5

WHEN (FAKE) NEWS FEELS TRUE

Intuitions of truth and the acceptance and correction of misinformation

Norbert Schwarz and Madeline Jalbert

An analysis of 2.8 million episodes of news sharing on Twitter found that 59% of the news items were shared without having been opened (Gabielkov, Ramachandran, Chaintreau, & Legout, 2016). Apparently, six out of ten readers found the headline compelling enough to share the piece without reading it. In this chapter, we review what makes a message "feel" true, even before we have considered its content in any detail. We first discuss the basic psychological processes involved in assessing the truth of a message and illustrate them with select experiments. Subsequently, we address the implications of these processes for information sharing on social media and the correction of misinformation.

Evaluating truth

While retweeting something without reading it may strike many readers as surprising and irresponsible, it is not distinctly different from how we communicate in everyday life. In daily conversations, we proceed on the tacit assumption that the speaker is a cooperative communicator whose contributions are relevant to the ongoing conversation, truthful, informative, and clear (Grice, 1975; Sperber & Wilson, 1986). Unless we have reason to doubt that the speaker observes these tacit rules of conversational conduct, we accept the content of the utterance without much questioning and treat it as part of the common ground of the conversation. These conversational processes contribute to many errors in human judgment (for reviews, see Schwarz, 1994, 1996). Some research even suggests that comprehension of a statement requires at least temporary acceptance of its truth (Gilbert, 1991) before it can be checked against relevant evidence.

While suspension of belief is possible (Hasson, Simmons, & Todorov, 2005; Schul, Mayo, & Burnstein, 2008), it requires implausibility of the message or distrust at the time it is received. Hence, the deck is usually stacked in favor

of accepting information rather than rejecting it, provided there are no salient markers that call the speaker's cooperativeness into question. Going beyond the default of information acceptance requires motivation and cognitive resources, which we are most likely to invest when the topic is important to us and there are few competing demands and distractions. In the absence of these conditions, information is likely to be accepted – and sometimes passed on – without much scrutiny.

When people do evaluate whether information is likely to be true, they typically consider some (but rarely all) of the five criteria shown in Table 5.1 (Schwarz, 2015). Is the claim compatible with other things they know? Is it internally consistent and coherent? Does it come from a trustworthy source? Do other people agree with it? Is there much evidence to support it? Each of these criteria is sensible and does, indeed, bear on the likely truth of a message. These criteria can be assessed by considering relevant knowledge, which is a relatively slow and effortful process and may require extensive information search. The same criteria can also be assessed by relying on one's intuitive response, which is faster and less taxing. When the initial intuitive response suggests that something may be wrong, people are likely to turn to the more effortful analysis, provided time and circumstances allow for it. This makes initial intuitive assessments of truth a key gatekeeper for whether people will further engage with the message using a critical eye or just nod along in agreement. These assumptions are compatible with a long history of research in social (e.g., Petty & Cacioppo, 1986) and cognitive (e.g., Kahneman, 2011; Stanovich, 1999) psychology, where the slow and effortful strategy is often referred to as "analytic", "systematic", or "system 2"

TABLE 5.1 Truth criteria

Criterion	*Analytic evaluation*	*Intuitive evaluation*
Compatibility: Is it compatible with other things I know?	Is this compatible with knowledge retrieved from memory or obtained from trusted sources?	Does this make me stumble or does it flow smoothly?
Coherence: Is it internally coherent?	Do the elements fit together in a logical way? Do the conclusions follow from what is presented?	Does this make me stumble or does it flow smoothly?
Credibility: Does it come from a credible source?	Does the source have the relevant expertise? Does the source have a vested interest? Is the source trustworthy?	Does the source feel familiar and trustworthy?
Consensus: Do other people believe it?	What do my friends say? What do the opinion polls say?	Does it feel familiar?
Evidence: Is there supporting evidence?	Is there supportive evidence in peer-reviewed scientific articles or credible news reports? Do I remember relevant evidence?	Does some evidence easily come to mind?

processing and the fast and intuitive strategy as "intuitive", "heuristic", or "system 1" processing.

Key to intuitive assessments of truth is the ease with which the message can be processed. For example, when something is incompatible with other things we know or the story we are told is incoherent, we stumble and backtrack to make sure we understood it correctly (Johnson-Laird, 2012; Winkielman, Huber, Kavanagh, & Schwarz, 2012). This makes the subjective experience of ease of processing, often referred to as processing fluency, a (fallible) indicator of whether the message may have a problem that needs closer attention. Similar considerations apply to the other truth criteria, as discussed later in the chapter. Throughout, difficult processing marks the message for closer scrutiny, whereas easy processing favors message acceptance.

If ease or difficulty of processing was solely determined by attributes substantively associated with whether a message is likely to be true, relying on one's processing experience would not pose a major problem. However, messages can be easy or difficult to process for many reasons – reading may be slow because the message is incoherent (a relevant criterion) or because the print font is hard to read (which is unrelated to truth). Because people are more sensitive to their subjective experiences than to the source of those experiences (Schwarz, 2012), many incidental influences that have no bearing on the substance of the message can influence its perceived truth. We discuss these incidental influences and their role in media consumption after reviewing the five dominant truth criteria. As will become apparent, when thoughts flow smoothly, people are likely to agree without much critical analysis (see also Oyserman & Dawson, this volume).

The "big five" of truth judgment: analytic and intuitive processes

A claim is more likely to be accepted as true when it is *compatible* with other things one knows than when it is at odds with other knowledge. Compatibility can be assessed analytically by checking the information against one's knowledge, which requires motivation and time (Petty & Cacioppo, 1986). A less demanding indicator is provided by one's metacognitive experiences and affective responses. When something is inconsistent with existing beliefs, people tend to stumble – they take longer to read it, and have trouble processing it (e.g., Taber & Lodge, 2006; Winkielman et al., 2012). Moreover, information that is inconsistent with one's beliefs produces a negative affective response, as shown in research on cognitive consistency (Festinger, 1957; Gawronski & Strack, 2012). Accordingly, one's processing experiences and affective responses can serve as (fallible) indicators of whether a proposition is consistent with other things one believes.

A given claim is also more likely to be accepted as true when it fits a broader story that lends *coherence* to its individual elements, as observed in research on mental models (for a review, see Johnson-Laird, 2012) and analyses of jury decision making (Pennington & Hastie, 1993). Coherence can be determined

through a systematic analysis of the relationships between different pieces of declarative information. Alternatively, it can be assessed by attending to one's processing experience: coherent stories are easier to process than stories with internal contradictions (Johnson-Laird, 2012), which makes ease of processing a (fallible) indicator of coherence. Indeed, people draw on their fluency experience when they evaluate how well things "go together" (Topolinski, 2012), as observed in judgments of semantic coherence (Topolinski & Strack, 2008, 2009) and syllogistic reasoning (Morsanyi & Handley, 2012).

Information is also more likely to be accepted as true when it comes from a credible and trustworthy source. As decades of persuasion research illustrates, evaluations of *source credibility* can be based on declarative information that bears, for example, on the communicator's expertise, education, achievement, or institutional affiliation and the presence or absence of conflicting interests (for reviews, see Eagly & Chaiken, 1993; Petty & Cacioppo, 1986). However, credibility judgments can also be based on feelings of familiarity. In daily life, people trust familiar others more than strangers (Luhmann, 1979), from personal interactions to e-commerce (Gefen, 2000). Familiarity resulting from previous encounters or even just repeatedly seeing pictures of a face is sufficient to increase perceptions of honesty and sincerity as well as agreement with what the person says (Brown, Brown, & Zoccoli, 2002; Weisbuch & Mackie, 2009). Similarly, the mere repetition of a name can make an unknown name seem familiar, making its bearer "famous overnight" (Jacoby, Woloshyn, & Kelley, 1989), which may also increase perceived expertise. Familiar people are also easier to recognize and remember, and their names become easier to pronounce with repeated encounters. Variables that influence the ease with which source information can be processed can therefore enhance the perceived credibility of the source. Indeed, a given claim is more likely to be judged true when the name of its source is easy to pronounce (Newman et al., 2014).

To assess the likely truth of a claim, people also consider whether others believe it – if many people agree, there's probably something to it. This *social consensus* (Festinger, 1954) criterion is central to many social influence processes and is sometimes referred to as the principle of "social proof" (Cialdini, 2009). As numerous studies indicated, people are more confident in their beliefs if they are shared by others (Newcomb, 1943; Visser & Mirabile, 2004), more likely to endorse a message if many others have done so as well (Cialdini, 2009), and place more trust in what they remember if others remember it similarly (Harris & Hahn, 2009; Ross, Buehler, & Karr, 1998). Conversely, perceiving dissent reliably undermines message acceptance, which makes reports on real or fabricated controversies an efficient strategy for swaying public opinion (Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012; Lewandowsky, Gignac, & Vaughan, 2013). To assess the extent of consensus, people may consult public opinion polls or ask their friends. Alternatively, they may rely on how "familiar" the belief feels – after all, one should have encountered popular beliefs, shared by many, more frequently than unpopular beliefs, held by few. Empirically, familiar

information is easier to read, understand, and remember than unfamiliar information, which makes ease of processing a (fallible) indicator of familiarity and popularity. Accordingly, incidental changes in ease of processing can influence perceived consensus.

Finally, people's confidence in a belief increases with the *amount of supporting evidence*. Support can be assessed through an external search, as in a scientific literature review or through recall of pertinent information from memory; in either case, confidence increases with the amount of supportive information. Alternatively, support can be gauged from how easy it is to find supportive evidence – the more evidence there is, the easier it should be to find some (in memory or in the literature). This lay theory is at the heart of Tversky and Kahneman's (1973) availability heuristic. Unfortunately, this heuristic can be misleading. If the only supportive piece of information comes to mind easily because it has been endlessly repeated or is very vivid and memorable, we may erroneously conclude that support is strong. Moreover, attention to *what* comes to mind and attention to the *ease* with which it does so will often lead to different conclusions. On the one hand, reliance on the substantive arguments brought to mind results in higher confidence the more arguments one retrieves or generates. On the other hand, reliance on ease of recall results in lower confidence the more arguments one tries to come up with because finding many arguments is difficult, which suggests that there probably aren't many (Haddock, Rothman, Reber, & Schwarz, 1999; for reviews, see Schwarz, 1998; Schwarz & Vaughn, 2002).

Regardless of which truth criteria people draw on, easily processed information enjoys an advantage over information that is difficult to process: it feels more familiar, more compatible with one's beliefs, more internally consistent, more widely held, better supported, and more likely to have come from a credible source. These inferences reflect that familiar, frequently encountered information and information that is coherent and compatible with one's knowledge is indeed easier to process than information that is not. Hence, ease of processing provides heuristically useful – but fallible – information for assessing how well a claim meets major truth criteria.

Making claims "feel" true

So far, our discussion highlighted that ease or difficulty of processing can result both from variables that are meaningfully related to key criteria of truth or from incidental influences. This is important for two reasons. From a research perspective, it allows researchers to manipulate processing fluency in ways that are independent of substantive characteristics of a message and its source. From an applied perspective, it highlights that claims can "feel" true merely because they are easy to process, which provides many opportunities for manipulation. Next, we review some of the most important variables that influence the ease or difficulty of message processing.

Repetition

Demagogues have known for millennia that truth can be created through frequent repetition of a lie – as Hitler put it, "Propaganda must confine itself to a few points and repeat them over and over again" (cited in Toland, 1976, p. 221). Empirical research supports demagogues' intuition. Studying wartime rumors, Allport and Lepkin (1945) found that the best predictor of whether people believed a rumor was the number of times they were exposed to it. Testing this observation in the laboratory, Hasher, Goldstein, and Toppino (1977) asked participants to rate their confidence that each of 60 statements was true. Some statements were factually correct (e.g., "Lithium is the lightest of all metals"), whereas others were not (e.g., "The People's Republic of China was founded in 1947"). Participants provided their ratings on three occasions, each two weeks apart. Across these sessions, some statements were repeated once or twice, whereas others were not, resulting in one, two, or three exposures. As expected, participants were more confident that a given statement was true the more often they had seen it, independent of whether it was factually true or false. Numerous follow-up studies confirmed the power of repetition across many content domains, from trivia statements (e.g., Bacon, 1979) to marketing claims (e.g., Hawkins & Hoch, 1992) and political beliefs (e.g., Arkes, Hackett, & Boehm, 1989), with the time delay between exposure and judgment ranging from minutes (e.g., Begg & Armour, 1991) to months (Brown & Nix, 1996). Dechêne, Stahl, Hansen, and Wänke (2010) provide a comprehensive meta-analysis of this "illusory truth" effect.

The influence of repetition is most pronounced for claims that people feel uncertain about, but is also observed when more diagnostic information about the claims is available (Fazio, Rand, & Pennycook, 2019; Unkelbach & Greifeneder, 2018). Worse, repetition even increases agreement among people who actually know that the claim is false – if only they thought about it (Fazio, Brashier, Payne, & Marsh, 2015). For example, repeating the statement "The Atlantic Ocean is the largest ocean on Earth" increased its acceptance even among people who knew that the Pacific is larger. When the repeated statement felt familiar, they nodded along without checking it against their knowledge. Even warning people that some of the claims they will be shown are false does not eliminate the effect, although it attenuates its size. More importantly, warnings only attenuate the influence of repetition when they *precede* exposure to the claims – warning people *after* they have seen the claims has no discernable influence (Jalbert, Newman, & Schwarz, 2019).

Repetition also increases perceived social consensus, that is, the perception that a belief is shared by many others. Weaver, Garcia, Schwarz, and Miller (2007) had participants read opinion statements purportedly taken from a group discussion in which a given opinion was presented once or thrice. Each opinion statement was attributed to a group member. Not surprisingly, participants assumed that more people shared the opinion when they read it three times from

three different group members (72%) than when they read it only once (57%). However, reading the opinion three times from the *same* group member was almost as influential, resulting in a consensus estimate of 67% – apparently, the single repetitive source sounded like a chorus. Later studies showed that people trust an eyewitness report more the more often it is repeated, even when all repetitions come from the same single witness (Foster, Huthwaite, Yesberg, Garry, & Loftus, 2012). Similarly, newspaper readers are more confident in the accuracy of a report when the same message is presented in several newspapers, even if all newspapers solely rely on the same single interview with the same speaker (Yousif, Aboody, & Keil, 2019). Such findings suggest that frequent repetition of the same soundbite in TV news can give the message a familiarity that increases its perceived popularity and truth. This concern also applies to social media, where the same message keeps showing up as friends and friends of friends like it and repost it, resulting in many exposures within a network.

Beyond repetition

Despite its popularity with past and present demagogues, repetition is just one of many variables that can facilitate easy processing of a statement, making the statement appear more popular, credible, and true. Next, we review some of these other variables.

Reber and Schwarz (1999) manipulated the ease of reading through the *color contrast* of the print font. Depending on condition, some statements (e.g., 'Orsono is a city in Chile') were easy to read due to high color contrast (e.g., dark blue print on a white background), whereas others were difficult to read due to low color contrast (e.g., light blue print on a white background). As predicted, the same statement was more likely to be judged true when it was easy rather than difficult to read. Similarly, the readability of *print fonts* can influence intuitive assessments of truthfulness and the extent to which we closely scrutinize a message. For example, when asked, "How many animals of each kind did Moses take on the Ark?" most people answer "two" even though they know that the biblical actor was Noah, not Moses. Song and Schwarz (2008) presented this Moses question (taken from Erickson & Mattson, 1981) in one of the fonts shown in Figure 5.1. They warned participants that some of the questions may be misleading, in which case they should answer "Can't say". When the Moses question was presented in the easy to read black Arial font, 88% failed to notice a problem and answered "two", whereas only 53% did so when the question was presented in the more difficult to read gray Brush font.

Other variables that influence ease of processing have similar effects. For example, handwritten essays are more compelling when the *handwriting* is easy to read (Greifeneder et al., 2010) and so are spoken messages when the speaker's *accent* is easy to understand (Levy-Ari & Keysar, 2010). Similarly, the same conference talk is less impressive when its video recording has low *audio quality*, and a

Print font	**~ ~ n~ ~ ~ rin~ ~ it~ o' noti~ in~ ~ rro**
How many animals of each kind did Moses take on the Ark?	88%
How many animals of each kind did Moses take on the Ark?	~ ~%

FIGURE 5.1 Print font and the detection of misleading information

Source: Adapted from Song and Schwarz (2008), Experiment 1.

poor phone connection during a researcher's radio interview can impair listeners' impression of the quality of her research program (Newman & Schwarz, 2018). People also find a statement to be more true when presented with a version of it that *rhymes* rather than one that doesn't, even when the two versions are substantively equivalent (McGlone & Tofighbakhsh, 2000). Even a *photo* without any probative value can increase acceptance of a statement, provided the photo makes it easier to imagine what the statement is about (for a review, see Newman & Zhang, this volume).

Merely having a name that is easy to *pronounce* is sufficient to endow the person with higher credibility and trustworthiness. For example, consumers trust an online seller more when the seller's eBay username is easy to pronounce – they are more likely to believe that the product will live up to the seller's promises and that the seller will honor the advertised return policy (Silva, Chrobot, Newman, Schwarz, & Topolinski, 2017). Similarly, the same claim is more likely to be accepted as true when the name of its source is easy to pronounce (Newman et al., 2014).

As this selective review indicates, any variable that can influence ease of processing can also influence judgments of truth. This is the case because people are very sensitive to their processing experience but insensitive to where this experience comes from. When their attention is directed to the incidental source of their experience, the informational value of the experienced ease or difficulty is undermined and its influence attenuated or eliminated, as predicted by feelings-as-information theory (for reviews, see Schwarz, 2012, 2018).

Analytic versus intuitive processing

As in other domains of judgment, people are more likely to invest the time and effort needed for careful information processing when they are sufficiently motivated and have the time and opportunity to do so (for reviews, see Greifeneder, Bless, & Pham, 2011; Greifeneder & Schwarz, 2014). One may hope that this favors careful processing whenever the issue is important. However, this optimism may not be warranted. In the course of everyday life, messages about issues we consider personally important may reach us when we have other things

on our minds and lack the opportunity to engage with them. Over repeated encounters, such messages may become familiar and fluent enough to escape closer scrutiny even when the situation would allow us to engage with them. As reviewed previously, telling recipients that some of the information shown to them is false is only protective when the warning precedes the first exposure; later warnings show little effect (Jalbert et al., 2019). Similarly, the motivation and opportunity to examine a message critically may exert only a limited influence once the message has been encoded (for a review, see Lewandowsky et al., 2012).

Implications for social media

The dynamics of truth judgment have important implications for the acceptance and correction of false information in the real world. Beginning with the proliferation of cable TV and talk radio, citizens in democracies enjoyed ever more opportunities to selectively expose themselves to media that fit their worldview. The advent of social media is the latest step in this development and, in many ways, one might think that social media were designed to make questionable messages seem true. To begin with, most social media messages are short, written in simple language, and presented in optics that are easy to read, which satisfies many of the technical prerequisites for easy processing. These fluent messages are posted by one's friends, a credible source. The content they post is usually compatible with one's own beliefs, given the similarity of opinions and values in friendship networks (for a review of network homophily, see McPherson, Smith-Lovin, & Cook, 2001). Posted messages are liked by other friends, thus confirming social consensus, and reposted, thus ensuring multiple repeated exposures. With each exposure, processing becomes easier and perceptions of social consensus, coherence, and compatibility increase. Comments and related posts provide additional supporting evidence and further enhance familiarity. At the same time, the accumulating likes and reposts ensure that the filtering mechanism of the feed makes exposure to opposing information less and less likely. The *Wall Street Journal*'s "Blue Feed/Red Feed" site illustrates how Facebook's filtering mechanism resulted in profoundly different news feeds for liberals and conservatives during the 2016 elections in the United States, and a growing body of research traces how opinion homophily within networks contributes to controversies between networks (Del Vicario et al., 2016; Gargiulo & Gandica, 2017). The observed narrowing of recipients' information diet on social media is enhanced through the personalization of internet offerings outside of social media, where internet providers and search engines track users' interests to tailor information delivery (Pariser, 2011).

These processes not only increase the acceptance of claims that feel increasingly familiar and compatible with what else one knows but also foster a high sense of expertise and confidence. After all, much of what one sees in one's feed is familiar, which suggests that one knows most of what there is to know about

the topic. It has also been seen without much opposing evidence, suggesting that the arguments are undisputed. This enhances what Ross and Ward (1996) described as "naïve realism" – the belief that the world is the way I see it and whoever disagrees is either ill-informed (which motivates persuasion efforts) or ill-intentioned (if persuasion fails). These beliefs further contribute to polarization and the mutual attribution of malevolence.

Implications for the correction of misinformation

That people can arrive at judgments of truth by relying more on analytic or more on intuitive strategies poses a major challenge for public information campaigns aimed at correcting false beliefs. Extensive research in education shows that students' misconceptions can be corrected by confronting them with correct information, showing students step by step why one idea is wrong and another one right, preferably repeating this process multiple times (for reviews, see Vosniadou, 2008). This works best when the recipient wants to acquire the correct information and is sufficiently motivated to pay attention, think through the issues, and remember the new insights (for a review, see Sinatra & Pintrich, 2003). Public information campaigns often follow these procedures by confronting the "myths" with "facts", consistent with content-focused theories of message learning (McQuail, 2000; Rice & Atkin, 2001). While this works in the classroom, with motivated recipients, sufficient time, and the benefit of incentives, the reality of public information campaigns is starkly different. For any given topic, only a small segment of the population will care enough to engage with the details; most are likely to notice the message only in passing, if at all, and will process it superficially while doing something else. Even if they remember the corrective message as intended when tested immediately, it may fade quickly from memory.

Under such conditions, repeating false information in order to correct it may mostly succeed in spreading the false information to disinterested recipients who may otherwise never have encountered it. Not having processed the message in detail, they may now find the false claims a bit more familiar and easier to process when they hear or see them again. This way, the attempt to correct the erroneous beliefs of a few may prepare numerous others to accept those beliefs through repeated exposure (for a review, see Schwarz, Sanna, Skurnik, & Yoon, 2007). For example, Skurnik, Yoon, Park, and Schwarz (2005) exposed older and younger adults once or thrice to product statements like "Shark cartilage is good for your arthritis", and these statements were explicitly marked as "true" or "false". When tested immediately, the corrections seemed successful – all participants were less likely to accept a statement as true the more often they were told that it is false. This is the hoped-for success and most studies stop at this point. But after a three-day delay, repeated warnings backfired and older adults were now more likely to consider a statement "true", the more often they had been explicitly told that it is false. Presumably, the recipients could no longer recall whether the statement had been originally marked as true or false, but still

experienced repeated statements as easier to process and more familiar, which made the statements "feel" true.

Even exposing people to only true information can make it more likely that they accept a false version of that information as time passes. Garcia-Marques, Silva, Reber, and Unkelbach (2015) presented participants with ambiguous statements (e.g., "crocodiles sleep with their eyes closed") and later asked them to rate the truth of statements that were either identical to those previously seen or that directly contradicted them (e.g., "crocodiles sleep with their eyes open"). When participants made these judgments immediately, they rated repeated identical statements as more true, and contradicting statements as less true, than novel statements, which they had not seen before. One week later, however, identical as well as contradicting statements seemed more true than novel statements. Put simply, as long as the delay is short enough, people can recall the exact information they just saw and reject the opposite. As time passes, however, the details get lost and contradicting information feels more familiar than information one has never heard of – yes, there was something about crocodiles and their eyes, so that's probably what it was.

As time passes, people may even infer the credibility of the initial source from the confidence with which they hold the belief. For example, Fragale and Heath (2004) exposed participants two or five times to statements like "The wax used to line Cup-o-Noodles cups has been shown to cause cancer in rats". Next, participants learned that some statements were taken from the *National Enquirer* (a low credibility source) and some from *Consumer Reports* (a high credibility source) and had to assign the statements to their likely sources. The more often participants had heard a statement, the more likely they were to attribute it to *Consumer Reports* rather than the *National Enquirer.* In short, frequent exposure not only increases the apparent truth of a statement, it also increases the belief that the statement came from a trustworthy source. Similarly, well-intentioned efforts by the Centers for Disease Control and the *Los Angeles Times* to debunk a rumor about "flesh-eating bananas" morphed into the belief that the *Los Angeles Times* had warned people not to eat those dangerous bananas, thus reinforcing the rumor (Emery, 2000). Such errors in source attribution increase the likelihood that people convey the information to others, who themselves are more likely to accept (and spread) it, given its alleged credible source (Rosnow & Fine, 1976).

Such findings illustrate that attempts to correct misinformation can backfire when they focus solely on message content at the expense of the message's impact on recipients' later processing experience. Even when a corrective message succeeds in changing the beliefs of recipients who deeply care about the topic and process the message with sufficient attention, it may spread the false information to many others who don't care about the topic. Unfortunately, the latter are likely to outnumber the former. In those cases, the successful correction of a few false believers may come at the cost of misleading many bystanders. To avoid such backfire effects, it will usually be safer to refrain from any reiteration

of false information and to focus solely on the facts. The more the facts become familiar and fluent, the more likely it is that they will be accepted as true and serve as the basis of judgments and decisions (Lewandowsky et al., 2012; Schwarz et al., 2007, 2016).

Unfortunately, the truth is usually more complicated than false stories, which often involve considerable simplification. This puts the truth at a disadvantage because it is harder to process, understand, and remember. It is therefore important to present true information in ways that facilitate its fluent processing. This requires clear step-by-step exposition and the avoidance of jargon. It also helps to pay close attention to incidental influences on ease of processing. Making the font easy to read and the speaker's pronunciation easy to understand, adding photos and repeating key points are all techniques that should not be left to those who want to mislead – they can also give truth a helping hand and should be used.

Finally, at the individual level, the best protection against the influence of misinformation is skepticism at the time the information is first encountered (for a review, see Lewandowsky et al., 2012). Once people have processed the false information, warnings exert little influence. In addition to explicit warnings, general feelings of suspicion and distrust increase message scrutiny and decrease message acceptance (for reviews, see Mayo, 2017; Schwarz & Lee, 2019). Explicit warnings as well as suspicion and distrust entail that the communicator may not adhere to the norms of cooperative conversational conduct (Grice, 1975), thus flagging the message for closer scrutiny. Unfortunately, in a polarized public opinion climate, merely realizing that a message supports the "other" side is itself likely to elicit suspicion and distrust, further impairing correction attempts in polarized contexts.

Acknowledgments

Preparation of this chapter was supported by the Linnie and Michael Katz Endowed Research Fellowship Fund through a fellowship to the second author and funds of the USC Dornsife Mind and Society Center to the first author.

References

Allport, F. H., & Lepkin, M. (1945). Wartime rumors of waste and special privilege: Why some people believe them. *Journal of Abnormal and Social Psychology*, *40*, 3–36.

Arkes, H. R., Hackett, C., & Boehm, L. (1989). The generality of the relation between familiarity and judged validity. *Journal of Behavioral Decision Making*, *2*, 81–94.

Bacon, F. T. (1979). Credibility of repeated statements: Memory for trivia. *Journal of Experimental Psychology: Human Learning and Memory*, *5*, 241–252.

Begg, I., & Armour, V. (1991). Repetition and the ring of truth: Biasing comments. *Canadian Journal of Behavioural Science*, *23*, 195–213.

Brown, A. S., Brown, L. A., & Zoccoli, S. L. (2002). Repetition-based credibility enhancement of unfamiliar faces. *The American Journal of Psychology*, *115*, 199–2009.

Brown, A. S., & Nix, L. A. (1996). Turning lies into truths: Referential validation of falsehoods. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1088–1100.

Cialdini, R. B. (2009). *Influence: Science and practice*. Boston: Pearson Education.

Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review*, *14*, 238–257.

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., . . . Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, *113*(3), 554–559.

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

Emery, D. (2000, February 23). The great banana scare of 2000. Retrieved May 24, 2002, from http://urbanlegends.about.com/library/weekly/aa022302a.htm

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*(5), 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*(5), 1705–1710.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 123–146.

Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.

Foster, J. L., Huthwaite, T., Yesberg, J. A., Garry, M., & Loftus, E. (2012). Repetition, not number of sources, increases both susceptibility to misinformation and confidence in the accuracy of eyewitnesses. *Acta Psychologica*, *139*, 320–326.

Fragale, A. R., & Heath, C. (2004). Evolving information credentials: The (mis)attribution of believable facts to credible sources. *Personality and Social Psychology Bulletin*, *30*, 225–236.

Gabielkov, M., Ramachandran, A., Chaintreau, A., & Legout, A. (2016). Social clicks: What and who gets read on Twitter? *ACM SIGMETRICS Performance Evaluation Review*, *44*, 179–192. http://dx.doi.org/10.1145/2896377.2901462

Garcia-Marques, T., Silva, R. R., Reber, R., & Unkelbach, C. (2015). Hearing a statement now and believing the opposite later. *Journal of Experimental Social Psychology*, *56*, 126–129.

Gargiulo, F., & Gandica, Y. (2017). The role of homophily in the emergence of opinion controversies. *Journal of Artificial Societies and Social Simulation*, *20*(3), 8. doi: 10.18564/jasss.3448. Retrieved from http://jasss.soc.surrey.ac.uk/20/3/8.htm

Gawronski, B., & Strack, F. (Eds.). (2012). *Cognitive consistency: A fundamental principle in social cognition*. New York: Guilford Press.

Gefen, D. (2000). E-commerce: The role of familiarity and trust. *Omega*, *28*, 725–737.

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, *46*, 107–119.

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*, 230–237.

Greifeneder, R., Bless, H., & Pham, M. T. (2011). When do people rely on cognitive and affective feelings in judgment? A review. *Personality and Social Psychology Review*, *15*, 107–141.

Greifeneder, R., & Schwarz, N. (2014). Metacognitive processes and subjective experience. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (pp. 314–327). New York, NY: Guilford Press.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics, vol. 3: Speech acts* (pp. 41–58). New York: Academic Press.

Haddock, G., Rothman, A. J., Reber, R., & Schwarz, N. (1999). Forming judgments of attitude certainty, importance, and intensity: The role of subjective experiences. *Personality and Social Psychology Bulletin*, *25*, 771–782.

Harris, A. J. L., & Hahn, U. (2009). Bayesian rationality in evaluating multiple testimonies: Incorporating the role of coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1366–1372.

Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning & Verbal Behavior*, *16*, 107–112.

Hasson, U., Simmons, J. P., & Todorov, A. (2005). Believe it or not: On the possibility of suspending belief. *Psychological Science*, *16*, 566–571.

Hawkins, S. A., & Hoch, S. J. (1992). Low-involvement learning: Memory without evaluation. *Journal of Consumer Research*, *19*, 212–225.

Jacoby, L. L., Woloshyn, V., & Kelley, C. M. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General*, *118*, 115–125.

Jalbert, M., Newman, E. J., & Schwarz, N. (2019). *Only half of what I tell you is true: How experimental procedures lead to an underestimation of the truth effect*. Manuscript under review.

Johnson-Laird, P. N. (2012). Mental models and consistency. In B. Gawronski & F. Strack (Eds.), *Cognitive consistency: A fundamental principle in social cognition* (pp. 225–243). New York: Guilford Press.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Macmillan.

Levy-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*, 1093–1096.

Lewandowsky, S., Ecker, U. K. H., Seifert, C., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, *13*, 106–131.

Lewandowsky, S., Gignac, G. E., & Vaughan, S. (2013). The pivotal role of perceived scientific consensus in acceptance of science. *Nature Climate Change*, *3*, 399–404.

Luhmann, N. (1979). *Trust and power*. Chichester, UK: Wiley.

Mayo, R. (2017). Cognition is a matter of trust: Distrust tunes cognitive processes. *European Review of Social Psychology*, *26*, 283–327.

McGlone, M. S., & Tofighbakhsh, J. (2000). Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. *Psychological Science*, *11*, 424–428.

McPherson, M., Smith-Lovin, L., & Cook. J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*, 415–444.

McQuail, D. (2000). *McQuail's mass communication theory*. Newbury Park, CA: Sage Publications.

Morsanyi, K., & Handley, S. J. (2012). Logic feels so good: I like it! Evidence for intuitive detection of logicality in syllogistic reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 596–616.

Newcomb, T. M. (1943). *Personality and social change*. New York: Holt, Rinehart, & Winston.

Newman, E. J., Sanson, M., Miller, E. K., Quigley-McBride, A., Foster, J. L., Bernstein, D. M., & Garry, M. (2014). People with easier to pronounce names promote truthiness of claims. *PLoS One*, *9*(2). doi: 10.1371/journal.pone.0088671

Newman, E. J., & Schwarz, N. (2018). Good sound, good research: How audio quality influences perceptions of the researcher and research. *Science Communication*, *40*(2), 246–257.

Newman, E. J., & Zhang, L. (2020). Truthiness: How nonprobative photos shape beliefs. 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, UK: Routledge.

Oyserman, D., & Dawson, A. (2020). Your fake news, our fakes: Identity-based motivation shapes what we believe, share, and accept. In R. Greifeneder, M. Jaffé, E. J. Newman, & N. Schwarz (Eds.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 173–195). London, UK: Routledge.

Pariser, E. (2011). *The filter bubble: How the new personalized web is changing what we read and how the think*. New York: Penguin Books.

Pennington, N., & Hastie, R. (1993). The story model for juror decision making. In R. Hastie (Ed.), *Inside the juror* (pp. 192–223). New York: Cambridge University Press.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, *19*, 123–205.

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, *8*, 338–342.

Rice, R., & Atkin, C. (Eds.). (2001). *Public communication campaigns* (3rd ed.). Newbury Park, CA: Sage Publications.

Rosnow, R. L., & Fine, G. A. (1976). *Rumor and gossip: The social psychology of hearsay*. New York: Elsevier.

Ross, L., & Ward, A. (1996). Naive realism in everyday life: Implications for social conflict and misunderstanding. In E. S. Reed, E. Turiel, & T. Brown (Eds.), *Values and knowledge* (pp. 103–135). Hillsdale, NJ: Lawrence Erlbaum.

Ross, M., Buehler, R., & Karr, J. W. (1998). Assessing the accuracy of conflicting autobiographical memories. *Memory and Cognition*, *26*, 1233–1244.

Schul, Y., Mayo, R., & Burnstein, E. (2008). The value of distrust. *Journal of Experimental Social Psychology*, *44*, 1293–1302.

Schwarz, N. (1994). Judgment in a social context: Biases, shortcomings, and the logic of conversation. *Advances in Experimental Social Psychology*, *26*, 123–162.

Schwarz, N. (1996). *Cognition and communication: Judgmental biases, research methods, and the logic of conversation*. Hillsdale, NJ: Erlbaum.

Schwarz, N. (1998). Accessible content and accessibility experiences: The interplay of declarative and experiential information in judgment. *Personality and Social Psychology Review*, *2*, 87–99.

Schwarz, N. (2012). Feelings-as-information theory. In P. A. Van Lange, A. W. Kruglanski, & E. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 289–308). Thousand Oaks, CA: Sage Publications.

Schwarz, N. (2015). Metacognition. In M. Mikulincer, P. R. Shaver, E. Borgida, & J. A. Bargh (Eds.), *APA handbook of personality and social psychology: Attitudes and social cognition* (pp. 203–229). Washington, DC: APA.

Schwarz, N. (2018). Of fluency, beauty, and truth: Inferences from metacognitive experiences. In J. Proust & M. Fortier (Eds.), *Metacognitive diversity: An interdisciplinary approach* (pp. 25–46). New York: Oxford University Press.

Schwarz, N., & Lee, S. W. S. (2019). The smell of suspicion: How the nose curbs gullibility. In J. P. Forgas & R. F. Baumeister (Eds.), *The social psychology of gullibility: Fake news, conspiracy theories, and irrational beliefs* (pp. 234–252). New York: Routledge and Psychology Press.

Schwarz, N., Newman, E., & Leach, W. (2016). Making the truth stick and the myths fade: Lessons from cognitive psychology. *Behavioral Science & Policy*, *2*(1), 85–95.

Schwarz, N., Sanna, L. J., Skurnik, I., & Yoon, C. (2007). Metacognitive experiences and the intricacies of setting people straight: Implications for debiasing and public information campaigns. *Advances in Experimental Social Psychology*, *39*, 127–161.

Schwarz, N., & Vaughn, L. A. (2002). The availability heuristic revisited: Ease of recall and content of recall as distinct sources of information. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 103–119). Cambridge: Cambridge University Press.

Silva, R. R., Chrobot, N., Newman, E., Schwarz, N., & Topolinski, S. (2017). Make it short and easy: Username complexity determines trustworthiness above and beyond objective reputation. *Frontiers in Psychology*, *8*, 2200.

Sinatra, G. M., & Pintrich, P. (2003). The role of intentions in conceptual change learning. In G. M. Sinatra & P. R. Pintrich (Eds.), *Intentional conceptual change*. Mahwah, NJ: Lawrence Erlbaum Associates.

Skurnik, I., Yoon, C., Park, D. C., & Schwarz, N. (2005). How warnings about false claims become recommendations. *Journal of Consumer Research*, *31*, 713–724.

Song, H., & Schwarz, N. (2008). Fluency and the detection of misleading questions: Low processing fluency attenuates the Moses illusion. *Social Cognition*, *26*, 791–799.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Cambridge, MA: Harvard University Press.

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah: Erlbaum.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*(3), 755–769.

Toland, J. (1976). *Adolf Hitler*. Garden City, NY: Doubleday.

Topolinski, S. (2012). Nonpropositional consistency. In B. Gawronski & F. Strack (Eds.), *Cognitive consistency: A fundamental principle in social cognition* (pp. 112–131). New York: Guilford Press.

Topolinski, S., & Strack, F. (2008). Where there's a will: There's no intuition: The unintentional basis of semantic coherence judgments. *Journal of Memory and Language*, *58*, 1032–1048.

Topolinski, S., & Strack, F. (2009). The architecture of intuition: Fluency and affect determine intuitive judgments of semantic and visual coherence and judgments of grammaticality in artificial grammar learning. *Journal of Experimental Psychology: General*, *138*, 39–63.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207–232.

Unkelbach, C., & Greifeneder, R. (2018). Experiential fluency and declarative advice jointly inform judgments of truth. *Journal of Experimental Social Psychology*, *79*, 78–86.

Visser, P. S., & Mirabile, R. R. (2004). Attitudes in the social context: The impact of social network composition on individual-level attitude strength. *Journal of Personality and Social Psychology*, *87*, 779–795.

Vosniadou, S. (Ed.). (2008). *International handbook of research on conceptual change*. New York, NY: Routledge.

Weaver, K., Garcia, S. M., Schwarz, N., & Miller, D. T. (2007). Inferring the popularity of an opinion from its familiarity: A repetitive voice can sound like a chorus. *Journal of Personality and Social Psychology*, *92*, 821–833.

Weisbuch, M., & Mackie, D. (2009). False fame, perceptual clarity, or persuasion? Flexible fluency attribution in spokesperson familiarity effects. *Journal of Consumer Psychology*, *19*(1), 62–72.

Winkielman, P., Huber, D. E., Kavanagh, L., & Schwarz, N. (2012). Fluency of consistency: When thoughts fit nicely and flow smoothly. In B. Gawronski & F. Strack (Eds.), *Cognitive consistency: A fundamental principle in social cognition* (pp. 89–111). New York: Guilford Press.

Yousif, S. R., Aboody, R., & Keil, F. C. (2019). The illusion of consensus: A failure to distinguish between true and false consensus. *Psychological Science*, *30*(8), 1195–1204.