Science Communication
2018, Vol. 40(2) 246–257

DOI: 10.1177/1075547018759345
journals.sagepub.com/home/scx
SAGE

Good Sound, Good Research: How Audio Quality Influences Perceptions of the Research and Researcher

Eryn J. Newman[1,2] and Norbert Schwarz[1]

Abstract

Increasingly, scientific communications are recorded and made available online. While researchers carefully draft the words they use, the quality of the recording is at the mercy of technical staff. Does it make a difference? We presented identical conference talks (Experiment 1) and radio interviews from NPR's Science Friday (Experiment 2) in high or low audio quality and asked people to evaluate the researcher and the research they presented. Despite identical content, people evaluated the research and researcher less favorably when the audio quality was low, suggesting that audio quality can influence impressions of science.

Keywords

fluency, science communication, audio quality, truth

Scientific communications—conference talks, academic job talks, radio interviews—are increasingly recorded and made available online. As a speaker, should you be worried about the often modest technical quality of the recording? May the skill of the technical staff affect a viewer's impression of the

[1]University of Southern California, Los Angeles, CA, USA
[2]Australian National University, Canberra, ACT, Australia

Corresponding Author:
Eryn J. Newman, Research School of Psychology, Australian National University, Acton, ACT 0200, Australia.
Email: eryn.newman@anu.edu.au

quality of your research? During a radio interview, does the sound quality of your phone connection influence a listener's evaluation of you and your research? Two experiments with the audio quality of conference talk videos and radio interviews suggest the answer to these questions is a resounding, Yes!

Psychological research shows that when messages are difficult to process people think they are less compelling (Schwarz, 2015; Schwarz, Newman, & Leach, 2016). This is sensible when the processing difficulty arises because the message is convoluted or has poor logic. But incidental variables that are unrelated to the content of a message can also produce a feeling of difficulty. A large body of work on *cognitive fluency* shows that people monitor cognitive processing from one moment to the next and are sensitive to changes in the ease or difficulty of ongoing cognitive activity (for a review of cognitive fluency, see Alter & Oppenheimer, 2009; Schwarz, 2015). Because people are more sensitive to their processing experience than to where this experience comes from, they tend to misread difficulty in processing arising from incidental variables as bearing on the quality of the content of a message (e.g., Schwarz, 2015; Schwarz et al., 2016). That is, people can misread difficulty arising from presentation variables—such as print fonts, color contrast, or noise—as bearing on whether a message makes sense, holds true or is from a trustworthy source.

For example, the same message is less likely to be accepted as true when the print font is difficult to read. In one study, Reber and Schwarz (1999) asked subjects to evaluate the truth of trivia claims presented in high color contrast (e.g., dark blue text on a white background) and low color contrast font (e.g., yellow text on a white background). Although the content of the messages was the same, people rated the same claims as less likely to be true when presented in low color contrast. Phonetic elements of a message can produce similar effects—people are less likely to believe messages delivered in an accent that is difficult to understand (Lev-Ari & Keysar, 2010). And people are less likely to believe a message when it is attributed to a speaker with a difficult to pronounce name (Newman et al., 2014). While print color, accent, and the pronunciation of a name have no bearing on the truth of a message, these variables can create a kind of cognitive stumble that people tend to misread as a sign that there is something wrong with the content of the message.

The variables that influence our evaluations of a message also influence our evaluations of a messenger. People who bear difficult to pronounce names also seem unfamiliar and less trustworthy (Laham, Koval, & Alter, 2012; Newman et al., 2014). And essays that are difficult to read—either because the handwriting is complicated or the words are unusual—seem like they were written by a less intelligent author (Greifeneder et al., 2010; see also Oppenheimer, 2006). Taken together, it is well established that incidental variables that create a cognitive stumble (referred to as a feeling of disfluency) can systematically lower our impressions of a message and a messenger. What is less clear is whether poor sound quality would exert the same effects.

Moreover, research in social and cognitive psychology has yet to examine such effects in the context of science communication. The majority of studies examining the effects of cognitive fluency on truth and credibility draw on trivia claims, not findings from scientific research. It is therefore unclear whether people's perceptions of a messenger who is a scientist and a message derived from science—two sources of credibility—would be susceptible to a tangential variable such as sound: Would the disfluency resulting from poor sound quality be sufficient to hurt listeners' impressions of the science and scientist, essentially putting the researcher at the mercy of the recording staff?

We examined these questions in two experiments. In Experiment 1, we selected two conference talks from YouTube and altered their acoustic quality. Participants watched both talks, one presented with high audio quality and one presented with low audio quality, and rated the quality of the talks, the intelligence of the speaker, how much they liked the speaker, and the importance of the research. In Experiment 2, we selected two radio interviews from NPR's Science Friday and altered the apparent phone line quality of the researcher. Participants listened to both interviews, one presented with high and one with low audio quality, and rated the quality of the research, the competence of the researcher, how good the interview was, and whether they would share the interview on social media. If poor audio quality creates a feeling of disfluency for the listener, it should hurt the substantive evaluation of the research and researcher.

Experiment 1

Method

Participants and Design. Ninety-seven Amazon Mechanical Turk workers listened and responded to both segments. We manipulated audio quality within-participants, who either saw (1) the engineering talk in good and physics talk in poor audio quality or (2) the physics talk in good and the engineering talk in poor audio quality. Participants were randomly assigned to conditions using a randomizer tool on Turkitron.com.

Materials and Procedure. We selected two conference talks (in physics and engineering) from YouTube and altered their acoustic features using iMovie software. The good audio quality version of each talk was created with an audio filter called "small room," which reduces the echo and increases the clarity of the speaker; the poor audio quality version was created with an audio filter called "Large Room," which does the opposite, increasing the echo and decreasing the clarity of the speaker. We trimmed the talks to 2- to 3-minute segments.

The materials were presented via Qualtrics and the conference talks were embedded in a YouTube video format within the questionnaire. Immediately after watching each clip, participants evaluated the talk and the speaker using 5-point rating scales (with 5 being the most positive evaluation) in response to the following questions: (1) How good do you think the talk was? (2) How smart do you think the speaker is? (3) How much do you like the speaker? (4) How important do you think this research is?

Results and Discussion

As predicted, the audio quality of the video influenced viewers' evaluations of the research and researcher (top panel of Figure 1). When the video was difficult to hear, viewers thought that the talk was worse, the speaker less intelligent and less likeable, and the research less important. This influence of audio quality was obtained for both presentations (bottom panel of Figure 1). Moreover, its impact was sufficient to reverse viewers' preference order. As a comparison of the dark grey bars with the light grey bars across each panel Figure 1 indicates, whoever had the better audio quality was considered the better speaker and researcher, working on a more important project.

Technically, a 2 (audio quality: high, low) × 4 (ratings: talk, smart, like, important) repeated-measures analysis of variance (ANOVA) showed that when people heard a talk in good audio quality, they were more impressed with the researcher and the research ($F[1, 96] = 40.30$, $p < .001$, raw mean scale difference .60, 95% confidence interval [CI; .41, .79]; see top panel of Figure 1). This pattern held across each rating but was slightly less pronounced for judgments about intelligence (mean difference between fluent and disfluent clips, Talk Good = .93, 95% CI [.68, 1.17], Speaker Smart = .33, 95% CI [.12, .54], Like Speaker = .69, 95% CI [.44, .94], Research Important = .51, 95% CI [.28, .74]). These differences in effect size are reflected in an interaction between audio quality and ratings, $F(3, 94) = 7.12$, $p < .001$.

We also examined these effects separately for each of the talks—would the effect of audio quality hold for the physicist and the engineer? The answer is yes (lower panel of Figure 1). People who heard the physicist in good audio quality ($M = 3.83$, 95% CI [3.63, 4.03]) rated the physicist as a better speaker, presenting better research than people who heard the physicist in poor audio quality (M

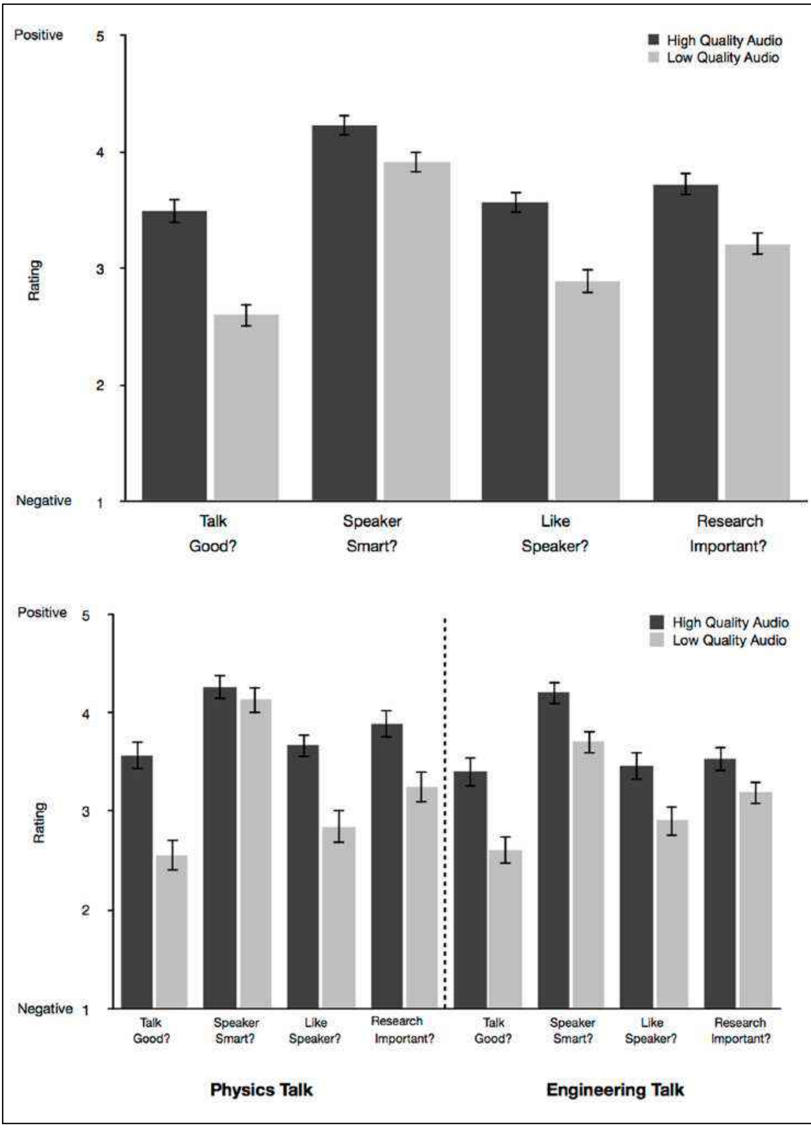

Figure 1. The top panel displays mean ratings of the talk and researcher by audio quality (High vs. Low). The lower panel displays these same means split by video. This lower panel is a between-subject comparison; participants either saw the High Quality Audio Physics Talk + Low Quality Audio Engineering Talk or the Low Quality Audio Physics Talk + High Quality Audio Engineering Talk. Note that error bars represent 1 SE.

= 3.21, 95% CI [2.98, 3.45]). People who heard the engineer in good audio quality (M = 3.64, 95% CI [3.42, 3.86]) rated the engineer more highly than people who heard the engineer in poor audio quality (M = 3.08, 95% CI [2.87, 3.30]).

The means and confidence intervals further indicate systematic preference reversals between the physics and engineering research: Viewers' evaluations always favored the research and researcher presented in better audio quality (raw mean preference for physicist, when physicist was high quality and engineer was low quality .61, 95% CI [.31, .92]); raw mean preference for engineer, when engineer was high quality and physicist was low quality .57, 95% CI [.26, .87]). A 2 (talk: Physicist, Engineer) × 4 (ratings: talk, smart, like, important) × 2 (Fluency condition: Physicist fluent/Engineer disfluent, Engineer fluent/Physicist disfluent) repeated-measures ANOVA further showed the Talk × Fluency Condition interaction, $F(1,95) = 38.75$, $p < .001$, underlying these patterns. A Talk × Fluency Condition × Ratings interaction, $F(1,93) = 7.45$, $p < .001$, showed as in Figure 1 (lower panel), that the preference for better audio held for both talks and across all ratings, but did not reach statistical significance for the rating of how smart the physics speaker was (raw mean difference .09, 95% CI [–.29, .47]).

In short, substantive evaluations of the research and researcher were at the mercy of the technical quality of the recording.[1] However, the videos used in Experiment 1 showed graduate students presenting at a nondescript conference. That is, participants had very few cues about the status and quality of the researcher. Would audio quality also influence people's evaluations when the context is well known and highly credible? It is possible that audio quality would have little bearing on people's judgements when they can draw on the context in which a message is delivered to evaluate the research and researcher. Indeed, variables that usually influence the perceived truth of a message such as repetition have little bearing on people's judgments of truth when the credibility of the source is known (e.g., Begg, Anas, & Farinacci, 1992; Henkel & Mattson, 2011; Unkelbach & Stahl, 2009). To replicate the effects of Experiment 1 in this alternative context with more cues to credibility, we manipulated the audio quality of two radio interviews on NPR's Science Friday, a highly respected science journalism broadcast that features well credentialed researchers. If the audio effect holds under these conditions, it would suggest that audio quality can overshadow not only content, but also a highly credible source.

Experiment 2

Method

Participants and Design. Ninety-nine Amazon Mechanical Turk workers listened and responded to both radio segments. We manipulated audio quality within-participants, who either heard (1) the physics interview in good and the genetics interview in poor audio quality or (2) the genetics interview in good and the physics interview in poor audio quality. Participants were randomly assigned to conditions using a randomizer tool on Turkitron.com.

Materials and Procedure. We selected two Science Friday interviews from NPR.com (in physics and genetics) and altered their acoustic features using iMovie software. The good audio quality version of each talk was created with no audio filters so that participants heard the interview as it was originally recorded. The poor audio quality version was created with audio filters that made it sound as if the researcher had called in on a bad phone line. We trimmed the talks to 2- to 3-minute segments.

The materials were presented via Qualtrics and the NPR interviews were embedded in a YouTube video format within the questionnaire (the screen remained black, while they listened to each clip). Immediately after listening to each clip, participants evaluated the talk and the speaker using 5-point rating scales (with 5 being the most positive evaluation) in response to the following questions: (1) How good do you think the research is? (2) How competent do you think the researcher is? (3) How good do you think the interview was? (4) How likely is it that you would share this interview on social media?

Results and Discussion

Although Science Friday is a reputable source of high quality science news, listeners' evaluations of the researcher and the research were again less favorable when the audio quality was poor (Figure 2). This observation held for each speaker and each evaluation (bottom panel Figure 2).

A 2 (audio quality: high, low) × 4 (ratings: research, competence, interview, share) repeated-measures ANOVA showed that when people heard an interview in good audio quality, they were more impressed with the researcher and the research ($F(1, 98) = 5.29$, $p = .02$, raw mean scale difference .19, 95% CI [.03, .36]; see top panel of Figure 2). This pattern held across each rating—how good the research was, competence of the researcher, how good the interview was, and whether they would share on social media.[2] We also found that, in general, people had the highest ratings for the researcher's competence (M = 4.20, 95% CI [4.06, 4.34]), followed by how good the research was (M = 3.86, 95% CI [3.69, 4.02]), followed by how good the interview was (M = 3.47, 95% CI [3.30, 3.66]), and the lowest rating was the

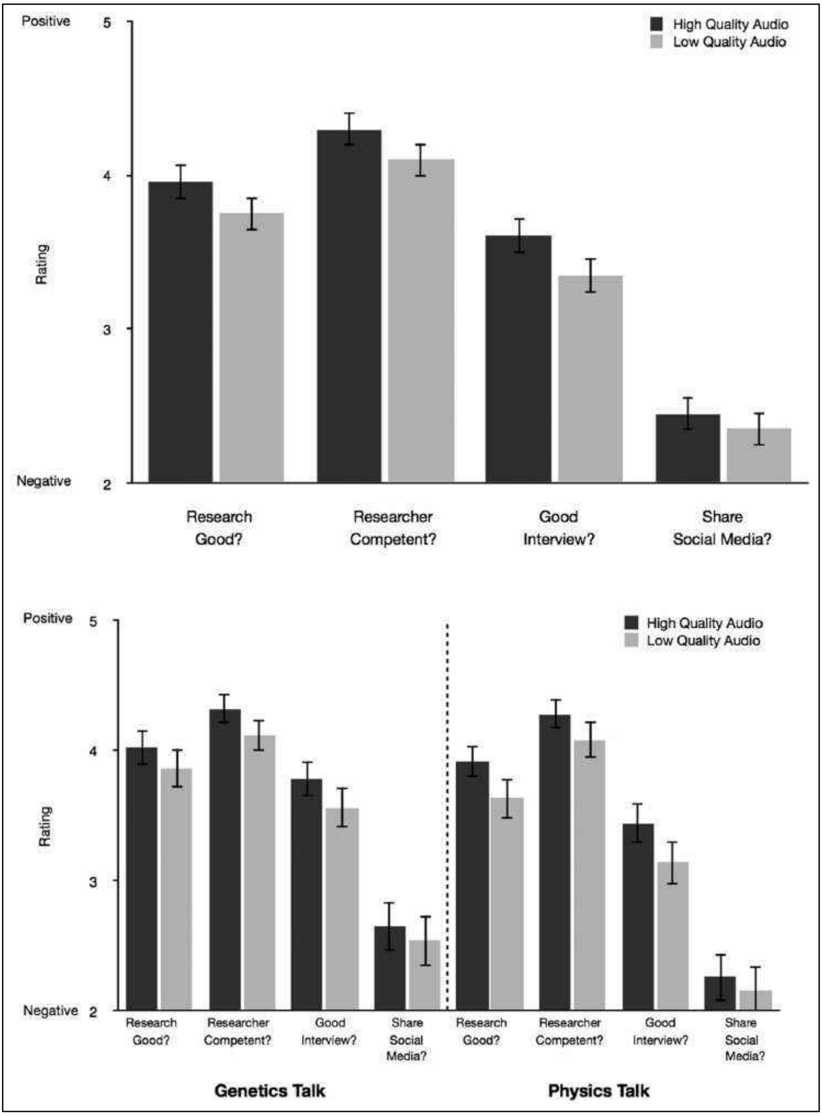

Figure 2. The top panel displays mean ratings of the research and researcher by audio quality (High Quality vs. Low Quality). The lower panel displays these same means split by interview. This lower panel is a between-subjects comparison; participants either saw the High Quality Audio Physics Interview + Low Quality Audio Genetics Interview or the Low Quality Audio Physics Interview + High Quality Audio Genetics Interview. Note that error bars represent 1 SE.

likelihood that they would actually share the interview on social media (M = 2.40, 95% CI [2.17, 2.64]); $F(3, 96) = 58.48, p < .01$.

As in Experiment 1, and as the lower panel of Figure 2 shows, a between-participants comparison showed that this directional pattern held for both radio interviews (physicist in good audio quality, M = 3.48, 95% CI [3.26, 3.69]; physicist in poor audio quality, M = 3.26, 95% CI [2.99, 3.52]; geneticist in good audio quality, M = 3.69, 95% CI [3.48, 3.90]; geneticist in low quality, M = 3.52, 95% CI [3.29, 3.75]). A 2 (Interview: Physicist, Geneticist) × 4 (ratings: good, competent, interview good, share) × 2 (Fluency condition: Physicist fluent/Geneticist disfluent, Geneticist fluent/Physicist disfluent) repeated-measures ANOVA showed a Talk × Fluency Condition interaction, $F(1, 97) = 5.86, p = .02$.

We also found a Talk × Ratings interaction, $F(3, 95) = 4.13, p = .01$, while people rated the interviews similarly on research importance (raw mean difference M = .16, 95% CI [−.37, .04]) and researcher competence (raw mean difference M = .04, 95% CI [−.14, .22]), they rated the Genetics interview as being a better interview (raw mean difference M = .37, 95% CI [.15, .60]) and said they were more likely to share the interview on social media (raw mean difference M = .38, 95% CI [.15, .62]).

General Discussion

These findings converge with an extensive body of research into the role of processing fluency in evaluative judgment. The more difficult a statement is to process, the less likely it is to be judged true and compelling—even when the difficulty derives solely from incidental features, such as the color contrast of the print font (Reber & Schwarz, 1999), the accent of the speaker (Lev-Ari & Keysar, 2010), the ease with which the name of a source can be pronounced (Newman et al., 2014), or whether the material rhymes (McGlone & Tofighbakhsh, 2000). Our findings add to this body of work and show that the technical quality of recordings can profoundly affect the substantive evaluation of the research and researcher, even to the extent of reversing their relative standing when pitted against a presentation with better audio quality. Clearly, ensuring good audio quality is in the researcher's and reporter's interest.

Unfortunately, ensuring good audio quality is not always possible and sometimes phone connections simply are bad. What can be done in such cases? Experimental research shows that people are less likely to misinterpret processing disfluency when they are aware of it and attribute it to the correct source (for a review, see Schwarz, 2012). For example, in one study, an essay was difficult to read because the printer was low on toner (Oppenheimer, 2006). Explicitly making readers aware of this technical problem reduced the otherwise observed negative impact of poor visual quality on readers' essay evaluation. This suggests that alerting listeners to poor audio quality may similarly limit its adverse impact.

These considerations converge on a straightforward but important message to science communicators: Next time you are recorded, make sure that the audio quality is good. Listeners are likely to attribute any difficulty they have in understanding you to the quality of your presentation, and the quality of your research. Sometimes it may be better not to be recorded at all than to accept the adverse consequences of a poor recording.

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. Obviously, poor sound quality may also hurt perceptions of the researcher and research because recipients cannot understand the content. To rule out this possibility we asked people at the end of each audio clip what the clip was about. When we limit analysis to those who correctly responded about the content of both talks (64% of our sample), we replicate the same pattern of results—poor sound quality hurts impressions of the research and researcher. That is, the effect of sound quality is not simply driven by people who had trouble understanding the content of the talk.
2. The finding that people said they would be *more likely to share an interview* with high sound quality could be driven by factors beyond a feeling of fluency—it is also plausible that this preference for the high quality clip is partly driven by a concern that other people on social media can hear easily.

References

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, *13*, 219-235. doi:10.1177/1088868309341564

Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General*, *121*, 446-458.

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.

Henkel, L. A., & Mattson, M. E. (2011). Reading is believing: The truth effect and source credibility. *Consciousness and Cognition*, *20*, 1705-1721.

Laham, S. M., Koval, P., & Alter, A. L. (2012). The name-pronunciation effect: Why people like Mr. Smith more than Mr. Colquhoun. *Journal of Experimental Social Psychology*, *48*, 752-756.

Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, *46*, 1093-1096.

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

Newman, E. J., Sanson, M., Miller, E., 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*, e88671. doi:10.1371/journal.pone.0088671

Oppenheimer, D. M. (2006). Consequences of erudite vernacular utilized irrespective of necessity: Problems with using long words needlessly. *Applied Cognitive Psychology*, *20*, 139-156.

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

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

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: American Psychological Association.

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

Unkelbach, C., & Stahl, C. (2009). A multinomial modeling approach to dissociate different components of the truth effect. *Consciousness and Cognition*, *18*(1), 22-38.

Author Biographies

Eryn J. Newman (PhD, Victoria University of Wellington, New Zealand, 2013) is a lecturer at the Australian National University and a research fellow and former post-doc at the USC Dornsife Mind & Society Center. Her research addresses distortions of cognition and memory and ways to correct them. She currently focuses on how anecdotes, photographs, and other tangential, nonprobative information can influence judgments of truth. Her work has appeared in leading journals of the field and has been funded by Fulbright.

Norbert Schwarz (DrPhil, University of Mannheim, Germany, 1980) is provost professor of psychology and marketing at the University of Southern California and co-director of the USC Mind & Society Center. He investigates the interplay of feeling and thinking and the context-sensitive and embodied nature of judgment. He has been elected to the American Academy of Arts and Sciences and German National Academy of Sciences.