for combatting fake news are based on the large literature on how people misremember specific events from their pasts (e.g., strategies for thinking about the particular person, or source, who presented the information at a particular time and place, and whether that person is trustworthy or has adequate expertise; Lewandowsky, Ecker, & Cook, 2017; Rapp & Salovich, 2018). These strategies might work in some contexts, but we suggest that a different set of strategies might be better suited for correcting false beliefs instilled by fake news.¹

Consider the so-called "birtherism" movement in which many people came to falsely believe that Barack Obama was born *outside* of the United States. Of course, it is possible that people have false memories about a specific event in which they heard the claim that Barack Obama was (or was not) born outside of the United States. However, we suggest that whether people can accurately and completely recollect the event in which they were exposed to this claim matters less than whether they currently hold the belief that the statement "Barack Obama was born outside of the United States" is true. It is the false belief that needs to be corrected, not one's memory of when one heard it. This suggests that we should look to the literature on belief correction rather than to the literature aimed at helping people to accurately remember specific past events.

In short, in this chapter we briefly describe several properties of the knowledge base that have implications for processing fake news, before turning to implications for combatting the problem.

Principle #1

We are biased to believe information is true. All else being equal, upon initially encountering information, people are biased to assume that it is true rather than false. Gilbert (1991) argues that this bias is cognitively efficient, and it means we only need to tag information when it is false (and otherwise we can assume it is true). This bias allegedly reflects a disposition to believe in a generally truthful world, where any given piece of information is more likely to be true than false. But in this model, tagging a piece of information as false is effortful and requires a second step of unbelieving, setting up a problem if the process of unbelieving is disrupted.

Multiple studies support this prediction, namely that interrupted falsehoods are later remembered as true (e.g., Gilbert, Tafarodi, & Malone, 1993). In one study (Gilbert, Krull, & Malone, 1990), participants simulated learning a foreign language in an immersion setting that required them to routinely guess the meanings of words before receiving feedback. Participants read a series of Hopi-English translations, each of which was immediately followed by a true or false label; critically, a subset of trials were interrupted with a tone that required participants to press a key. Later, participants took a true/false test. As predicted, interrupted false trials were often later misremembered as true, consistent with the idea that the tone interfered with the additional, effortful process of “unbelieving” the false translations.

A second example from Gilbert and colleagues (1990) involved participants viewing videos and judging whether smiles were real or fake. After seeing each face, participants were explicitly told whether the smile was real or fake. On some of the trials, cognitive load was induced by requiring participants to respond as quickly as possible to a tone or to identify which of two tones was higher or lower in pitch. During the subsequent test phase, participants were more likely to make errors and claim that individuals in the videos had expressed real smiles if the real/fake label had been paired with the secondary task. As with the Hopi-English translations, the explanation was that increased cognitive load resulted in the effortful unbelieving stage being bypassed. Consequently, judgments on the test reflected only the initial belief-assenting stage (see also, Mandelbaum, 2014).

If Gilbert and colleagues are correct, then people might generally be biased to believe in headlines like “Clint Eastwood Refuses to Accept Presidential Medal of Freedom from Obama, Says He Is Not My President” with unbelieving only occurring after a second, effortful step.² The implication is that that people who are distracted or otherwise busy will be less likely to “unbelieve” false information. Lending credence to Gilbert and colleagues’ view, people who are being “cognitively lazy” (i.e., fail to engage more reflective, effortful reasoning

processes when encountering new information) are more likely to believe in the veracity of relatively implausible fake news headlines, whereas those who tend to engage more reflective, effortful reasoning processes are less likely to believe in those same implausible fake news headlines (Bronstein, Pennycook, Bear, Rand, & Cannon, 2019; Pennycook & Rand, 2019). That is, people who gave the obvious but incorrect answer to questions such as “How many cubic feet of dirt are there in a hole that is three feet deep by three feet wide by three feet long?” were also more likely to accept the fake news headlines (note: there is no dirt in a hole). Relatedly, those individuals who are more likely to engage reflective, effortful reasoning processes when encountering new information are more likely to reject conspiracy theories than those who fail to engage those same processes (Swami, Voracek, Stieger, Tran, & Furnham, 2014).

One caveat to Principle #1 must be made: This effect occurs when people are forced to choose between truth and falseness. When remaining uncertain is an option, the effect disappears (Street & Richardson, 2015; see also, Street & Kingstone, 2017). In other words, Principle #1 describes a heuristic that is used in a specific situation. When a true/false judgment must be made, the heuristic holds because people assume that most people tell the truth most of the time (DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996).

Principle #2: we use shortcuts to judge truth

More generally, people use heuristics to judge the truth of incoming information, judging easy-to-process information as more likely to be true than harder-to-process information. That is, statements that are easy to process (to read, to hear, to understand) are more likely to be judged as true relative to claims that are harder to process. Consequently, statements written in an easy-to-read font are more likely to be judged as true than those written in more difficult-to-read fonts (Reber & Schwarz, 1999). Relatedly, rhyming aphorisms (*woes unite foes*) are judged as more likely to be true than ones that do not rhyme (“fools live poor to die rich”; McGlone, & Tofigbakhsh, 2000).

Repetition is perhaps the easiest way to make some statements feel easier to process than others, as a statement will be read more easily (more quickly) the second time it is seen. Experimentally, this is typically demonstrated by having participants read a series of trivia statements in the first phase and then make truth ratings about a larger set of statements in the second phase (commonly one-half old and one-half new). Repeated statements are judged as more likely to be true than new statements, a finding called the *illusory truth effect*. Originally demonstrated in the 1970s (Hasher, Goldstein, & Toppino, 1977), this effect has been replicated many times, is robust across procedural variations and materials, and can last for many weeks (see Dechêne, Stahl, Hansen, & Wänke, 2010, for a meta-analysis).

The illusory truth effect is not merely a laboratory phenomenon nor is it restricted to obscure trivia statements: The effect has been obtained for fake

news headlines that were actually seen on Facebook (Pennycook, Cannon, & Rand, 2018). Pennycook and colleagues (2018) showed that just a single, brief exposure to a fake news headline (e.g., ‘Donald Trump Sent His Own Plane to Transport 200 Stranded Marines’) increased the likelihood that people believed the headline was true. This boost in the perceived accuracy of the fake news headlines persisted after a week-long delay, and additional exposures further increased belief in the veracity of the headlines. Surprisingly, these same effects were obtained even when the participants were told that the articles were disputed by an independent group of fact checkers. This finding suggests that labeling information as false or disputed (as is done by third-party fact-checking organizations like PolitiFact and Snopes) may not be using the most effective means for combatting fake news and misinformation, a point we return to later in this chapter.

Principle #3: we interpret incoming information

Fake news stories are interesting because they are typically about real people and places, like Hillary Rodham Clinton and Russia; they are rarely about people or places we have never heard of before. The familiar actors and places mean we draw on our prior knowledge when processing fake news stories, making them easy to understand and to elaborate upon. Furthermore, we have a script for news stories – a set of expectations about what kinds of information they will typically contain: who, what, when, where, why, and a little human interest. We have some idea about what to expect when reading this type of story.

Because our default goal is normally to extract meaning and to fit new information into our existing schemas and belief structures, we do not store information verbatim, and we seldom remember the exact wording of conversations or form of other communications. For example, the sentence “the karate expert *hit* the block” is misremembered as “the karate expert *broke* the block” (Brewer, 1977). A similar effect occurs in the visual domain; people falsely remember having seen causal actions (e.g., a foot hitting a ball) that were only implied by the surrounding video footage (Strickland & Keil, 2011).

Furthermore, we are wired to go beyond what is given, to make inferences, and to fill in the gaps. In other words, the system is cognitively efficient, leveraging the past to prepare for the future. We learn things that are not explicitly stated. For example, young children do not need to be explicitly told that “dolphins communicate by squeaking” after learning that “dolphins travel in pods” and that “pods communicate by squeaking” (Bauer & San Souci, 2010). We transfer our knowledge to new situations and to new instances. For example, identifying the artist of a painting we are seeing for the first time, based on our past experiences with that artist’s work (Kornell & Bjork, 2008).

Our ability to interpret allows us to handle imperfect inputs, so long as they are “close enough”. Speech is notoriously disfluent – we stutter, speak ungrammatically, and use ambiguous referents. Accordingly, the system is designed to

handle imperfect inputs, allowing a certain amount of “deviation” in messaging (Ferreira, Bailey, & Ferraro, 2002). A word is recognized in many different fonts, accents, contexts – such flexibility is advantageous as we encounter new variations of stored information. On the other hand, a flexible system has the side effect of missing inaccuracies that are “close enough”. For example, people often fail to detect incorrect presuppositions in questions, even when warned to do so. Many people will attempt to provide a numerical answer to the question “How many animals of each kind did Moses take on the ark?” even when warned about this type of trick, and later demonstrate knowledge that the reference should be to Noah, not to Moses (the ‘Moses Illusion’; Bottoms, Eslick, & Marsh, 2010; Erickson & Mattson, 1981).

Many different strategies have been implemented in efforts to reduce or eliminate the “Moses Illusion” and related phenomena. Ironically, however, many of these implemented strategies backfire, making people even more likely to remember or rely upon the misinformation. For example, Eslick, Fazio, and Marsh (2011) attempted to draw participants’ attention to false information in claims by highlighting it in red font. But this only made participants more likely to repeat those errors on a later test. One strategy that does seem to reduce the influence of misinformation, however, is to encourage individuals to play the role of a professional fact checker when reading new information (Rapp, Hinze, Kohlhepp, & Ryskin, 2014; Brashier, Eliseev, & Marsh, 2020). Doing so makes people more likely to identify misinformation and less likely to rely on that misinformation later on.

Principle #4: knowledge feels different

The phenomenology of retrieving knowledge is very different from what one experiences when retrieving a personal memory (which will involve mentally traveling back in time and place to a specific event). Knowledge “pops to mind”, and is labeled as “known” rather than “remembered” (Tulving, 1972). When retrieving the fact that George Washington was the first president of the United States, for example, most people do not think back to a specific instance of learning that information; instead, they simply know it. Of course, there can be exceptions, where one remembers the circumstances of learning information (e.g., Conway, Gardiner, Perfect, Anderson, & Cohen, 1997), but in most cases the things we know are relatively decontextualized from original learning. While people are less likely to be persuaded by sources perceived as untrustworthy (Guillory & Geraci, 2013; Mills & Jellison, 1967; Priester & Petty, 1995), the source of information is typically forgotten over even relatively brief periods of time (e.g., Underwood & Pezdek, 1998) and as one encounters information in multiple different contexts (Watkins & Kerkar, 1985).

At the moment, a popular suggestion for dealing with fake news involves training people to better evaluate and remember sources of information (Rapp &

Salovich, 2018; Swire & Ecker, 2018). This advice is based on the large literature on how people misremember the past, which can often be traced to problems in *source monitoring* (Johnson, Hashtroudi, & Lindsay, 1993): a presupposition embedded in a question is misattributed to experience; a statement in a narrative is misattributed to a video; or a photo provides details that are incorporated into one's memory. The natural solution is to focus people's attention on the origin of their memories, either at the time of learning (e.g., discounting information from a low credibility source) or specifically discriminating sources at test (Lindsay & Johnson, 1989). This approach is endorsed by professional organizations such as IFLA (the International Federation of Library Associations and Institutions), which has issued a number of recommendations for spotting fake news, including researching the quality of the hosting website and the credibility of the author.

But source monitoring is unlikely to solve the problem that fake news poses (Marsh & Yang, 2017). Aside from the fact that people do not always agree about whether a source is credible (e.g., Fox News is more credible to some than others), source credibility is often inferred from cues that can potentially mislead, such as the number of citations included (Putnam & Phelps, 2017) or the ease with which a message is read (Fragale & Heath, 2004). In fact, ease of processing information is typically more likely to influence truth judgments than is the source of the information: Statements that are initially labeled as false or obtained from an unreliable source are nevertheless more likely to be judged as true relative to novel statements (Begg & Armour, 1991; Begg, Anas, & Fari-nacci, 1992; Stanley, Yang, & Marsh, 2019).

Even if sources are correctly evaluated and that information is stored in memory, it is likely to have no impact on later judgments, as our bias toward cognitive efficiency means that our default is to use information without spontaneously thinking about where it came from (reflecting the fact that much of what we know is decontextualized, Tulving, 1972). The result is that people may be able to identify statements as having come from a reliable or unreliable source (two weeks after exposure, if explicitly prompted), and yet fail to apply that knowledge when asked to judge the truth of the statements (Henkel & Mattson, 2011).

Furthermore, because information often is associated with multiple sources, remembering that something came from a low credibility source does not necessarily mean it should be ignored (Marsh, Meade, & Roediger, 2003). That is, it is possible that Donald Trump made the same claim as Nancy Pelosi, and thus, the low credibility source is ignored given the memory of (or misattribution to) the higher credibility source. Third, a major problem is that source information simply isn't well remembered over time. This is the problem captured in the classic "sleeper effect" – a low credibility source is not persuasive initially, but over time the message has more impact as the source information is forgotten (e.g., Underwood & Pezdek, 1998).

Principle #5: we do not always retrieve and/or apply what we know

People often think of knowledge as something that is always used if stored in memory, but knowing something does not guarantee that it will come to mind, or be applied, when needed. The most common example of this phenomena is the tip-of-the-tongue state (TOT state; see Brown, 1991). Most commonly encountered when trying to remember specific terms or names, a TOT state involves the frustrating feeling that what we know is “just out of reach”. When in a TOT state, one is hyper-aware of the problem; however, at other times the rememberer may not be aware that they are having a retrieval failure. More generally, information that has not been used recently is less likely to be spontaneously retrieved (Cantor, Eslick, Marsh, Bjork, & Bjork, 2015).

We already gave one example of the general principle that knowledge isn’t always utilized, noting how our system accepts inputs that are “close enough” even when we know better (e.g., failing to notice the error in the question “How many animals of each kind did Moses take on the ark?”). This problem occurs even with well-learned information. Cantor and Marsh (2017) created biology (e.g., “Water contains two atoms of Helium and how many atoms of oxygen?”) and history questions with incorrect presuppositions (“The British congress imposed fees on the colonies’ sugar, tea, and stamps that were called what?”) and posed them to PhD students in biology and history. All participants were warned not to answer questions with incorrect presuppositions, as illustrated in an example. These disciplinary experts performed slightly better in their thesis domain, but they still missed about 30% of the incorrect presuppositions in questions that targeted their field of graduate study (Cantor & Marsh, 2017).

A second example involves the illusory truth effect; the assumption in the literature has been that this fluency heuristic comes into play in the face of ignorance, when someone does not know better. But recent evidence suggests that this is not always the case. For example, exposure to blatantly false facts (e.g., “The short-pleated skirt worn by Scotsmen is the sari”) increases the likelihood they are later judged true, as compared to new statements. This effect occurs regardless of whether knowledge is estimated based on norms of general knowledge or measured in a given individual (Fazio, Brashier, Payne, & Marsh, 2015). It is not clear how far one can push these effects; initial results suggested a boundary condition, with no illusory truth effect for implausible statements like “The earth is a perfect square” (Pennycook et al., 2018) – but more recent work suggests this finding is due to a floor effect in belief, rather than the nature of the items themselves (Fazio, Rand, & Pennycook, under review).

A third example involves reliance on information clearly labeled as false. In our own studies (Stanley et al., 2019), participants studied qualified statements (e.g., “It is *unlikely* that that the body of a rotten tree is called a daddock” or “It is *impossible* that the study of snakes is called ophiology”). In the critical experiment, all qualifiers were negative, meaning that every single statement

FIGURE 8.1 Average truth rating for statements previously qualified as improbable, unlikely, or impossible, as compared to new statements

was qualified as being impossible, improbable, or unlikely. Two days later, participants rated the truth of those statements without the qualifiers, as well as new statements. Previously qualified statements were rated as more likely to be true than new statements (see Figure 8.1), even though every single one of the original qualifiers cast doubt on the veracity of the statements. When prompted, the majority of subjects correctly identified the qualifiers presented in the first part of the experiment; they knew that the statements were part of a set that was very unlikely to be true, and yet exposure in that set increased the likelihood of calling a negatively qualified statement “true”.

Implications for the fake news problem

Thus far, we have focused on the processing of fake news stories, emphasizing that we are not optimistic about interventions that focus solely on source credibility. While it is appealing to offload the problem to third parties (e.g., to fact checkers like PolitiFact or Snopes) or to algorithms that diagnose content and warn if sources are low credibility, such solutions will likely be incomplete ones. To the extent that such labels prevent people from reading any of the content, they may help; however, once information is processed, a source-based strategy is not likely to help, especially as time passes.

Instead, we focus on the aftermath of fake news, and the problem of correcting the misinformation conveyed by fake news stories. One issue is the distribution of corrections – it cannot be assumed that people see them and process

them. On twitter, at least, the distribution of corrections is quite discouraging. Shin and colleagues analyzed over 400,000 tweets about 57 different political rumors, and reached the following conclusion: “Overall, there were few tweets rejecting any rumor in our dataset, be it true or false” (Shin, Jian, Driscoll, & Bar, 2017, p. 1222). They found that less than 4% of tweets rejected rumors, as opposed to endorsing them. And to the extent that fact checks are shared, they are likely to be ones that confirm one’s own opinion. An examination of Twitter shares in 2012 showed that people were more likely to share fact-checks (e.g., from the Washington Post’s Fact Checker) that matched their political affiliations. That is, Democrats were responsible for sharing fact-checks that favored Obama, whereas fact-checks favoring Romney were more likely to be shared by Republicans (Shin & Thorson, 2017). Finally, we note that even if people encounter the corrections, they may not be successful. Newspaper retractions, for example, do not always stick; some Americans continued to hold false beliefs about the existence of weapons of mass destruction in Iraq, even though they also remembered reading retractions of those claims (Lewandowsky, Stritzke, Oberauer, & Morales, 2005).

One question is whether debunking fake news is as difficult as debunking serious misconceptions (e.g., about the vaccine-autism link). Misconceptions such as the vaccine-autism link are notoriously difficult to correct (Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012), with corrections often losing their effectiveness over time and sometimes even backfiring and increasing belief (Nyhan & Reifler, 2010; cf. Wood & Porter, 2019). It is an open question whether fake news stories are in this category. In one experimental study, for example, the researchers examined belief in statements such as “Anthony Scaramucci subject of Senate Russian Investigation” and “President Obama fakes birth certificate” and then corrected them (Porter, Wood, & Kirby, 2018). Corrections were quite successful – although it is not clear how much participants believed the stories in the first place (participants who never received the corrections were described as ‘credulous’ of the original headlines). Furthermore, long-term effects were not examined, and it is quite possible that corrections would lose their efficacy as time passed (Butler, Fazio, & Marsh, 2011).

We believe the prudent path is to follow the insights generated from the larger literature and assume that debunking is not as easy as it might seem initially. Lewandowsky and colleagues (2012) have outlined a number of principles to follow when debunking; we briefly review those before turning to some additional recommendations based on our own work. The clearest advice is to avoid repeating the fake news claim when debunking it (Lewandowsky et al., 2012). That is, every message that says “Pizzagate is false” is increasing the fluency of the word “Pizzagate”, which is problematic since increasing the fluency of a concept can increase the likelihood of entire statements being labeled as true.

Other recommendations include using simple language that is easy to read and also to add photographs. Photographs are interesting as people show a bias to consider statements paired with photos as truer than statements presented

alone (Newman, Garry, Bernstein, Kantner, & Lindsay, 2012). That is, showing a picture of a turtle paired with the claim “turtles are deaf” increases belief in the veracity of that statement, as opposed to simply reading the statement “turtles are deaf” (perhaps because the photos help readers to generate evidence to support the claim; Newman et al., 2015). It is interesting to note that the Pennycook and colleagues’ finding of an illusory truth effect for fake news paired photos with the fake headlines, because that is the way such headlines typically appear on social media.

In our own recent work, we have identified an asymmetry in the likelihood that corrections induce reliable belief revision. False beliefs are more likely to be corrected when the correction involves instilling a new positive belief, compared to rescinding an existing belief (see also, Lewandowsky et al., 2012). More specifically, people are better at correcting false beliefs when they initially believe in a negated claim like “A poem written for a bride is *not* an epithalamium” but then receive the correction that “A poem for a bride is an epithalamium”. In contrast, false beliefs are more difficult to correct when the initial, false belief is an affirmation (e.g., “Michelangelo’s statue of David is located in Venice”) and the correction takes the form of a negation (e.g., “Michelangelo’s statue of David is *not* located in Venice”). This general strategy of changing beliefs by instilling new beliefs (affirmations) instead of rescinding existing beliefs (negations) may be a particularly promising strategy for combatting fake news and misinformation. In fact, Horne and colleagues (2015) were able to successfully change some individuals’ false beliefs and intentions about vaccinations by providing individuals with information about the dangers of diseases like measles and mumps (i.e., instilling a new, affirmative belief about disease risk) rather than just stating that there is no link between vaccines and autism (i.e., belief negation).

Conclusions and future directions

Humans are capable of learning an impressive amount of information and can retain much of it over very long retention intervals. In this chapter, we have described how this system has properties that facilitate the acquisition of new information and allow inferences and transfer to new situations. For example, the system is flexible, both in terms of inputs (accommodating noise) and outputs (going beyond encountered information to generate new insights). The system is cognitively efficient, with shortcuts that reflect the state of the world (e.g., bias to assume truth, using fluency as a heuristic to judge truth). Finally, the representations tend to be decontextualized (sourceless, retrieved without a sense of reliving) in a way that facilitates their usage across contexts and applications.

But as described here, the properties that make the system powerful also facilitate the learning of misinformation from fake news. Upon encountering a fake news source, people may or may not recognize the source as low credibility. Once reading commences, people are often biased to believe incoming

information is true; in fact, they may have to believe it to comprehend it (and take a second step to unbelieve it). Readers of fake news will draw on a wealth of prior knowledge when processing the article, likely generating inferences that go beyond the stated text. And people may repeatedly encounter the same claims, increasing their fluency and, thus, the likelihood that they believe the claims are true.

These issues are particularly problematic today, which some have dubbed the “post-truth” era (Lewandowsky, Ecker, & Cook, 2017) due to the sheer amount of misinformation being circulated widely and quickly through social media platforms. These social media platforms can foster the creation of cliques, or “echo chambers” comprised of people who come to hold the same (often false) beliefs and who collectively seek out and share information that reinforces their shared beliefs (i.e., “echo chambers”). Such echo chambers have been identified on Twitter (Barberá, Jost, Nagler, Tucker, & Bonneau, 2015; Williams, McMurray, Kurz, & Lambert, 2015), Facebook (Del Vicario, et al., 2016), and various blogs (Suhay, Blackwell, Roche, & Bruggeman, 2015). Members may not realize they are receiving information from low-credibility sources; they are biased to believe that information from other community members is true. And when the same claims and information are repeated without correction, belief in those claims is likely to increase, and potentially spawn new misinformation as people make inferences that go beyond what has already been provided, thereby generating even more misinformation.

It is likely impossible to stop the spread of and belief in fake news; instead, we are forced to develop viable strategies for correcting this misinformation. Such efforts should not focus on the source of the original misinformation. Instead, our recommendations focus on the content of the messaging: The content should be simple and accompanied by pictures (to increase fluency), should avoid qualifying claims, and should affirm correct information rather than negating the myth (and more generally the myth should not be stated). Furthermore, the propagation of retractions/debunking messages must be considered, given that people seem much more interested in the initial claims than the corrections.

In this chapter, we argued that combatting errors in the knowledge system (false beliefs, illusions of knowledge, misconceptions) may require different strategies than those recommended to prevent the misremembering of specific events. There may, however, be viable approaches to correction that go beyond the focus on source that also make use of episodic memory. For example, a particularly memorable event (e.g., an emotional, charged press conference that is witnessed in person) in which some false claim is corrected (or retracted) might be more likely to instill lasting belief change than a less memorable event that involves the correction. In this way, knowledge and episodic memory might *interact* in certain contexts to increase the chances that a correction sticks over time. Whether this kind of interaction is useful for making corrections stick over time warrants further research.

Notes

- 1 Of course, it is possible for a false belief to be driven by a false memory, as demonstrated in Polage (2012); but we believe this is much less likely than illusions of knowledge.
- 2 Although controversial, more recent work argues for a different position, namely that validation may sometimes accompany comprehension (Richter, 2015; Richter, Schroeder, & Wöhrmann, 2009). On this view, validating incoming information tends to be a relatively automatic and effortless process that occurs with comprehension, instead of a separate, effortful step after comprehension.

References

- Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science*, 26(10), 1531–1542.
- Bauer, P. J., & San Souci, P. (2010). Going beyond the facts: Young children extend knowledge by integrating episodes. *Journal of Experimental Child Psychology*, 107, 452–465.
- Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General*, 121(4), 446.
- Begg, I. M., & Armour, V. (1991). Repetition and the ring of truth: Biasing comments. *Canadian Journal of Behavioural Science*, 23(2), 195.
- Bottoms, H. C., Eslick, A. N., & Marsh, E. J. (2010). Memory and the Moses illusion: Failures to detect contradictions with stored knowledge yield negative memorial consequences. *Memory*, 18(6), 670–678.
- Brashier, N. M., Eliseev, E. D., & Marsh, E. J. (2020). An initial accuracy focus prevents illusory truth. *Cognition*, 194, 104054.
- Brewer, W. F. (1977). Memory for the pragmatic implications of sentences. *Memory & Cognition*, 5, 673–678.
- Bronstein, M. V., Pennycook, G., Bear, A., Rand, D. G., & Cannon, T. D. (2019). Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking. *Journal of Applied Research in Memory and Cognition*, 8(1), 108–117.
- Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin*, 109, 204–223.
- Butler, A. C., Fazio, L. K., & Marsh, E. J. (2011). The hypercorrection effect persists over a week, but high-confidence errors return. *Psychonomic Bulletin & Review*, 18, 1238–1244.
- Cantor, A. D., Eslick, A. N., Marsh, E. J., Bjork, R. A., & Bjork, E. L. (2015). Multiple-choice tests stabilize access to marginal knowledge. *Memory & Cognition*, 43, 193–205.
- Cantor, A. D., & Marsh, E. J. (2017). Expertise effects in the Moses illusion: Detecting contradictions with stored knowledge. *Memory*, 25, 220–230.
- Conway, M. A., Gardiner, J. M., Perfect, T. J., Anderson, S. J., & Cohen, G. M. (1997). Changes in memory awareness during learning: The acquisition of knowledge by psychology undergraduates. *Journal of Experimental Psychology: General*, 126, 393–413.
- Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review*, 14, 238–257.

- Del Vicario, M., Vivaldo, G., Bessi, A., Zollo, F., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). Echo chambers: Emotional contagion and group polarization on facebook. *Scientific Reports*, 6, 37825.
- DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology*, 70(5), 979–995.
- Erickson, T. D., & Mattson, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning & Verbal Behavior*, 20, 540–551.
- Eslick, A. N., Fazio, L. K., & Marsh, E. J. (2011). Ironic effects of drawing attention to story errors. *Memory*, 19(2), 184–191.
- Fazio, L. K., Brashier, N. M., Payne, B. K., & Marsh, E. J. (2015). Knowledge does not protect against illusory truth. *Journal of Experimental Psychology: General*, 144(5), 993–1002.
- Fazio, L. K., Rand, D. G., & Pennycook, G. (under review). Repetition increases perceived truth equally for plausible and implausible statements.
- Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, 11, 11–15.
- Fragale, A. R., & Heath, C. (2004). Evolving informational credentials: The (mis) attribution of believable facts to credible sources. *Personality and Social Psychology Bulletin*, 30(2), 225–236.
- Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, 46, 107–119.
- Gilbert, D. T., Krull, D. S., & Malone, P. S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology*, 59, 601–613.
- Gilbert, D. T., Tafarodi, R. W., & Malone, P. S. (1993). You can't not believe everything you read. *Journal of Personality and Social Psychology*, 65, 221–233.
- Guillory, J. J., & Geraci, L. (2013). Correcting erroneous inferences in memory: The role of source credibility. *Journal of Applied Research in Memory and Cognition*, 2(4), 201–209.
- Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior*, 16, 107–112.
- Henkel, L. A., & Mattson, M. E. (2011). Reading is believing: The truth effect and source credibility. *Consciousness and Cognition*, 20(4), 1705–1721.
- Horne, Z., Powell, D., Hummel, J. E., & Holyoak, K. J. (2015). Countering antivaccination attitudes. *Proceedings of the National Academy of Sciences*, 112(33), 10321–10324.
- Isberner, M. B., & Richter, T. (2014). Does validation during language comprehension depend on an evaluative mindset? *Discourse Processes*, 51(1–2), 7–25.
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, 114, 3–28.
- Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the “enemy of induction”? *Psychological Science*, 19(6), 585–592.
- Lewandowsky, S., Ecker, U. K., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the “post-truth” era. *Journal of Applied Research in Memory and Cognition*, 6(4), 353–369.
- Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13, 106–131.
- Lewandowsky, S., Stritzke, W. G., Oberauer, K., & Morales, M. (2005). Memory for fact, fiction, and misinformation: The Iraq War 2003. *Psychological Science*, 16, 190–195.
- Lindsay, D. S., & Johnson, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition*, 17, 349–358.

- Mandelbaum, E. (2014). Thinking is believing. *Inquiry: An Interdisciplinary Journal of Philosophy*, 57(1), 55–96.
- Marsh, E. J., Cantor, A. D., & Brashier, N. M. (2016). Believing that humans swallow spiders in their sleep: False beliefs as side effects of the processes that support accurate knowledge. *Psychology of Learning and Motivation*, 64, 93–132.
- Marsh, E. J., Meade, M. L., & Roediger III, H. L. (2003). Learning facts from fiction. *Journal of Memory and Language*, 49(4), 519–536.
- Marsh, E. J., & Yang, B. W. (2017). A call to think broadly about information literacy. *Journal of Applied Research in Memory and Cognition*, 6, 401–404.
- McGlone, M. S., & Tofigbakhsh, J. (2000). Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. *Psychological Science*, 11, 424–428.
- Mills, J., & Jellison, J. M. (1967). Effect on opinion change of how desirable the communication is to the audience the communicator addressed. *Journal of Personality and Social Psychology*, 6(1), 98.
- Newman, E. J., Garry, M., Bernstein, D. M., Kantner, J., & Lindsay, D. S. (2012). Non-probative photographs (or words) inflate truthiness. *Psychonomic Bulletin & Review*, 19, 969–974.
- Newman, E. J., Garry, M., Unkelbach, C., Bernstein, D. M., Lindsay, D. S., & Nash, R. A. (2015). Truthiness and falsiness of trivia claims depend on judgmental contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41, 1337–1348.
- Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32, 303–330.
- Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, 147(12), 1865–1880.
- Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188, 39–50.
- Polage, D. C. (2012). Making up history: False memories of fake news stories. *Europe's Journal of Psychology*, 8, 245–250.
- Porter, E., Wood, T. J., & Kirby, D. (2018). Sex trafficking, Russian infiltration, birth certificates, and Pedophilia: A survey experiment correcting fake news. *Journal of Experimental Political Science*, 5, 159–164.
- Priester, J. R., & Petty, R. E. (1995). Source attributions and persuasion: Perceived honesty as a determinant of message scrutiny. *Personality and Social Psychology Bulletin*, 21(6), 637–654.
- Putnam, A. L., & Phelps, R. J. (2017). The citation effect: In-text citations moderately increase belief in trivia claims. *Acta Psychologica*, 179, 114–123.
- Rapp, D. N., & Donovan, A. M. (2017). Routine processes of cognition result in routine influences of inaccurate content. *Journal of Applied Research in Memory and Cognition*, 6(4), 409–413.
- Rapp, D. N., Hinze, S. R., Kohlhepp, K., & Ryskin, R. A. (2014). Reducing reliance on inaccurate information. *Memory & Cognition*, 42, 11–26.
- Rapp, D. N., & Salovich, N. A. (2018). Can't we just disregard fake news? The consequences of exposure to inaccurate information. *Policy Insights from the Behavioral and Brain Sciences*, 5(2), 232–239.
- Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, 8, 338–342.
- Richter, T. (2015). Validation and comprehension of text information: Two sides of the same coin. *Discourse Processes*, 52(5–6), 337–355.

- Richter, T., Schroeder, S., & Wöhrmann, B. (2009). You don't have to believe everything you read: Background knowledge permits fast and efficient validation of information. *Journal of Personality and Social Psychology*, 96(3), 538.
- Shin, J., Jian, L., Driscoll, K., & Bar, F. (2017). Political rumoring on Twitter during the 2012 US presidential election: Rumor diffusion and correction. *New Media & Society*, 19, 1214–1235.
- Shin, J., Jian, L., Driscoll, K., & Bar, F. (2018). The diffusion of misinformation on social media: Temporal pattern, message, and source. *Computers in Human Behavior*, 83, 278–287.
- Shin, J., & Thorson, K. (2017). Partisan selective sharing: The biased diffusion of fact-checking messages on social media. *Journal of Communication*, 67, 233–255.
- Stanley, M. L., Yang, B. W., & Marsh, E. J. (2019). When the unlikely becomes likely: Qualifying language does not influence later truth judgments. *Journal of Applied Research in Memory and Cognition*, 8(1), 118–129.
- Street, C. N., & Kingstone, A. (2017). Aligning Spinoza with Descartes: An informed Cartesian account of the truth bias. *British Journal of Psychology*, 108(3), 453–466.
- Street, C. N., & Richardson, D. C. (2015). Descartes versus Spinoza: Truth, uncertainty, and bias. *Social Cognition*, 33(3), 227–239.
- Strickland, B., & Keil, F. (2011). Event completion: Event based inferences distort memory in a matter of seconds. *Cognition*, 121, 409–415.
- Suhay, E., Blackwell, A., Roche, C., & Bruggeman, L. (2015). Forging bonds and burning bridges: Polarization and incivility in blog discussions about Occupy Wall Street. *American Politics Research*, 43(4), 643–679.
- Swami, V., Voracek, M., Stieger, S., Tran, U. S., & Furnham, A. (2014). Analytic thinking reduces belief in conspiracy theories. *Cognition*, 133, 572–585.
- Swire, B., & Ecker, U. K. H. (2018). Misinformation and its correction: Cognitive mechanisms and recommendations for mass communication. *Misinformation and Mass Audiences*, 195–211.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). New York, NY: Academic Press.
- Underwood, J., & Pezdek, K. (1998). Memory suggestibility as an example of the sleeper effect. *Psychonomic Bulletin & Review*, 5, 449–453.
- Watkins, M. J., & Kerkar, S. P. (1985). Recall of a twice-presented item without recall of either presentation: Generic memory for events. *Journal of Memory and Language*, 24, 666–678.
- Williams, H. T., McMurray, J. R., Kurz, T., & Lambert, F. H. (2015). Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Global Environmental Change*, 32, 126–138.
- Wood, T., & Porter, E. (2019). The elusive backfire effect: Mass attitudes' steadfast factual adherence. *Political Behavior*, 41, 135–163.