American Psychological Association
American Psychology-Law Society
Law and Human Behavior

ISSN: 0147-7307
2021, Vol. 45, No. 5, 481–495
https://doi.org/10.1037/lhb0000466

Sound and Credibility in the Virtual Court: Low Audio Quality Leads to Less Favorable Evaluations of Witnesses and Lower Weighting of Evidence

Elena Bild[1], Annabel Redman[1], Eryn J. Newman[1], Bethany R. Muir[1], David Tait[2], and Norbert Schwarz[3]
[1] Research School of Psychology, Australian National University
[2] School of Humanities and Communication Arts, Western Sydney University
[3] Mind and Society Center, University of Southern California

OPEN DATA OPEN MATERIALS PREREGISTERED

Objectives: Recent virtual court proceedings have seen a range of technological challenges, producing not only trial interruptions but also cognitive interruptions in processing evidence. Very little empirical research has focused on how the subjective experience of processing evidence affects evaluations of trial participants and trial decisions. Metacognitive research shows that the subjective ease or difficulty of processing information can affect evaluations of people, belief in information, and how a given piece of information is weighted in decision making. ***Hypotheses:*** We hypothesized that when people experienced technological challenges (e.g., poor audio quality) while listening to eyewitness accounts, the difficulty in processing evidence would lead them to evaluate a witness more negatively, influence their memory for key facts, and lead them to weigh that evidence less in final trial judgments. ***Method:*** Across three experiments (total $N = 593$), participants listened to audio clips of witnesses describing an event, one presented in high-quality audio and one presented in low-quality audio. ***Results:*** When people heard witnesses present evidence in low-quality audio, they rated the witnesses as less credible, reliable, and trustworthy (Experiment 1, $d = 0.32$; Experiment 3, $d = 0.55$); had poorer memory for key facts presented by the witness (Experiment 2, $d = 0.44$); and weighted witness evidence less in final guilt judgments (Experiment 3, $\eta_p^2 = .05$). ***Conclusion:*** These results show that audio quality influences perceptions of witnesses and their evidence. Because these variables can contribute to trial outcomes, audio quality warrants consideration in trial proceedings.

Public Significance Statement

Low audio quality led people to evaluate witnesses more negatively and reduced their memory for presented evidence. Even when memory for key facts was held constant, poor audio quality led people to weight evidence less heavily in final judgments of guilt. These findings suggest several policy implications regarding the conditions under which remote testimony is presented in the courtroom and the critical importance of controlling for a high-quality acoustic experience in the physical courtroom more generally.

Keywords: virtual courtrooms, audio quality, cognitive fluency, witness credibility

Supplemental materials: https://doi.org/10.1037/lhb0000466.supp

Jennifer S. Hunt served as Action Editor.

Elena Bild https://orcid.org/0000-0002-8758-4293
Eryn J. Newman https://orcid.org/0000-0001-8663-7173
Bethany R. Muir https://orcid.org/0000-0002-1981-6945
Norbert Schwarz https://orcid.org/0000-0002-8868-7067
Elena Bild is now at the School of Psychology, University of Queensland.

We have no known conflict of interest to disclose.

The data are available at https://osf.io/vf34x/?view_only=3cffa073590749e69415a32878be7146.
The experiment materials are available at https://osf.io/vf34x/?view_only=3cffa073590749e69415a32878be7146.
The preregistered design is available at https://osf.io/vf34x/?view_only=3cffa073590749e69415a32878be7146.

Correspondence concerning this article should be addressed to Eryn J. Newman, Research School of Psychology, Australian National University, Building 39, Science Road, Canberra, ACT 2601, Australia. Email: eryn.newman@anu.edu.au

From intermittent internet connections, to video freezing, to full audio dropout—these are just some of the reasons trials have been paused, witnesses recalled, and remote hearings abandoned in virtual court settings (e.g., Lapinski et al., 2020; Todd et al., 2020). These interruptions have obvious consequences for human judgment and cognition: If the internet drops out, an eyewitness presenting evidence may end up repeating themselves, and if the audio disappears, jury members will be missing out on critical evidence. But more insidious technological glitches may also affect trial proceedings. An echo or static on the audio may be enough to influence jury decision making, even when the jury is not actually missing out on content and merely experiences some difficulty in listening to evidence. Such smaller glitches are rarely enough to cause a pause to a trial, or to have a witness recalled, but may nonetheless systematically affect human judgment and warrant consideration by the courts.

Over the last several years, technology has been increasingly incorporated into the criminal justice system. One great benefit is that technology allows various actors to contribute remotely when they cannot attend the court in person. In 2020, courts all over the world rapidly turned to remote testimony, with more than 60 countries implementing some form of virtual court during coronavirus disease (COVID-19) restrictions (Remote Courts Worldwide, 2020). In England and Wales, for instance, shortly after lockdown procedures were in place in early 2020, there was a 500% increase in audio hearings and a 340% increase in video hearings across all courts and tribunals (Ryan et al., 2020). There are many advantages to this remote or virtual court shift, including lowered costs, more flexibility, and greater access to resources for rural and disadvantaged communities (e.g., Gourdet et al., 2020). But remote attendance changes the courtroom experience for all involved, and its impact on decision making is not yet well understood (McIntyre et al., 2020; Rossner et al., 2021). Moreover, policy and procedures are in their infancy, with significant variations in how remote procedures are conducted (e.g., Scotland using cinemas to accommodate juries who watched the trial on the large screen; Barrie, 2020; also see Norton Rose Fulbright, 2020). Responses to disruptions in audiovisual links, such as procedures for asking parties to disconnect and reconnect when their audio connection is poor, are also likely to differ between trials (Byrom et al., 2020). Although audio–visual links (AVL) and Zoom platforms allow for the use of both video and audio inputs, we focus on audio because it is also an alternative form of court attendance: One may join remotely using audio only in some jurisdictions (e.g., Australia, Federal Circuit Court of Australia, 2020; United Kingdom, Byrom et al., 2020; United States, Totenberg, 2020), and whereas one can turn off video if a connection is poor, audio evidence is critical in contributing content in a case. Anecdotal examples capture the challenges with low-quality audio:

> On repeated occasions throughout the proceedings that afternoon one or more of the parties "dropped out," necessitating a communication between them and my tipstaff advising of the steps they should take to "dial back in." Reconnection was successful on each occasion, although not without interruption to the course of the proceedings. From time to time counsel were also difficult to hear and on other occasions their submissions were fractured or time delayed. Despite the valiant endeavours of the court reporters, the integrity of the transcript suffered as a result. (Boys & Sams, 2020, The Effectiveness of the Virtual Courtroom section)

Indeed, the third author experienced these audio issues recently when presenting expert evidence remotely. Although the video connection was intact, the quality of audio in the physical courtroom was such that evidence had to be repeated and clarified several times over (see also Rowden & Wallace, 2019, for an extended critique of the complexities involved in remote expert testimony). Beyond the context of a worldwide pandemic, courts rely on remote testimony in a number of contexts. Witnesses can join hearings remotely via AVL for good cause, for example (for a discussion, see Fobes, 2020), and when children present evidence in a closed-circuit context, they are also at the mercy of the quality of court technology in the physical courtroom. Complete audio dropout and prolonged freezing are strong diagnostic cues to a judge that information is being lost, having potentially significant effects for human judgment. However, smaller glitches in audio—some static, some echo—may go undetected or may seem of little concern even when detected. Whether there is a smooth connection is not easily controlled by a speaker, who is usually at the mercy of their Wi-Fi and network, but this variable may have significant consequences for how a speaker is perceived and how their evidence is received by a jury. Audio quality is also relevant in in-person settings, where the architecture, technology in the court, and distance can all bear on the quality of the listener's experience. Although audio is independent of the reliability of a given witness and the content of a witness report, it may nonetheless affect a listener's perceptions of a witness and what they say and the extent to which their evidence is used in decision making. The present research tested this possibility.

Metacognitive Experiences

A large body of research in social and cognitive psychology shows that processing the content of a message is accompanied by subjective experiences of ease or difficulty, which can shape how much people trust the communicator, agree with the message, remember its details, and rely on it when making a decision (for reviews, see Alter & Oppenheimer, 2009; Schwarz, 2015; Schwarz et al., 2021). These metacognitive experiences of ease or difficulty are often referred to as fluency experiences, denoting how "fluently" some cognitive operation can be executed. The inferences that people draw from their metacognitive experiences are often warranted. For example, convoluted messages are more difficult to process than logical ones, and arguments that are at odds with one's knowledge are more difficult to follow than arguments that are not (Schwarz, 2015). But unfortunately, people are more sensitive to their own fluency experience than to the source of that experience—they notice that something is easy or difficult but do not realize where that difficulty comes from. For example, is a message difficult to process because its content makes no sense or because the audio is difficult to understand? In most cases, people attribute their difficulty to the content they are focused on rather than to the influence of incidental background variables. Hence, numerous incidental variables can influence recipients' evaluation of the content. For example, in one study, people saw claims printed in easy-to-perceive high-color-contrast fonts or in difficult-to-perceive low-color-contrast fonts. When participants were asked to rate the truth of those claims, they were less likely to believe a given claim when it was presented in low-color-contrast font than in high-color-contrast font (Reber & Schwarz, 1999). Similar effects are found with phonetic or audio-perceptual experiences. Accents are one example—when provided audio information through a foreign accent, people rate it less likely to be true (Lev-Ari & Keysar, 2010), provide harsher sentences to a defendant (Romero-Rivas et al., 2021), and find eyewitnesses less credible (Frumkin, 2007). Other research has shown that people evaluate academic conference talks more negatively when there is a (simulated) slight echo on the microphone (Newman & Schwarz, 2018; see also Fiechter et al., 2018). In short, whenever information is difficult to perceive, understand, or imagine, when processing is clunky or strained, recipients evaluate the substantive content of the information more negatively.

The same holds for the impressions people form of a speaker. For example, people evaluate speakers with a difficult-to-pronounce name more negatively than speakers with an easy-to-pronounce name (even from the same world region)—those with complicated names are rated less trustworthy, less familiar, and more dangerous (Laham et al., 2012; Newman et al., 2014). In one study, people were more concerned about making a purchase when an online seller had a complicated or difficult-to-pronounce username—they were less confident that the product description was accurate, wondered whether the seller would honor the return policy, and found the seller less trustworthy overall (Silva et al., 2017). These effects of pronunciation and usernames were robust and held even when people had access to more objective information about the seller's reputation. Similarly, faces that have been seen less often and are thus less easy to process compared to repeatedly seen faces, seem relatively less sincere and honest (Brown et al., 2002; see also Weisbuch & Mackie, 2009). And even scientific experts are perceived to be less competent when there is background noise in a radio interview (Newman & Schwarz, 2018). Taken together, incidental variables that produce cognitive difficulty in processing information about people, events, and products can have systematic consequences in human judgment: When processing is difficult, people arrive at more negative evaluations. This holds across all modes of information presentation, from the auditory presentations discussed here to the readability of written material (for reviews, see Alter & Oppenheimer, 2009; Schwarz et al., 2021).

Fluency in the Courtroom

Assessments of credibility, reliability, and trustworthiness are crucial factors within the context of the criminal justice system (Martire et al., 2020). Indeed, a vast literature has examined the psychological variables that influence these assessments—for instance, whether a trial participant is perceived to have high or low socioeconomic status (Espinoza & Willis-Esqueda, 2008), what clothing they are wearing (Fontaine & Kiger, 1978), or whether they are perceived to be in custody (Rossner et al., 2017), to name just a few. This existing research shows that relatively tangential attributes of a witness can reliably influence perceived witness credibility (e.g., McKimmie et al., 2013; see also Koehler et al., 2016). But the metacognitive research reviewed above suggests that the same will hold for variables that are even more distal from the witness, such as the audio quality in which a witness statement is recorded or the print font in which it is transcribed. We provide a first test of this possibility by focusing on a variable that is important in physical court spaces and can vary depending on technology but is particularly relevant in the virtual courtrooms that proliferated in response to the COVID-19 pandemic—namely, audio quality.

In Experiment 1, we asked people to listen to two children giving testimony about an innocuous experience (going to the doctor or going to the movies) and then to rate the credibility, reliability, and trustworthiness of each witness. For one testimony, the audio was altered such that there was a slight delay creating an echo when the witness spoke; for the other, the audio was enhanced so that the witness was particularly easy to hear. In Experiment 2, we used the same materials, but instead of testing perceptions of the eyewitness, we tested people's memory for the evidence. Finally, in Experiment 3, we created new materials from adult eyewitnesses and extended the design such that people first read a full trial summary and estimated the guilt of the defendant, then encountered additional eyewitness testimony (in high- or low-quality audio) and were asked to evaluate the witness and estimate the guilt of the defendant a second time. We expected that across all experiments, audio quality would influence evaluations of the witness and their evidence—specifically, that people would evaluate testimony less favorably when it was presented in low rather than high audio quality. In addition, we expected that low audio quality may reduce people's memory for the content of the evidence and, when we controlled for memory, low audio quality may still reduce the extent to which people weighted that evidence in their impressions of guilt.

Experiment 1

In this first experiment, the primary research question was whether audio quality would influence impressions of an eyewitness and their testimony. We expected that the same testimony, provided by the same witness, would be evaluated less favorably when presented in low- rather than high-quality audio. Note that audio quality is not only irrelevant to the content of the testimony but is also a technical variable over which the witness has little control. The key dependent variables were ratings of credibility, trust, and reliability of the eyewitness.

Method

Participants and Design

We posted 200 slots on Amazon Mechanical Turk, requiring a 90% approval rate, and 195 participants fully completed the study (70 female; age range = 18–67, $M = 34.57$, $SD = 10.67$). We manipulated audio quality (high or low) within participants, who heard descriptions of two (unrelated) target events. Key dependent variables were ratings of credibility, reliability, and trust. We preregistered our sample size as 200, assuming a small effect size, power of .80, and α level of .05. We also conducted sensitivity analyses for each experiment. Setting the α level at .05 and the sample size at 195, we had 80% power to detect an effect size (d) of 0.20. When we considered this sensitivity analysis with more conservative exclusions, to be more confident that people had listened carefully to testimony, with a reduced sample of 157 (after excluding 38 participants) and α level of .05, we had 80% power to detect an effect size (d) of 0.23. As noted below, we find the same significant pattern of results with these participants included or excluded. (All preregistrations, materials, and data can be found on the Open Science Framework https://osf.io/vf34x/?view_only=3cffa073590749e69415a32878be7146.)

Materials and Procedure

In this experiment, participants were asked to act as though they were jurors and imagine they were listening to a child's testimony in a courtroom. They were told that they

> will hear some short snippets of testimony from children describing everyday events. Children's testimony is often presented via an audio link or video link—so what you hear today is in a similar format to what jurors might encounter in a real courtroom setting.

The key manipulation was that one of the audio clips was presented in high-quality audio and one was presented in low-quality audio. In developing the audio clips, we used an iPhone to record two male American children, aged 10 and 11 years, reading scripts describing two everyday events: a trip to the movies and a routine doctor's checkup. Both clips were just over 2 min in length. The recordings were edited using GarageBand software to create high- and low-quality versions. The low-quality version was created using a "delay" effect, which increased the echo, as though the speakers were in a large room. The high-quality version was created using a "noise gate" effect, which removed the background noise and distortion, and a "bright vocal" effect, which enhanced the child's voice. The recordings were pilot tested with a few volunteers who listened to the audio recordings and were able to accurately transcribe (allowing pauses for writing) the low-quality testimony, suggesting that the audio manipulations did not obscure the content. More importantly, we tested this directly in the study by asking participants to provide a short description of the witnessed event. Across counterbalanced conditions, participants heard about (a) a doctor's checkup in high-quality audio and a movie visit in low-quality audio or (b) a movie visit in high-quality audio and a doctor's checkup in low-quality audio. The order of the clips was also counterbalanced between participants. Participants were randomly assigned to conditions via Qualtrics.

After listening to each clip, participants were asked to evaluate the child using 5-point rating scales in response to three key questions: "How credible was the child?" "How reliable was the child?" and "How trustworthy was the child?" (e.g., 1 = *not credible at all*, 5 = *very credible*). Although we combined these ratings to reflect participants' general impressions of the witness in the high-quality and low-quality audio conditions (both Cronbach's αs > .92), we found the same significant pattern of results when we analyzed ratings of credibility, reliability, and trust separately. Participants were also asked to rate how easy it was to understand the child and provide a brief description of the child's testimony. The manipulation check of how easy it was to understand the child showed that high-audio-quality evidence was rated as much easier to understand ($M = 4.07$, $SD = 0.95$) than low-audio-quality evidence ($M = 2.01$, $SD = 1.12$), $t(194) = 21.64$, $p < .001$, $d = 1.98$, 95% CI [1.72, 2.25] (Hedges' g_{av}, to correct for bias in d, as reported throughout). Similar audio effects were used across studies, and we performed this manipulation check of audio only in Experiment 1.

Results and Discussion

Our primary research question in this study was whether audio quality influenced people's perceptions of a witness's statement and impressions of a witness. As Figure 1 shows, the answer is yes. When the audio was difficult to hear, participants rated the witness less favorably than when the audio quality was clear. At an item level of analysis, this audio effect was found for both fluency conditions.

To ensure data quality, we examined whether participants could describe the nature of each testimony. Of the 195 participants, 38 were not able to explain in a short text box what the testimony was about for at least one of the witness accounts. We excluded those participants from our analysis presented here, though we found the same significant pattern of results when we examined the full sample. Participants evaluated witnesses more favorably when they heard the witness present testimony in high-quality audio ($M = 4.07$, $SD = 0.84$) rather the low-quality audio ($M = 3.72$, $SD = 0.89$), $t(156) = 5.44$, $p < .001$, $d = 0.32$, 95% CI [0.25, 0.56].

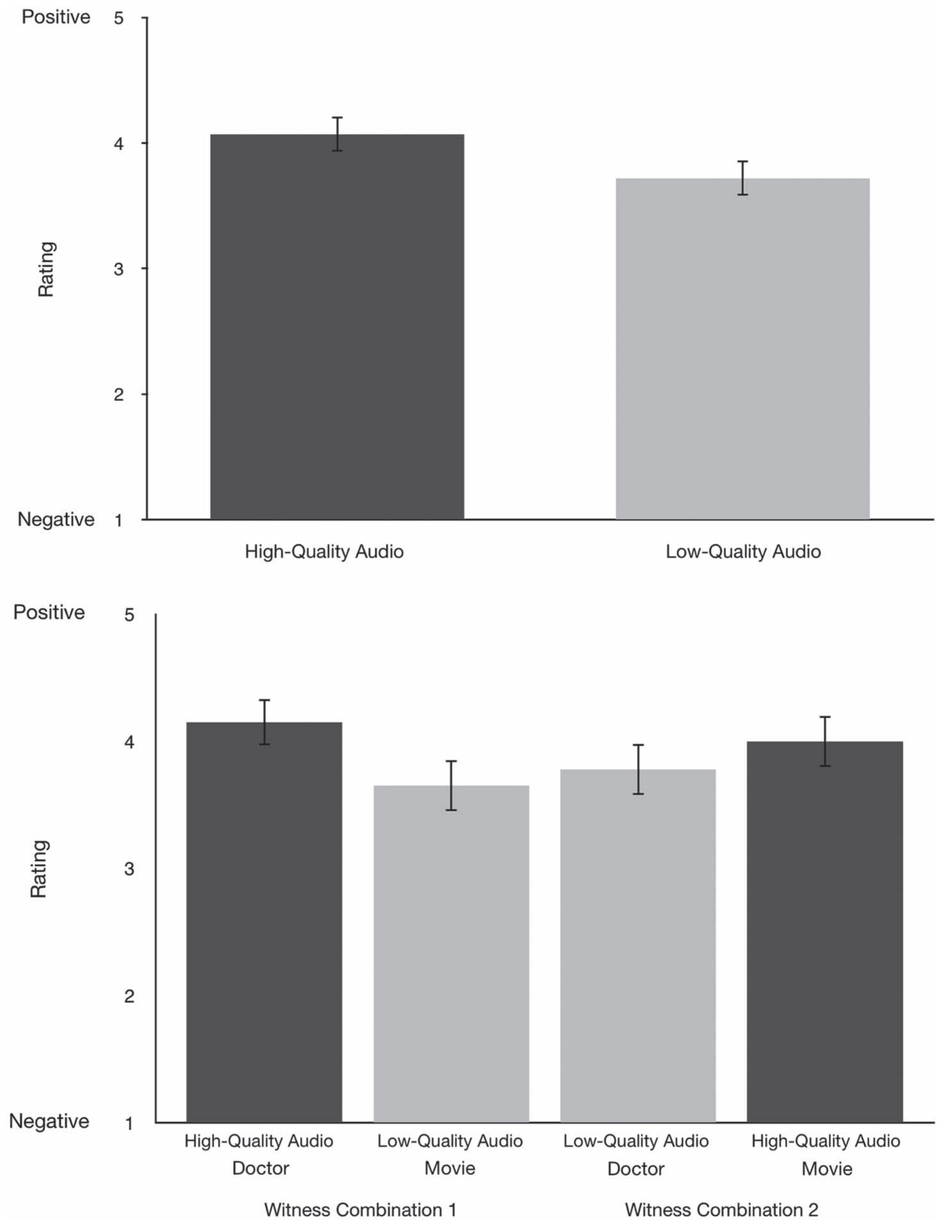
Replicating the pattern above, when we included witness combination in our analysis as a between-subjects variable (whether participants heard the doctor description or movie description in high-quality audio), a 2 (audio quality: high, low) × 2 (witness combination: doctor high/movie low, doctor low/movie high) repeated-measures mixed analysis of variance (ANOVA) showed the expected main effect for audio quality (as shown in the paired t test above), $F(1, 155) = 31.49$, $p < .001$, $\eta_p^2 = .17$, 90% CI [.09, .25], and an interaction with the witness-combination condition emerged, $F(1, 155) = 4.54$, $p = .035$, $\eta_p^2 = .03$, 90% CI [.001, .08]. Follow-up analyses showed that the audio effect held within both witness-combination conditions, but it was larger for the fluent doctor/disfluent movie condition. This may be because the doctor event was relatively more schematic than the movie event, but such a conclusion warrants further investigation. Participants evaluated the high-audio-quality witness more favorably (doctor: $M = 4.15$, $SD = 0.79$; movie: $M = 4.00$, $SD = 0.88$) than the low-audio-quality witness (doctor: $M = 3.78$, $SD = 0.95$; movie: $M = 3.65$, $SD = 0.83$). Paired comparisons supported the pattern observed in the means—high-quality doctor/low-quality movie: $t(73) = 5.09$, $p < .001$, $d = 0.61$, 95% CI [0.36, 0.87]; low-quality doctor/high-quality movie: $t(82) = 2.65$, $p = .010$, $d = 0.24$, 95% CI [0.04, 0.44]. This effect of audio quality appeared to be more robust for the fluent doctor/disfluent movie condition, as seen in the means, effect sizes, and paired comparisons. There was no main effect for the witness-combination condition, $F(1, 155) = 0.02$, $p = .899$, $\eta_p^2 = .00$, 90% CI [.00, .01].

In addition, evaluations of witnesses were dependent on the technical quality of the audio. This effect was independent of the content of the testimony, with a similar pattern being found when the child was testifying about a visit to the movies or to the doctor. These results highlight the importance of audio fluency for perceptions of witnesses.

Experiment 2

In Experiment 2, we explored the possibility that audio quality may influence not only people's perceptions of a witness but also how well they remember facts presented in testimony. Some research has suggested that difficulty in processing can act as a "problem signal" that conveys that something may be off and requires closer attention (Oppenheimer, 2008; Pieger et al., 2017; Song & Schwarz, 2008). This kind of signal could have produced the pattern of results we found in Experiment 1, with participants rating the witness less favorably. However, this problem signal can

Figure 1
Mean Ratings of Witnesses Across High- and Low-Audio-Quality Conditions



Note. The top panel is collapsed across witnessed event; the bottom panel presents means separately for the doctor- and movie-related witness statements. All figures present the data with exclusions present; however, data patterns and significance are the same with the full or reduced sample. Error bars represent 95% confidence intervals.

also influence the processing approach that people apply in a given task. Indeed, an experience of disfluency can lead people to adopt a more analytical processing strategy. For instance, in one study, participants were better at detecting misinformation in questions when those questions were presented in a disfluent format (e.g., in a difficult-to-read font), which made the participants less likely to rely on their intuition (Song & Schwarz, 2008; see also Alter et al., 2007; Liu et al., 2020). Arguably, this more analytical processing style should result in a better memory for materials presented in a disfluent format, as has been found in some studies in educational contexts (for a review, see Oppenheimer & Alter, 2014). Although there is some support for superior memory for disfluent information, other studies have not found this memory advantage (Eitel et al., 2014; Rummer et al., 2016). Indeed, a negative influence of disfluency on the perceived credibility of a witness (as observed in Experiment 1) may even undermine the listener's motivation to pay close attention to the details. If so, memory for the testimony should be worse under disfluent processing. We addressed these possibilities in Experiment 2, examining whether disfluency influenced participants' memory of the testimony and, if so, which direction this influence took. Jurors' memory for witness testimony is crucial to trial verdicts, and indeed, trials may be paused or testimony restarted when audio completely drops out or freezes (e.g., Byrom et al., 2020). But smaller background disfluency may not lead to such procedural interventions in practice. Here, we consider whether these trial interventions are indeed warranted, given the consequences of more general audio disfluency (rather than complete dropout or freezing) on memory for testimony.

Method

Participants and Design

We posted 200 slots on Amazon Mechanical Turk, requiring a 90% approval rate, and 193 participants fully completed the study (82 female; age range = 18–69, $M = 33.02$, $SD = 9.19$). Participants who had completed Experiment 1 were unable to participate in Experiment 2. We preregistered our sample size as 200, assuming a small effect size. We also conducted a sensitivity analysis: Setting an α level of .05 and sample size of 193, we had 80% power to detect an effect size (d) of 0.20. When we considered this sensitivity analysis with more conservative exclusions (requiring that for both witness accounts, participants had clicked to the next page only after finishing the audio clip), to be more confident that participants had listened to the full testimony, with a reduced sample of 128 (after 65 participants were excluded) and α level of .05, we had 80% power to detect an effect size (d) of 0.25. We found the same significant pattern of results with these participants included or excluded. The audio quality was manipulated within participants as in Experiment 1. The key dependent variable was memory for factual evidence, measured as discrimination (d') and response bias (c) using signal detection analysis (Stanislaw & Todorov, 1999).

Materials and Procedure

Experiment 2 was identical to Experiment 1 with the following exceptions. After listening to the audio clip, instead of completing ratings about the eyewitnesses, participants were asked to complete a recognition test consisting of 20 statements about the testimony, designed to test their memory for key facts. Test items consisted of statements such as "The doctor took the boy's height." Participants had to decide whether a statement was "old" (an accurate description of content included in the testimony) or "new" (a related but inaccurate description of content included in the testimony; e.g., "The doctor listened to the boy's heart"). There were 10 old and 10 new statements included in the recognition test, all provided in the Online Supplemental Materials.

Results and Discussion

The primary research question in this study was whether audio quality would influence people's memory for information presented in testimony. The answer was yes, with participants being significantly better at discriminating between old and new test items when the audio quality was high than when it was low. That is, disfluency did not lead to better memory for testimony, but rather, reduced memory for key facts.

To test this research question, we calculated signal detection theory parameters to measure participants' sensitivity on the recognition tests (i.e., their ability to accurately discriminate between old and new items). Response rates were classified either as hits, where an item was included in the testimony and correctly identified as old, or false alarms, where an item was not included in testimony and was incorrectly identified as old at test. From the existing literature, it was unclear whether participants would have higher or lower discrimination after listening to low-quality audio stimuli. In line with the fluency literature more generally, we expected that participants would have an "old" bias (familiarity bias) for high-quality audio and a "new" bias for low-quality audio.

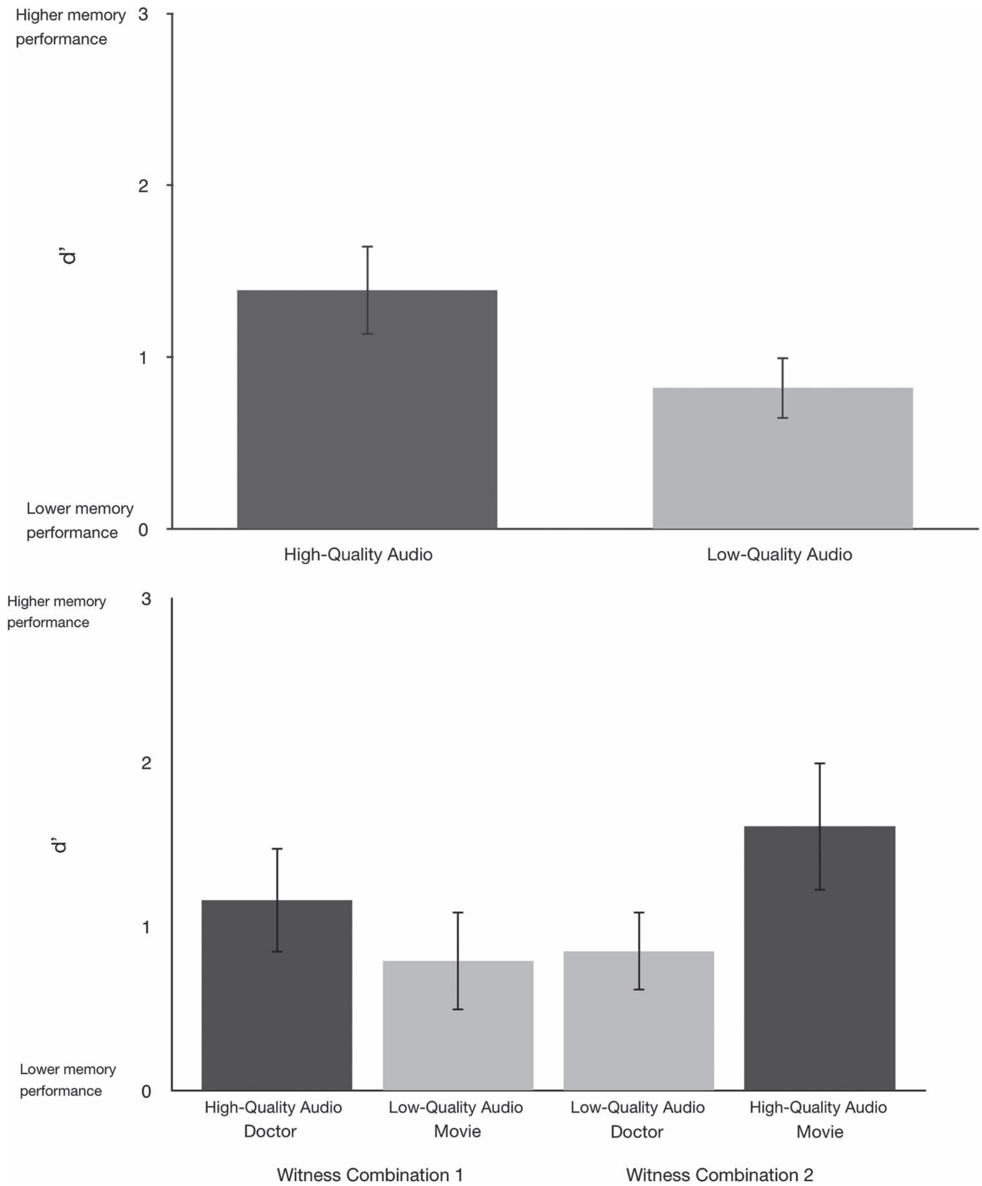
As Figure 2 shows, a paired-samples t test between the high- and low-quality audio conditions showed that participants had higher discrimination when they listened to testimony in high-quality audio, $t(127) = 3.81$, $p < .001$, $d = 0.44$, 95% CI [0.21, 0.68]. This pattern held when we included witness combination (whether the doctor description or movie description was in high-quality audio) in our analysis. A 2 (audio quality: high, low) × 2 (witness combination: doctor high/movie low, doctor low/movie high) repeated-measures mixed ANOVA with witness-combination condition as a between-subjects variable showed the expected main effect for audio quality, $F(1, 126) = 14.29$, $p < .001$, $\eta_p^2 = .10$, 90% CI [.03, .19], and no main effect of witness combination, $F(1, 126) = 2.32$, $p = .130$, $\eta_p^2 = .02$, 90% CI [.00, .07]. Although it is tempting to conclude that there was an interaction with witness combination, this did not reach significance, $F(1, 126) = 1.66$, $p = .200$, $\eta_p^2 = .01$, 90% CI [.00, .06], indicating that the effect of audio quality held regardless of which witnessed event appeared in high or low audio quality.

Closer inspection revealed that high-quality audio led to higher hits ($M = 0.72$, $SD = 0.24$) than low-quality audio ($M = 0.66$, $SD = 0.20$), $t(127) = 2.75$, $p = .010$, $d = 0.27$, 95% CI [0.08, 0.47], and low-quality audio led to more false alarms ($M = 0.38$, $SD = 0.19$) than high-quality audio ($M = 0.32$, $SD = 0.19$), $t(127) = 2.87$, $p = .005$, $d = 0.31$, 95% CI [0.11, 0.52]. In short, poor audio quality impaired memory performance on both measures, reducing the correct identification of old items and increasing the erroneous acceptance of new items as previously seen. Participants' bias to respond new or old (as captured by c values) was not significantly different when the audio quality was low ($M = -0.06$, $SD = 0.64$) and when the audio quality was high ($M = -0.07$, $SD = 0.72$), $t(127) = 0.11$, $p = .911$.

In addition, the results of Experiment 2 show that participants had better memory for testimony when it was presented in high-quality audio. This is consistent with the robust finding that disfluency impairs the perceived trustworthiness of messages, as reviewed in Schwarz et al. (2021) and observed in Experiment 1, as well as Newman and Schwarz (2018), for manipulations of audio quality. Hence, participants may have attended less to witnesses who seemed to be less credible, reliable, and accurate, resulting in reduced memory. A negative impact of disfluent processing on memory performance has been observed in several studies (Eitel et al., 2014; Kühl et al., 2014; Rummer et al., 2016), whereas others have observed improved memory under conditions of disfluency (Besken & Mulligan, 2014; Diemand-Yauman et al., 2011; French et al., 2013; Sungkhasettee et al., 2011). At present, the

Figure 2

Mean Discrimination for Testimony Across High- and Low-Audio-Quality Conditions



Note. The top panel is collapsed across witnessed event; the bottom panel presents means separately for the doctor- and movie-related witness statements. Error bars represent 95% confidence intervals.

conditions that determine these diverging effects of fluency on memory are poorly understood (for discussions, see Oppenheimer & Alter, 2014; Schwarz, 2015). We suggest that in the absence of explicit motivation to attend to a message (e.g., as in educational settings), people are less likely to attend to messages of questionable credibility, which results in poorer memory when disfluency hurts message perception (as in the present studies). Compatible with this conjecture, we found a systematic effect across counterbalances in Experiment 2. When we reran the key audio comparison as a repeated-measures analysis including audio quality and counterbalance as factors, we detected an interaction effect of counterbalance in Experiment 2, $F(3, 124) = 3.66$, $p = .014$, $\eta_p^2 = .08$, 90% CI [.01, .15] (not observed in Experiment 1, $F(3, 153) = 2.11$, $p = .101$, $\eta_p^2 = .04$, 90% CI [.00, .09]). Disfluency (poor audio) had the most robust negative effect on memory performance in the first task, when people did not know that a memory test would follow; it was less influential in the second task, when people knew that a memory test would be administered. Specifically, memory (as measured using d') performance across counterbalances showed that conditions with disfluent testimony first, when participants had no knowledge of an upcoming memory test, had a raw mean difference of 0.90, 95% CI [0.59, 1.21]; in contrast, conditions with disfluent testimony second, when participants could anticipate an upcoming memory test, had a raw mean difference of 0.12, 95% CI [−0.41, 0.17].

Experiment 3

In Experiments 1 and 2, people evaluated short witness descriptions with little contextual trial information. In Experiment 3, we aimed to replicate the results of Experiment 1 but incorporated an initial trial summary so we could capture people's impressions about a case before introducing audio testimony. Extending beyond Experiments 1 and 2, this adapted paradigm allowed us to examine the degree to which people updated their impressions about a case and, in particular, how variation in audio quality influenced the way participants weighted witness testimony in making assessments of guilt.

To date, there is limited evidence regarding the extent to which processing fluency guides evidence weighting in decision making. One series of studies that looked directly at this question considered how people used fluency as a cue to resolve contrasting information about a judgment target in a marketing context (Shah & Oppenheimer, 2007). The key hypothesis was that fluency may influence which information seems right: Fluent information is more likely to be judged as true and should hence be more heavily weighted in people's judgments. Across those studies, Shah and Oppenheimer found that fluency affected the extent to which people relied on information in decision making, with people making decisions in line with the fluent cues. For example, when participants were exposed to conflicting cues, one fluent and one disfluent, they used information from the fluent cue more when making final judgments (Shah & Oppenheimer, 2007, Experiment 2). That is, fluency indirectly affected judgments and operated as a kind of domain-general basis for cue weighting (see Schwarz, 2015, for a similar discussion on truth assessment).

In the courtroom, use of one of two conflicting accounts could shape belief in a particular story or, inevitably, assessments of guilt. Thus, in Experiment 3, we tested not only whether the audio effect on witness impressions occurred against a backdrop of a more contextually rich trial description but also whether audio quality influenced the extent to which testimony was weighted in assessments of guilt.

Method

Participants and Design

Two hundred five Prolific workers fully completed the study (103 female; age range = 18–68, $M = 31.46$, $SD = 12.45$). As in Experiments 1 and 2, we manipulated audio quality within subjects. The key dependent variables were witness impressions (reliability, trust, and accuracy) and ratings of guilt before and after hearing the audio testimony. We preregistered our sample size as 200, assuming a small-to-medium effect size. We also conducted a sensitivity analysis; setting an α level of .05 and a sample size of 205, we had 80% power to detect an effect size (d) of 0.20. With a final sample of 118, following later manipulation checks, and α level of .05, we had 80% power to detect an effect size (d) of 0.26. We used Prolific as an online data collection platform for Experiment 3 given the specific metric measurements we used in the trial description, allowing restriction to a U.K. sample. It also allowed a replication outside of the Amazon Mechanical Turk platform, with different materials. (Prolific has been used as a core online data collection platform in experimental psychological studies and has high data quality; see, e.g., Peer et al., 2017.)

Materials and Procedure

Initial Trial Summary and Initial Guilt Ratings. Participants completed the study via Qualtrics, in which they were asked to act as though they were jurors assessing a case. After being welcomed to the study, participants first read a trial transcript. We used a transcript depicting a case in which a child was hit by a car outside a school at pickup time by one of the other parents at the school. The transcript contained 2,147 words and should have taken about 5 min to read. After reading the transcript, participants were asked to assess the extent to which the driver was guilty of negligent driving and/or dangerous driving. *Negligent driving* was defined as not paying proper attention to their driving, and *dangerous driving* was defined as driving that was dangerous or under the influence of alcohol or drugs. We included two different charges in line with plausible charges in this particular case.

Immediately after reading the trial summary, participants provided initial ratings on how guilty they found the defendant in response to the charges of negligent driving and dangerous driving, using two separate slider scales from 0 to 100, with higher numbers indicating higher guilt.

Eyewitness Statements and Subsequent Guilt Ratings. The key contention in the case was how far the defendant had been from the child when the child ran out to the road. Thus, after reading the initial case description, participants heard two witnesses (one from the defense and one from the prosecution) who spoke to this issue of distance in brief audio clips.

Two female undergraduate students were recorded reading each witness account, thus allowing us to counterbalance for witness voice. The recordings were edited using iMovie (Version 10.1.14) to create high- and low-quality audio versions. The high-quality version was unedited, and the low-quality version was created using the "large room" filter, which increases the echo and decreases the clarity of the speaker. As in Experiment 1, the recordings were pilot tested on a small sample ($N = 19$) naive to the hypothesis, and volunteers were able to accurately transcribe the audio, showing that the crucial content (distance from the child) was not lost with depleted audio quality. More importantly, we tested this with our study participants. Across conditions, participants heard either (a) the prosecution witness in high-quality audio and the defense witness in low-quality audio or (b) the defense witness in high-quality audio and the prosecution witness in low-quality audio. The order of the witness statements was counterbalanced. Participants were randomly assigned to conditions via Qualtrics.

After listening to each witness statement, participants were asked to evaluate three questions: "How accurate do you think the witness was?" "How reliable did you feel the witness was?" and "How much do you trust the witness?" (e.g., 1 = *not very accurate at all*, 5 = *very accurate*). Although we have combined these ratings into a general impression of the witness for high-quality and low-quality audio (both Cronbach's αs > .87), we found the same significant pattern of results when we analyzed accuracy, reliability, and trust separately. To check that participants had encoded the crucial information from each witness, we asked them to report the distance the witness said the defendant's car was from the child. This process was repeated with the second witness statement. Finally, participants were asked to make final guilt ratings on both charges using the scales described above.

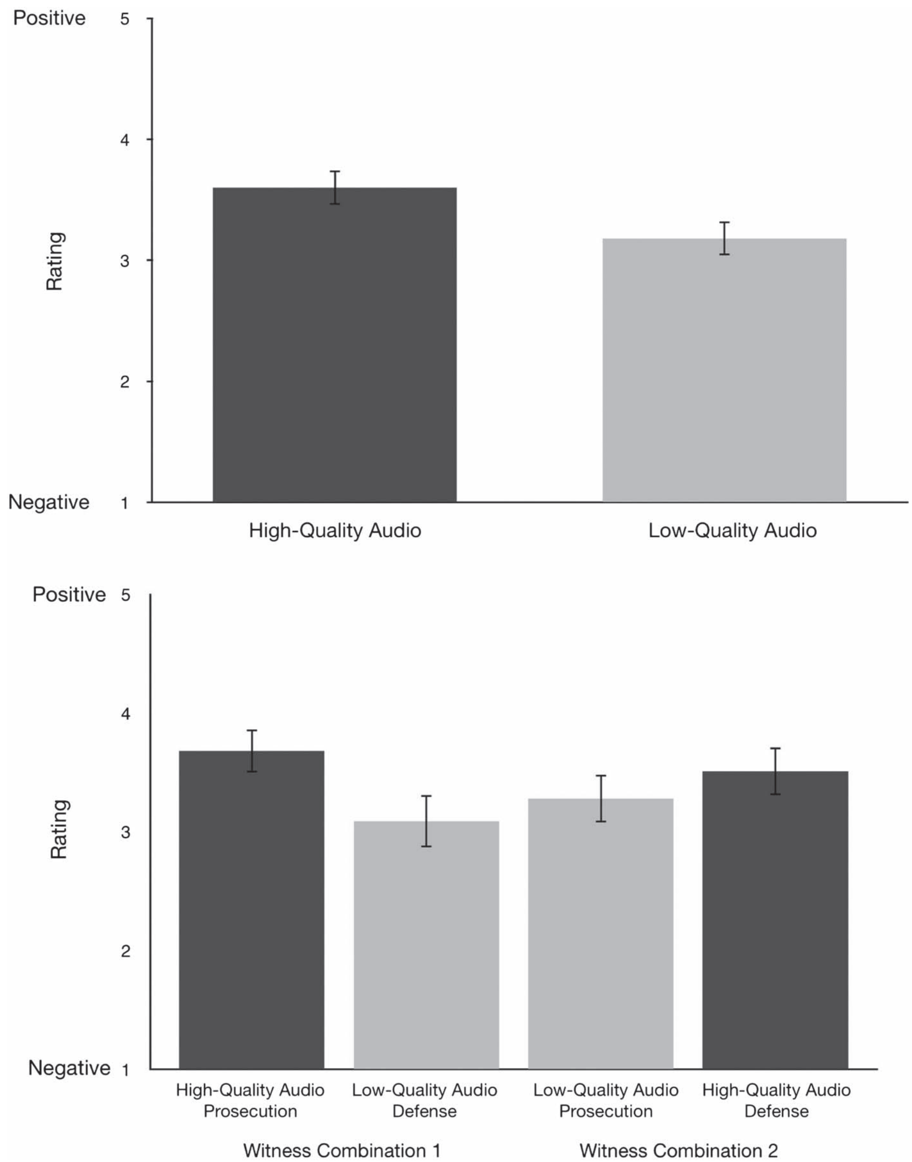
Results and Discussion

As in Experiments 1 and 2, audio quality influenced evaluations of the testimony. When the audio quality was high, participants evaluated the witness more favorably (top panel of Figure 3). At an item level, the effect of audio quality was consistent for both the prosecution and defense witness statement combinations (bottom panel of Figure 3). Further, audio quality influenced evidence weighting, with participants shifting their guilt ratings toward the witness statement that was processed fluently (Figure 4). In the following analyses, we included only those participants who were able to accurately state the distance reported in each of the witness accounts (n = 118)—an exclusion that is particularly important for the shift in guilt analysis, where we had to ensure that participants had encoded the critical information from each witness account.

Witness Ratings

As above, a paired-samples t test of the overall ratings of the high- and low-audio-quality witnesses showed that participants rated the

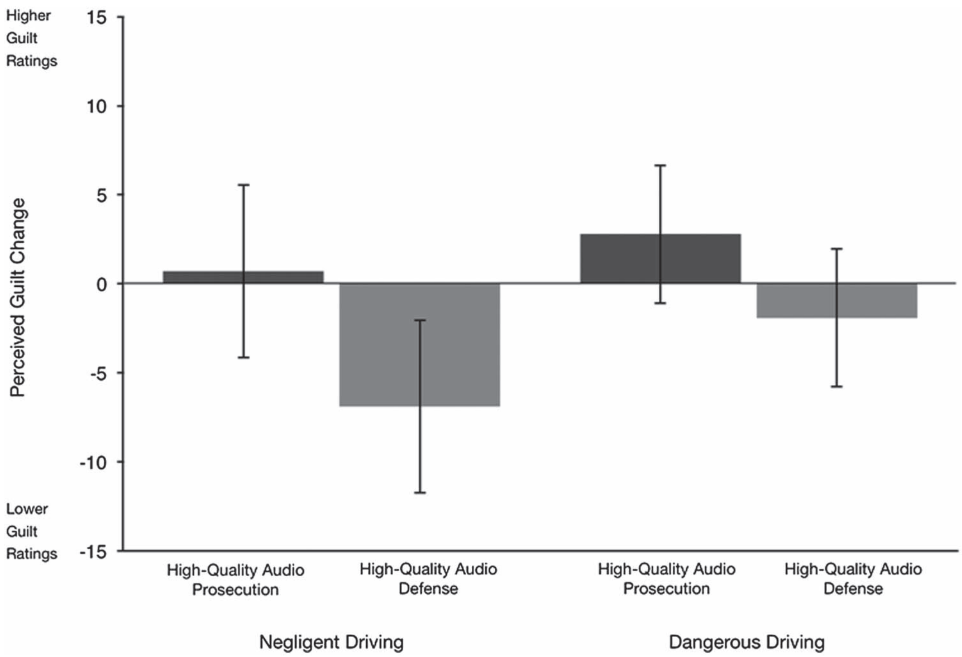
Figure 3
Mean Ratings of Witnesses Across the High- and Low-Audio-Quality Conditions



Note. The top panel is collapsed across the witness event; the bottom panel presents means separately for the prosecution and defense witness statements. Error bars represent 95% confidence intervals.

Figure 4

Mean Change From Initial to Final Guilt Ratings for Between-Subjects Conditions of Witness Statements Presented in High Audio Quality (Prosecution High/Defense Low; Prosecution Low/Defense High)



Note. Higher scores indicate a shift toward higher final guilt estimates, and lower scores indicate a shift toward lower final guilt estimates. The left panel shows negligent-driving guilt ratings, and the right panel shows dangerous-driving guilt ratings. Error bars represent 95% confidence intervals.

high-audio-quality witness (M = 3.60, SD = 0.72) higher than the low-audio-quality witness (M = 3.18, SD = 0.78), $t(117) = 4.54$, $p < .001$, d = 0.55, 95% CI [0.31, 0.81]. Notably, we found the same significant pattern with the full sample included.

This pattern held when we included witness combination (i.e., whether people heard the prosecution witness's description or the defense witness's description in high-quality audio) in our analysis. A 2 (audio quality: high, low) × 2 (witness combination: prosecution high/defense low, prosecution low/defense high) repeated-measures mixed ANOVA with witness combination as the between-subjects variable showed the expected main effect for audio quality (as shown in the paired t test above), $F(1, 116) = 20.51$, $p < .001$, $\eta_p^2 = .15$, 95% CI [.06, .25]. The interaction for witness combination did not meet the threshold of significance, $F(1, 116) = 3.83$, $p = .053$, $\eta_p^2 = .03$, 95% CI [.00, .12], suggesting that the effect of audio quality held regardless of which witness description appeared in high- or low-quality audio (see Figure 3). There was also no main effect for the witness-combination condition, $F(1, 116) = 0.01$, $p = .926$, $\eta_p^2 = .00$, 90% CI [.00, .01].

Guilt Ratings

After reading the initial case, participants rated the defendant's guilt on the negligent-driving charge as very close to midpoint of the 0–100 slider scale, M = 53.8, 95% CI [49.07, 58.52], but participants were less willing to conclude that the driver was guilty of dangerous driving, M = 36.34, 95% CI [31.38, 41.31].

To address the question of whether fluent evidence was weighted more heavily in final ratings of guilt, we measured the change in guilt ratings. More specifically, we measured the change of initial guilt ratings (after reading the trial summary) to final guilt ratings (after listening to both witness accounts). If fluent evidence was weighted more heavily, people's guilt ratings would update in line with the fluent evidence: Participants who heard the prosecution evidence in high-quality audio would shift toward higher estimates of guilt, and those who heard the defense evidence in high-quality audio would shift toward lower estimates of guilt. This is what we found, but as Figure 4 shows, the shift largely occurred in the defense condition. Given that each participant encountered exactly the same content and reported it correctly for each witness, this shift cannot be attributed to the influence of content; instead, it suggests that people took their metacognitive experience of ease or difficulty into account when weighting the evidence.

As Figure 4 shows, participants updated their guilt ratings in line with the witness who had high-quality audio. Hearing from the defense in fluent audio led participants to reduce their guilt assessments, relative to listening to the prosecution in fluent audio. In a 2 (witness combination: prosecution high/defense low, prosecution low/defense high) × 2 (charge: negligence, dangerous) repeated-measures ANOVA, a main effect emerged for witness combination—participants reduced their impressions of guilt when they heard the defense in high audio quality, relative to participants who heard the prosecution in high audio quality, $F(1, 116) = 5.85$, $p = .017$, $\eta_p^2 = .05$, 90% CI [.00, .12] (Figure 4)—and a main effect of charge, such that participants tended to shift to lower ratings of guilt for the negligent-driving charge relative to the dangerous-driving charge, $F(1, 116) = 4.32$, $p = .040$, $\eta_p^2 = .04$, 90% CI [.00, .12]. There was no interaction between charge and witness audio quality, $F < 1.00$, suggesting the same pattern of results for both charges, though see the separate analysis in the Supplemental Materials suggesting more robust effects for negligence than dangerous driving; these analyses compared both initial and final guilt scores for both guilt ratings separately.

Witness Ratings and Final Guilt Ratings

Next, we examined final guilt ratings and the extent to which the influence of audio quality on guilt ratings was explained by participants' impressions of the prosecution and defense witnesses. We conducted a mediation analysis for both the defense witnesses and the prosecution witnesses separately and expected that impressions of the witnesses would partially, if not fully, mediate the relationship between audio condition and guilt. That is, when witness impressions were added as a mediator, we expected that the direct relationship between audio quality and guilt ratings would be significantly reduced.

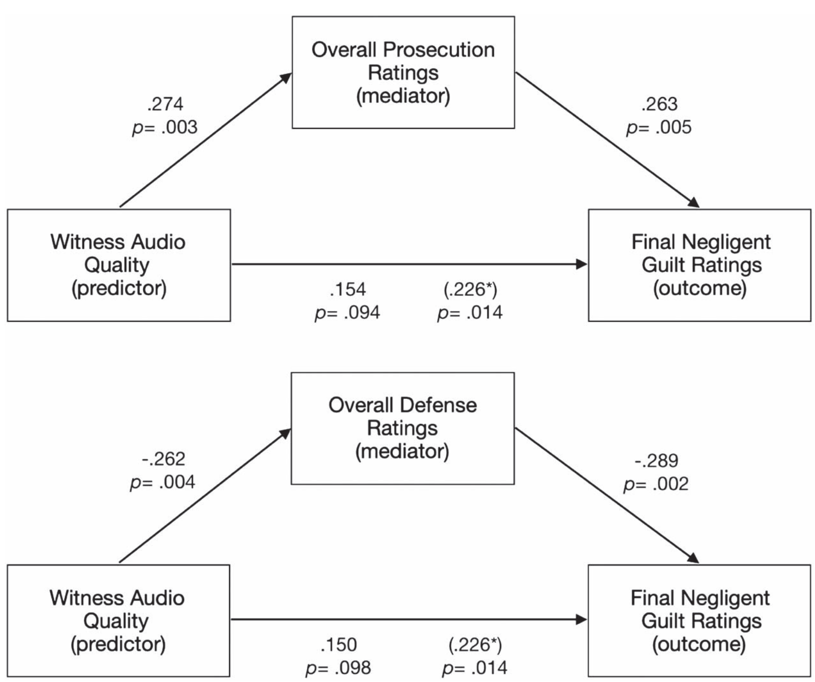
We used a hierarchical regression for the negligent-driving charge (there was no direct relationship with audio condition for dangerous driving; thus, we ran the analysis only for negligent driving here). The outcome variables for both prosecution and defense negligent-driving analyses were final guilt ratings. The predictor variable for both analyses was witness audio quality (whether the prosecution or defense witness testimony was presented in high- or low-quality audio). Witness ratings (collapsed across accuracy, reliability, and trust) of the prosecution and defense witnesses were entered as mediators into the two models. As Figure 5 shows, for both prosecution and defense analyses, witness ratings fully mediated the relationship between fluency condition and final guilt ratings for the negligent-driving charge.

Prosecution. Our analyses showed that indeed, as reported above, participants who heard the prosecution witness in high-quality audio rated the prosecution witness more favorably, whereas participants who heard the prosecution witness in low-quality audio rated the prosecution witness less favorably. The mediation analysis showed that the effect of audio quality on guilt was fully mediated by impressions of the witnesses: Participants who rated the prosecution witness more favorably gave higher final negligent-driving guilt ratings than participants who rated the prosecution witness less favorably.

As shown in the top panel of Figure 5, the regression predicting final negligent-driving guilt ratings from fluency condition was significant, $\beta = 0.226$, $p = .014$. The regression predicting overall prosecution ratings from audio condition was also significant, $\beta = 0.274$, $p = .003$. Overall prosecution ratings, controlling for audio, significantly predicted final negligent-driving guilt ratings, $\beta = 0.263$, $p = .005$, whereas the relationship between witness audio quality condition and final negligent-driving guilt ratings became nonsignificant once overall prosecution ratings were entered into the model, $\beta = 0.154$, $p = .094$. A Sobel test was conducted and confirmed full mediation in the model ($z = 2.04$, $p = .041$).

Figure 5

Mediation Analyses for Final Negligent-Driving Guilt Ratings, by Witness Audio Quality and With Overall Prosecution and Defense Witness Ratings as Mediators



Note. Standardized β coefficients are reported. Witness audio quality was coded as 1 = prosecution high-quality/defense low-quality; −1 = prosecution low-quality/defense high-quality.

Defense. Replicating results presented above, participants who heard the defense witness in low-quality audio rated the defense witness less favorably, whereas participants who heard the defense witness in high-quality audio rated the defense witness more favorably. Our results from the mediation analysis showed that the effect of audio quality on guilt was fully mediated by impressions of the witnesses: The more favorably participants rated the defense witness, the lower their final guilt rating for negligent driving.

As shown in the bottom panel of Figure 5, the regression predicting final negligent-driving guilt ratings from fluency condition was significant, $\beta = 0.226$, $p = .014$. The regression predicting overall defense ratings from witness audio condition was also significant, $\beta = -0.262$, $p = .004$. Overall defense ratings, controlling for audio, significantly predicted final negligent-driving guilt ratings, $\beta = -0.289$, $p = .002$, whereas the relationship between witness audio condition and final negligent-driving guilt ratings became nonsignificant once overall defense ratings were entered into the model, $\beta = 0.150$, $p = .098$. A Sobel test was conducted and confirmed full mediation in the model ($z = 2.10$, $p = .035$).

In addition, audio quality influences perceptions of testimony and witnesses even when there is rich background context, such as a trial summary. Further, we found that participants weighted fluently processed evidence more heavily when evaluating guilt, adjusting their assessments in line with fluent testimony. Our findings also suggest that the guilt ratings that supported the witness whose testimony was presented in high-quality audio were driven by higher ratings of accuracy, reliability, and trust.

General Discussion

While courts adapt to incorporate new forms of technology and obvious errors are addressed by procedural interventions, more insidious technical glitches may hurt the integrity of a trial by systematically biasing people's impressions of witnesses and how they use their evidence. Across three experiments, we found that low-quality audio systematically led to less favorable evaluations of witnesses, poorer memory for factual evidence, and reduced weighting of evidence in decision making. These findings highlight that audio quality is a technological feature of evidence that warrants procedural consideration.

Our findings are consistent with the literature on cognitive fluency and extend that literature to decision making in the courtroom. Although many studies have examined how the content of evidence or attributes of a witness can influence people's judgments, less empirical research has focused on how the phenomenological experience of processing information can alter people's decisions (see Newman et al., 2019). In our studies, participants listened to the same witness making the same statement—and arrived at different impressions of the witness and different guilt ratings, depending on whether the statement was played to them with high or low audio quality. The audio quality also influenced their memory for what the witness said and their weighting of the evidence. These effects were observed even though the speaker had no control over the quality of the audio delivery and reflect that people are sensitive to their metacognitive experience of ease or difficulty when processing information, but insensitive to the variables that cause the difficulty (for reviews, see Schwarz, 2015; Schwarz et al., 2021). People typically misread processing difficulty as a sign that something is wrong with the content they are processing, which results in less favorable impressions of the trustworthiness of the communicator (e.g., Newman et al., 2014) and the credibility and truth of the message (e.g., Reber & Schwarz, 1999).

As seen in Experiment 3, the observed influence of audio quality was robust, even when participants had other relevant and diagnostic information to use in their decisions. Indeed, people's initial assessments of guilt, provided after reading case information and before hearing witnesses, were important in shaping their final judgments, as observed in earlier work (Scurich & John, 2018). Nevertheless, the ease of processing the witness audio influenced their final ratings of guilt (see the Supplemental Analyses). These findings square with research on cognitive fluency and repetition showing that prior beliefs or declarative content and an experience of processing fluency can contribute to decisions (e.g., Fazio et al., 2015; Unkelbach & Greifeneder, 2018). Following this line of logic, in a given individual case, the extent to which audio variations influence case outcomes may operate in concert with more declarative variables, such as the strength of the evidence and whether people are close to the threshold for switching verdicts. In such scenarios, small experiential changes in the fluency of processing evidence may be particularly consequential.

The influence of processing difficulty is usually most pronounced when its source is subtle and decreases when people become aware that processing may be difficult solely because of some incidental influence, which usually requires that the source of disfluency is explicitly brought to their attention (Schwarz et al., 2021). Future research may examine whether bringing the audio variation to people's attention moderates the size of these effects. Further, one may also consider the extent to which bringing audio quality to people's attention requires explicit cues—in the moment—about technological disruption to effectively reduce the effect of audio quality on people's impressions about witnesses and their evidence. These are empirical questions worthy of future research.

In the fluency literature more generally, there has been a debate regarding whether difficulty in processing can aid or impede learning and memory. In Experiment 2, we tested whether disfluency increased memory for information, as has been observed in educational settings (Alter et al., 2007), or reduced memory for evidence, perhaps because people tuned out, experienced strained cognitive resources, or did not trust the witness (e.g., Experiment 1; Eitel et al., 2014; Kühl et al., 2014). Our results supported the latter notion, with participants performing worse on a memory test after hearing the child testify in low-quality audio. The cognitive or social-cognitive mechanisms producing this effect may be interesting avenues for future research, but the decreased number of correct answers in the disfluent condition suggests that audio disfluency in the courtroom may produce cognitive consequences that reach beyond the evaluation of a given witness. Notably, the effects we found for this difference in memory performance were most robust when people did not know that a memory test would follow disfluent audio, suggesting that motivation to remember may moderate this effect. This is an interesting question to pursue in future research.

Our findings on evidence weighting extended prior work by Shah and Oppenheimer (2007): People relied more heavily on fluently processed information in decision making, with impressions of guilt being influenced by audio quality. This pattern fits with previous findings; however, our paradigm extended this previous research in two ways. First, because we gave participants an initial trial summary, we could measure how decisions were revised with new evidence. Second, the mediation analyses provided insights into why participants shifted their judgments in the direction of the fluent information. Our mediation analyses suggested that participants rated the prosecution and defense witnesses higher or lower on assessments of accuracy, reliability, and trust depending on who had high- or low-quality audio recordings and shifted their ratings of guilt in the direction of the side they rated more favorably, suggesting that fluency can operate via social cues, interacting with final evidence evaluations.

Of course, there are several possible sources of audio disfluency in the courtroom and in virtual court contexts. Wilson and Sasse (2000) identified at least five major sources of audio distortions other than echoes that may influence the experience of the listener, with excess loudness, bad microphone, and significant dropout generally having the greatest effect on both subjective assessments (e.g., ratings of comfort and fatigue) and physiological responses compared to only minor dropout. Although existing research has shown that audio disruptions influence perceptions of others even when presented with video content (Newman & Schwarz, 2018), this study did not examine the role of visual displays. Whereas visual cues can enhance conceptual processing in legal contexts and influence judgments (e.g., Derksen et al., 2020; Sanson et al., 2020), visual discrepancies can be a source of cognitive disfluency. For instance, research subjects find it harder to detect emotions when two visual distortions are introduced—reductions in the image refresh rate and resolution (Wallbott, 1992). These additional forms of distortion raise interesting questions for future research and are important for policy development going forward. Although more pronounced disruptions might lead trial participants to talk over each other, require repeated testimony, and cause general confusion, judges presiding over such trials will likely intervene (Lapinski et al., 2020). Empirically derived estimations of how more nuanced technological features of evidence, such as low-quality audio, may affect decision making in the court will help to inform conditions under which procedural interventions are necessary.

Trials in the common law tradition largely rely on oral testimony, as used in this study, from defendants or eyewitnesses, so it is important that decision makers can focus on the content of the information without having to adjust for audio distortion or deal with the extra stress involved in listening to poor quality sound (see Wilson & Sasse, 2000). The issue is relevant not just to juries but to judges who may be hearing applications for bail or passing sentence. It may also affect the ability of witnesses to tell their story in a considered way: If the voice of the lawyer asking them questions is distorted, the witness could feel less confident they know what they are being asked and appear less credible in their response.

The implications of this research also stretch to the sociocultural domain: It is possible that the testimony that is affected by poor internet connection may be from more disadvantaged trial participants who have less access to up-to-date technology. There is evidence that some trial participants have been joining from remote locations and using Zoom on their phones (Raczynski, 2020). Although a possible solution is to ensure that all participants have the same technology, such as court appointed devices, installing new Wi-Fi routers to boost their internet connection is an expensive step. This is relevant to remote witnesses, who will continue to take part in court hearings from home or other remote sites now that the pandemic has highlighted remote hearing/virtual court potential (including cost savings). It is unlikely that remote juries will be used in the long term, apart from in a very limited number of civil jury matters. However, audio problems can also arise from the physical attributes and arrangements of regular courtrooms, particularly older ones; those problems may be more salient for older jurors and, sometimes, judges. So, the issues this article raises are potentially relevant for in-person hearings in physical courtrooms. Echoes on calls from remote witnesses may, this study suggests, make the witnesses less credible, but audio distortions from other participants within the courtroom may also have an impact on how the speaker is perceived.

Long pauses and disconnections may produce significant procedural interruptions in trial proceedings, but the experiments reported here suggest that low-quality audio may warrant interventions similar to those often used in cases of complete dropout or long delays. Procedures in the courtroom are carefully attuned to content, but these experiments suggest that experienced ease of processing information is a subtle feature of evidence that can bias juror decision making.

References

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, *13*(3), 219–235. https://doi.org/10.1177/1088868309341564

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General*, *136*(4), 569–576. https://doi.org/10.1037/0096-3445.136.4.569

Barrie, D. (2020, August 6). Coronavirus in Scotland: Cinemas could be the key for a virtual jury. *The Times*. https://www.thetimes.co.uk/article/coronavirus-in-scotland-cinemas-could-be-the-key-for-a-virtual-jury-dznmfcth0

Besken, M., & Mulligan, N. W. (2014). Perceptual fluency, auditory generation, and metamemory: Analyzing the perceptual fluency hypothesis in the auditory modality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 429–440. https://doi.org/10.1037/a0034407

Boys, T., & Sams, A. (2020, October 14). *COVID-19 and the virtual courtroom—Is technology a friend or foe?* Holding Redlich. https://www.holdingredlich.com/covid-19-and-the-virtual-courtroom-is-technology-a-friend-or-foe

Brown, A. S., Brown, L. A., & Zoccoli, S. L. (2002). Repetition-based credibility enhancement of unfamiliar faces. *The American Journal of Psychology*, *115*(2), 199–209. https://doi.org/10.2307/1423435

Byrom, N., Beardon, S., & Kendrick, A. (2020). *Rapid review: The impact of COVID-19 on the civil justice system*. Civil Justice Council and the Legal Education Foundation. https://www.judiciary.uk/wp-content/uploads/2020/06/FINAL-REPORT-CJC-4-June-2020.v2-accessible.pdf

Derksen, D. G., Giroux, M. E., Connolly, D. A., Newman, E. J., & Bernstein, D. M. (2020). Truthiness and law: Nonprobative photos bias perceived credibility in forensic contexts. *Applied Cognitive Psychology*, *34*(6), 1335–1344. https://doi.org/10.1002/acp.3709

Diemand-Yauman, C., Oppenheimer, D. M., & Vaughan, E. B. (2011). Fortune favors the bold (and the Italicized): Effects of disfluency on educational outcomes. *Cognition*, *118*(1), 111–115. https://doi.org/10.1016/j.cognition.2010.09.012

Eitel, A., Kühl, T., Scheiter, K., & Gerjets, P. (2014). Disfluency meets cognitive load in multimedia learning: Does harder-to-read mean better-to-understand? *Applied Cognitive Psychology*, *28*(4), 488–501. https://doi.org/10.1002/acp.3004

Espinoza, R. K. E., & Willis-Esqueda, C. (2008). Defendant and defense attorney characteristics and their effects on juror decision making and prejudice against Mexican Americans. *Cultural Diversity & Ethnic Minority Psychology*, *14*(4), 364–371. https://doi.org/10.1037/a0012767

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. https://doi.org/10.1037/xge0000098

Federal Circuit Court of Australia. (2020, April 22). *Practitioner and litigant guide to virtual hearings and Microsoft Teams*. https://qldbar.asn.au/baq/v1/viewDocument?documentId=1210

Fiechter, J. L., Fealing, C., Gerrard, R., & Kornell, N. (2018). Audiovisual quality impacts assessments of job candidates in video interviews: Evidence for an AV quality bias. *Cognitive Research: Principles and Implications*, *3*(1), Article 47. https://doi.org/10.1186/s41235-018-0139-y

Fobes, C. (2020). Rule 43(A): Remote witness testimony and a judiciary resistant to change. *Lewis & Clark Law Review*, *24*(1), 299–324. https://law.lclark.edu/live/files/29604-lcb241article7fobespdf

Fontaine, G., & Kiger, R. (1978). The effects of defendant dress and supervision on judgments of simulated jurors: An exploratory study. *Law and Human Behavior*, *2*(1), 63–71. https://doi.org/10.1007/BF01047503

French, M. M. J., Blood, A., Bright, N. D., Futak, D., Grohmann, M. J., Hasthorpe, A., Heritage, J., Poland, R. L., Reece, S., & Tabor, J. (2013). Changing fonts in education: How the benefits vary with ability and dyslexia. *The Journal of Educational Research*, *106*(4), 301–304. https://doi.org/10.1080/00220671.2012.736430

Frumkin, L. (2007). Influences of accent and ethnic background on perceptions of eyewitness testimony. *Psychology, Crime & Law*, *13*(3), 317–331. https://doi.org/10.1080/10683160600822246

Gourdet, C., Witwer, A. R., Langton, L., Banks, D., Planty, M. G., Woods, D., & Jackson, B. A. (2020). *Court appearances in criminal proceedings through telepresence*. RAND Corporation. https://doi.org/10.7249/RR3222

Koehler, J. J., Schweitzer, N. J., Saks, M. J., & McQuiston, D. E. (2016). Science, technology, or the expert witness: What influences jurors' judgments about forensic science testimony? *Psychology, Public Policy, and Law*, *22*(4), 401–413. https://doi.org/10.1037/law0000103

Kühl, T., Eitel, A., Damnik, G., & Körndle, H. (2014). The impact of disfluency, pacing, and students' need for cognition on learning with multimedia. *Computers in Human Behavior*, *35*(4), 189–198. https://doi.org/10.1016/j.chb.2014.03.004

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*(3), 752–756. https://doi.org/10.1016/j.jesp.2011.12.002

Lapinski, J., Hirschhorn, R., & Blue, L. (2020, September 29). Zoom jury trials: The idea vastly exceeds the technology. *Texas Lawyer*. https://www.law.com/texaslawyer/2020/09/29/zoom-jury-trials-the-idea-vastly-exceeds-the-technology/

Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, *46*(6), 1093–1096. https://doi.org/10.1016/j.jesp.2010.05.025

Liu, Y., Liu, R., Star, J. R., Wang, J., & Tong, H. (2020). The effect of perceptual fluency on overcoming the interference of the More A-More B intuitive rule among primary school students in a perimeter comparison task: The perspective of cognitive load. *European Journal of Psychology of Education*, *35*(4), 357–380. https://doi.org/10.1007/s10212-019-00424-w

Martire, K. A., Edmond, G., & Navarro, D. (2020). Exploring juror evaluations of expert opinions using the expert persuasion expectancy framework. *Legal and Criminological Psychology*, *25*(2), 90–110. https://doi.org/10.1111/lcrp.12165

McIntyre, J., Olijnyk, A., & Pender, K. (2020). Civil courts and COVID-19: Challenges and opportunities in Australia. *Alternative Law Journal*, *45*(3), 195–201. https://doi.org/10.1177/1037969X20956787

McKimmie, B. M., Newton, S. A., Schuller, R. A., & Terry, D. J. (2013). It's not what she says, it's how she says it: The influence of language complexity and cognitive load on the persuasiveness of expert testimony. *Psychiatry, Psychology and Law*, *20*(4), 578–589. https://doi.org/10.1080/13218719.2012.727068

Newman, E. J., Jalbert, M., & Feigenson, N. (2019). Cognitive fluency in the courtroom. In R. Bull & I. Blandón-Gitlin (Eds.), *The Routledge international handbook of legal and investigative psychology* (pp. 102–115). Routledge. https://doi.org/10.4324/9780429326530-8

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), Article e88671. https://doi.org/10.1371/journal.pone.0088671

Newman, E. J., & Schwarz, N. (2018). Good sound, good research: How audio quality influences perceptions of the research and the researcher. *Science Communication*, *40*(2), 246–257. https://doi.org/10.1177/1075547018759345

Norton Rose Fulbright. (2020). *COVID-19 and the global approach to further court proceedings, hearings*. https://www.nortonrosefulbright.com/en/knowledge/publications/bbfeb594/covid-19-and-the-global-approach-to-further-court-proceedings-hearings

Oppenheimer, D. M. (2008). The secret life of fluency. *Trends in Cognitive Sciences*, *12*(6), 237–241. https://doi.org/10.1016/j.tics.2008.02.014

Oppenheimer, D. M., & Alter, A. L. (2014). The search for moderators in disfluency research. *Applied Cognitive Psychology*, *28*(4), 502–504. https://doi.org/10.1002/acp.3023

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163. https://doi.org/10.1016/j.jesp.2017.01.006

Pieger, E., Mengelkamp, C., & Bannert, M. (2017). Fostering analytic metacognitive processes and reducing overconfidence by disfluency: The role of contrast effects. *Applied Cognitive Psychology*, *31*(3), 291–301. https://doi.org/10.1002/acp.3326

Raczynski, J. (2020, July 22). *The current status of the (virtual) courts*. Thomson Reuters. https://www.legalexecutiveinstitute.com/virtual-courts/

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, *8*(3), 338–342. https://doi.org/10.1006/ccog.1999.0386

Remote Courts Worldwide. (2020). *Remote courts worldwide country*. https://remotecourts.org/

Romero-Rivas, C., Morgan, C., & Collier, T. (2021). Accentism on trial: Categorization/stereotyping and implicit biases predict harsher sentences for foreign-accented defendants. *Journal of Language and Social Psychology*. Advance online publication. https://doi.org/10.1177/0261927X211022785

Rossner, M., Tait, D., & McCurdy, M. (2021). Justice reimagined: Challenges and opportunities with implementing virtual courts. *Current Issues in Criminal Justice*, *33*(1), 94–110. https://doi.org/10.1080/10345329.2020.1859968

Rossner, M., Tait, D., McKimmie, B., & Sarre, R. (2017). The dock on trial: Courtroom design and the presumption of innocence. *Journal of Law and Society*, *44*(3), 317–344. https://doi.org/10.1111/jols.12033

Rowden, E., & Wallace, A. (2019). Performing expertise: The design of audiovisual links and the construction of the remote expert witness in court. *Social & Legal Studies*, *28*(5), 698–718. https://doi.org/10.1177/0964663918802991

Rummer, R., Schweppe, J., & Schwede, A. (2016). Fortune is fickle: Null-effects of disfluency on learning outcomes. *Metacognition and Learning*, *11*(1), 57–70. https://doi.org/10.1007/s11409-015-9151-5

Ryan, M., Harker, L., & Rothera, S. (2020). *Remote hearings in the family justice system: A rapid consultation*. Nuffield Family Justice Observatory. https://www.nuffieldfjo.org.uk/resource/remote-hearings-rapid-consultation

Sanson, M., Crozier, W. E., & Strange, D. (2020). Court case context and fluency-promoting photos inflate the credibility of forensic science. *Zeitschrift für Psychologie mit Zeitschrift für Angewandte Psychologie*, *228*(3), 221–225. https://doi.org/10.1027/2151-2604/a000415

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). American Psychological Association. https://doi.org/10.1037/14341-006

Schwarz, N., Jalbert, M. C., Noah, T., & Zhang, L. (2021). Metacognitive experiences as information: Processing fluency in consumer judgment and decision making. *Consumer Psychology Review*, *4*(1), 4–25. https://doi.org/10.1002/arcp.1067

Scurich, N., & John, R. S. (2018). Jurors' presumption of innocence. *The Journal of Legal Studies*, *46*(1), 187–206. https://doi.org/10.1086/690450

Shah, A. K., & Oppenheimer, D. M. (2007). Easy does it: The role of fluency in cue weighting. *Judgment and Decision Making*, *2*(6), 371–379. https://psycnet.apa.org/record/2008-00191-005

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*, Article 2200. https://doi.org/10.3389/fpsyg.2017.02200

Song, H., & Schwarz, N. (2008). Fluency and the detection of misleading questions: Low processing fluency attenuates the Moses illusion. *Social Cognition*, *26*(6), 791–799. https://doi.org/10.1521/soco.2008.26.6.791

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149. https://doi.org/10.3758/BF03207704

Sungkhasettee, V. W., Friedman, M. C., & Castel, A. D. (2011). Memory and metamemory for inverted words: Illusions of competency and desirable difficulties. *Psychonomic Bulletin & Review*, *18*(5), Article 973. https://doi.org/10.3758/s13423-011-0114-9

Todd, R., Lim, W., Cheeseman, J., Strachan, J., Talas, T., & Kearney, M. (2020, April 20). *The remote courtroom: Tips and tricks for online hearings*. Ashurst. https://www.ashurst.com/en/news-and-insights/legal-updates/the-remote-courtroom-tips-and-tricks-for-online-hearings/

Totenberg, N. (2020, May 4). *Supreme Court arguments resume—But with a twist*. National Public Radio. https://www.npr.org/2020/05/04/847785015/supreme-court-arguments-resume-but-with-a-twist

Unkelbach, C., & Greifeneder, R. (2018). Experiential fluency and declarative advice jointly inform judgments of truth. *Journal of Experimental Social Psychology*, *79*, 78–86. https://doi.org/10.1016/j.jesp.2018.06.010

Wallbott, H. G. (1992). Effects of distortion of spatial and temporal resolution of video stimuli on emotion attributions. *Journal of Nonverbal Behavior*, *16*(1), 5–20. https://doi.org/10.1007/BF00986876

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. https://doi.org/10.1016/j.jcps.2008.12.009

Wilson, G. M., & Sasse, M. A. (2000). Investigating the impact of audio degradations on users: Subjective vs. objective assessment methods. In *Proceedings of the Annual Conference of CHISIG, OZCHI 2000: Interfacing reality in the new millennium* (pp. 135–142). The Computer Human Interaction Special Interest Group of the Ergonomics Society of Australia. https://core.ac.uk/download/pdf/1785691.pdf

Received February 11, 2021
Revision received August 6, 2021
Accepted September 7, 2021 ■