

*In press at Annual Review of Psychology*

**Judging Truth**

Nadia M. Brashier

Harvard University

Elizabeth J. Marsh

Duke University

Correspondence concerning this article should be addressed to Nadia M. Brashier, Department of Psychology, Harvard University, Cambridge, MA. Email: [nbrashier@fas.harvard.edu](mailto:nbrashier@fas.harvard.edu)

### Abstract

Deceptive claims surround us, embedded in fake news, advertisements, political propaganda, and rumors. How do people know what to believe? Truth judgments reflect inferences drawn from three types of information: base rates, feelings, and consistency with information retrieved from memory. First, people exhibit a bias to accept incoming information, because most claims in our environments are true. Second, people interpret feelings, like ease of processing, as evidence of truth. And third, people can (but do not always) consider whether assertions match facts and source information stored in memory. This three-part framework predicts specific illusions (e.g., truthiness, illusory truth), offers ways to correct stubborn misconceptions, and suggests the importance of converging cues in a post-truth world, in which falsehoods travel further and faster than the truth.

*Keywords:* inference, truth bias, illusory truth, fluency, knowledge, source

### Judging Truth

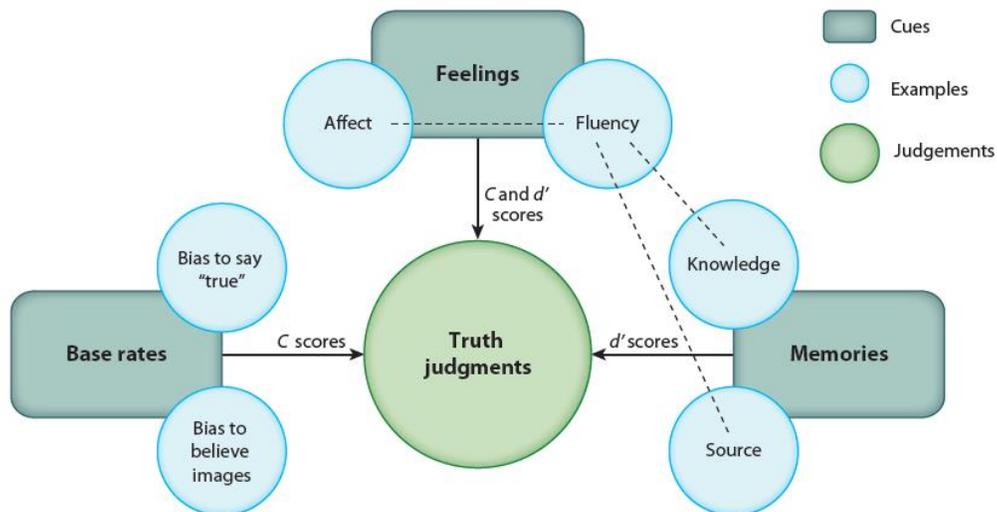
How do we know what to believe? *A camel's hump stores water. Albert Einstein failed math in school. Suicide rates peak during the holidays.* Most people believe these to be facts, though camel humps store fat, Einstein excelled at math, and suicide rates rise in the spring. Conversely, people also reject veridical information. *An octopus has three hearts. Anne Frank and Martin Luther King, Jr were born in the same year. The unicorn is Scotland's national animal.* Why do these facts *feel* false? This review describes the cognitive and affective cues that signal truth, how well these cues track reality, and the ways that they leave people vulnerable to specific and predictable illusions.

Public interest in the psychology of judging truth intensified after the 2016 U.S. presidential election; *post-truth*, *fake news*, and *misinformation* have each appeared as the word of the year in prominent dictionaries. Many worry that we now live in a world where opinions outweigh facts (see Lewandowsky, Ecker, & Cook, 2017). Misconceptions spread rapidly in the digital age – falsehoods are 70% more likely to be retweeted than truths (Vosoughi, Roy, & Aral, 2018) – but misleading headlines are as old as the printing press. In 1912, the *St Louis Dispatch* reported that most Titanic passengers survived (in fact, over 1,500 died). In 1981, the *New York Times* described a “rare cancer” seen in gay men (a disease now known as AIDS). In 1990, the *New York Times* also predicted an epidemic of “crack-exposed children” (unlike fetal alcohol syndrome, no discernible “crack baby” phenotype exists). Fortunately, scientific interest in the spread of misinformation also predates the fake news crisis, allowing the application of basic psychological science to contemporary controversies.

Decades of work from distinct research traditions describe basic processes that people use to distinguish true from false claims. By “true” and “false,” we refer to people’s judgments about objective truth, not attitudes, and thus we do not cover the persuasion literature (e.g., Albarracín & Shavitt, 2018; Falk & Scholz, 2018). Past reviews summarize how feelings, like fluency (Dechêne, Stahl, Hansen, & Wänke, 2010; Unkelbach, 2019) and affect (Forgas, 2019), shape truth. Our synthesis offers a different level of analysis, cutting across social and cognitive psychology to consider the many ways in which people construct truth.

### The Construction of Truth

Existing research reveals a common theme: Truth judgments are *constructed*. People draw inferences from relevant cues, some of which qualify as *heuristics* (i.e., cognitive shortcuts that save time and effort). We organize the literature into three main kinds of inference (see Figure 1). First, and most fundamentally, truth judgments reflect **inferences from base rates**. In line with Bayesian models of cognition (Tenenbaum, Griffiths, & Kent, 2006), people do not start from a position of complete ignorance, but instead begin with prior probabilities. Second, people draw **inferences from feelings**. Feelings-as-information theory suggests that people interpret their own subjective experiences as evidence of truth (Schwarz, 2012). Finally, people draw **inferences from consistency** with existing knowledge and source information stored in memory, complementing a referential theory of truth (Unkelbach & Rom, 2017).



*Figure 1.* Truth judgments reflect inferences from base rates, feelings, and consistency with memories. *C* scores measure the general bias to say “true,” while *d'* scores reflect the ability to distinguish true from false claims. Dashed lines indicate interactions described by previous research. Additional cues and connections may exist.

Each of these three inferences increases accuracy in general. A Bayesian-like reliance on base rates (i.e., calling claims “true” more often than “false”) is often effective, given that most information encountered in daily life is true. Statements accompanied by a feeling of ease are more likely to be true than those that feel strange or difficult to understand. And probabilistically, assertions consistent with information stored in memory tend to be more accurate than mismatches.

Specific kinds of errors follow from each class of inference. Assuming truth due to base rates can leave people credulous. Relying on feelings may predispose people to cognitive illusions, where statements that feel good are accepted irrespective of truth. And preferring consistency can make it difficult to update beliefs when the facts change (e.g., spicy foods ease, rather than exacerbate, digestive inflammation) or when learning something new (e.g., ink from tattoos migrates to the lymph nodes).

### **Inferring Truth from Base Rates**

*To get divorced, Ghanaians have to wear their wedding clothes to court.* Without any expertise about Ghana’s legal system, how do readers judge this assertion? We might predict that people accept ambiguous claims as true as often as they reject them as false. Instead, people are more likely to assume a statement is true than to call it false, reflecting the base rates in daily life (where most references are mundane and true). From a Bayesian perspective, it is rational to assume that incoming information is true, and then to revise in light of new evidence. Gilbert (1991) makes similar predictions, but argues that, to be comprehended, a claim must first be represented as true. In a second, resource-demanding step, it can be rejected as false.

“Unbelieving” (i.e., revising or updating) requires cognitive effort, leaving people vulnerable to distraction. In one study, participants learned supposed translations of Hopi words (e.g., *A monishna is a star*), each followed by a “true” or a “false” tag. A tone played just after some of these tags appeared. Asymmetrical errors emerged: Interruption led participants to misremember false claims as “true,” but the opposite did not occur (Gilbert, Krull, & Malone, 1990). A similar effect occurred when participants read crime reports and then recommended prison terms for robberies (Gilbert, Tatarodi, and Malone, 1993). The reports contained true and false information, printed in black and red font, respectively. Half

of the participants searched for digits while reading them. Distracted participants assigned terms consistent with details (e.g., *The robber had a gun*) that they should have ignored (i.e., that appeared in red). Multi-tasking appears to disrupt the second, unbelieving stage.

Without explicit tags, people exhibit a modest bias to accept new claims. In our own experiments, participants rarely endorse ambiguous claims seen for the first time as “definitely true” (e.g., Fazio, Brashier, Payne, & Marsh, 2015; Brashier, Eliseev, & Marsh, under review; Brashier, Umanath, Cabeza, & Marsh, 2017). However, their judgments consistently skew towards “true” (across experiments, the mean deviation from the middle of a 6-point scale was 0.30). Signal detection analysis provides a more precise estimate, where bias scores indicate participants’ tendency to say “true” (Unkelbach, 2007). A lenient criterion leaves listeners vulnerable to deceit, since people lie regularly (on average, 1.65 times a day; Serota, Levine, & Boster, 2010). Most people detect dishonesty at a rate barely better than chance (Bond & DePaulo, 2006; Hartwig & Bond, 2011), due to a bias to regard others’ statements as truthful (Bond & DePaulo, 2008).

People also observe and draw on base rates when considering images, as most things seen in daily life are real. Movies, virtual reality, and other visual media all take advantage of this reality principle. As a result, doctored photos can implant memories of childhood events (e.g., hot air balloon rides) that never happened (Wade, Garry, Read, & Lindsay, 2002). The fiction about Ghanaian divorce law mentioned earlier went viral, perhaps due to the accompanying photoshopped picture of stressed couples in tuxedos and wedding gowns. The power of pictures extends to truth, as demonstrated by Newman and colleagues (2012). They presented people with claims like *The first windmills were built in Persia*. Half appeared with a photo (e.g., of a windmill in a nondescript field) that provided no actual evidence about the veracity of the claim. Signal detection analysis revealed a bias to accept statements as true (i.e., lower *C* scores) when they appeared alongside a picture. *Truthiness* persists for days (Fenn, Newman, Pezdek, & Garry, 2013) and occurs unconsciously: Most participants (90%) fail to notice that pictures boost perceptions of truth (Newman, Azad, Lindsay, & Garry, 2018). In addition to making claims more believable, irrelevant photos increase people’s desire to share both true and false information on social

media (Fenn, Ramsay, Kantner & Pezdek 2019). Manipulated visuals exploit our reliance on images; at the extreme, “deepfakes” use artificial intelligence to depict events that never happened (e.g., speeches by world leaders; Yang, Li, & Lyu, 2018).

People can learn to disbelieve their eyes, such as in experiments in which people wear goggles that render the world upside down, but it takes extensive practice (Kohler, 1962). The aphorism that *seeing is believing* captures the Bayesian notion that what people see is usually true. However, there are situations in which this contingency is broken. For example, only semantically related photos make statements seem truer; pairing *The liquid metal inside a thermometer is mercury* with a picture of a lizard actually encourages a bias to say “false” (Newman et al., 2015). Obvious mismatches (e.g., thermometer – lizard) override inferences from base rates.

### **Inferring Truth from Feelings**

*Vitamin C prevents common colds.* Consumers continue to buy unnecessary supplements, partly because they hear this claim over and over. Repetition works because people behave like *cognitive misers*, using shortcuts to avoid computation (Fiske & Taylor, 1991). Heuristics minimize time and effort, whether predicting where serial offenders live (Snook, Taylor, & Bennell, 2004) or who will win tennis matches (Serwe & Frings, 2006). Despite ignoring some information, they can match or exceed the accuracy of statistical methods (i.e., *less-is-more effects*; Gigerenzer & Gaissmaier, 2011). People draw on adaptive toolboxes suited to a given domain, not general algorithms (Gigerenzer, 2002). For example, cooperating first and then imitating a social partner’s last behavior (i.e., going *tit for tat*; Axelrod, 1984) facilitates cooperation, but this strategy cannot discriminate truths from falsehoods. So how do people make “fast and frugal” truth judgments? One well-documented shortcut involves ease of processing.

*Just do it. Think different. Eat fresh.* These slogans immediately bring Nike, Apple, and Subway to mind. Marketers realize the power of repetition, a strategy that Hasher, Goldstein, and Toppino (1977) corroborated. In their seminal experiment, participants judged the truth of claims like *Divorce is only found in technically advanced societies*. Crucially, they saw some statements three times (*repeated*) and

others for the first time at test (*new*). Repeated claims seemed truer than new ones, a phenomenon coined *illusory truth*. According to a meta-analysis of over 50 studies, illusory truth is a medium effect (Cohen's  $d_s = 0.39-0.49$ ; Dechêne et al., 2010). Notably, this may be an underestimate, as most researchers instruct participants that they will encounter true and false information. Removing this simple warning, which rarely occurs in everyday life, doubles the effect of repetition (Jalbert, Newman, & Schwarz, unpublished manuscript).

Even a single previous exposure to a claim proves powerful. Illusory truth persists over time, emerging months later (Brown & Nix, 1996). Repetition boosts credibility of trivia (e.g., *House mice can run an average of four miles per hour*, Bacon, 1979), product claims (e.g., *Crest toothpaste removes caffeine stains from teeth*, Johar & Roggeveen, 2007), sociopolitical opinions (e.g., *Judges are far too lenient on criminals*; Arkes, Hackett, & Boehm, 1989), rumors (e.g., *A professor was giving a student good grades because he found out the professor plagiarized*; DiFonzo, Beckstead, Stupak, & Walders, 2016), and fake headlines (e.g., *Mike Pence: Gay conversion therapy saved my marriage*, Pennycook, Cannon, & Rand, 2018). Merely seeing concepts (e.g., hen's body temperature) increases later belief in detailed claims about those topics (e.g., *The temperature of a hen's body is about 104°F*; Begg, Armour, & Kerr, 1985). Reading statements like *Crocodiles sleep with their eyes open* even leads people to believe the opposite: A week later, direct contradictions (e.g., *Crocodiles sleep with their eyes closed*) seem more truthful than new items (Garcia-Marques, Silva, Reber, & Unkelbach, 2015). People forget the details of the initial claim, instead basing their judgments on how easy it feels to process key concepts (e.g., crocodiles, sleep).

Bolstering this fluency account, illusory truth occurs without repetition. Statements presented in high contrast (e.g., *The capital of Madagascar is Toamasina*) seem truer than those presented in low contrast (e.g., *The capital of Madagascar is Toamasina*; Reber & Schwarz, 1999). Aphorisms that rhyme (e.g., *What sobriety conceals, alcohol reveals*) seem more apt than those that do not (e.g., *What sobriety conceals, alcohol unmasks*; McGlone, & Tofighbakhsh, 2000). Finally, claims (e.g., *A giraffe can go without water longer than a camel can*) made by native speakers seem truer than those spoken with a

foreign accent (Lev-Ari & Keysar, 2010). Neuroimaging provides converging evidence; illusory truth reflects increased activity in the perirhinal cortex, a region implicated in other fluency effects like conceptual priming (Wang, Brashier, Wing, Marsh, & Cabeza, 2016).

This illusion reflects a *relative* metacognitive experience. Illusory truth effects are largest when people judge mixed lists that include both new and repeated statements (Dechêne, Stahl, Hansen, & Wänke, 2009; Garcia-Marques, Silva, Mello, & Hansen, 2019). We might expect the opposite to be true, given that a within-subjects design gives participants the opportunity to explicitly notice differences between fluent and disfluent items. When judging liking (Bornstein & D'Agostino, 1994) and frequency (Oppenheimer, 2004), for example, people spontaneously explain away fluency when manipulations are too heavy-handed. However, seeing repeated and new items side by side does not lead to *discounting* while evaluating truth. Participants do not simply forget claims seen earlier; they recognize statements from the exposure phase with accuracy exceeding 90% (Begg, Anas, & Farinacci, 1992). Rather, the effect is so strong that they fail to discount even after explicit warnings describing illusory truth (Nadarevic & Aßfalg, 2017).

How do people develop a heuristic this pernicious? In some ways, the term *illusory truth* is a misnomer – there can be “wisdom in feelings” (Schwarz, 2002). Fluency naturally correlates with truth in our daily lives. On average, people hear the single true version of a statement (e.g., *The capital of Argentina is Buenos Aires*) more often than any one of its many possible falsifications (e.g., *The capital of Argentina is La Paz*, *The capital of Argentina is Lima*, *The capital of Argentina is Montevideo*). With experience, people learn that fluency typically leads to the correct judgment in less time than other strategies (Unkelbach, 2007), allowing them to judge repeated statements more quickly than new ones (e.g., Scholl, Greifeneder, & Bless, 2014; Unkelbach & Greifeneder, 2018). Given that the fluency heuristic is learned, it can also be reversed. Participants update when feedback challenges the direction (Unkelbach, 2007) or validity (Scholl et al., 2014) of the relationship between color contrast and truth.

Cognitive, social, and consumer psychology converge on fluency's potential to lead us astray (Unkelbach, Koch, Silva, & Garcia-Marques, 2019). This empirical spotlight implies that fluency serves

as a single “clever” cue for truth. Indeed, some judgments reflect “one good reason” (Gigerenzer & Gaissmaier, 2011). Fluency is an appealing candidate – easy processing informs many judgments (Alter & Oppenheimer, 2009), including liking (Iyengar & Lepper, 2000), beauty (Reber, Schwarz, & Winkielman, 2004), and confidence (Schwartz & Metcalfe, 1992). University professors even assign higher grades in later offerings of a course, and judges give higher ratings to professional dancers in later seasons of *Dancing with the Stars* (O’Connor & Cheema, 2018). Fluency serves as a powerful cue, but other subjective feelings provide shortcuts to truth.

Most notably, emotion “bleeds into” many judgments (Greifeneder, Bless, & Pham, 2011; Lerner, Li, Valdesolo, & Kassam, 2015), from risk estimates (Johnson & Tversky, 1983) to judgments of learning (Hourihan & Bursey, 2017). Affect can literally make mountains out of molehills: Sadness leads people to estimate that hills are steeper (Riener, Stefanucci, Proffitt, & Clore, 2011). Mood cued by weather correlates with real-world outcomes, such as medical school admissions (Redlmeier & Baxter, 2009) and stock market returns (Hirshleifer & Shumway, 2003).

We ask ourselves *How do I feel about this?* and then attribute our feelings to the target (Schwarz, 2012). Easy processing feels good, according to both self-reports (Monahan, Murphy, & Zajonc, 2000) and facial myography over the zygomaticus “smiling” muscle (Carr, Rotteveel, & Winkielman, 2016; Winkielman & Cacioppo, 2001). As a result, faces seen recently appear happier than new ones (Carr, Brady & Winkielman, 2017). The reverse is also true: Smiling faces feel more familiar (an interpretation of fluency) than neutral ones (Garcia-Marques, Mackie, Claypool, & Garcia-Marques, 2004). Happy, but not sad, people endorse attitudes based on easy-to-retrieve (i.e., fluent) arguments (Ruder & Bless, 2003). Compared to a neutral mood, happiness leads people to falsely recognize words (Claypool, Hall, Mackie, & Garcia-Marques, 2008). Moreover, people enjoy repeating activities (e.g., visiting a museum, watching a movie) more than expected (O’Brien, 2019).

Given this warm glow of familiarity, Unkelbach and colleagues (2011) investigated whether illusory truth merely reflects positivity. Participants evaluated statistical claims with positive (*The divorce rate in Grenada is lower than in the rest of Spain*) and negative (*The divorce rate in Grenada is*

*higher than in the rest of Spain*) framings. Illusory truth emerged for both types of statements, confirming that fluency drives the effect. But claims that elicited more positive feelings also seemed truer, hinting that positivity independently cues truth. Conversely, sadness wipes out illusory truth (Koch & Forgas, 2012), which is an impressive finding given how robust the illusion is. After watching sad film clips, claims (e.g., *Instead of iron, horseshoe crabs have copper in their blood*) printed in an easy-to-read font seem just as true as those in a hard-to-read font. Negative mood also reduces gullibility, allowing people to spot deception (Forgas & East, 2008). In our own studies, people tend to judge claims as truthful when they appear beside a neutral, but not an angry or fearful, face (Brashier & Marsh, unpublished manuscript). In short, a bad mood may discourage people from going with their guts (Forgas, 2019).

### **Inferring Truth from Consistency with Memory**

*Barack Obama was the first black president of the United States.* Accepting this statement as true does not dictate whether you *like* Obama. Unlike subjective evaluations (e.g., confidence, liking), there is usually a right answer when making truth judgments; they exist “before the background of an objective value” (Dechêne et al., 2010, p. 254). A rational model conceives of truth judgments as evidence-based: People should accept information as true when it matches content retrieved from memory, including relevant facts (semantic memories) or details about source (episodic memories).

Humans know a lot – on average, the meaning of 42,000 words (Brysbaert, Stevens, Paweł Mandra, & Keuleers, 2016) and the faces of 5,000 people (Jenkins, Dowsett, & Burson, 2018). We draw on knowledge to understand the world, whether resolving ambiguities (e.g., Bransford & Johnson, 1972) or simulating events in the future (e.g., Benoit, Szpunar, & Schacter, 2014). Similarly, people retrieve what they know to evaluate whether a claim is true. They tend to accept claims that fit with facts stored in memory and to reject mismatches. For example, participants perceive *Ojos del Salado is the highest mountain in South America* to be truer than *The Nile is the longest river in South America* (e.g., Brashier et al., 2017; Fazio et al., 2015). Both statements are false, but people know less about the Aconcagua than the Amazon. Of course, people also hold misconceptions about the world (e.g., that the Great Wall of

China is visible from space), confuse opinions and facts (Pew Research Center, 2018), and claim to know impossible things. This *overclaiming* predicts belief in fake news (Pennycook & Rand, 2019d) and illusions of understanding (e.g., about genetics) co-occur with extreme beliefs (e.g., fear of genetically modified organisms; Fernbach, Light, Scott, Inbar, & Rozin, 2019).

Even when knowledge is objective and accurate, people may neglect it (Marsh & Umanath, 2014). Participants offer solutions to impossible problems (e.g., where to bury survivors of a plane crash; Barton & Sanford, 1993), overlook errors in stories (e.g., St. Petersburg as Russia's capital; Marsh & Fazio, 2006), and answer questions containing false premises (e.g., *How many animals of each kind did Moses take on the ark?*; Erickson & Mattson, 1981). Listeners pass over falsehoods that resemble the truth (Hinze, Slaten, Horton, Jenkins, & Rapp, 2014; van Oostendorp & de Mul, 1990) because errors plague ordinary speech. Messages only need to be "good enough" (Ferreira, Bailey, & Ferraro, 2002), so people accept partial matches between statements and the contents of memory (Reder & Kusbit, 1991).

Fluency further impairs our ability to catch inconsistencies. People notice fewer errors when questions like *How many animals of each kind did Moses take on the ark?* are easy to read (Song & Schwarz, 2008). Moreover, repetition makes contradictions of well-learned facts seem more credible. Intuitively, repeating *Deer meat is called veal* should not increase belief; most people know that deer meat is called venison, not veal. The literature reflects this assumption that knowledge protects us; a meta-analysis of illusory truth notes that statements must be ambiguous for the illusion to occur (Dechêne et al., 2010). Similarly, Unkelbach and Stahl (2009) and others used obscure trivia, "minimize knowledge" (e.g., *Cactuses can procreate via pathogenesis*), assuming that knowledge would wipe out the effect of repetition. Our work demonstrates the opposite. Participants read statements that contradicted well-known (e.g., *Newton proposed the theory of relativity*) and obscure facts (e.g., *Bell invented the wireless radio*), then they rated these and new statements' truthfulness. A final knowledge check determined which specific facts each participant knew. Repetition inflated judgments of falsehoods, regardless of whether or not they contradicted stored knowledge (Fazio, Brashier, Payne, & Marsh, 2015).

Multinomial modeling confirmed that fluency supersedes knowledge (in agreement with a *fluency-conditional model*). The model assumed in the literature, where fluency only comes into play when people lack knowledge (*knowledge-conditional model*), fit poorly. In short, people sometimes rely on fluency when they “know better” (Fazio et al., 2015), to the point where repetition may even increase belief in implausible claims, like *Smoking cigarettes is good for your lungs* and *The Earth is a perfect square* (Fazio, Rand, & Pennycook, in press). Thus, education only offers a partial solution to the misinformation crisis; people might learn new facts (e.g., waterboarding impairs memory) only to disregard them later (e.g., believing that torture works).

The picture looks just as dismal when we consider general intellect, rather than knowledge of specific facts: Illusory truth is immune to individual differences in fluid intelligence and cognitive style (De Keersmaecker et al., 2019). Developmental evidence also refutes the idea that cognitive resources protect people: Fluid intelligence declines by late adulthood (Hartshorne & Germine, 2015), yet older adults are equally or more discerning than their young counterparts. Repeating obscure claims like *The smallest insect species is the adelgid* misleads young and older adults to a similar extent. However, older adults spontaneously stick with what they know. They reject *Deer meat is called veal*, even when it feels fluent (Brashier, Umanath, Cabeza, & Marsh, 2017). Older adults take longer to judge truth, raising the possibility that simply slowing down benefits them.

Carefully considering claims helps in some situations and backfires in others. In a classic levels-of-processing experiment, participants reported whether statements appeared on the left or right side of the screen (*shallow processing*), indicated where verbs were missing (*deep processing*), or related statements to a personal event or feeling (*deepest processing*, associated with self-reference). Later, they judged the truth of these and new items. Surprisingly, illusory truth *increased* with depth of encoding (Unkelbach & Rom, 2017). Elaborative processing not only failed to reverse the illusion, it actually enhanced it. A more useful approach prompts people to behave like “fact checkers.” In our studies, participants initially judged how interesting or truthful claims were. An initial accuracy focus (initial truth ratings) eliminated illusory truth later, but only when participants had relevant knowledge stored in

memory. Without knowledge, people fell back on fluency regardless of how they processed statements at exposure (Brashier et al., 2020).

Much like stored knowledge, episodic memories (of specific past experiences) provide a basis for comparison when judging truth. For example, recalling that a claim came from a low-credibility person or publication (i.e., *source memory*) is informative. Correspondingly, statements that participants remember hearing from an “untrustworthy” voice seem less true than new ones (i.e., reverse illusory truth effect; Begg, Anas, and Farinacci, 1992). This pattern complements a *referential theory* of truth, where believability reflects activation of nodes in an information network. When the perceived credibility of a statement matches the credibility of its source, these coherent references increase perceived truth (Unkelbach & Rom, 2017). Occasionally, though, people neglect source information. As examples, participants exhibit illusory truth while receiving deterministic advice (from a person labeled as 100% accurate; Unkelbach & Greifeneder, 2018) and for statements they actually identify as coming from an unreliable source (Henkel & Mattson, 2011).

Of course, sources rarely appear with clear labels in daily life. People struggle to assess the quality of sources (Mitchell & Johnson, 2009), relying on shortcuts like the presence of in-text citations (Putnam & Phelps, 2017), the pronounceability of a stranger’s name (Newman et al., 2014), and repetition (with fluent claims misattributed to reputable publications; Fragale & Heath, 2004). We also tend to trust others, even complete strangers (Dunning, Fetchenhauer, & Schlösser, in press). Our willingness to trust poses a problem when social partners have goals other than accuracy – for example, participants are willing to share fake news that they identify as false (Pennycook et al., 2018). Furthermore, even if people notice that a source is questionable, they often forget these episodic details. Without recollective encoding, supported by activity in the hippocampus and ventrolateral prefrontal cortex, claims initially tagged as “false” can appear credible later (Mitchell, Dodson, & Schacter, 2005). Once people fail to recollect whether a statement came from a trustworthy source, they draw inferences from fluency (Unkelbach & Stahl, 2009).

Disregarding knowledge or forgetting source information poses new dangers in the digital age (Marsh & Rajaram, 2019). Typing keywords into the Google search bar is fast and easy; relying on the Internet becomes so habitual that people search for answers to easy questions instead of simply retrieving them from memory (e.g., *What is the center of a hurricane called?* Storm, Stone, & Benjamin, 2017). But search algorithms return content based on keywords, not truth. If you search *flat Earth*, for example, Google dutifully returns photoshopped pictures of a 150-foot wall of ice that keeps us from slipping off the planet. To make matters worse, users rarely read articles before sharing – 59% of shared links on Twitter are not clicked on first (Gabiolkov, Ramachandran, Chaintreau, & Legout, 2016). When readers do make it to the actual article, subtle misinformation in the title shapes their impressions (Ecker, Lewandowsky, Chang, & Pillai, 2014). Even mainstream news outlets sometimes use clickbait headlines (e.g., *Power causes brain damage* published by *The Atlantic*) that mislead readers.

### **Summary**

We argue that the data patterns across very different literatures point to *constructive processes* in judging truth. People rely on base rates, which improves accuracy in general, but can increase gullibility. They interpret subjective experiences like fluent processing and affect as evidence – feelings that correlate with truth but can also prove illusory. And they draw inferences from consistency with information stored in memory. Stored knowledge and memories for source can be completely diagnostic of truth, but people often neglect, misremember, or forget them. These inferences bear on the real world, where misinformation causes disease (e.g., measles outbreaks due to vaccine hesitation), harms the environment (e.g., poaching due to myths about rhino horns curing cancer), and encourages violence (e.g., against rumored child abductors described on Whatsapp). Next, we consider the implications of our three-part constructive approach to truth for correcting stubborn misconceptions.

### **Correcting Misconceptions**

Ideally, we would prevent misconceptions from taking hold in the first place, as they are notoriously difficult to correct (Cohen's  $d_s = 0.75-1.06$ ; Chan, Jones, Jamieson, & Albarracín, 2017). Even when people successfully correct myths (e.g., *Playing Mozart can improve a baby's intelligence*) in

the short term, they struggle to do so after time passes (Swire, Ecker, & Lewandowsky, 2017). The trouble is that people concurrently store corrections and the original misinformation, as indicated by activity in the left angular gyrus and bilateral precuneus (Gordon, Quadflieg, Brooks, Ecker, & Lewandowsky, 2019); the newer correction is forgotten at a faster rate than the older misconception (according to Jost's law; Wixted, 2004). In addition, debunking messages reach fewer people than the original misinformation. For example, Snopes debunked the claim that Nancy Pelosi agreed to a border wall in exchange for a gun ban, but the fake news story received nearly 20 times more engagements on Facebook than the correction. This problem is a familiar one for psychologists: Original studies continue to be cited, despite high-profile failures to replicate them.

Given that correction is so tricky, what suggestions can our framework offer? First, we can appeal to people's experiences with *base rates*. Adding visuals to corrections may make them more compelling, inducing truthiness. Second, messages should avoid reinforcing *feelings* associated with myths. Simply negating misinformation (or presenting it alongside the truth in a myth-fact format) makes it fluent (Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012). In addition, people do not fully process negatives (e.g., *not*). In a classic example, an experimenter poured sugar into two jars, one labeled *sucrose* and the other labeled *not sodium cyanide*. Participants preferred not to drink Kool-Aid made with sugar from the *not sodium cyanide* jar (Rozin, Markwith, & Ross, 1990). Thus, the Centers for Disease Control and Prevention's well-meaning message that *vaccines do not cause autism* may reinforce a vaccine-autism link. However, the advice to *replace rather than repeat* makes corrections less salient (Ecker, Hogan, & Lewandowsky, 2017) and poses a challenge when the jury is still out (e.g., advanced paternal age is only one possible cause of autism).

In these situations, we can leverage people's desire for *consistency* with their knowledge and beliefs. Exposure to opposing viewpoints (e.g., about gay rights) can polarize, rather than moderate, views (Bail et al., 2018). To get around this concern, Gehlbach, Robinson, and Vriesema (2019) manipulated whether or not participants answered questions like *How credible is the medical data that germs are a primary cause of disease?* before judging the credibility of climate science data. Belief in

medicine exceeds belief in climate science; most people agree that germs cause disease. Once they acknowledged the value of other scientific fields, conservatives were more likely to endorse climate science.

Another intriguing possibility is to use disfluency to cue analytic thinking (Alter, Oppenheimer, Epley, & Erye, 2007). People experience knee jerk reactions when information favors their opinions (Gilead, Sela, & Maril, 2019) and lazy thinking sometimes prevents them from rejecting intuitive, but incorrect, responses. In contrast, partisans who perform well on the cognitive reflection test (i.e., analytic thinkers) more accurately discern fake from real headlines, even when they align with their politics (Pennycook & Rand, 2019b). So how do we encourage analytic thinking? In one study, participants read an essay favoring capital punishment in an easy- or hard-to-read font, then judged whether the message seemed reliable, intelligent, and believable. After easy reading, participants' personal beliefs determined their impressions. This preference for consistency disappeared after difficult reading (Hernandez & Preston, 2013). Thus, presenting myths in a disfluent format may help people to set aside their opinions.

Finally, the mainstream advice to *consider the source* misses the mark (Marsh & Yang, 2017). Trust indicators piloted by Google, Facebook, and Twitter tell readers about the quality of publications – but these well-intentioned projects overestimate people's ability to keep track of sources. Efforts by fact checkers (e.g., Politifact, FactCheck.org) to flag viral claims (e.g., with a Truth-o-Meter) may even backfire; tagging some fake news stories as false boosts the perceived accuracy of untagged ones (*implied truth effect*; Pennycook & Rand, 2019c). A better solution uses crowd-sourced judgments about source trustworthiness, which prove accurate (Pennycook & Rand, 2019a) and could serve as inputs to social media ranking algorithms that determine what people see. Of course, censoring or manipulating content may anger users (see public outcry to Kramer, Guillory, & Hancock, 2014).

## **Conclusion**

Gartner, Inc. estimates that people will consume more false than true information by 2022, a frightening possibility that is consistent with trends on social media; falsehoods already outrace the truth

on Twitter (Vosoughi et al., 2018). But if fluency no longer naturally correlates with truth, what other cues will people turn to? Psychologists know a lot about fluency – it shapes perceived truth over long delays, among intelligent people, despite contradictory knowledge, for claims coming from “unreliable” sources, and in the face of diagnostic advice. But fluency is only one way to infer truth; people also draw inferences from other feelings, base rates, and consistency with what they know, remember, and believe.

This broader framework suggests the need for more complete models of truth. Current multinomial models pit two cognitive processes against each other. In one influential model, Unkelbach and Stahl (2009) demonstrated that people rely on fluency when they forget source information. We also focus on two processes in our *fluency-conditional model*, positing that people search memory for relevant knowledge when they experience disfluency (Fazio et al., 2015). While valuable, these models offer little insight into situations in which many cues converge. Inside a car dealership, for example, a buyer might encounter familiar slogans like *Engineered to move the human spirit (fluency)* accompanied by photos (*base rate*), retrieve facts about fuel economy (*knowledge*), note that the salesman wants to make commission (*source*), and feel excitement (*affect*). How do they evaluate whether claims about a car are true, and ultimately decide whether to buy it? If researchers consider multiple, simultaneous cues, superordinate heuristics might appear. As examples, the customer could weigh all five cues equally or *tally* reasons that a claim seems true versus false (Bobadilla-Suarez & Love, 2018); alternatively, they might “take the best” (Newell & Shanks, 2003), falling back on one good reason (e.g., their own emotions) and ignoring other cues (e.g., the salesman’s motives). Minimizing information may be especially tempting under distraction or for older adults, who consider less information (Meyer, Russo, & Talbot, 1995) and prefer fewer options (Reed, Mikels, & Simon, 2008) when making decisions. Moving forward, the field needs to consider the intersecting cognitive, affective, and social processes that make falsehoods believable in a post-truth world.

### Summary Points

1. In everyday life, accurate claims outnumber inaccurate ones. People exhibit a bias to accept incoming information (i.e., judge claims to be “true”) that reflects these base rates.
2. Subjective feelings convey useful information about the world. Thus, we infer truth from feelings like easy processing and our own affect.
3. Relevant facts and memories about the source of a claim can be completely diagnostic of the truth. As a result, people believe statements that match information retrieved from memory and reject inconsistencies.
4. Considering these three classes of inference – from base rates, feelings, and consistency – can improve approaches to correcting misconceptions.

## Future Issues

1. Shortcuts other than fluency (e.g., affect), as well as reliance on external memory aids like Google, are understudied.
2. Interactions between cues (e.g., affect and source) also deserve more empirical attention (see Figure 1).
3. Cues for truth and falseness may not be simple inverses, existing on a continuum. Strategies analogous to *recalling-to-reject*, for example, lend themselves to labeling claims as false, but not true.
4. Few data speak to how older adults judge truth, and even fewer address childhood. A lifespan perspective is crucial, as older adults shared the most fake news in the 2016 election (Grinberg et al., 2019; Guess, Nagler, & Tucker, 2019).
5. Most existing studies use behavioral measures, but cognitive neuroscience can probe processes unavailable to conscious awareness.
6. Current multinomial models of truth pit two processes (e.g., fluency and knowledge) against each other. Modeling three or more simultaneous cues could reveal higher order heuristics (e.g., *tallying*, *take-the-best*).

Acknowledgments

We thank Keith Payne and Emmaline Drew Eliseev for their comments on an earlier draft. A National Science Foundation SBE Postdoctoral Research Fellowship supports N.M.B.

## Literature Cited

- Albarracin D, Shavitt S. 2018. Attitudes and attitude change. *Annu. Rev. Psychol.* 69:299-327
- Alter AL, Oppenheimer DM. 2009. Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review* 13:219-35
- Alter AL, Oppenheimer DM, Epley N, Eyre RN. 2007. Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General* 136:569-76
- Arkes HR, Hackett C, Boehm L. 1989. The generality of the relation between familiarity and judged validity. *Journal of Behavioral Decision Making* 2:81-94
- Axelrod R. (1984). *The evolution of cooperation*. New York: Basic Books
- Bacon FT. 1979. Credibility of repeated statements: Memory for trivia. *Journal of Experimental Psychology: Human Learning and Memory* 5:241-52
- Bail CA, Argyle LP, Brown TW, Bumpus JP, Chen H, Fallin Hunzaker MB, ... Volfovsky A. 2018. Exposure to opposing views on social media can increase political polarization. *PNAS* 115:9216-21
- Barton SB, Sanford AJ. 1993. A case study of anomaly detection: Shallow semantic processing and cohesion establishment. *Memory & Cognition* 21:477-87
- Begg I, Anas A, Farinacci S. 1992. Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General* 121:446-58
- Begg I, Armour V, Kerr T. 1985. On believing what we remember. *Canadian Journal of Behavioural Science* 17:199-214
- Benoit RG, Szpunar KK, Schacter DL. 2014. Ventromedial prefrontal cortex supports affective future simulation by integrating distributed knowledge. *PNAS* 111:16550-5
- Bobadilla-Suarez S, Love BC. 2018. Fast or frugal, but not both: Decision heuristics under time pressure. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 44:24-33
- Bond CF, DePaulo, BM. 2006. Accuracy of deception judgments. *Personality and Social Psychology Review*, 10, 214-34

- Bond CF, DePaulo BM. 2008. Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin* 134:477-92
- Bornstein RF, D'Agostino PR. 1994. The attribution and discounting of perceptual fluency: Preliminary tests of a perceptual fluency/ attributional model of the mere exposure effect. *Social Cognition* 12:103-28
- Bransford JD, Johnson MK. 1972. Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior* 11:717-26
- Brashier NM, Eliseev ED, Marsh EJ. 2020. An initial accuracy focus prevents illusory truth. *Cognition*
- Brashier NM, Marsh EJ. unpublished manuscript. Positivity does not cue truth.
- Brashier NM, Umanath S, Cabeza R, Marsh, EJ. 2017. Competing cues: Older adults rely on knowledge in the face of fluency. *Psychology and Aging* 32:331-7
- Brown AS, Nix LA. 1996. Turning lies into truths: Referential validation of falsehoods. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22:1088-100
- Brybaert M, Stevens M, Mandera P, Keuleers E. 2016. How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. *Frontiers in Psychology* 7:116
- Carr EW, Brady TF, Winkielman P. 2017. Are you smiling or have I seen you before? Familiarity makes faces look happier. *Psychological Science* 28: 1087-102
- Carr EW, Rotteveel M, Winkielman P. 2016. Easy moves: Perceptual fluency facilitates approach-related action. *Emotion*, 16: 540-52
- Chan MS., Jones CR, Jamieson KH, Albarracín D. 2017. Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science* 28:1531-46
- Claypool HM, Hall CE, Mackie DM, Garcia-Marques T. 2008. Positive mood, attribution, and the illusion of familiarity. *Journal of Experimental Social Psychology* 44:721-8

- De Keersmaecker J, Dunning D, Pennycook G, Rand DG, Sanchez C, Unkelbach C, Roets A. 2019. Investigating the robustness of the illusory truth effect across individual differences in cognitive ability, need for cognitive closure, and cognitive style. *Personality and Social Psychology Bulletin*
- Dechêne A, Stahl C, Hansen J, Wänke M. 2009. Mix me a list: Context moderates the truth effect and the mere-exposure effect. *Journal of Experimental Social Psychology* 45:1117-22
- 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* 14:238-57
- DiFonzo N, Beckstead JW, Stupak N, Walders K. 2016. Validity judgments of rumors heard multiple times: The shape of the truth effect. *Social Influence* 11:22-39
- Dunning D, Fetchenhauer D, Schlösser T. 2019. Why people trust: Solved puzzles and open mysteries. *Current Directions in Psychological Science* 28:366-71
- Ecker UKH, Hogan JL, Lewandowsky S. 2017. Reminders and repetition of misinformation: Helping or hindering its retraction? *Journal of Applied Research in Memory and Cognition* 6:185-92
- Ecker UK, Lewandowsky S, Chang EP, Pillai R. 2014. The effects of subtle misinformation in news headlines. *Journal of Experimental Psychology: Applied* 20:323-5
- Erickson TD, Mattson ME. 1981. From words to meaning: A semantic illusion. *Journal of Verbal Learning & Verbal Behavior*, 20:540-51
- Falk E, Scholz C. 2018. Persuasion, influence, and value: Perspectives from communication and social neuroscience. *Annual Review of Psychology* 69:329-56
- Fazio LK, Brashier NM, Payne BK, Marsh EJ. 2015. Knowledge does not protect against illusory truth. *Journal of Experimental Psychology: General* 144:993-1002
- Fazio LK, Rand DG, Pennycook G. in press. Repetition increases perceived truth equally for plausible and implausible statements. *Psychonomic Bulletin & Review*
- Fenn E, Newman EJ, Pezdek K, Garry M. 2013. The effect of nonprobative photographs on truthiness persists over time. *Acta Psychologica* 144:207-11

- 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*, 8:131-8
- Fernbach PM, Light N, Scott SE, Inbar Y, Rozin P. 2019. Extreme opponents of genetically modified foods know the least but think they know the most. *Nature Human Behavior*
- Ferreira F, Bailey KGD, Ferraro V. 2002. Good-enough representations in language comprehension. *Current Directions in Psychological Science* 11:11-5
- Fiske ST, Taylor SE. 1991. *Social cognition* (2nd ed). New York: McGraw-Hill
- Forgas JP. 2019. Happy believers and sad skeptics? Affective influences on gullibility. *Current Directions in Psychological Science* 28:306-13
- Forgas JP, East R. 2008. On being happy and gullible: Mood effects on skepticism and the detection of deception. *Journal of Experimental Social Psychology*, 44:1362-7
- Fragale AR, Heath C. 2004. Evolving information credentials: The (mis)attribution of believable facts to credible sources. *Personality and Social Psychology Bulletin* 30:225-36
- Gabielkov M, Ramachandran A, Chaintreau A, Legout A. 2016. Social clicks: What and who gets read on Twitter? *ACM SIGMETRICS Performance Evaluation Review* 44:179-92
- Garcia-Marques T, Mackie DM, Claypool HM, Garcia-Marques L. 2004. Positivity can cue familiarity. *Personality and Social Psychology Bulletin* 30:585-93
- Garcia-Marques T, Silva RR, Mello J, Hansen J. 2019. Relative to what? Dynamic updates of fluency standards and between-participants illusions of truth. *Acta Psychologica* 195:71-9
- Garcia-Marques T, Silva RR, Reber R, Unkelbach C. 2015. Hearing a statement now and believing the opposite later. *Journal of Experimental Social Psychology* 56:126-29
- Gehlbach H, Robinson CD, Vriesema CC. 2019. Leveraging cognitive consistency to nudge conservative climate change beliefs. *Journal of Environmental Psychology* 61:134-37
- Gigerenzer G. 2002. The adaptive toolbox. In G. Gigerenzer & R. Selten (Eds), *Bounded rationality: The adaptive toolbox* (pp. 37-50). Cambridge: MIT Press

- Gigerenzer G, Gaissmaier W. 2011. Heuristic decision making. *Annual Review of Psychology* 62:451-82
- Gilbert DT. 1991. How mental systems believe. *American Psychologist* 46:107–19
- Gilbert DT, Krull D, Malone P. 1990. Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology* 59:601–13
- Gilbert DT, Tafarodi RW, Malone PS. 1993. You can't not believe everything you read. *Journal of Personality and Social Psychology* 65:221–33
- Gilead M, Sela M, Maril A. 2019. That's my truth: Evidence for involuntary opinion confirmation. *Social Psychological and Personality Science* 10:393-401
- Gordon A, Quadflieg S, Brooks JCW, Ecker UKH, & Lewandowsky S. 2019. Keeping track of 'alternative facts': The neural correlates of processing misinformation corrections. *Neuroimage* 93:46-56
- Greifeneder R, Bless H, Pham MT. 2011. When do people rely on affective and cognitive feelings in judgment? A review. *Personality and Social Psychology Review* 15:107-41
- Grinberg N, Joseph K, Friedland L, Swire-Thompson B, Lazer D. 2019. Fake news on Twitter during the 2016 U.S. presidential election. *Science* 25:374-8
- Guess A, Nagler J, Tucker J. 2019. Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances* 5
- Hartshorne JK, Germine JT. 2015. When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychological Science* 26:433-443
- Hartwig M, Bond CF. 2011. Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin*, 137:643-59
- Hasher L, Goldstein D, Toppino T. 1977. Frequency and the conference of referential validity. *Journal of Verbal Learning & Verbal Behavior* 16:107-12
- Henkel LA, Mattson ME. 2011. Reading is believing: The truth effect and source credibility. *Consciousness and Cognition* 20:1705-21

- Hernandez I, Preston JL. 2013. Disfluency disrupts the confirmation bias. *Journal of Experimental Social Psychology* 49:178-182
- Hinze SR, Slaten DG, Horton WS, Jenkins R, Rapp DN. 2014. Pilgrims sailing the Titanic: Plausibility effects on memory for misinformation. *Memory & Cognition* 42:305-24
- Hirshleifer D, Shumway T. 2003. Good day sunshine: Stock returns and the weather. *Journal of Finance* 58:1009-32
- Hourihan KL, Bursey E. 2017. A misleading feeling of happiness: Metamemory for positive emotional and neutral pictures. *Memory* 25: 35-43
- Iyengar SS, Lepper MR. 2000. When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology* 79:995-1006
- Jalbert M, Newman EJ, Schwarz N. unpublished manuscript. Only half of what I'll tell you is true: How experimental procedures lead to an underestimation of the truth effect.
- Jenkins R, Dowsett AJ, Burton AM. 2018. How many faces do people know? *Proceedings of the Royal Society B* 285:20181319
- Johar GV, Roggeveen AL. 2007. Changing false beliefs from repeated advertising: The role of claim-refutation alignment. *Journal of Consumer Psychology* 17:118-27
- Johnson E, Tversky A. 1983. Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology* 45:20-31
- Koch A, Forgas JP. 2012. Feeling good and feeling truth: The interactive effects of mood and processing fluency on truth judgments. *Journal of Experimental Social Psychology* 48:481-5
- Kohler I. 1962. Experiments with goggles. *Scientific American* 206:62-86
- Kramer ADI, Guillory JE, Hancock JT. 2014. Experimental evidence of massive-scale emotional contagion through social networks. *PNAS* 111:8788-90
- Lerner JS, Li Y, Valdesolo P, Kassam KS. 2015. Emotion and decision making. *Annual Review of Psychology* 66:799-823

- 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* 46:1093-6
- Lewandowsky S, Ecker UKH, Cook J. 2017. Beyond misinformation: Understanding and coping with the "post-truth" era. *Journal of Applied Research in Memory and Cognition* 6: 353-69
- Lewandowsky S, Ecker UKH, Seifert CM, Schwarz N, Cook J. 2012. Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest* 13: 106-31
- Marsh EJ, Fazio LK. 2006. Learning errors from fiction: Difficulties in reducing reliance on fictional stories. *Memory & Cognition* 34:1140-9
- Marsh EJ, Rajaram S. 2019. The digital expansion of the mind: Implications of Internet usage for memory and cognition. *Journal of Applied Research in Memory and Cognition*
- Marsh EJ, Umanath S. 2014. Knowledge neglect: Failures to notice contradictions with stored knowledge. In D.N. Rapp and J. Braasch (Eds), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 161-80). Cambridge: MIT Press
- Marsh EJ, Yang BW. 2017. A call to think broadly about information literacy. *Journal of Applied Research in Memory and Cognition* 6:401-4
- McGlone MS, Tofigbakhsh J. 2000. Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. *Psychological Science* 11:424-8
- Meyer BJ, Russo C, Talbot A. 1995. Discourse comprehension and problem solving: Decisions about the treatment of breast cancer by women across the life span. *Psychology and Aging* 10:84-103
- Mitchell JP, Dodson CS, Schacter DL. 2005. fMRI evidence for the role of recollection in suppressing misattribution errors: The illusory truth effect. *Journal of Cognitive Neuroscience* 17:800-10
- Mitchell KJ, Johnson MK. 2009. Source monitoring 15 years later: What have we learned from fMRI about the neural mechanisms of source memory? *Psychological Bulletin* 135:638-77

- Monahan JL, Murphy ST, Zajonc RB. 2000. Subliminal mere exposure: Specific, general, and diffuse effects. *Psychological Science* 11:462-6
- Nadarevic L, Aßfalg A. 2016. Unveiling the truth: Warnings reduce the repetition-based truth effect. *Psychological Research* 81:1-13
- Newell BR, Shanks DR. 2003. Take the best or look at the rest? Factors influencing “one reason” decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29:53-65
- Newman EJ, Azad T, Lindsay DS, Garry M. 2018. Evidence that photos promote rosiness for claims about the future. *Memory & Cognition* 46:1223-33
- Newman EJ, Garry M, Bernstein DM, Kantner J, Lindsay DS. 2012. Nonprobative photographs (or words) inflate truthiness. *Psychonomic Bulletin & Review* 19: 969-74
- Newman EJ, Garry M, Unkelbach C, Bernstein DM, Lindsay DS, Nash RA. 2015. Truthiness and falsiness of trivia claims depend on judgmental contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 41: 1337-48
- Newman EJ, Sanson M, Miller EK, Quigley-McBride A, Foster JL, Bernstein DM, Garry M. 2014. People with easier to pronounce names promote truthiness of claims. *PLoS One* 9:e88671
- O’Brien E. 2019. Enjoy it again: Repeat experiences are less repetitive than people think. *Journal of Personality and Social Psychology* 116:519-40
- O’Connor K, Cheema, A. 2018. Do evaluations rise with experience? *Psychological Science* 29:779-90
- Oppenheimer D. 2004. Spontaneous discounting of availability in frequency judgment tasks. *Psychological Science* 15:100-5
- Pennycook G, Cannon TD, Rand, DG. 2018. Prior exposure increases perceived accuracy of fake news but has no effect on entirely implausible statements. *Journal of Experimental Psychology: General* 147: 1865-80
- Pennycook G, Rand DG. 2019a. Fighting misinformation on social media using crowdsourced judgments of news source quality. *PNAS*

- Pennycook G, Rand DG. 2019b. Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*
- Pennycook G, Rand DG. 2019c. The implied truth effect: Attaching warnings to a subset of fake news stories increases perceived accuracy of stories without warnings. *Management Science*
- Pennycook G, Rand DG. 2019d. Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of Personality*
- Pew Research Center. 2018. Distinguishing between factual and opinion Statements in the News. Retrieved from [www.journalism.org/2018/06/18/distinguishing-between-factual-and-opinion-statements-in-the-news](http://www.journalism.org/2018/06/18/distinguishing-between-factual-and-opinion-statements-in-the-news)
- Putnam AL, Phelps RJ. 2017. The citation effect: In-text citations moderately increase belief in trivia claims. *Acta Psychologica* 179:114-123
- Reber R, Schwarz N. 1999. Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition* 8:338-42
- Reber R, Schwarz N, Winkielman P. 2004. Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review* 8:364-82
- Reder LM, Kusbit GW. 1991. Locus of the Moses illusion: Imperfect encoding, retrieval, or match? *Journal of Memory and Language* 30:385-406
- Redlmeier DA, Baxter, SD. 2009. Rainy weather and medical school admission interviews. *Canadian Medical Association Journal* 181:933
- Reed AE, Mikels JA, Simon KI. 2008. Older adults prefer less choice than young adults. *Psychology and Aging* 23:671-5
- Riener CR, Stefannuci JK, Proffitt DR, Clore G. 2011. An effect of mood on the perception of geographical slant. *Cognition and Emotion* 25:174-82
- Rozin P, Markwith M, Ross B. 1990. The sympathetic magical law of similarity, nominal realism and neglect of negatives in response to negative labels. *Psychological Science* 1:383-4

- Ruder, M., & Bless, H. 2003. Mood and the reliance on the ease of retrieval heuristic. *Journal of Personality and Social Psychology* 85:20-32
- Scholl SG, Greifeneder R, Bless H. 2014. When fluency signals truth: Prior successful reliance on fluency moderates the impact of fluency on truth judgments. *Journal of Behavioral Decision Making* 27:268–80
- Schwartz BL, Metcalfe J. 1992. Cue familiarity but not target retrievability enhances feeling-of-knowing judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 18:1074-83
- Schwarz N. 2002. Situated cognition and the wisdom in feelings: Cognitive tuning. In L. F. Barrett & P. Salovey (Eds), *Emotions and social behavior. The wisdom in feeling: Psychological processes in emotional intelligence* (pp. 144-66). New York: The Guilford Press
- Schwarz N. 2012. Feelings-as-information theory. In P. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds), *Handbook of theories of social psychology* (pp. 289-308). Thousand Oaks, CA: Sage Publications
- Serota KB, Levine TR, Boster FJ. 2010. The prevalence of lying in America: Three studies of self-reported lies. *Human Communication Research* 36:2-25
- Serwe S, Frings C. 2006. Who will win Wimbledon? The recognition heuristic in predicting sports events. *Journal of Behavioral Decision Making* 19:321-32
- Snook B, Taylor PJ, Bennell C. 2004. Geographic profiling: The fast, frugal, and accurate way. *Applied Cognitive Psychology* 18:105-21
- Song HJ, Schwarz N. 2008. Fluency and the detection of misleading questions: Low processing fluency attenuates the Moses Illusion. *Social Cognition* 26: 791-9
- Storm BC, Stone SM, Benjamin AS. 2017. Using the Internet to access information inflates future use of the Internet to access other information. *Memory* 25:717-23
- Swire B, Ecker UKH, Lewandowsky S. 2017. The role of familiarity in correcting inaccurate information. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 43:1948-61

- Tenenbaum JB, Griffith TL, Kemp C. 2006. Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Science* 10:309-18
- 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* 33:219-30
- Unkelbach C, Bayer M, Alves H, Koch A, Stahl C. 2011. Fluency and positivity as possible causes of the truth effect. *Consciousness and Cognition* 20:594-602
- Unkelbach C, Greifeneder R. 2018. Experiential fluency and declarative advice jointly inform judgments of truth. *Journal of Experimental Social Psychology* 79:78-86
- Unkelbach C, Koch A, Silva RA, Garcia-Marques T. 2019. Truth by repetition: Explanations and implications. *Current Directions in Psychological Science*
- Unkelbach C, Rom SC. 2017. A referential theory of the repetition-induced truth effect. *Cognition* 160:110-26
- Unkelbach C, Stahl C. 2009. A multinomial modeling approach to dissociate different components of the truth effect. *Consciousness and Cognition* 18:22-38
- van Oostendorp H, de Mul S. 1990. Moses beats Adam: A semantic relatedness effect on a semantic illusion. *Acta Psychologica* 74:35-46
- Vosoughi S, Roy D, Aral S. 2018. The spread of true and false news online. *Science* 359:1146-51
- Wade KA, Garry M, Read JD, Garry M. 2002. A picture is worth a thousand lies: Using false photographs to create false childhood memories. *Psychonomic Bulletin & Review* 9:597-603
- Wang W, Brashier NM, Wing EA, Marsh EJ, Cabeza R. 2016. On known unknowns: Fluency and the neural mechanisms of the illusory truth effect. *Journal of Cognitive Neuroscience* 28:739-46
- Winkielman P, Cacioppo JT. 2001. Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology* 81:989-1000

Wixted JT. 2004. On common ground: Jost's (1897) law of forgetting and Ribot's (1881) law of retrograde amnesia. *Psychological Review* 111:864-79

Yang X, Li Y, Lyu S. 2018. Exposing deep fakes using inconsistent head poses. arXiv:1811.00661 [cs.CV]

## Annotated References

- Bond CF, DePaulo, BM. 2006. Meta-analysis concludes that people detect deception at rates close to chance.
- Brashier NM, Umanath S, Cabeza R, Marsh EJ. 2017. Older adults reject fluent falsehoods that contradict their knowledge.
- Dechêne A, Stahl C, Hansen J, Wänke M. 2010. Meta-analysis confirms that repetition increases perceived truth.
- Fazio LK, Brashier NM, Payne BK, Marsh EJ. 2015. Illusory truth occurs even when claims contradict well-known facts.
- Forgas JP, East R. 2008. Sad mood wipes out illusory truth.
- Gilbert DT. 1991. Proposes that people accept claims in order to comprehend them.
- Hasher L, Goldstein D, Toppino T. 1977. First demonstration of illusory truth.
- Pennycook G, Rand DG. 2019b. Lazy thinking, not motivated reasoning, explains belief in fake news.
- Unkelbach. 2007. People can reverse the fluency heuristic with feedback.
- Vosoughi S, Roy D, Aral S. 2018. Falsehoods diffuse faster than the truth on Twitter.