

# When Debunking Scientific Myths Fails (and When It Does Not): The Backfire Effect in the Context of Journalistic Coverage and Immediate Judgments as Prevention Strategy

Science Communication  
2016, Vol. 38(1) 3–25  
© The Author(s) 2015  
Reprints and permissions:  
[sagepub.com/journalsPermissions.nav](http://sagepub.com/journalsPermissions.nav)  
DOI: 10.1177/1075547015613523  
[scx.sagepub.com](http://scx.sagepub.com)  


Christina Peter<sup>1</sup> and Thomas Koch<sup>1</sup>

## Abstract

When reporting scientific information, journalists often present common myths that are refuted with scientific facts. However, correcting misinformation this way is often not only ineffective but can increase the likelihood that people misremember it as true. We test this backfire effect in the context of journalistic coverage and examine how to counteract it. In a web-based experiment, we find evidence for a systematic backfire effect that occurs after a few minutes and strengthens after five days. Results show that forming judgments immediately during reception (in contrast to memory-based) can reduce backfire effects and prevent erroneous memory from affecting participants' attitudes.

## Keywords

backfire effect, judgment-formation, memory, experiment

---

<sup>1</sup>LMU Munich, Germany

## Corresponding Author:

Christina Peter, Department of Communication Studies and Media Research, LMU Munich, Oettingenstr. 67, Munich 80538, Germany.  
Email: [peter@ifkw.lmu.de](mailto:peter@ifkw.lmu.de)

Once people have been confronted with a piece of information, it is hard to erase it—even if it turns out that the information was incorrect. One example of this is the so-called MMR (measles, mumps, and rubella) vaccine controversy: At the end of the 1990s, a study erroneously reported that the combined MMR vaccination could cause autism. Though the study turned out to be fraudulent and many disclaimer campaigns were launched, a lot of people still believe that this risk exists (Hargreaves, Lewis, & Speers, 2003; Nyhan, Reifler, Richey, & Freed, 2014). From a normative perspective, it is, especially in the fields of politics and health, important that citizens are well and correctly informed, because they base far-reaching, momentous decisions on information that they assume to be true (for the distinction between being misinformed and being uninformed, see Kuklinski, Quirk, Jerit, Schwieder, & Rich, 2000). Hence, it is in the public interest to reduce the “widespread prevalence and persistence of misinformation in contemporary societies” (Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012, p. 106).

Therefore, many information campaigns try to correct misinformation, often by picking up on certain myths and rectifying them immediately. Such debunking campaigns are not only common in the context of health information but also can be found regarding political issues (“Top 10 Myths About Immigration,” Anchondo, 2010; “The Top 10 Myths About TTIP,” European Union, 2015). Apart from that, these kinds of “myths and facts stories” have become increasingly present in journalistic coverage (e.g., when reporting new scientific results that contradict certain prevalent myths); for instance, the *Washington Post* regularly publishes “Five myths about . . .” articles, where it debunks myths about all kind of topics.

Myths and facts stories have a common underlying structure: usually, the myth is presented in form of a highlighted statement (e.g., “We only use 10% of our brains” or “Earth is closer to the sun during summer”), followed by a longer passage that debunks it and contains scientific data about the actual situation. In many cases, the myth is directly followed by a short debunking claim such as “False!”.

Although myths and facts stories are clearly designed to debunk common myths existing in the population, research shows that this procedure might have the opposite effect, because disclaimers do not always lead to a rectification, but sometimes to a solidification of the myth: “Attempts to warn people about false information can backfire and unintentionally increase people’s acceptance of the false information as true” (Skurnik, Yoon, & Schwarz, 2007, p. 4). This backfire effect can be momentous, especially when people base judgments on their erroneous memory.

While there is some research describing and analyzing the backfire effect, to our knowledge, there are no studies examining how to counteract it. In the

present study, we propose that the judgment-formation strategy might help reduce backfire error. Research on judgment-formation distinguishes between two types of strategies (Hastie & Park, 1986; Matthes, Wirth, & Schemer, 2007): *Immediate judgments* are formed during the reception of information, meaning that people form judgments about an issue the moment they encounter information about it and store this judgment. If people using this strategy are asked for their opinion at a later point in time, they recall this previously formed judgment, but not the arguments that led to it. *Memory-based judgments*, on the other hand, are only formed when people are asked for them; therefore, people using this strategy do not form judgments during the reception process. We assume that forming judgments during reception can reduce the risk of backfire effects, since recipients can use their attitude toward that specific issue as a retrieval cue to determine the truth of a statement. This article first explains the backfire effect as well as the judgment-formation strategies and reports on the conducting of an experiment that tested our assumptions.

## The Backfire Effect

The backfire effect describes the phenomenon that rectifying misinformation can have the opposite effect and lead recipients to misremember the false information as true (Lewandowsky et al., 2012; Nyhan & Reifler, 2010; Schwarz, Sanna, Skurnik, & Yoon, 2007; Skurnik, Yoon, Park, & Schwarz, 2005). Thus, debunking false information can actually strengthen the belief in its truth (Johnson & Seifert, 1994; Wilkes & Leatherbarrow, 1988). Skurnik et al. (2005) demonstrate this paradoxical effect with warnings about false statements: They presented participants with claims about health and nutrition, with each claim labeled as true or false directly after the presentation (e.g., “Aspirin destroys tooth enamel,” p. 715). Later, participants were asked to indicate whether each statement (and several statements that were not actually presented in the stimulus) was true, false, or new. The authors compared backfire errors (labeling originally false statements as true) with the opposite errors (labeling originally true statements as false). If people solely made random errors, they should have equally often remembered facts as false and myths as true. Yet, participants significantly more often misremembered myths as facts than the other way around. This effect was even more pronounced after a 3-day delay. Skurnik et al. (2007) replicated these findings in the context of health campaigns: They showed their participants “Myth&Facts” flyers about the flu vaccine. While participants made hardly any mistakes directly after exposure to the flyer, the backfire effect emerged after a 30-minute delay, with people misremembering significantly more myths as true than facts as false.

Why do people, after some time has passed, tend to believe that a statement is true, even though it was clearly indicated as being false? The occurrence of this backfire effect is based on two processes: first, on increasing processing fluency through the (repeated) presentation of a statement, which leads to higher credibility (Hasher, Goldstein, & Toppino, 1977; Parks & Toth, 2006; Reber & Schwarz, 1999). Second, the backfire effect is based on the fact that memory for contextual details fades off faster than for the information itself.

### *Process 1: Enhanced Feeling of Familiarity*

If people were asked to judge the validity of the statement, “The mortality rate of bowel cancer is 18%,” only a few experts would be able to tell without any doubt whether this claim was really true. Those without any expert knowledge could check whether they had heard this statement before (Arkes, Boehm, & Xu, 1991; Arkes, Hackett, & Boehm, 1989; Schwartz, 1982): Does this claim sound familiar? If recipients believe that they have encountered it before, they are more inclined to believe it. This phenomenon is known as the *truth effect* (e.g., Arkes et al., 1989; Bacon, 1979; Begg, Anas, & Farinacci, 1992; Boehm, 1994; Gigerenzer, 1984; Hasher et al., 1977; Koch & Zerback, 2013; a meta-analysis by Dechêne, Stahl, Hansen, & Wänke, 2010, confirms the stability of this effect).

Familiarity is considered as the key determinant of the truth effect. The familiarity hypothesis is based on two mechanisms following each other: In a first step, every contact with a stimulus increases its processing fluency, which means it can be processed more easily after the first reception (Bornstein, 1989; Reber, Schwarz, & Winkielman, 2004; Reber, Winkielman, & Schwarz, 1998). In a second step, this enhanced processing fluency triggers a feeling of familiarity and leads participants to believe that they have heard or seen this statement before (Parks & Toth, 2006; Reber & Schwarz, 1999; Unkelbach, 2007). Unkelbach (2006, 2007) offers an explanation as to why processing fluency makes a statement seem more credible. He shows that recipients learn in everyday life that processing fluency correlates with the validity of messages. People trust in the heuristic that true statements have a higher probability of being repeated, whereas false statements are not passed on (Unkelbach, 2007; Unkelbach, Bayer, Alves, Koch, & Stahl, 2011). Thus, the feeling of familiarity finally leads to higher credibility (Parks & Toth, 2006; Reber & Schwarz, 1999).

For the occurrence of these processes, it is irrelevant whether the statements are factually true or false (Gigerenzer, 1984; Hasher et al., 1977); the truth effect occurs regardless of whether statements were initially rated as

credible or questionable (Arkes et al., 1989). The effect even occurs when the credibility of statements is actively put into question beforehand (Begg, Armour, & Kerr, 1985). Generally, the truth effect is more pronounced if respondents are unsure about the validity of statements (Roggeveen & Johar, 2002; Unkelbach, 2007).

### *Process 2: Short-Lasting Memory for Context Information*

The truth effect explains why the (repeated) presentation of a statement leads recipients to believe that it is true. However, it does not explain why this effect is also triggered when people are explicitly warned that a certain piece of information is not true. Here, a second process comes into play: The speed of forgetting differs between context information (e.g., additional information that a statement is true or false) and the information itself. Even if some time has passed, people implicitly sense that they have encountered a statement before, while the memory for context information vanishes more quickly (Mandler, 1980; Skurnik et al., 2005). Thus, recognizing that a statement has been encountered before is easier than remembering the exact context in which the statement was presented (Mandler, 1980). The existence of this process is confirmed in several studies showing that warnings about false claims or correcting misinformation have the desired effect directly after exposure to the information, but that after a delay, contextual information fades, “leaving only enhanced familiarity” (Skurnik et al., 2005, p. 722).

Taking both processes together, after some time and/or distraction, the specific recall of the context vanishes and recipients start forgetting whether a claim was actually labeled as true or false. Now, they rely on the heuristic that a claim that sounds familiar will probably be true (Hasher et al., 1977); the backfire effect is triggered. Based on these considerations, we expect that rectifying false claims enhances backfire effects: Reading statements explicitly labeled as true or false leads people to more often misremember false information as true than true information as false (Hypothesis 1a). In line with the studies reported above, we expect this effect to strengthen after a delay (Hypothesis 1b).

Backfire errors can be momentous when people base attitudes and/or decisions on misinformation they erroneously remember as true. Thus, it is not only necessary to investigate whether the rectification of misinformation can backfire but also how this affects people’s attitudes. The only study to investigate participants’ judgments in the context of the backfire effect, conducted by Skurnik et al. (2007; see above), yielded ambiguous results: While the different flyers (the “Facts-and-Myths” flyer vs. the “Only-Facts” flyer) did not affect participants’ intentions, it affected participants’ risk perceptions.

Furthermore, the authors only checked for differences in risk perceptions between the flyer groups, but they did not examine whether the number of backfire errors actually caused these effects. In the present study, we want to focus on this connection: We predict a positive correlation between the valence of the false statements misremembered as facts and people's attitudes (Hypothesis 1c). For example, the more the myth with a negative valence people misremember as true (e.g., "Side effects of the flu vaccination are worse than the flu"), the more negative their attitude toward the flu vaccination should become.

This relationship should be especially pronounced for people who strongly base their attitudes on the information they (mis-)remember when forming a judgment. However, research on social cognition shows that people do not always form their attitudes like this (e.g., Hastie & Park, 1986): Depending on the judgment-formation strategy people apply, they either form judgments during reception or only at the time the judgment is actually needed; thus, people rely more or less on the single arguments that they remember. Consequently, we assume that the judgment-formation strategy could play an important role in the occurrence of backfire errors and their prevention.

## **Judgment-Formation Strategy: Immediate Versus Memory-Based Judgments**

A large body of research has dealt with how people process information to form judgments. In this context, researchers have developed different models to describe two modes of cognitive processing and decision making (e.g., Chaiken, 1980; Kahneman & Frederick, 2002; Petty & Cacioppo, 1986), which are commonly labeled as dual process theories (Evans, 2008). While these theories usually deal with *how* information is processed when judgments are formed (fast/heuristic/automatic vs. slow/systematic/controlled), research on judgment-formation strategy investigates *when* judgments are actually formed (Mackie & Asuncion, 1990).

When asked for a judgment on an issue, in some cases people recall information that they have encountered on the topic and then base their judgment on the arguments that they remember. This strategy is called forming memory-based judgments (Hastie & Park, 1986; Lavine, 2002). Since the judgment is based predominantly on arguments that people can recall at the time the judgment is needed, the resulting attitudes can be rather unstable: "Memory-based judgments are based on whatever information comes to mind at the time of judgment" (Matthes et al., 2007, pp. 247-248). This can result in different attitudes on the same issue at different points in time since memory-based judgments are prone to recency, availability, and salience effects (Hastie

& Park, 1986; Tversky & Kahneman, 1973). The main characteristic of this judgment-formation strategy is that people encode and store the information presented, but they do not form judgments at that point in time (Matthes et al., 2007). This strategy is especially applied when the individual is unaware that the information is needed for a future judgment (Hastie & Park, 1986).

Apart from these memory-based judgments, there is a second judgment-formation strategy that has been referred to as on-line judgments (Bizer, Tormala, Rucker, & Petty, 2006; Hastie & Park, 1986). These judgments are formed immediately during the reception of information; we will refer to this strategy as *immediate judgments*.<sup>1</sup> They are formed and stored in memory when information on an issue is actually encountered. When people applying this strategy are asked for their attitudes later, they only recall the attitude itself but not the single arguments that led to it (Srull & Wyer, 1989). Consequently, the attitude and the recalled arguments do not have to correlate (Hastie & Park, 1986; Matthes et al., 2007). The reason for this is what Anderson (1981) called the two-memory hypothesis: Single arguments are stored in a different memory system than the actual judgment. For this reason, even if memory for the single argument fades, people can still recall the judgment that they have formed. This also means that attitudes formed during reception are more stable over time and independent of what arguments can be remembered at the time a judgment is needed (Bizer et al., 2006). Nevertheless, Hastie and Park (1986) stress that there can be a relationship between the valence of the judgment and the recalled arguments for people with immediate judgments since they can use their existing attitudes as retrieval cues (“judgment causes memory,” p. 259). For example, a more positive attitude might lead people to remember predominantly positive arguments through biased encoding.

Most of the research on judgment-formation strategies has been conducted using experimental settings: Participants are presented with the same stimulus, yet the judgment-formation strategy is manipulated via different tasks. As one of the first studies to compare both judgment strategies, Hastie and Park (1986) presented their participants with a description of a fictitious person. Participants in the immediate judgment group were told that they should form an impression of the person (e.g., his personality and likeability). In contrast, participants that were not to form a judgment during reception (= memory-based) were presented with a distraction task: They were asked to judge the grammaticality of each sentence. According to Hastie and Park (1986), a distraction task is necessary since an experimental situation almost automatically evokes immediate judgments—a problem they encountered in previous experiments. Results showed strong correlations between the recall

of specific information about the fictitious person and attitudes toward him for the memory-based judgment strategy, but not for the immediate judgment strategy.

Subsequent studies adopted Hastie and Park's operationalization of the judgment-formation tasks (Bizer et al., 2006; Hamilton, Sherman, & Maddox, 1999; Mackie & Asuncion, 1990; Tormala & Petty, 2001). Tormala and Petty (2001), for instance, showed participants statements about a person and asked participants in the immediate judgment condition to form an impression of the person, whereas the memory-based task was to focus on the sentences, and whether they were simple or complex in nature. Similarly, Mackie and Asuncion (1990) presented participants with arguments about standardized tests for admission to college generated by fellow students. To evoke immediate judgments, participants were asked to "pay attention to the strength of arguments the student uses and to how well this student can support his or her position" (p. 7). For a memory-based judgment strategy, participants were told "that we were interested in how dynamically the other student expressed his or her opinion" (p. 7). All these studies manipulated the judgment-formation strategies in experimental settings; however, Matthes et al. (2007) were the first to actually measure judgment-formation strategies via self-reports. They developed items for both strategies (e.g., immediate judgment: "Intuitively, I knew from the beginning on how I stood on that issue," p. 253).

In summary, the main difference between both judgment-formation strategies is the point in time when a judgment is formed. This has several consequences that might be relevant for the occurrence of the backfire effect: It is plausible that memory-based judgments will lead to more backfire errors than immediate judgments since people with immediate judgments can use their attitude as a retrieval cue for remembering if statements were true or false (Hastie & Park, 1986). If, for example, a person with a positive attitude on flu vaccination does not remember if the statement "The side effects are worse than the flu" was true or false, he or she can use the recalled attitude as an anchor: Why would I hold a positive attitude toward flu vaccination if side effects were worse than the flu? Consequently, if they have formed a favorable attitude in the first place, it is likely that they will judge this statement as false. Based on that, we predict that people applying a memory-based judgment strategy will make more backfire errors than people forming immediate judgments (Hypothesis 2a). Furthermore, results by Skurnik et al. (2005, 2007) suggest that this difference will be more pronounced after a delay (Hypothesis 2b).

Moreover, the relationship between memory and attitudes predicted in Hypothesis 1c could depend on the judgment-formation strategy that is applied: People that have already formed and stored an attitude during the



reception of specific information (= immediate judgment) can recall this very attitude when asked for it; this implies that their judgments should be stable over time (Hastie & Park, 1986). People with a memory-based judgment strategy, however, have not formed and stored an attitude during reception; when asked for it, they need to recall information about the issue and base their attitude on the arguments that they can remember (Matthes et al., 2007). Thus, their attitudes should be less stable and affected by (erroneous) memory for information about the issue. If people were to assess their attitudes shortly after reception and again after a delay, we predict that there will be a stronger correlation between both judgments for people applying the immediate judgment strategy than for those with the memory-based judgment strategy (Hypothesis 2c). In line with this, we assume that after a delay, the attitudes of people with a memory-based judgment strategy will be affected by backfire effects, while the attitudes of people forming immediate judgments will not be affected by backfire effects (Hypothesis 2d).

## Method

### *Participants*

We recruited 335 participants (57.9% female; age:  $M = 41.34$  years,  $SD = 16.23$ ) through an online access panel for social science research (SoSci Panel; Leiner, 2012, 2014). The panel is noncommercial, and its members agreed to participate in scientific surveys. It is managed by a researcher of the University of Munich and includes around 100,000 panelists from Germany, Austria, and the German-speaking part of Switzerland. For the present experiment, we only invited panelists from Germany; this sample is, however, not representative of the German population, as members are younger and better educated. Nevertheless, compared to traditional student samples, the SoSci Panel offers more heterogeneous samples regarding age, education, geography, and personal interest. For more information on the composition and limitations of the SoSci Panel, see Leiner (2014). Participation was unpaid and voluntary. Respondents were informed that all of their responses were confidential.

### *Procedure and Stimulus*

Participants were told that the study was about journalistic myths and facts stories, a specific type of article, where journalists pick up different claims circulating about a certain issue and explain whether these claims are false (myths) or true (facts). After this short introduction, all respondents were

informed that they would see three newspaper articles, which they should read carefully. The first and the last article served as distraction articles and reported about building a new city hall in a small town (Article 1) and about powering down public lighting systems at night in a German municipality (Article 3). The stimulus was Article 2, which reported about a new bowel cancer test that could be done at home. Thus, we wanted to ensure that participants did not focus closely on the stimulus article and that some time had passed between reading the stimulus article and the questions relating to memory and attitudes.

The stimulus article started with a short introduction explaining that there is a new bowel cancer test to detect tumors of the colon and rectum. The test was described as being easy to do at home: one has to place small samples of a stool on special cards and send them to a laboratory for analysis. After this introduction, the journalist reported eight scientific statements about the test and identified each statement immediately as true or false, followed by a short explanation (e.g., “The bowel cancer test is more effective than conventional screenings”—“False! The colonoscopy is the most effective screening method available”). Four of the statements were identified as being false; four of them were identified as being true. We made sure that statements labeled as true were indeed objectively true and statements labeled as false were indeed objectively false (we checked the truth on the website of a national institute for health care).

The article presented the new test in a rather negative light, for example, by identifying three of the statements in favor of the test as being false. This is an important difference to the study by Skurnik et al. (2007), as they labeled only negative statements as being myths (e.g., “Side effects of the flu vaccination are worse than the flu”). Yet, studies in the context of framing research have shown that negative statements are judged to be more credible than positive ones (Hilbig, 2009, 2012; Koch, Peter, & Obermaier, 2013). After the stimulus presentation, we tested participants’ memory for the statements as well as their attitudes toward the test. Additionally, we collected sociodemographic data and items to check whether our manipulation was successful (see Measures). Subsequently, participants viewed a debriefing message with additional information about the real purpose of the experiment.

## *Design*

To test our hypotheses, we varied two factors: first, we manipulated the two different judgment-formation strategies (immediate vs. memory-based judgments). Analogously to other experiments on judgment-formation (e.g., Bizer et al., 2006; Hastie & Park, 1986; Mackie & Asuncion, 1990; Tormala & Petty,

2001), we showed participants different instructions prior to the stimulus: For immediate judgments (Condition 1), we told participants that we were interested in their opinion about the topics of the three articles and that we would ask questions about their opinion afterward; we asked them to decide how they felt about each topic while reading the articles. For memory-based judgments (Condition 2), we told participants that we were interested in the linguistic and journalistic quality of the articles, and that we would ask questions about the quality of the article afterward; we asked them to decide how they felt about the journalistic quality while reading the articles. Both groups were to read the stimulus thoroughly. Participants were randomly assigned to one of the two conditions. Both groups did not differ significantly regarding gender,  $\chi^2(N = 335) = 0.55, p = .51$ , age,  $t(333) = 0.56, p = .58$ , education,  $\chi^2(N = 335) = 7.55, p = .11$ , and regarding the time they spent reading the article,  $t(333) = 0.86, p = .39$ .

Second, to test whether effects strengthen over time, we employed a repeated measurement: participants answered questions a few minutes after the stimulus presentation and again after a delay of a few days (the minimum delay was 2 days,  $M = 5.20, SD = 1.47$ ). Since the study was designed as a web-based experiment, the time period between both measurements varied to some extent—we will consider this in the Results section. In both questionnaires, we asked for memory regarding the truth of the statements, attitudes toward the test, as well as questions regarding the judgment-formation strategy.

## Measures

To measure *memory for the truth of the statements*, we presented the eight statements shown in our stimulus along with four new claims. For each claim, we asked participants to indicate whether it was labeled as “true” or “false” in the article or whether it had not been presented in the article (“new”).

*Attitudes toward the new bowel cancer test* were measured by a three-item semantic differential: “I believe that the presented bowel cancer test is . . .” “not appropriate vs. appropriate,” “bad vs. good,” and “not reasonable vs. reasonable” (t1:  $M = 3.10, SD = 1.11, \alpha = .92$ ; t2:  $M = 3.24, SD = 1.14, \alpha = .94$ ). To avoid sequence effects, we randomized the sequence of the memory and attitude measure for both questionnaires.

Participants’ *judgment-formation strategy* was assessed using items derived from the scale developed by Matthes et al. (2007). Since this scale was constructed to measure immediate and memory-based judgment-formation in surveys and, hence, for (political) issues people have already heard about, we had to change the items slightly. We took two items asking for immediate judgment-formation strategy (“I knew immediately how I stood on that

new bowel cancer test” and “While reading that article, I had already formed a concrete opinion”) and two items asking for memory-based judgment-formation strategy (“When asked about the bowel cancer test, I had to recall all arguments first” and “Not until I was asked about my opinion on the bowel cancer test, had I considered what arguments spoke in favor of it and what spoke against it”; t1:  $M = 3.94$ ,  $SD = 0.91$ ,  $\alpha = .82$ ; t2:  $M = 3.53$ ,  $SD = 0.96$ ,  $\alpha = .78$ ).

## Results

The backfire effect occurs when people erroneously remember misinformation as true: they actually remember having heard a statement before, but are wrong about its original truth value. Consequently, in line with Skurnik et al. (2005), we only included statements in the analysis that had been presented in the stimulus article *and* that participants correctly remembered as being part of the stimulus (meaning that they did not label them as “new”). Thus, there are two different types of errors that can occur: labeling an originally false statement as true (backfire error) or labeling an originally true statement as false (fact-false error). If participants randomly guessed at the truth of statements, then both errors would occur in equal measure. However, in line with previous research, we expected backfire errors to occur more often than fact-false errors (Hypothesis 1a) and that this effect would strengthen after a delay (Hypothesis 1b). We calculated a 2 (type of error)  $\times$  2 (short vs. long delay) mixed ANOVA on the error rates. The analysis revealed a significant main effect for the type of error: participants more often remembered originally false statements as true (backfire error;  $M = 0.18$ ,  $SD = 0.24$ ) than originally true statements as false (fact-false error;  $M = 0.05$ ,  $SD = 0.13$ ),  $F(1, 329) = 113.18$ ,  $p < .001$ ,  $\eta^2_{\text{partial}} = .26$ ; Hypothesis 1a was thus confirmed. Furthermore, there was a significant main effect for delay,<sup>2</sup> with participants making more errors after 5 days (t1:  $M = 0.08$ ,  $SD = 0.16$ ; t2:  $M = 0.15$ ,  $SD = 0.21$ ),  $F(1, 329) = 91.84$ ,  $p < .001$ ,  $\eta^2_{\text{partial}} = .22$ . The significant interaction effect confirmed Hypothesis 1b: While backfire errors increased considerably after a 5-day delay (t1:  $M = 0.12$ ,  $SD = 0.21$ ; t2:  $M = 0.23$ ,  $SD = 0.28$ ), fact-false errors only slightly increased (t1:  $M = 0.03$ ,  $SD = 0.10$ ; t2:  $M = 0.07$ ,  $SD = 0.15$ ),  $F(1, 329) = 14.84$ ,  $p < .001$ ,  $\eta^2_{\text{partial}} = .04$ .

To test whether there was a relationship between backfire errors and attitude, we calculated correlations between the two variables. After the short delay, we found a small but significant correlation,  $r(334) = .18$ ,  $p < .001$ ; this means that the more originally false statements participants erroneously remembered as true (e.g., thinking that the false statement “The bowel cancer test is more effective than conventional screenings” was true), the more

favorable their attitudes toward the bowel cancer test were. After 5 days, this correlation was even more pronounced,  $r(328) = .31, p < .001$ . Thus, Hypothesis 1c can be confirmed. However, the difference between the two correlation coefficients did not reach statistical significance,  $z$  score =  $-1.77, p = .08$ .

In a second step, we examined the influence of the judgment-formation strategy. In line with prior research, we manipulated immediate versus memory-based judgments via instructions. For a manipulation check, we calculated an index of the four items that measured the judgment-formation strategy via self-report (Matthes et al., 2007). Higher values indicated that participants formed immediate judgments; lower values indicated a memory-based judgment-formation strategy. An independent  $t$  test revealed significant results between participants that received an instruction for immediate judgments ( $M = 4.07, SD = 0.85$ ) and participants that read an instruction for memory-based judgments ( $M = 3.80, SD = 0.96$ ),  $t(333) = 2.79, p < .01, d = .30$ . Yet, although scoring significantly lower than participants who formed immediate judgments, those in the memory-based group also reported that they mostly formed judgments during reception. To deal with this, we used the self-report measurement to distinguish between immediate and memory-based judgments. Since we asked for the judgment-formation strategy shortly after the stimulus as well as after a 5-day delay, we divided the groups based on both measures: The immediate judgment group consisted of participants that scored higher than 3 on both judgment-formation strategy measures ( $n = 167$ ); participants in the memory-based judgment group had to score 3 or lower on both measures ( $n = 46$ ). All other participants were excluded from further analysis. Both groups did not differ significantly regarding gender,  $\chi^2(N = 213) = 0.21, p = .73$ , age,  $t(211) = -0.55, p = .58$ , education,  $\chi^2(N = 213) = 2.45, p = .65$ , and with regard to the time they spent reading the article,  $t(211) = -0.82, p = .41$ .

We calculated a 2 (judgment-formation strategy)  $\times$  2 (short vs. long delay) mixed ANOVA on the backfire error rate. Results showed a significant main effect for judgment-formation strategy: Participants that formed memory-based judgments erroneously remembered originally false statements as true about twice as often ( $M = 0.31, SD = 0.32$ ) as participants that formed immediate judgments did ( $M = 0.16, SD = 0.23$ ),  $F(1, 208) = 17.63, p < .001, \eta^2_{\text{partial}} = .08$ . Hypothesis 2a was thus confirmed. Furthermore, there was a significant interaction effect, indicating that backfire errors increased considerably for participants with memory-based judgments after 5 days (t1:  $M = 0.20, SD = 0.30$ ; t2:  $M = 0.43, SD = 0.34$ ), but that these were less for people who formed immediate judgments (t1:  $M = 0.12, SD = 0.20$ ; t2:  $M = 0.20, SD = 0.25$ ),  $F(1, 208) = 11.46, p < .001, \eta^2_{\text{partial}} = .05$ . Thus, there was also support for Hypothesis 2b.

However, we wanted to make sure that participants with memory-based judgments were systematically more prone to backfire errors and were not just making more errors in general. Therefore, we calculated a separate 2 (judgment-formation strategy)  $\times$  2 (short vs. long delay) mixed ANOVA on the fact-false error rate. There was no significant main effect for judgment-formation strategy: Participants with memory-based judgments did not remember facts as false ( $M = 0.15$ ,  $SD = 0.22$ ) more often than participants with immediate judgments did ( $M = 0.12$ ,  $SD = 0.19$ ),  $F(1, 205) = 1.03$ ,  $p = .31$ ,  $\eta^2_{\text{partial}} = .01$ . Furthermore, there was no significant interaction effect, since fact-false errors increased for both strategies equally after the delay,  $F(1, 208) = 0.13$ ,  $p = .72$ ,  $\eta^2_{\text{partial}} = .001$ .

Finally, we expected the attitudes of participants with memory-based judgments to be rather unstable and depend on the number of backfire errors they made, while the attitudes for people with immediate judgments should not depend on backfire errors (Hypothesis 2d). Consequently, the correlation between attitudes reported after a short versus long delay should be significantly stronger for participants with immediate judgments compared to participants with memory-based judgments (Hypothesis 2c). To test these assumptions, we calculated separate regression models for immediate and memory-based judgment-formation strategies, with the attitude t2 as the outcome variable, and attitude t1 as well as the backfire errors t2 as predictors. As expected, for participants with immediate judgments, their attitudes formed shortly after reception were the only significant predictor for the attitudes after 5 days,  $\beta = .82$ ,  $p < .001$ ; the number of backfire errors made after 5 days scarcely affected the attitudes reported,  $\beta = .08$ ,  $p = .08$ ; in total, 71.4% of variance could be explained by this model,  $F(2, 163) = 207.17$ ,  $p < .001$ . For participants with memory-based judgments, a different pattern emerged: Here, both the attitude t1,  $\beta = .47$ ,  $p < .001$ , as well as the backfire errors t2,  $\beta = .28$ ,  $p < .05$ , were significant predictors for the attitude reported after 5 days,  $F(2, 24) = 14.04$ ,  $p < .001$ ,  $R^2 = 37.2$ . Furthermore, we checked whether the slopes differed significantly (Cohen, Cohen, West, & Aiken, 2003): As assumed, attitudes at t1 predicted significantly stronger attitudes at t2 when people formed immediate judgments ( $\beta = .82$ ) than when people formed memory-based judgments ( $\beta = .47$ ),  $t(205) = 3.08$ ,  $p = .002$ . Thus, the results confirmed both Hypotheses 2c and 2d. Furthermore, the fact that the attitudes of participants with memory-based judgments reported after 5 days were influenced by backfire errors led them to hold more favorable attitudes toward the bowel cancer screening test ( $M = 3.45$ ,  $SD = 0.72$ ) than participants with immediate judgments did ( $M = 3.16$ ,  $SD = 1.32$ ),  $t(135) = -1.97$ ,  $p < .05$ ,  $d = .27$ .

## Discussion

Myths existing in the population can have serious consequences, since (important) decisions are guided by intentions and attitudes, which are based on past experiences, and on information that we have gathered in our everyday life (Ajzen, 1985, 1991). Although myths and facts stories seem like an elegant way of communicating scientific information to the public, the results of the present study provide evidence for detrimental effects of correcting false information. In line with prior research, we found evidence that the backfire effect is a rather robust and systematic error (Lewandowsky et al., 2012; Nyhan & Reifler, 2010; Schwarz et al., 2007; Skurnik et al., 2005). The study shows that after only a few minutes, people start to misremember originally false information as true, but only rarely misremember facts as false. After a delay of several days, about one out of four originally false statements is erroneously remembered as a fact, leading people to believe, for example, that the presented bowel cancer test is recommended by independent IGeL monitoring—although the presented article not only identified this statement as false but also pronounced that the IGeL monitoring rates the benefit of this test as uncertain due to a lack of scientific evidence.

Furthermore, we were able to confirm a connection between backfire errors and participants' attitudes that also slightly increased over time: The more false statements (such as "The bowel cancer test is recommended by independent IGeL monitoring") participants misremembered as true, the more favorable their attitudes toward the test became. This result shows that the backfire effect has momentous consequences: People not only systematically misremember the truth of misinformation but also change their attitudes accordingly. To further investigate this connection, we integrated research on the judgment-formation strategy. Our results show that participants that did not form an attitude during reception (memory-based judgments) make significantly more backfire errors than people with existing attitudes do, even after a delay of only a few minutes. After 5 days, people that did not form attitudes during reception wrongfully remember almost half of the false statements as true. Apart from that, after 5 days, the attitudes of participants who formed immediate judgments were not influenced by the backfire mistakes they made but by their attitude they reported directly after the reception of the stimulus; it accounted for almost 70% of the variance. The fact that these participants make backfire errors after 5 days (although only half as much as the memory-based group do), but that these errors did not affect their attitude, provides further evidence that they simply recall the attitude they formed during reading the stimulus without including further information that they (mis-)remember. At the same time, the fact that they make only half as many

backfire mistakes as persons with memory-based attitudes do indicates that they can use their attitudes stored in memory as an anchor for assessing the truth of the statements. In contrast, after a few days, participants in the memory-based group based their attitude both on the attitude they stated directly after the stimulus as well as on the backfire mistakes they made.

### *Limitations*

It has to be noted that the current study is of course limited in various ways. First, the manipulation of the judgment-formation strategy has not worked adequately, even though we used the same instructions as prior experiments did. These experiments, however, did not control whether their manipulation was indeed successful (Bizer et al., 2006; Hamilton et al., 1999; Hastie & Park, 1986; Mackie & Asuncion, 1990; Tormala & Petty, 2001). We, for the first time, checked the effectiveness of this common manipulation with the scale developed by Matthes et al. (2007) and revealed a problem that should be considered in future studies: Though our treatment guided participants slightly toward the respective judgment-formation strategies (and the difference between the two conditions was highly significant), participants generally had a tendency for forming immediate judgments. As Hastie and Park (1986) already remarked, this might be a problem with the experimental situation: Respondents are quite aware that they will be asked questions after, for example, reading an article. Yet, as discussed, a memory-based judgment-formation strategy is applied especially when people are unaware that they will have to report on a judgment later (Hastie & Park, 1986). Thus, it seems that a distraction task cannot fully prevent participants from forming attitudes toward an issue presented during an experiment. We dealt with this problem by splitting participants post hoc in the respective groups; by doing so, we ensured that we indeed compared respondents that formed their attitudes during reception of the stimulus presentation with participants that formed their attitudes only when asked for them. This approach, however, gives rise to a second limitation of the current study: the question of causality. In contrast to the random assignment to both groups, the post hoc split makes it impossible to demonstrate with certainty that differences between the two groups are due to a causal link between the treatment and our observed variables. Differences between participants who formed immediate judgments and those with memory-based judgments might also be due to another systematic factor. However, we showed that the two groups did at least not differ significantly regarding age, gender, education, and the time they spent on reading the article. This problem of causal inference also applies to the question as to whether the backfire effect affects attitudes or whether this relationship is the other



way round and attitudes affect backfire effects. We tried to minimize this threat to internal validity by basing our assumptions on well-justified and plausible theoretical considerations.

Third, conducting this study online caused the problem that we could not force participants to answer the second questionnaire exactly after 5 days. Hence, the amount of time between the first and the second measure differed somewhat. However, we ensured that at least 2 days had elapsed between stimulus presentation and the second data collection. Moreover, around half of the respondents filled out the questionnaire exactly after 5 days. Finally, we were able to show that the number of errors was independent of how many days had passed between both measurements.

### *Implications and Future Research Directions*

The present study again confirms that the backfire effect is a systematic error. It not only occurs in the context of information campaigns (as tested by Skurnik et al., 2007) but also in the context of journalistic coverage. Furthermore, the fact that the participants in our study read more than one article (as they normally would during newspaper reception) causes the effect to occur even after a short delay of only a few minutes. Taking into account how much information on different topics people are confronted with every day, we might even have underestimated the strength of this phenomenon to date.

In addition to the study by Skurnik et al. (2007), we were able to show that the backfire effect not only occurs when statements with a rather negative valence (e.g., “Side effects of the flu vaccination are worse than the flu”) are labeled as false but also when positive statements are identified as myths (e.g., “The bowel cancer test is recommended by independent IGeL monitoring”). This is important insofar as studies in the context of framing research have shown that negative statements are per se more likely to be rated as true than positive ones are (Hilbig, 2009, 2012; Koch et al., 2013). Admittedly, to control for this bias, the valence of statements needs to be manipulated within an experimental design.

The occurrence of the backfire effect in the context of journalistic myths and facts stories implies that the risk of corrected misinformation backfiring is not limited to the medical sector, since this type of journalistic article deals with a variety of issues. Recently, a quality newspaper tried to rectify common myths about the Sinti and Romanies: It did so by printing out several myths (e.g., “Sinti and Romanies are often criminals”) in bold letters and then correcting them with scientific data (e.g., “There is scientific evidence that Sinti and Romanies are not more often criminals than the majority of the population”). Based on the results of the present study, there is reason to fear

that such an article might even strengthen prejudice against minorities. Furthermore, we were able to show that erroneous memory in terms of misinformation can have severe consequences, since it influences people's attitudes toward an issue. As prior research has consistently shown that people base their actions on attitudes (for an overview, see Fishbein & Ajzen, 2005), the backfire effect might be far more problematic than previously assumed.

The results for the judgment-formation strategy have two important implications: first, forming a judgment during reception significantly reduces backfire effects and leads people to be better at telling which information was actually true and which was false. Second, and even more important, the fact that people with existing attitudes erroneously remember misinformation as true in some cases does not affect their attitudes: If they have formed a positive attitude toward a presented issue, they will keep this attitude regardless of what they remember. This fact is crucial for information campaigns about medical myths: For example, if people read an information campaign that corrects misinformation about the threats of the cervical cancer vaccination and form a rather positive attitude about that vaccination during reception, it is likely that they will keep this positive attitude even if they are unsure about the truth of single arguments later.

Taking these results together, there are two main recommendations for designing information campaigns that deal with immediate correction of misinformation or—as in our stimuli—for journalistic myths and facts stories: First, journalists could try to work only with facts and not repeat common myths about the issue at all (Skurnik et al., 2007). Yet this might bear the risk that facts and myths will coexist in people's memory. A second strategy could benefit from the results regarding judgment-formation strategies: If campaign designers or journalists repeat myths to correct them, they should encourage readers to form attitudes during reception, for example, by inserting claims like "What is your opinion?" or "Make up your mind!" If such indications are in fact sufficient in triggering immediate judgments, this needs to be investigated by future research. However, it has to be noted that such claims may only be helpful in debunking myths for journalistic myths and facts stories or information campaigns where myths are immediately rectified. If myths are published and only rectified after some time has passed (as in our introductory example regarding the MMR vaccination), triggering immediate judgments may not have any benefits or might even be counterproductive by contributing to the solidification of a myth.

### **Declaration of Conflicting Interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## Notes

1. We thank one of the reviewers for pointing out that the original term “on-line” can be misleading since nowadays it refers almost exclusively to Internet usage. To avoid confusion, we decided to relabel this judgment-formation strategy in the present article.
2. To make sure that the different time spans between t1 and t2 did not affect the results, we calculated correlations between days passed and errors. Results show that the number of days passed between t1 and t2 does not affect how many errors participants make after a delay; backfire:  $r(332) = .001, p = .99$ ; fact-false:  $r(332) = .07, p = .24$ .

## References

- Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckmann (Eds.), *Action control: From cognition to behavior* (pp. 11-39). Berlin, Germany: Springer.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179-211. doi:10.1016/0749-5978(91)90020-T
- Anchondo, L. (2010). *Top 10 myths about immigration*. Retrieved from <http://www.immigrationpolicy.org/high-school/top-10-myths-about-immigration>
- Anderson, N. H. (1981). *Foundations of information integration theory*. New York, NY: Academic Press.
- Arkes, H. R., Boehm, L., & Xu, G. (1991). Determinants of judged validity. *Journal of Experimental Social Psychology*, 27, 576-605. doi:10.1016/0022-1031(91)90026-3
- Arkes, H. R., Hackett, C., & Boehm, L. (1989). The generality of the relation between familiarity and judged validity. *Journal of Behavioral Decision Making*, 2(2), 81-94. doi:10.1002/bdm.3960020203
- Bacon, F. T. (1979). Credibility of repeated statements: Memory for trivia. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 241-252. doi:10.1037/0278-7393.5.3.241
- 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, 446-458. doi:10.1037/0096-3445.121.4.446
- Begg, I., Armour, V., & Kerr, T. (1985). On believing what we remember. *Canadian Journal of Behavioural Science*, 17, 199-214. doi:10.1037/h0080140
- Bizer, G. Y., Tormala, Z. L., Rucker, D. D., & Petty, R. E. (2006). Memory-based versus on-line processing: Implications for attitude strength. *Journal of Experimental Social Psychology*, 42, 646-653. doi:10.1016/j.jesp.2005.09.002

- Boehm, L. E. (1994). The validity effect: A search for mediating variables. *Personality and Social Psychology Bulletin*, 20, 285-293. doi:10.1177/0146167294203006
- Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968-1987. *Psychological Bulletin*, 106, 265-289. doi:10.1037/0033-2909.106.2.265
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39, 752-766. doi:10.1037/0022-3514.39.5.572
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/ correlation analysis for the behavioral sciences*. Mahwah, NJ: Lawrence Erlbaum.
- 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. doi:10.1177/1088868309352251
- European Union (2015). *The top 10 myths about TTIP: Separating fact from fiction* [Online information flyer]. Retrieved from [http://trade.ec.europa.eu/doclib/docs/2015/march/tradoc\\_153266.pdf](http://trade.ec.europa.eu/doclib/docs/2015/march/tradoc_153266.pdf)
- Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59, 255-278. doi:10.1146/annurev.psych.59.103006.093629
- Fishbein, M., & Ajzen, I. (2005). The influence of attitudes on behavior. In D. Allbaracín, B. T. Johnson & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 173-222). Mahwah, NJ: Erlbaum.
- Gigerenzer, G. (1984). External validity of laboratory experiments: The frequency-validity relationship. *American Journal of Psychology*, 97, 185-195. doi:10.2307/1422594
- Hamilton, D. L., Sherman, S. J., & Maddox, K. B. (1999). Dualities and continua: Implications for understanding perceptions of persons and groups. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 606-629). New York, NY: Guilford Press.
- Hargreaves, I., Lewis, J., & Speers, T. (2003). *Towards a better map: Science, the public and the media*. London, England: Economic and Social Research Council.
- Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior*, 16, 107-112. doi:10.1016/S0022-5371(77)80012-1
- Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review*, 93, 258-268. doi:10.1037/0033-295X.93.3.258
- Hilbig, B. E. (2009). Sad, thus true: Negativity bias in judgments of truth. *Journal of Experimental Social Psychology*, 45, 983-986. doi:10.1016/j.jesp.2009.04.012
- Hilbig, B. E. (2012). Good things don't come easy (to mind). Explaining framing effects in judgments of truth. *Experimental Psychology*, 59, 38-46. doi:10.1027/1618-3169/a000124
- Johnson, H. M., & Seifert, C. M. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental*

- Psychology: Learning, Memory, and Cognition*, 20, 1420-1436. doi:10.1037/0278-7393.20.6.1420
- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49-81). Cambridge, England: Cambridge University Press.
- Koch, T., Peter, C., & Obermaier, M. (2013). Optimisten glaubt man nicht. Wie sich valenzbasiertes Framing auf die Glaubwürdigkeit von Aussagen und deren Kommunikator auswirkt [Optimists do not believe you. How valence based framing affects the credibility of statements and their communicator]. *Medien & Kommunikationswissenschaft*, 61, 551-567.
- Koch, T., & Zerback, T. (2013). Helpful or harmful? How frequent repetition affects perceived statement credibility. *Journal of Communication*, 63, 993-1010. doi:10.1111/jcom.12063
- Kuklinski, J. H., Quirk, P. J., Jerit, J., Schwieder, D., & Rich, R. F. (2000). Misinformation and the currency of democratic citizenship. *Journal of Politics*, 62, 790-816. doi:10.1111/0022-3816.00033
- Lavine, H. (2002). On-line versus memory-based process models of political evaluation. In K. R. Monroe (Ed.), *Political psychology* (pp. 225-274). Mahwah, NJ: Erlbaum.
- Leiner, D. (2012). *SoSci Panel: The noncommercial online access panel*. Poster presented at the General Online Research Conference, Mannheim, Germany.
- Leiner, D. (2014). *Convenience samples from online respondent pools: A case study of the SoSci Panel* (Working Paper). Retrieved from [http://www.researchgate.net/publication/259669050\\_Convenience\\_Samples\\_from\\_Online\\_Respondent\\_Pools\\_A\\_case\\_study\\_of\\_the\\_SoSci\\_Panel](http://www.researchgate.net/publication/259669050_Convenience_Samples_from_Online_Respondent_Pools_A_case_study_of_the_SoSci_Panel)
- Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13, 106-131. doi:10.1177/1529100612451018
- Mackie, D. M., & Asuncion, A. G. (1990). On-line and memory-based modification of attitudes: Determinants of message recall-attitude change correspondence. *Journal of Personality and Social Psychology*, 59, 5-16. doi:10.1037/0022-3514.59.1.5
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, 87, 252-271. doi:10.1037/0033-295X.87.3.252
- Matthes, J., Wirth, W., & Schemer, C. (2007). Measuring the unmeasurable? Toward operationalizing on-line and memory-based political judgments in surveys. *International Journal of Public Opinion Research*, 19, 247-257. doi:10.1093/ijpor/edm001
- Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32, 303-330. doi:10.1007/s11109-010-9112-2
- Nyhan, B., Reifler, J., Richey, S., & Freed, G. L. (2014). Effective messages in vaccine promotion: A randomized trial. *Pediatrics*, 133, 835-842. doi:10.1016/j.vaccine.2011.11.112

- Parks, C. M., & Toth, J. P. (2006). Fluency, familiarity, aging, and the illusion of truth. *Aging, Neuropsychology, and Cognition, 13*, 225-253. doi:10.1080/138255890968691
- Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 123-205). New York, NY: Academic Press.
- Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition, 8*, 338-342. doi:10.1006/ccog.1999.0386
- 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*, 364-382. doi:10.1207/s15327957pspr0804\_3
- Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science, 9*, 45-48. doi:10.1111/1467-9280.00008
- Roggeveen, A. L., & Johar, G. V. (2002). Perceived source variability versus familiarity: Testing competing explanations for the truth effect. *Journal of Consumer Psychology, 12*, 81-91. doi:10.1207/S15327663JCP1202\_02
- Schwartz, M. (1982). Repetition and rated truth value of statements. *American Journal of Psychology, 95*, 393-407. doi:10.2307/1422132
- 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. doi:10.1016/S0065-2601(06)39003-X
- Skurnik, I., Yoon, C., Park, D. C., & Schwarz, N. (2005). How warnings about false claims become recommendations. *Journal of Consumer Research, 31*, 713-724. doi:10.1086/426605
- Skurnik, I., Yoon, C., & Schwarz, N. (2007). "Myths & Facts" about the flu: Health education campaigns can reduce vaccination intentions. Retrieved from [http://webuser.bus.umich.edu/yoonc/research/Papers/Skurnik\\_Yoon\\_Schwarz\\_2005\\_Myths\\_Facts\\_Flu\\_Health\\_Education\\_Campaigns\\_JAMA.pdf](http://webuser.bus.umich.edu/yoonc/research/Papers/Skurnik_Yoon_Schwarz_2005_Myths_Facts_Flu_Health_Education_Campaigns_JAMA.pdf)
- Srull, T. K., & Wyer, R. S. (1989). Person memory and judgment. *Psychological Review, 96*, 58-83. doi:10.1037/0033-295X.96.1.58
- Tormala, Z. L., & Petty, R. E. (2001). On-line versus memory-based processing: The role of "need to evaluate" in person perception. *Personality and Social Psychology Bulletin, 27*, 1599-1612. doi:10.1177/01461672012712004
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207-232. doi:10.1016/0010-0285(73)90033-9
- Unkelbach, C. (2006). The learned interpretation of cognitive fluency. *Psychological Science, 17*, 339-345. doi:10.1111/j.1467-9280.2006.01708.x
- Unkelbach, C. (2007). Reversing the truth effect: Learning the interpretation of processing fluency in judgments of truth. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 219-230. doi:10.1037/0278-7393.33.1.219

- 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. doi:10.1016/j.concog.2010.09.015
- Wilkes, A. L., & Leatherbarrow, M. (1988). Editing episodic memory following the identification of error. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 40*, 361-387. doi:10.1080/02724988843000168

### **Author Biographies**

**Christina Peter**, PhD, is a postdoctoral research fellow at the Department of Communication Studies and Media Research, LMU Munich, Germany. Her research focuses on media usage and effects, persuasion, and research methods.

**Thomas Koch**, PhD, is a postdoctoral research fellow at the Department of Communication Studies and Media Research, LMU Munich, Germany. His research focuses on persuasion, public relations, and media effects.