

DR. PETER WHITEHEAD (Orcid ID : 0000-0002-1966-0848)

Article type : Research Dialogue

The Cognitive Processes Underlying False Beliefs

Matthew L. Stanley, Peter S. Whitehead, and Elizabeth J. Marsh

Duke University

Author Note

Matthew L. Stanley, Department of Psychology and Neuroscience, Center for Cognitive Neuroscience, Fuqua School of Business, Duke University

Peter S. Whitehead, Department of Psychology and Neuroscience, Center for Cognitive Neuroscience, Duke University

Elizabeth J. Marsh, Department of Psychology and Neuroscience, Duke University

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1002/JCPY.1289](https://doi.org/10.1002/JCPY.1289)

This article is protected by copyright. All rights reserved

Accepted Article

Correspondence should be addressed to Elizabeth J. Marsh, Department of Psychology and Neuroscience, 417 Chapel Drive, Box 90086, Durham, NC 27708-0086
Phone: (919) 660-5796. E-mail: emarsh@psych.duke.edu

Abstract

Why do consumers sometimes fall for spurious claims – e.g., brain training games that prevent cognitive decline, toning sneakers that sculpt one's body, flower essence that cures depression – and how can consumers protect themselves in the modern world where information is shared quickly and easily? As cognitive scientists, we view this problem through the lens of what we know, more generally, about how people evaluate information for its veracity, and how people update their beliefs. That is, the same processes that support true belief can also encourage people to sometimes believe misleading or false information. Anchoring on the large literature on truth and belief updating allows predictions about consumer behavior; it also highlights possible solutions while casting doubt on other possible responses to misleading communications.

Keywords: misinformation, heuristics, truth, belief, updating

Introduction

Corporate marketing teams sometimes promote misleading, unverified, or utterly false claims. For example, Goop has faced considerable criticism for marketing numerous products (e.g., vaginal steaming, coffee enemas, jade eggs) that do not yield their advertised medical benefits. New Balance falsely claimed that their technologically advanced sneakers strategically activate the glutes, quads, hamstrings, and calves, enabling their wearers to tone and burn extra calories. Advertisements for the herbal supplement, Airborne, claimed that it helps to ward off harmful bacteria and germs, preventing the flu and common cold; such benefits were, however, unestablished. Misleading and false advertisements such as these are not a new phenomenon; famous examples include Stanley's Snake Oil and Listerine's claims in the 1970s that their mouthwash cured colds. But recent developments in digital media (and social media, in particular) mean that such falsehoods spread more rapidly and more widely than ever before. A single tweet can, for example, reach millions of people in an instant. In this context, it has become increasingly important to understand how consumers make judgments about the veracity of information, and how they update their beliefs in the face of retractions or other feedback.

What processes drive consumers' judgments about the veracity of information, and what processes play a role in belief revision (i.e., the correction of false beliefs)? Oftentimes, advertisers appeal to people's emotions, insecurities, and motivations. This "hot cognition" occurs within a system tuned to learn and update beliefs about the world. People's belief in falsehoods (and their ability to update their beliefs) depends on the same processes that give rise to accurate beliefs, not upon unique, special processes or strategies specific to any particular domain of research (Isberner & Richter, 2014; Marsh, Cantor, & Brashier, 2016; Marsh & Stanley, 2021; Rapp & Donovan, 2017; Rapp & Salovich, 2018).

In this paper, we describe four psychological principles that underlie belief. These principles are agnostic as to the veracity of information; that is, incorrect beliefs are often the byproduct of a system that normally leads people to the right answers. For example, to preview, people are biased to accept information as true, and most of the time that bias leads them to the correct answer,

quickly and more efficiently than other more strategic routes; but this shortcut will occasionally be wrong, leading to systematic errors. Throughout, we comment on how the digital world has exacerbated and/or changed the playing field for consumers and advertisers alike. We then discuss how the four principles play out in a number of different situations, explaining why some situations or consumer mindsets have the effects that they do on belief. Finally, we consider belief updating, especially in the context of FCC-issued corrections.

Psychological Principles of Belief

Principle #1: Truth Bias. People are biased to believe that information is true, and unbelieving information requires a second, effortful step in processing. A consumer reading a claim about toning sneakers, for example, will automatically believe that the claim is true before (potentially) deciding it is untrue. This bias reflects a general disposition to believe in a truthful world, as well as the belief that others are more likely to tell the truth than to lie (Gilbert, 1991; Bond & DePaulo, 2006; Vrij, 2008). It is also cognitively efficient, meaning that we only need to “tag” information when a claim is false; otherwise, we can assume it is true (Gilbert, 1991).

What evidence is there to support the claim that people are biased to believe information is true? Critically, this account predicts an asymmetry in memory errors. If unbelieving is disrupted, the information will not be tagged as false and thus will later be remembered as true. However, an interruption will not have a parallel impact on how truths are later remembered, since truth is the default. Several experimental studies support these ideas (e.g., Gilbert, Krull, & Malone, 1990; Gilbert, Tafarodi, & Malone, 1993). In one of Gilbert and colleagues’ original studies (1993), participants played the role of a trial judge making sentencing decisions for a criminal defendant. They read aloud a series of statements about a case, after being instructed that they should only believe statements printed in black ink (and that red statements were false). Sometimes a string of digits appeared underneath the text and participants had to press a key when the number five appeared, thus interrupting their processing of the text. On a later recognition test, true statements were rarely misremembered as false, but more than 20% of false statements were misremembered as true. Furthermore, interrupted subjects sentenced the defendant to longer prison sentences; that

is, in addition to misremembering false statements as true, participants acted on this information during sentencing. These results are consistent with the idea that interruption interferes with the subsequent step of tagging the false information.

Similar effects have been obtained in experiments where participants judged the veracity of smiles (Gilbert et al., 1990) and learned translations of Hopi words (Gilbert et al., 1993). In the former, participants judged whether each of a series of videos depicted real or fake smiles; they received feedback on each guess. On some of the trials, participants also responded as quickly as possible to a tone. On a later test, participants made more mistakes when judging fake smiles, incorrectly believing that they were real smiles. Such errors were particularly likely for trials paired with the secondary task; these disruptions selectively impaired participants' abilities to "unbelieve" (i.e., tag as false) the fake smile.

The bias towards truth wins out over people's expectations in certain situations. In one study, all participants saw the same videos of people talking about their holidays, and were told to figure out which speakers were lying. In reality, 50% of speakers were lying, but across experimental conditions participants were told that 20%, 50%, or 80% of speakers were lying. This information had a large impact on judgments early in the study, with people in the 80% liars condition believing fewer videos – but over the course of the experiment, they gradually believed more and more of the speakers, with participants in the 80% condition no longer showing a lie-bias by the last 20% of trials in the experiment (Street & Richardson, 2015a). On the other hand, the same authors also demonstrated a boundary condition for the truth bias (Street & Richardson, 201b; see also, Street & Kingstone, 2017). The bias to assume truth disappeared when participants were allowed to report being unsure about the credibility of the incoming information. Participants listened to a stream of incoming information; in one condition, they categorized information as true or false, in another they choose between true, false, and unknown. The bias to assume truth only occurred in the former condition.

Principle #2: Bias to Extract Meaning. People's default is to process incoming information for meaning, using what they already know to interpret new information. Humans' ability to extract meaning means that we can draw inferences and do not need everything to be explicitly stated.

Even young children do not need to be explicitly told that “dolphins communicate by squeaking,” if they learn that “dolphins travel in pods” and that “pods communicate by squeaking” (Bauer & San Souci, 2010). They easily make the inference to believe that dolphins do, in fact, communicate by squeaking.

This ability to make meaning and extract inferences, however, can also affect how events are remembered. Listerine did not have to actually state that “Listerine kills colds” for people to remember that as the critical takeaway of the advertisement (Dyer & Kuehl, 1978; Mazis, McNeill, & Bernhardt, 1983). This idea is captured in a classic study (Brewer, 1977) where participants who studied claims such as “the karate champion hit the cinder block” later remembered that the karate champion actually *broke* the cinder block. Similarly, a video that strongly implies an action will later be remembered as depicting that action. People who watched a video that showed a person swinging their foot towards a ball later remembered having seen the moment of contact, even though it was not shown (Strickland & Keil, 2011). The bottom line is that an advertiser or a corporation may not have to explicitly state their desired message, as consumers will connect the dots for them, with consequences for what consumers believe.

This focus on meaning allows humans to handle imperfect inputs, glossing over errors in communication, so long as they are “close enough” to the meaning we expected to process. Speech is notoriously disfluent – speakers stutter, speak ungrammatically, and use ambiguous referents. Accordingly, we are able to figure out the meaning even if there is a certain amount of noise in the messaging (Ferreira, Bailey, & Ferraro, 2002). We can recognize a word 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 they are 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 if warned not to answer if the question contains an error. This happens even though the same people later demonstrate knowledge that the reference should be to Noah, not Moses (the Moses Illusion; Bottoms, Eslick, & Marsh, 2010; Erikson & Matson, 1981). In other words, consumers may often miss

errors in communications if those errors are closely related to the truth. Even phonological similarity is sometimes close enough to slip by unnoticed (Shafto & MacKay, 2000). That is, there is a phonological parallel to the Moses illusion, dubbed the Armstrong illusion. People miss the error in the question “*What was the famous line uttered by Louis Armstrong when he first set foot on the moon?*”. The incorrect presupposition (Louis Armstrong, in this case) sounds like the correct reference (Neil Armstrong) but is semantically unrelated. In other words, some consumers will not notice an error so long as it sounds like the expected input.

Most of the time, our focus on meaning means that we remember the gist of events, rather than the exact wording or text used. That gist captures the meaning of events, including inferences that were never stated. There is one exception, however: people are more likely to remember the exact wording of songs and poems (Rubin, 1977). That is, consumer may remember the surface characteristics of an advertisement (i.e., the exact wording) to the extent that the words are set to music, rhyme, or otherwise provide associative cues that support verbatim memory. The only other exceptions to the bias to extract meaning involve situations where the orienting task specifically directs the subject to non-semantic aspects (i.e., surface characteristics like font or sound), but such processing is not the default.

Principle #3: Source as a Heuristic for Truth. People often make very quick judgments about the truth of information, without doing extensive research. They do so by relying on cues that tend to be correlated with truth in the world, including the source of the information. For example, seminal research from Hovland and Weiss (1951) indicates that people are more likely to discount information from a gossip columnist (a low credibility source) than an article in the *New England Journal of Biology and Medicine* (a high credibility source; see also, Kelman & Hovland, 1953). Decades of psychological research have converged on the conclusion that people are more likely to believe information from sources they perceive to be credible, whether that credibility is gained through trustworthiness or expertise (Brinol, Petty, & Tormala, 2004; Eagly & Chaiken, 1993; Hovland & Weiss, 1951; Kelman & Hovland, 1953; McGinnies, 1973; McGinnies & Ward, 1980; Petty & Wegener, 1998; Pornpitakpan, 2004; Priester & Petty, 1995; Ziegler, 2010). Advertisers are clearly aware of this, spending billions of dollars annually on advertising campaigns that feature celebrities

and social media influencers peddling a variety of products. And the perceived credibility of celebrities and social media influencers who endorse products does, in fact, closely predict consumers' beliefs about those products and purchase intentions (Chu & Kamal, 2008; Goldsmith, Lafferty, & Newell, 2013; Harmon & Coney, 1982; Ismagilova, Slade, Rana, & Dwivedi, 2020; Lou & Yuan, 2019; Munnukka et al., 2019).

Of course, one challenge for the consumer is to accurately judge the credibility of a source. In some cases, we can draw on past experience to know whether a source is credible or not (such as when one encounters *The National Enquirer* in the check-out line). Other times, credibility is inferred from secondary cues. For example, people interpret citations as evidence that something is true, as the author is telling the reader where information came from. Increased number of citations – such as the citation at the end of this sentence – increase perceived credibility and belief (Putnam & Phelps, 2017). In-text citations are supposed to offer evidence (direct or indirect) that some claim is true, but the quality of cited material can, of course, vary widely. For example, the tobacco industry infamously funded questionable research published in questionable venues to offer “evidence” that tobacco and cancer were not linked (Oreskes & Conway, 2010). More generally, it is becoming more and more difficult to evaluate the credibility of sources in the digital age (Marsh & Yang, 2017). Consider, for example, *native advertisements*. Like infomercials and ‘advertorials’ that mimic more credible television programs, native advertisements closely resemble more credible web sites in content, format, and style, but are paid for by outside individuals or organizations. They have become commonplace across social media platforms, from Instagram to Facebook to Reddit. Recent research suggests that people have difficulty identifying these native advertisements. In one study, fewer than 10% of participants reported having seen an advertisement on a previously viewed web page, despite the advertisement being explicitly labeled as an “advertisement” or “sponsored content” (Wojdyski & Evans, 2016).

However, simply identifying a source as low credibility is not sufficient. What happens later is the key. The problem is that most information that is retrieved from memory feels source-less; 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. When a consumer looks at a display of orange juice, for example, information about the brand may pop to mind unaccompanied by its source (be it an advertisement or a friend).

One issue is that source information is forgotten quickly (Underwood & Pezdek, 1998), leading people to struggle to attribute remembered information to the right sources (Mather et al., 1999). This is 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 (Hannah & Sternthal, 1984; Hovland & Weiss, 1952; Underwood & Pezdek, 1998). This occurs because the claim and the source of the claim are encoded separately in memory (Underwood & Pezdek, 1998), and source information degrades faster than item information. A claim is more likely to remain coupled to its source if the information was elaborated upon during encoding (Hannah & Sternthal, 1984), especially if the audience encoded the information in reference to themselves (Mazursky & Schull, 1987). Less clear is whether or not advertisers want consumers to remember the source of information.

Principle #4: Ease of Processing (Fluency) as a Heuristic for Truth. Another heuristic for judging truth is how easy it is to process information, or *fluency*; that is, people tend to interpret easy processing as evidence for truth (Alter & Oppenheimer, 2009; Brashier & Marsh, 2020; Arkes, Hackett, & Boehm, 1989; Boehm, 1994; Schwarz, Jalbert, Noah, & Zhang, 2020; Unkelbach, 2007). The logic is that any one truth is more likely to have been experienced previously (and is thus more fluent) than any one of the infinite possible false versions of it. The association between truth and fluency is learned (Unkelbach, 2007), although it appears even in children (Fazio & Sherry, 2020).

Repetition is perhaps the simplest and most straightforward way to make information easier to process (Brashier & Marsh, 2020; Dechêne et al., 2010). That is, each time you read a sentence, you read it more quickly and easily. Experimentally this is captured in studies where participants read a series of obscure trivia statements in an initial phase of the experiment; after a brief delay,

participants then rate the truth of both previously seen (repeated) statements and new (unrepeated) statements. Repeated statements are consistently judged as more likely to be true than new statements, because repeated statements are easier to read (process) the second time. A single prior exposure is sufficient to boost truth judgments, and additional exposures can further boost truth judgments (but with rapidly diminishing returns; Hasher et al., 1977; Dechêne et al., 2010). Since the original demonstration in the 1970s (Hasher, Goldstein, & Toppino, 1977), this so-called *illusory truth effect* has consistently been shown across diverse knowledge domains, in different languages, and across long delay periods (e.g., months) (for a meta-analysis and review, see Dechêne et al., 2010). Of particular relevance, repetition also boosts judged truth for claims about consumer products (Hawkins & Hoch, 1992; Johar & Roggeveen, 2007; Sundar, Kardes, & Wright, 2015).

Critically, the illusory truth effect does not require verbatim repetition. Prior exposure to part of a statement—even when that piece of the statement cannot be judged as true or false (e.g., the phrase “a hen’s body temperature”)—is sufficient to boost truth judgments for the full claim (e.g., “the temperature of a hen’s body is about 104°F”; Begg, Armour, & Kerr, 1985; see also Arkes, Boehm, & Xu, 1991). This finding is broadly consistent with the larger priming literature, where the exact repetition of words and phrases is not necessary to speed up the subsequent processing of related content. For example, participants tend to process a verb more quickly when it follows exposure to a different form of that verb (e.g., *found* primes *find*; Marslen-Wilson & Lorraine, 1997).

Originally, it was assumed that fluency was a heuristic used in situations of uncertainty, and thus, most studies used stimuli that people did not have pre-existing knowledge about. But more recent work makes clear that fluency still plays a role even when people have relevant knowledge stored in memory. For example, even after people demonstrate that they know what a sari is, prior exposure to a false statement about a sari (i.e., *a sari is the short, pleated skirt worn by Scotsmen*) still boosted the judged truth of these claims (Fazio, Brashier, Rajaram, & Marsh, 2015). Related research has found that the size of the illusory truth effect is similar for plausible and implausible claims (once floor and ceiling effects are taken into account; Fazio, Rand, & Pennycook, 2019). Although repetition does not boost truth judgments for particularly extreme, blatantly false claims

(e.g., *The Earth is a perfect square*), a mere inkling of potential plausibility is sufficient for repetition to boost judgments of truth (Pennycook, Cannon, & Rand, 2018).

Critically, the key is making information easy to process, not simply making people feel like they have experienced that information before. Repetition is not required; visual and auditory features can be manipulated to systematically shift the ease with which information is processed, with consequences for truth judgments. In the visual domain, for example, stronger figure-ground contrasts (e.g., black text on a white background; Reber & Schwartz, 1999), easier to read fonts (e.g., *Times New Roman*, 14-pt font; Song & Schwarz, 2008), and clearer, more readable handwriting (Greifeneder et al., 2010) all make it easier to process claims, which, in turn, boosts truth judgments. In the auditory domain, clearer auditory presentations (Newman & Schwarz, 2018) and familiar accents (Lev-Ari & Keysar, 2010) make it easier to hear and understand the claims, again with consequences for truth judgments.

Evaluating Possible Interventions and Correcting False Beliefs

The principles just described often support accurate beliefs, even though our focus in this paper is on incorrect beliefs. With that in mind, we consider some specific examples where the four principles just described can be used to explain why particular interventions/behaviors have the effects that they do. Some of the examples focus on how people come to hold false beliefs, before we turn to correcting false beliefs. It should be noted that correcting false beliefs should not be confused with the correction of *false memories*. By false memories, we refer broadly to problems with episodic memory, whereby people misremember the specifics of a particular event (or even remember an event that never actually happened). We know a fair amount about how to correct people's memories of specific past events but doing so can be psychologically different from false beliefs. Event memories involve thinking back to a particular time and place but in contrast, as already discussed, knowledge often "pops to mind" and is not associated with a particular time and place (perhaps because it was encountered in many different places or because source information was forgotten).

We consider characteristics of the encoding situation (that is, at the time of encountering misleading claims) as well as correction of false beliefs already instilled in consumers. To preview,

people's strong reliance on fluency (ease of processing) does not make us optimistic, and in fact some typical suggestions may backfire if they increase the fluency of the claim. We consider qualities of the to-be-assessed communications (i.e., qualifying language, pictures) as well as the consumer's mindsight (including attentional focus and prior knowledge) before turning to the correction of false beliefs.

Warnings May Sometimes Help and Other Times have No effect. The simplest solution is to warn people – but warnings are not as simple as they might seem. A warning is likely to be helpful to the extent it prevents consumers from ever processing the information (i.e., such as when a warning gets people to skip viewing an ad), but the predictions become less clear when warnings are paired with the consumption of information. One point worth making here is that most of the studies described here do have warnings – people who fall for the Moses Illusion are explicitly warned that questions contain errors and that their job is to identify those questions; similarly, participants are explicitly told in most illusory truth studies that some statements will be true and others false. Jalbert and colleagues (2020) manipulated whether or not participants received this standard illusory truth instruction; participants who were warned showed a reduced illusory truth effect (but it was not eliminated). On the other hand, warnings have little effect on people's bias to assume information is true (Gilbert et al., 1990), consistent with the finding described earlier that people pay less attention to base-rates as time passes. One reason for the mixed findings may be that warnings effectively impose a prospective memory task on people: people have to remember to watch for and identify misleading information.

Assuming information is processed, the most effective warnings work because they encourage the critical consumption of information. The success of the warning, however, depends upon being able to critically evaluate the information and to notice errors. In an experiment where participants were told to evaluate every statement for its truth value (upon their first exposure), the illusory truth effect disappeared -- but only when participants had relevant knowledge stored in memory that could inform their truth judgments (Brashier, Eliseev, & Marsh, 2019). That is, the trial-by-trial decision helped participants to bring their stored knowledge to bear, creating an "accuracy focus" that counteracted the fluency signal that so often guides truth judgments. However, warnings

will do little if people do not have any relevant information stored in memory. And environments such as social media do not encourage an accuracy focus – a concern since people are increasingly encountering advertisements on social media (Gottfried & Shearer, 2016; Dunaway, Searles, Sui, & Paul, 2018).

Warnings encountered after information is processed are unlikely to be effective, because information has already been processed as if it were true – meaning that it is now being retrospectively tagged as “false”, a tag that will be forgotten at a faster rate than the information itself.

Qualifying Language Doesn’t Help. Precise language sounds like a good idea, but it may not be helpful. For example, consider the effects of using qualifiers such as “likely” or “unlikely” to convey a certain precision in one’s claim. These kinds of qualifiers are frequently employed and relied upon in everyday and professional life (Budescu & Wallsten, 1995; Wallsten, Budescu, Zwick, & Kemp, 1993). But exposure to qualified statements increases belief in later *unqualified* versions of the same statement – for example, reading “*It’s unlikely that the body of a rotten tree is called a daddock*” increases later belief in “*The body of a rotten tree is called a daddock*” just as much as if people read “*It is certain that the body of a rotten tree is called a daddock*” (Stanley, Yang, & Marsh, 2019). The size of the illusory truth effect was the same regardless of what version of the statement people saw earlier in the experiment.

Reliance on fluency likely contributes to this null effect. Because so many of the words are repeated across instances, a qualified version of a statement (e.g., *It’s unlikely that the study of snakes is called cynology*) makes it easier to read a different version of the statement later (e.g., *the study of snakes is called cynology*). In this way, qualifying claims with terms like “unlikely”, “improbable”, and “uncertain” can actually have the *opposite* effect of the speaker’s original intention: these qualified claims seem more likely to be true later on, as compared to new statements, even when qualifiers are used to induce doubt about the truth of the claims.

Similar effects were observed in an experiment where participants studied only statements that were negatively qualified. That is, every statement during study was qualified as “*improbable*”,

“impossible”, or “unlikely.” Even so, studied statements were rated as more likely to be true on a test two days later, as compared to new statements not studied in the first part of the experiment (Stanley et al., 2019). A parallel group of subjects were tested on their memories for the qualifiers they had seen; they were asked to select the three qualifiers from a list of six options, and then asked whether they had seen only negative or positive qualifiers. More than 80% of participants knew that all qualifiers were negative, yet this knowledge was not sufficient to eliminate the illusory truth effect.

“Photographic” evidence can mislead. Most debunking handbooks recommend including pictures in debunking campaigns, but not all pictures are equal. Simple pictures encountered at the same time as the misinformation can impact belief in it. In one study, Newman and colleagues (2012) presented people with statements like “*Macadamia nuts are in the same evolutionary family as peaches.*” Photos accompanied some of the statements; critically, the photos did not provide any evidence about the veracity of the claim (for example, simply providing a picture of a macadamia nut). Participants were biased to accept statements as true when they were accompanied by a photo (Fenn et al., 2013; Newman et al., 2015). This bias occurs outside of conscious awareness (Newman et al. 2018) and can persist for several days (Fenn et al. 2013). Pairing a statement with a photo also increased participants’ tendencies to like and share information in a simulated online environment (regardless of whether that information was true or false in reality; Fenn et al., 2019). This finding is particularly problematic as more and more Americans are using social media as their primary source of news and information about the world (Shearer & Gottfried, 2017) and as corporations are devoting more resources to advertising on these online platforms.

Non-probative photos are not themselves misleading; rather, they change the way people process accompanying text. They facilitate “the conceptual processing of the claim by making it easier to imagine and understand in the recipient’s mind” (Newman et al., 2012; Newman et al., 2020). Only *semantically related* photos make accompanying statements seem more likely to be true (Newman et al. 2015). That is, pairing trivia statements with *semantically unrelated* photos induces conceptual disfluency, meaning that the unrelated photos inhibit participants’ ability to generate thoughts and images associated with the trivia statement itself.

Of course, images may also be distorted or fake, raising a different set of issues. Bad actors can easily develop and use visual content (pictures or videos) that seems real despite being doctored – a problem given that people generally view visual content as the best kind of evidence for something being true (Kelly & Nace, 1994). People have tremendous difficulty detecting digitally doctored images, even when manipulations are meant to be obvious (Nightingale, Wade, & Watson, 2017). However, perhaps even more concerning is that deepfakes aren't necessarily more likely to cause problems than simple photos or text, highlighting how technological sophistication may not be a pre-requisite for misleading consumers (Murphy & Flynn, 2021; see also Nash, 2018).

Mind-wandering does not Protect Against Illusory Truth. A distracted consumer is not necessarily an uninfluenced one, which is problematic given that as much as half of our waking time is spent in periods of mind wandering (Killingsworth & Gilbert, 2010; see also McVay, Kane, & Kwapil, 2009 and Seli et al., 2018). Mind-wandering can be deliberate, as when we start thinking about dinner during a boring talk, or accidental, as when the student realizes he/she has no idea what they just read in the textbook. Here our focus is not on what the mind wanders to, but rather on what happens to the stream of information that the mind has wandered from. To the extent that unattended information still impacts later belief, it highlights how people rely on fluency (ease of processing) as a cue for truth rather than episodic memory or other more deliberative strategies.

We tested these ideas in experiments that used thought probes during the initial part of the study, during which participants were exposed to many statements. Briefly, as is standard in the mind-wandering literature, participants were periodically interrupted during the study phase to ask if they were on or off task; these self reports were used to categorize just-studied trials as instances of full awareness vs. mind-wandering (Stanley, Whitehead, Marsh, & Seli, 2021). On the final truth test, participants judged the truth of both new and studied statements, with the studied statements further divided into those that occurred while attending to the task vs. while mind-wandering. Critically, participants showed an illusory truth effect for statements that they reported not having paid attention to. The fluency signal boosted truth judgments later on (relative to new statements that had not been presented) even though participants reported that they had been mind-

wandering – an important finding given that people are often distracted, mind-wandering, or exposed to information in the background.

Knowing Better is Not Always Protective. We alluded earlier to the idea that people could have information stored in memory and yet fail to notice a contradiction with that stored knowledge (the *Moses illusion*). This problem can happen even when people are relatively expert in a domain, as illustrated in a study with History and Biology PhD students (who had years of experience in their respective disciplines). Participants were warned not to answer questions with incorrect presuppositions, and then were asked to answer questions like “*Water contains two atoms of helium and how many of oxygen?*” Even Biology PhD students sometimes missed that the reference should be to *hydrogen* and instead answered the question (Cantor & Marsh, 2017). Experts did better when the key term was bolded (i.e., **helium**), drawing their attention to the potential trick – but the effect was still not eliminated.

Similarly, source information can be stored in memory and not utilized at test. That is, while source is often forgotten (as described earlier), sometimes the problem is a failure of retrieval rather than forgetting. This distinction is illustrated in a study in which participants read statements from two different sources, one of whom was described as highly reliable and the other as less reliable, as determined by independent fact-checkers. The final test was either two or four weeks later; overall, source memory was low (consistent with the idea that people forget source information relatively quickly). However, the interesting point for present purposes is that people were better able to identify the source of information presented three times in the original session, as opposed to once – but those items were also rated as more likely to be true on a final test than the items presented once originally, regardless of which source they came from (Henkel & Mattson, 2011).

Correcting Misinformation. What happens after misinformation is distributed (and the cat is out of the bag)? The Federal Trade Commission (FTC) can require companies that produce false or deceptive advertising to issue corrective advertisements. These corrective advertisements are meant to rectify the false or misleading claims in the original advertisement, with the hope of leaving consumers with true beliefs about the products.

Unfortunately, corrective advertisements commonly repeat the original falsehood in the process of correcting it. Repeating a falsehood during an attempted correction (e.g., “It is a myth that *X*” or “It is false that *X*”) reduces the likelihood of successful belief revision, because repetition makes processing the falsehood easier (Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012; Skurnik, Yoon, Park, & Schwarz, 2005; Skurnik, Yoon, & Schwarz, 2007; cf. Ecker, Hogan, & Lewandowsky, 2017). A notable example involves the Listerine case from the 1970s, where the FTC forced Warner-Lambert, Listerine’s manufacturer at the time, to spend \$10 million dollars on corrective advertisements stating the following: “*Contrary to prior advertising, Listerine will not help prevent colds or sore throats or lessen their severity.*” The problem here is that the falsehood— “*help prevent colds or sore throats or less their severity*”—was repeated in the corrective advertisement. Later experimental research using the original deceptive advertisement along with the corrective advertisement found that the corrective advertisement did not actually reduce participants’ reliance on the falsehoods in the original Listerine advertisement (Dyer & Kuehl, 1978; Mazis, McNeill, & Bernhardt, 1983).

The Listerine correction also involves a *negation* of a belief, rather than an *affirmation* of a belief. In other words, it asked people to *unbelieve* something, which the truth bias suggests is more difficult than believing in something. We directly compared affirming versus negating corrections in experiments where people processed statements multiple times to ensure that they had learned them (Stanley, Yang, Stone, & Marsh, 2021). Critically, half were affirmations (i.e., *Michelangelo’s statue of David is located in Venice*) and half were negations (i.e., *A poem written for a bride is a not an epithalamium*); all were false. The key manipulation involved the corrections: sometimes subjects received corrections that required them to unbelieve information (i.e., *Michelangelo’s statue of David is **not** located in Venice*) and other times the corrections instilled belief (i.e., *A poem written for a bride **is** an epithalamium*). On a final test, participants were more successful when the correction instilled belief than when it had required them to unbelieve something (Stanley, Yang, Stone, & Marsh, 2021). Follow-up data showed that people viewed affirmations as more informative than negations; a negation like *Michaelangelo’s statue of David is **not** located in Venice* is not viewed as informative because it leaves open an infinite number of places where the status might be.

Negations are also harder to process, and in general, corrections should be easy to process, relying on simple language rather than complex arguments (Lewandowsky et al., 2012; Lewandowsky, Ecker, & Cook, 2017; Lorenz-Spreen, Lewandowsky, Sunstein, & Hertwig, 2020). The most effective corrections do more than negate false information: they offer alternative, concrete information to replace the negation that is easy to process (Lorenz-Spreen, Lewandowsky, Sunstein, & Hertwig, 2020; Marsh & Stanley, 2021). Of course, sometimes this advice can be difficult to follow in practice. For example, while the scientific consensus is that vaccines do *not* cause autism (DeStefano, Price, & Weintraub, 2013; Taylor, Swerdfeger, & Eslick, 2014), there is no simple explanation available to describe the complex genetic and environmental interactions that do (Marsh, Cantor, & Brashier, 2016). This means that there is no simple, easy-to-process alternative claim that could be repeated to help correct the widespread false belief that vaccines cause autism. Unfortunately, explanations may be similarly difficult in the case of misleading advertisements; corporations will not want to direct consumers to other brands or to explain why their claims were inflated.

Conclusions

Companies have a vested interest in spreading favorable information—true or false—about their products in ways that directly affect consumer behavior. The consumer psychology literature has advanced our understanding of social, affective, and motivational forces that give rise to people's beliefs, but less research has examined the *cognitive* processes that give rise to consumers' beliefs. Our review of cognitive science research examining how, why, and when people form beliefs showcases a different, and potentially fruitful, way to understand the rapid spread of falsehoods in the digital age that might affect consumer judgement and decision making.

With an eye toward addressing these new challenges, we have reviewed evidence for four general principles of cognition that impact when, why, and how people come to believe that information is true or false. Because they are domain general, these principles can play a critical role in understanding and predicting people's beliefs across different domains, from politics to entertainment to advertising. Further, we linked belief in different situations (i.e., while mind wandering, after encountering a warning) to these principles. These same principles also have

implications for the successful correction of false beliefs, which can be surprisingly challenging. There is a pressing need for field research and scalable interventions that utilize these general principles to help people acquire more true beliefs about the world.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

References

- Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, *13*(3), 219–235.
- Arkes, H. R., Boehm, L. E., & Xu, G. (1991). Determinants of judged validity. *Journal of Experimental Social Psychology*, *27*(6), 576-605.

- Arkes, H. R., Hackett, C., & Boehm, L. (1989). The generality of the relation between familiarity and judged validity. *Journal of Behavioral Decision Making*, 2(2), 81-94.
- Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, 54(7), 462-479.
- Bauer, P. J., & San Souci, P. (2010). Going beyond the facts: Young children extend knowledge by integrating episodes. *Journal of Experimental Child Psychology*, 107(4), 452-465.
- Begg, I., Armour, V., & Kerr, T. (1985). On believing what we remember. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 17(3), 199-214.
- Boehm, L. E. (1994). The validity effect: A search for mediating variables. *Personality and Social Psychology Bulletin*, 20(3), 285-293.
- Bond Jr, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, 10(3), 214-234.
- Brashier, N. M., & Marsh, E. J. (2020). Judging truth. *Annual Review of Psychology*, 71, 499-515.
- Brewer, W. F. (1977). Memory for the pragmatic implications of sentences. *Memory & Cognition*, 5(6), 673-678.
- Briñol, P., Petty, R. E., & Tormala, Z. L. (2004). Self-validation of cognitive responses to advertisements. *Journal of Consumer Research*, 30(4), 559-573.
- 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.
- Budescu, D. V., & Wallsten, T. S. (1995). Processing linguistic probabilities: General principles and empirical evidence. *Psychology of Learning and Motivation*, 32, 275-318.
- Cantor, A. D., & Marsh, E. J. (2017). Expertise effects in the Moses illusion: Detecting contradictions with stored knowledge. *Memory*, 25(2), 220-230.
- Chu, S. C., & Kamal, S. (2008). The effect of perceived blogger credibility and argument quality on message elaboration and brand attitudes: An exploratory study. *Journal of Interactive Advertising*, 8(2), 26-37.

Cohen, G., & Faulkner, D. (1989). Age differences in source forgetting: effects on reality monitoring and on eyewitness testimony. *Psychology and Aging, 4*(1), 10-17.

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*(4), 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.

DeStefano, F., Price, C. S., & Weintraub, E. S. (2013). Increasing exposure to antibody-stimulating proteins and polysaccharides in vaccines is not associated with risk of autism. *The Journal of Pediatrics, 163*(2), 561-567.

Dunaway, J., Searles, K., Sui, M., & Paul, N. (2018). News attention in a mobile era. *Journal of Computer-Mediated Communication, 23*(2), 107-124.

Dyer, R. F., & Kuehl, P. G. (1978). A longitudinal study of corrective advertising. *Journal of Marketing Research, 15*(1), 39-48.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.

Ecker, U. K., Hogan, J. L., & Lewandowsky, S. (2017). Reminders and repetition of misinformation: Helping or hindering its retraction?. *Journal of Applied Research in Memory and Cognition, 6*(2), 185-192.

Erickson, T. D., & Mattson, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning & Verbal Behavior, 20*, 540-551.

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*, 993-1002.

Fazio, L. K., Rand, D. G., & Pennycook, G. (2019). Repetition increases perceived truth equally for plausible and implausible statements. *Psychonomic Bulletin & Review, 26*, 1705-1710.

Fazio, L. K., & Sherry, C. L. (2020). The effect of repetition on truth judgments across development. *Psychological Science, 31*(9), 1150-1160.

- Fenn, E., Newman, E. J., Pezdek, K., & Garry, M. (2013). The effect of nonprobative photographs on truthiness persists over time. *Acta Psychologica, 144*(1), 207-211.
- Fenn, E., Ramsay, N., Kantner, J., Pezdek, K., & Abed, E. (2019). Nonprobative photos increase truth, like, and share judgments in a simulated social media environment. *Journal of Applied Research in Memory and Cognition, 8*(2), 131-138.
- Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science, 11*(1), 11-15.
- Forster, S., & Lavie, N. (2009). Harnessing the wandering mind: The role of perceptual load. *Cognition, 111*(3), 345-355.
- Gallo, D.A., Roberts, M. J., & Seamon, J. G. (1997). Remembering words not presented in lists: Can we avoid creating false memories? *Psychonomic Bulletin & Review, 4*, 271-276.
- 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*(4), 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*(2), 221-233.
- Gottfried, J. & Shearer, E. News use across social media platforms 2016. Pew Research Center (2016). <http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/>
- Greifeneder, R., Alt, A., Bottenberg, K., Seele, T., Zelt, S., & Wagener, D. (2010). On writing legibly: Processing fluency systematically biases evaluations of handwritten material. *Social Psychological and Personality Science, 1*(3), 230-237.
- Hannah, D. B., & Sternthal, B. (1984). Detecting and explaining the sleeper effect. *Journal of Consumer Research, 11*(2), 632-642.
- Harmon, R. R., & Coney, K. A. (1982). The persuasive effects of source credibility in buy and lease situations. *Journal of Marketing Research, 19*(2), 255-260.
- Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning & Verbal Behavior, 16*, 107-112.

- Hawkins, S. A., & Hoch, S. J. (1992). Low-involvement learning: Memory without evaluation. *Journal of Consumer Research*, 19(2), 212-225.
- Henkel, L. A., & Mattson, M. E. (2011). Reading is believing: The truth effect and source credibility. *Consciousness and Cognition*, 20(4), 1705-1721
- Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *Public Opinion Quarterly*, 15(4), 635-650.
- Hovland, C. I., & Weiss, W. (1952). The influence of source credibility on communication effectiveness. *Public Opinion Quarterly* 15, 635-650.
- Isberner, M. & Richter, T. (2014). Does validation during language comprehension depend on an evaluative mindset?, *Discourse Processes*, 51, 7-25.
- Ismagilova, E., Slade, E. L., Rana, N. P., & Dwivedi, Y. K. (2020). The effect of electronic word of mouth communications on intention to buy: A meta-analysis. *Information Systems Frontiers*, 22(5), 1203-1226.
- Jalbert, M., Newman, E., & Schwarz, N. (2020). Only half of what I'll tell you is true: Expecting to encounter falsehoods reduces illusory truth. *Journal of Applied Research in Memory and Cognition*, 9(4), 602-613.
- Johar, G. V., & Roggeveen, A. L. (2007). Changing false beliefs from repeated advertising: The role of claim-refutation alignment. *Journal of Consumer Psychology*, 17(2), 118-127.
- Kelly, J. E., & Nace, D. (1994). Digital imaging & believing photos. *Visual Communication Quarterly*, 1(1), 4-18.
- Kelman, H. C., & Hovland, C. I. (1953). " Reinstatement" of the communicator in delayed measurement of opinion change. *The Journal of Abnormal and Social Psychology*, 48(3), 327-335.
- Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, 330(6006), 932-932.
- Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, 46(6), 1093-1096.

- 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(3), 106-131.
- Loftus, E. F. (2003). Make-believe memories. *American Psychologist*, 58(11), 867-873.
- Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, 13(5), 585-589.
- Lorenz-Spreen, P., Lewandowsky, S., Sunstein, C. R., & Hertwig, R. (2020). How behavioural sciences can promote truth, autonomy and democratic discourse online. *Nature Human Behaviour*, 4(11), 1102-1109.
- Lou, C., & Yuan, S. (2019). Influencer marketing: How message value and credibility affect consumer trust of branded content on social media. *Journal of Interactive Advertising*, 19(1), 58-73.
- 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., & Stanley, M. (2020). False beliefs: Byproducts of an adaptive knowledge base? In R. Greifeneder, M. Jaffé, E. J. Newman, & N. Schwarz (Eds.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 131–146). London: Routledge.
- 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.
- Mazis, M. B., McNeill, D. L., & Bernhardt, K. L. (1983). Day-after recall of Listerine corrective commercials. *Journal of Public Policy & Marketing*, 2(1), 29-37.
- Mazursky, D., & Schul, Y. (1987). The effects of advertisement encoding on the failure to discount information: Implications for the sleeper effect. *Journal of Consumer Research*, 15, 24-35.

- McDermott, K. B., & Roediger, H. L., III (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory & Language*, 39, 508-520.
- McGinnies, E. (1973). Initial attitude, source credibility, and involvement as factors in persuasion. *Journal of Experimental Social Psychology*, 9(4), 285-296.
- McGinnies, E., & Ward, C. D. (1980). Better liked than right: Trustworthiness and expertise as factors in credibility. *Personality and Social Psychology Bulletin*, 6(3), 467-472.
- McGlone, M. S., & Tofiqbakhsh, J. (2000). Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. *Psychological Science*, 11(5), 424-428.
- McVay, J. C., Kane, M. J., & Kwapil, T. R. (2009). Tracking the train of thought from the laboratory into everyday life: An experience-sampling study of mind wandering across controlled and ecological contexts. *Psychonomic Bulletin & Review*, 16(5), 857-863.
- McNally, R. J. (2012). Searching for repressed memory. In R. F. Belli (Ed.), *True and false recovered memories: Toward a reconciliation of the debate*, Nebraska Symposium on Motivation (Vol. 58, pp. 121 – 147). New York: Springer
- Meyersburg, C. A., Bogdan, R., Gallo, D. A., & McNally, R. J. (2009). False memory propensity in people reporting recovered memories of past lives. *Journal of Abnormal Psychology*, 118, 399-404.
- Munnukka, J., Maity, D., Reinikainen, H., & Luoma-aho, V. (2019). “Thanks for watching”. The effectiveness of YouTube vlogendorsements. *Computers in human behavior*, 93, 226-234.
- Murphy, G., & Flynn, E. (2021). Deepfake false memories. *Memory*, 1-13.
- Nash, R. A. (2018). Changing beliefs about past public events with believable and unbelievable doctored photographs. *Memory*, 26(4), 439-450.
- Newman, E. J., Garry, M., Bernstein, D. M., Kantner, J., & Lindsay, D. S. (2012). Nonprobative photographs (or words) inflate truthiness. *Psychonomic Bulletin & Review*, 19(5), 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(5), 1337-1348.

- Newman, E. J., & Schwarz, N. (2018). Good sound, good research: How audio quality influences perceptions of the research and researcher. *Science Communication*, 40(2), 246-257.
- Newman, E. J., & Zhang, L. (2020). Truthiness: How non-probative photos shape belief. In R. Greifeneder, M. Jaffé, E. J. Newman, & N. Schwarz (Eds.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 90–114). London: Routledge.
- Nightingale, S. J., Wade, K. A., & Watson, D. G. (2017). Can people identify original and manipulated photos of real-world scenes?. *Cognitive Research: Principles and Implications*, 2(1), 1-21.
- Oreskes, N., & Conway, E. M. (2010). Defeating the merchants of doubt. *Nature*, 465(7299), 686-687.
- Parks, C. M., & Toth, J. P. (2006). Fluency, familiarity, aging, and the illusion of truth. *Aging, Neuropsychology, and Cognition*, 13(2), 225-253.
- Petty, R. E., & Wegener, D. T. (1998). Matching versus mismatching attitude functions: Implications for scrutiny of persuasive messages. *Personality and Social Psychology Bulletin*, 24(3), 227-240.
- 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.
- Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology*, 34(2), 243-281.
- 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., & 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(3), 338-342.

- Roozenbeek, J., Schneider, C. R., Dryhurst, S., Kerr, J., Freeman, A. L., Recchia, G., ... & Van Der Linden, S. (2020). Susceptibility to misinformation about COVID-19 around the world. *Royal Society Open Science*, 7(10), 201199.
- Rubin, D. C. (1977). Very long-term memory for prose and verse. *Journal of Verbal Learning and Verbal Behavior*, 16(5), 611-621.
- Schwarz, N., Jalbert, M., Noah, T., & Zhang, L. (2021). Metacognitive experiences as information: Processing fluency in consumer judgment and decision making. *Consumer Psychology Review*, 4(1), 4-25.
- Seli, P., Beaty, R. E., Cheyne, J. A., Smilek, D., Oakman, J., & Schacter, D. L. (2018). How pervasive is mind wandering, really?. *Consciousness and Cognition*, 66, 74-78.
- Shafto, M. & MacKay, D. G. (2000). The Moses, mega-Moses, and Armstrong illusions: Integrating language comprehension and semantic memory. *Psychological Science*, 11, 372-378.
- Skurnik, I., Yoon, C., Park, D. C., & Schwarz, N. (2005). How warnings about false claims become recommendations. *Journal of Consumer Research*, 31(4), 713-724.
- Stanley, M.L., Whitehead, P.S., Marsh, E. J., & Seli, P. (2021). *Prior exposure increases judged truth even during periods of mind wandering* [Manuscript submitted for publication]. Department of Psychology & Neuroscience, Duke University.
- 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.
- Stanley, M. L., Yang, B. W., Stone, A. R., & Marsh, E. J. (2021). *Asymmetry in belief revision* [Manuscript submitted for publication]. Department of Psychology & Neuroscience, Duke University.
- 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. (2015a). Lies, damn lies, and expectations: How base rates inform lie–truth judgments. *Applied Cognitive Psychology*, 29(1), 149-155.

Street, C. N., & Richardson, D. C. (2015b). 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*(3), 409-415.

Sundar, A., Kardes, F. R., & Wright, S. A. (2015). The influence of repetitive health messages and sensitivity to fluency on the truth effect in advertising. *Journal of Advertising, 44*(4), 375-387.

Taylor, L. E., Swerdfeger, A. L., & Eslick, G. D. (2014). Vaccines are not associated with autism: An evidence-based meta-analysis of case-control and cohort studies. *Vaccine, 32*(29), 3623-3629.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *The organization of memory* (pp. 381-403). Academic Press.

Underwood, J., & Pezdek, K. (1998). Memory suggestibility as an example of the sleeper effect. *Psychonomic Bulletin & Review, 5*(3), 449-453. Unkelbach, C. (2007). Reversing the truth effect: Learning the interpretation of processing fluency in judgments of truth. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 219–230.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. (2nd ed.). Chichester, England: Wiley.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380), 1146-1151.

Wallsten, T. S., Budescu, D. V., Zwick, R., & Kemp, S. M. (1993). Preferences and reasons for communicating probabilistic information in verbal or numerical terms. *Bulletin of the Psychonomic Society, 31*(2), 135-138.

Wojdynski, B. W., & Evans, N. J. (2016). Going native: Effects of disclosure position and language on the recognition and evaluation of online native advertising. *Journal of Advertising, 45*(2), 157-168.