4

Remote Testimonial Fact-Finding

Renee L. Danser, D. James Greiner, Elizabeth Guo, and Erik Koltun

Prior to the COVID-19 pandemic, debate mounted concerning the wisdom, and perhaps the inevitability, of moving swaths of the court system online. These debates took place within a larger discussion of how and why courts had failed to perform their most basic function of delivering (approximately) legally accurate adjudicatory outputs in cases involving unrepresented litigants[1] while inspiring public confidence in them as a method of resolving disputes compared to extralegal and even violent alternatives. Richard Susskind, a longtime critic of the status quo, summarized objections to a nearly 100 percent in-person adjudicatory structure: "[T]he resolution of civil disputes invariably takes too long [and] costs too much, and the process is unintelligible to ordinary people."[2] As Susskind noted, fiscal concerns drove much of the interest in online courts during this time, while reformists sought to use the move online as an opportunity to expand the access of, and services available to, unrepresented litigants.

[1] *See, e.g.*, Ellen Lee Degnan, Thomas Ferriss, D. James Greiner & Roseanna Sommers, *Using Random Assignment to Measure Court Accessibility for Low-Income Divorce Seekers*, Proc. Nat. Acad. Scis., Apr. 6, 2021 (documenting enormous differences in the experiences of lawyerless versus lawyered litigants in a substantively simple kind of legal proceeding).

[2] Richard Susskind, Online Courts and the Future of Civil Justice 27 (2019); *see also* Ayelet Sela, *e-Nudging Justice: The Role of Digital Choice Architecture in Online Courts*, 2019 J. Disp. Resol. 127, 127–28 (2019) (noting that proponents of online courts contend that they lessen problems of confusing and complex process, lack of access, voluminous case filings, and costliness, while perhaps increasing settlement rates); Harold Hongju Koh, *The "Gants Principles" for Online Dispute Resolution: Realizing the Chief Justice's Vision for Courts in the Cloud*, 62 B.C. L. Rev. 2768, 2773 (2021) (noting gains in efficiency from online courts but suggesting a cost in the form of reduced access for those without resources and sophistication); Ayelet Sela, *Diversity by Design: Improving Access to Justice in Online Courts with Adaptive Court Interfaces*, 15 L. & Ethics Hum. Rts. 125, 128 (2021) (arguing that it is "widely recognized" that online courts promote efficiency, effectiveness, accessibility, and fairness).

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press

Yet as criticisms grew, and despite aggressive action by some courts abroad,[3] judicial systems in the United States moved incrementally. State courts, for example, began to implement online dispute resolution (ODR), a tool that private-sector online retailers had long used.[4] But courts had deployed ODR only for specified case types and only at specified stages of civil proceedings. There was little judicial appetite in the United States for a wholesale transition.

The COVID-19 pandemic upended the judicial desire to move slowly. As was true of so many other sectors of United States society, the judicial system had to reconstitute online, and had to do so convulsively. And court systems did so, in the fashion in which systemic change often occurs in the United States: haphazardly, unevenly, and with many hiccups. The most colorful among the hiccups entered the public consciousness: audible toilet flushes during oral argument before the highest court in the land; a lawyer appearing at a hearing with a feline video filter that he was unable to remove himself; jurors taking personal calls during voir dire.[5] But there were notable successes as well. The Texas state courts had reportedly held 1.1 million proceedings online as of late February 2021, and the Michigan state courts' live-streaming on YouTube attracted 60,000 subscribers.[6]

These developments suggest that online proceedings have become a part of the United States justice system for the foreseeable future. But skepticism has remained about online migration of some components of the judicial function, particularly around litigation's most climactic stage: the adversarial evidentiary hearing,

[3] *See* Robert Lapper, *Access to Justice in Canada: The First On-line Court*, Commonwealth Laws. Ass'n, https://www.commonwealthlawyers.com/cla/access-to-justice-in-canada-the-first-on-line-court/ (describing British Columbia's move to mandatory online adjudication for certain matters).

[4] *See* Ayelet Sela, *The Effect of Online Technologies on Dispute Resolution System Design: Antecedents, Current Trends, and Future Directions*, 21 Lewis & Clark L. Rev. 635 (2017); *see also* Ethan Katsh & Leah Wing, *Ten Years of Online Dispute Resolution (ODR): Looking at the Past and Constructing the Future*, 38 U. Tol. L. Rev. 19, 41 (2006).

[5] *See* Daniel Victor, *"I'm Not a Cat," Says Lawyer Having Zoom Difficulties*, N.Y. Times (Feb. 9, 2021), https://www.nytimes.com/2021/02/09/style/cat-lawyer-zoom.html (discussing a hearing in which a lawyer began participation with a cat video filter on his Zoom account that he was unable to remove without the judge's guidance); Fred Barbash, *Oyez. Oy vey. Was That a Toilet Flush in the Middle of a Supreme Court Live-Streamed Hearing?* Wash. Post (May 7, 2020), https://www.washingtonpost.com/nation/2020/05/07/toilet-flush-supreme-court/ (the article title speaks for itself); Ashley Feinberg, *Investigation: I Think I Know Which Justice Flushed*, Slate (May 8, 2020), https://slate.com/news-and-politics/2020/05/toilet-flush-supreme-court-livestream.html (discussing the same Supreme Court oral argument); Eric Scigliano, *Zoom Court Is Changing How Justice Is Served, for Better, for Worse, and Possibly Forever*, The Atlantic (Apr. 13, 2021), https://www.theatlantic.com/magazine/archive/2021/05/can-justice-be-served-on-zoom/618392/ (describing juror informality during remote voir dire).

[6] Scigliano, *Zoom Court Is Changing*; *see also* Pew Charitable Trs., How Courts Embraced Technology, Met the Pandemic Challenge, and Revolutionized Their Operations (2021), https://www.pewtrusts.org/-/media/assets/2021/12/how-courts-embraced-technology.pdf (noting that by Nov. 2020, 82 percent of all courts in the United States were permitting remote proceedings in eviction matters).

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press

especially the jury trial. Perhaps believing that they possessed unusual skill in discerning truthful and accurate testimony from its opposite,[7] some judges embraced online bench fact-finding. But the use of juries, both civil and criminal, caused angst, and some lawyers and judges remained skeptical that the justice system could or should implement any kind of online jury trial. Some of this skepticism waxed philosophical. Legal thinkers pondered, for example, whether a jury trial is a jury trial if parties, lawyers, judges, and the jurors themselves do not assemble in person in a public space.[8] Some of the skepticism about online trials may have been simple fear of the unknown.[9]

Other concerns skewed more practical and focused on matters for which existing research provided few answers. Will remote trials yield more representative juries by opening proceedings to those with challenges in attending in-person events, or will the online migration yield less-representative juries by marginalizing those on the wrong side of the digital divide? Will remote trials decrease lawyer costs and, perhaps as a result, expand access to justice by putting legal services more within the reach of those with civil justice needs?[10] Time, and more data, can tell if the political will is there to find out.

But two further questions lie at the core of debate about whether online proceedings can produce approximately accurate and publicly acceptable fact-finding from testimonial evidence. The first is whether video hearings diminish the ability of fact-finders to detect witness deception or mistake. The second is whether video as a communication medium dehumanizes one or both parties to a case, causing less humane decision-making.[11] For these questions, the research provides an answer (for the first question) and extremely useful guidance (for the second).

[7] For discussions of beliefs deserving of similar levels of credibility, *see, e.g.*, Malcolm W. Browne, *Perpetual Motion?* N.Y. Times (June 4, 1985), https://www.nytimes.com/1985/06/04/science/perpetual-motion.html; *Astronomy: Geocentric Model*, Encyclopaedia Britannica, https://www.britannica.com/science/geocentric-model. We focus on this point further below.

[8] *See, e.g.*, Susan A. Bandes & Neal Feigenson, *Virtual Trials: Necessity, Invention, and the Evolution of the Courtroom*, 68 Buff. L. Rev. 1275 (2020); Christopher L. Dodson, Scott Dodson & Lee H. Rosenthal, *The Zooming of Federal Litigation*, 104 Judicature 12 (2020); Jenia Iontcheva Turner, *Remote Criminal Justice*, 53 Tex. Tech. L. Rev. 197 (2021).

[9] Ed Spillane, *The End of Jury Trials: COVID-19 and the Courts: The Implications and Challenges of Holding Hearings Virtually and in Person during a Pandemic from a Judge's Perspective*, 18 Ohio St. J. Crim. L. 537 (2021).

[10] David Freeman Engstrom, *Digital Civil Procedure*, 169 U. Pa. L. Rev. 7 (2021); David Freeman Engstrom, *Post-COVID Courts*, 68 UCLA L. Rev. Disc. 246 (2020); Richard Susskind, *Remote Courts*, Practice, July/Aug. 2020, at 1.

[11] The second of the two questions addressed herein is more often expressed with respect to criminal trials, with the worry being dehumanization of an accused. See Derwyn Bunton, Chair, Nat'l Ass'n for Pub. Def., NAPD Statement on the Issues with the Use of Virtual Court Technology (2020), https://www.publicdefenders.us/files/NAPD%20Virtual%20Court%20Statement%208_1.pdf; *see also* Turner, *Remote Criminal Justice*; Anne Bowen Poulin, *Criminal Justice and Videoconferencing Technology: The Remote Defendant*, 78 Tul. L. Rev. 1089 (2004); Cormac T. Connor, *Human Rights Violations in the Information Age*, 16 Geo. Immigr. L.J. 207 (2001). But dehumanization might apply equally to the civil context,

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press

This chapter addresses these questions by reviewing the relevant research. By way of preview, our canvass of the relevant literatures suggests that both concerns articulated above are likely misplaced. Although there is reason to be cautious and to include strong evaluation with any move to online fact-finding, including jury trials, available research suggests that remote hearings will neither alter the fact-finder's (nearly non-existent) ability to discern truthful from deceptive or mistaken testimony nor materially affect fact-finder perception of parties or their humanity. On the first point, a well-developed body of research concerning the ability of humans to detect when a speaker is lying or mistaken shows a consensus that human detection accuracy is only slightly better than a coin flip. Most importantly, the same well-developed body of research demonstrates that such accuracy is the same regardless of whether the interaction is in-person or virtual (so long as the interaction does not consist solely of a visual exchange unaccompanied by sound, in which case accuracy worsens). On the second point, the most credible studies from the most analogous situations suggest little or no effect on human decisions when interactions are held via videoconference as opposed to in person. The evidence on the first point is stronger than that on the second, but the key recognition is that for both points, the weight of the evidence is contrary to the concerns that lawyers and judges have expressed, suggesting that the Bench and the Bar should pursue online courts (coupled with credible evaluation) to see if they offer some of the benefits their proponents have identified. After reviewing the relevant literature, this chapter concludes with a brief discussion of a research agenda to investigate the sustainability of remote civil justice.

A final point concerning the scope of this chapter: As noted above, judges and (by necessity) lawyers have largely come to terms with online bench fact-finding based on testimonial evidence, reserving most of their skepticism for online jury trials. But as we relate below, the available research demonstrates that to the extent that the legal profession harbors concerns regarding truth-detection and dehumanization in the jury's testimonial fact-finding, it should be equally skeptical regarding that of the bench. The evidence in this area either fails to support or affirmatively contradicts the belief that judges are better truth-detectors, or are less prone to dehumanization, than laity. Thus, we focus portions of this chapter on online jury trials because unwarranted skepticism has prevented such adjudications from the level of use (under careful monitoring and evaluation) that the currently jammed court system likely needs. But if current research (or our analysis of it here) is wrong, meaning that truth-detection and dehumanization are fatal concerns for online jury trials,

where the concern might be a reduced human connection to, say, a plaintiff suing about a physical injury. Indeed, depending on one's perspective, in a case involving a human being against a corporation, dehumanization might eliminate an unfair emotional advantage inuring to the human party or further "skew" a system already tilted in favor of corporate entities. *See* Engstrom, *Digital Civil Procedure*.

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press

then online bench adjudications based on testimonial evidence should be equally concerning.

4.1 ADJUDICATION OF TESTIMONIAL ACCURACY

A common argument against the sustainability of video testimonial hearings is that, to perform its function, the fact-finder must be able to distinguish accurate from inaccurate testimony. Inaccurate testimony could arise via two mechanisms: deceptive witnesses, that is, those who attempt to mislead the trier of fact through deliberate or knowing misstatements; and mistaken witnesses, that is, those who believe their testimony even though their perception of historical fact was wrong.[12] The legal system reasons as follows. First, juries are good, better even than judges, at choosing which of two or more witnesses describing incompatible versions of historical fact is testifying accurately.[13] Second, this kind of "demeanor evidence" takes the form of nonverbal cues observable during testimony.[14] Third, juries adjudicate credibility best if they can observe witnesses' demeanor in person.[15]

Each component of this reasoning is false.

First, research shows that humans have just above a fifty-fifty chance ability to detect lies, approximately 54 percent overall, if we round up. For example, Bond and

[12] To clarify: we refer in this chapter to concerns of duplicitous or mistaken testimony about particular historical events and, relatedly, conflicts in testimony about such events. Imagine, in other words, one witness who testifies that the light was green at the time of an accident, and another witness who says that the light was red. Physics being what it is, one of these witnesses is lying or mistaken. In cases of genuine issues of material fact, a trier of fact (a jury, if relevant law makes one available) is ordinarily supposed to discern which witness is truthful and accurate. Triers of fact, including juries, may be able to identify mistakes or lies by considering the plausibility of testimony given other (presumably accurate and transparent) evidence; *see, e.g.*, Scott v. Harris, 550 U.S. 372 (2007), or background circumstances, or by considering internal inconsistencies in a witness's testimony. But to our knowledge, few in the United States legal profession argue that an in-person interaction is necessary to facilitate this latter exercise in truth detection.

[13] Markman v. Westview Instruments, 570 U.S. 370 (1996).

[14] *See, e.g.*, George Fisher, *The Jury's Rise as Lie Detector*, 107 YALE L.J. 576, 577 n.2, 703 n.597 (1997) (collecting case law adjudicating the advisability of judges promoting the use of nonverbal cues to detect witness deception); Cara Salvatore, *May It Please the Camera: Zoom Trials Demand New Skills*, LAW360 (June 29, 2020), https://www.law360.com/articles/1278361/may-it-please-the-camera-zoom-trials-demand-new-skills ("Being able to see the witnesses face-on instead of sideways, [a judge] said, vastly improves the main job of fact-finders – assessing credibility.").

[15] *See, e.g.*, Mattox v. United States, 156 U.S. 237, 243 (1895) (holding that in the criminal context, the Constitution gives the defendant the right to "compel [a prosecution witness] to stand face to face with the jury in order that they may look at him, and judge by his demeanor upon the stand and the manner in which he gives his testimony whether he is worthy of belief"). Note that there are challenges to the notion of what it means to be "face to face" with a witness other than the online environment. See Julia Simon-Kerr, *Unmasking Demeanor*, 88 GEO. WASH. L. REV. ARGUENDO 158 (2020) (discussing the challenges posed by a policy of requiring witnesses to wear masks while testifying).

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press

DePaulo, in their meta-analysis of deception detection studies, place human ability to detect deception at 53.98 percent, or just above chance.[16] Humans are really bad at detecting deception. That probably comes as a shock to many of us. We are pretty sure, for example, we can tell when an opponent in a game or a sport is fibbing or bending the truth to get an edge. We are wrong. It has been settled in science since the 1920s that human beings are bad at detecting deception.[17] The fact that judges and lawyers continue to believe otherwise is a statement of the disdain in which the legal profession holds credible evidence and empiricism more generally.[18]

Second, human (in)ability to detect deception does not change with an in-person or a virtual interaction. Or at least, there is no evidence that there is a difference between in-person and virtual on this score, and a fair amount of evidence that there

[16] Charles F. Bond Jr. & Bella M. DePaulo, *Accuracy of Deception Judgements*, 10 Personality & Soc. Psych. Rev. 214, 219 (2006).

[17] *See* William M. Marston, *Studies in Testimony*, 15 J. Crim. L. & Criminology 5, 22–26 (1924); Glenn E. Littlepage & Martin A. Pineault, *Verbal, Facial, and Paralinguistic Cues to the Detection of Truth and Lying*, 4 Personality & Soc. Psych. Bull. 461 (1978); John E. Hocking, Joyce Bauchner, Edmund P. Kamiski & Gerald R. Miller, *Detecting Deceptive Communication from Verbal, Visual, and Paralinguistic Cues*, 6 Hum. Comm. Rsch. 33, 34 (1979); Miron Zuckerman, Bella M. DePaulo & Robert Rosenthal, *Verbal and Nonverbal Communication of Deception*, 14 Adv. Exp. Soc. Psych. 1, 39–40 (1981); Bella M. DePaulo & Robert L. Pfeifer, *On-the-Job Experience and Skill at Detecting Deception*, 16 J. Applied Soc. Psych. 249 (1986); Ginter Kohnken, *Training Police Officers to Detect Deceptive Eyewitness Statements: Does It Work?* 2 Soc. Behav. 1 (1987). For experiments specifically addressing the effects of rehearsal, *see, e.g.*, Joshua A. Fishman, *Some Current Research Needs in the Psychology of Testimony*, 13 J. Soc. Issues 60, 64–65 (1957); Norman R. F. Maier, *Sensitivity to Attempts at Deception in an Interview Situation*, 19 Personnel Psych. 55 (1966); Norman R. F. Maier & Junie C. Janzen, *Reliability of Reasons Used in Making Judgments of Honesty and Dishonesty*, 25 Perceptual & Motor Skills 141 (1967); Norman R. F. Maier & James A. Thurber, *Accuracy of Judgments of Deception When an Interview Is Watched, Heard, and Read*, 21 Personnel Psych. 23 (1968); Paul Ekman & Wallace V. Friesen, *Nonverbal Leakage and Cues to Deception*, 32 Psychiatry 88 (1969); Paul Ekman & Wallace V. Friesen, *Detecting Deception from the Body or Face*, 29 J. Personality & Soc. Psych. 288 (1974); Pal Ekman, Wallace V. Friesen & Klaus R. Scherer, *Body Movement and Voice Pitch in Deceptive Interaction*, 16 Semiotica 23 (1976); Gerald R. Miller & Norman E. Fontes, The Effects of Videotaped Court Materials on Juror Response 11–42 (1978); Glenn E. Littlepage & Martin A. Pineault, *Detection of Deceptive Factual Statements from the Body and the Face*, 5 Personality & Soc. Psych. Bull. 325, 328 (1979); Gerald R. Miller, Mark A. deTurck & Pamela J. Kalbfleisch, *Self-Monitoring, Rehearsal, and Deceptive Communication*, 10 Hum. Comm. Rsch. 97, 98–99, 114 (1983) (reporting unpublished studies of others as well as the authors' work); Carol Toris & Bella M. DePaulo, *Effects of Actual Deception and Suspiciousness of Deception on Interpersonal Perceptions*, 47 J. Personality & Soc. Psych. 1063 (1984) (although most people cannot do better than chance in detecting falsehoods, most people confidently believe they can do so); Paul Ekman, Telling Lies 162–89 (1985); Charles F. Bond Jr. & William E. Fahey, *False Suspicion and the Misperception of Deceit*, 26 Br. J. Soc. Psych. 41 (1987).

[18] For another example, see Chief Justice John Roberts, who reacted to quantitative standards for measuring political gerrymandering as follows: "It may be simply my educational background, but I can only describe it as sociological gobbledygook." Gill v. Whitford Oral Argument, Oyez, https://www.oyez.org/cases/2017/16-1161.

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press

is no difference. Most likely, humans are really bad at detecting deception regardless of the means of communication. This also probably comes as a shock to many of us. In addition to believing (incorrectly) that we can tell when kids or partners or opponents are lying, we think face-to-face confrontation matters. Many judges and lawyers so believe. They are wrong.

Why are humans so bad at deception detection? One reason is that people rely on what they think are nonverbal cues. For example, many think fidgeting, increased arm and leg movement, and decreased eye contact are indicative of lying. None are. While there might be some verbal cues that could be reliable for detecting lies, the vast majority of nonverbal cues (including those just mentioned, and most others upon which we humans tend to rely) are unreliable, and the few cues that might be modestly reliable can be counterintuitive.[19] Furthermore, because we hold inaccurate beliefs about what is and is not reliable, it is difficult for us to disregard the unreliable cues.[20] In a study that educated some participants on somewhat reliable nonverbal cues to look for, with other participants not getting this information, participants with the reliable cues had no greater ability to detect lying.[21] We humans are not just bad at lie detection; we are also bad at being trained at lie detection.

While a dishonest demeanor elevates suspicion, it has little-to-no relation to actual deception.[22] Similarly, a perceived honest demeanor is not reliably associated with actual honesty.[23] That is where the (ir)relevance of the medium of communication matters. If demeanor is an unreliable indicator for either honesty or dishonesty, then a fact-finder suffers little from losing whatever the supposedly superior opportunity to observe demeanor an in-person interaction might provide.[24] For

[19] Bella M. DePaulo, James J. Lindsay, Brian E. Malone & Laura Muhlenbruck, *Cues to Deception*, 129 PSYCH. BULL. 74, 95 (2003).

[20] Lucy Akehurst, Gunter Kohnken, Aldert Vrij & Ray Bull, *Lay Persons' and Police Officers' Beliefs regarding Deceptive Behaviour*, 10 APPLIED COGNITIVE PSYCH. 468 (1996).

[21] Hannah Shaw & Minna Lyons, *Lie Detection Accuracy: The Role of Age and the Use of Emotions as a Reliable Cue*, 32 J. POLICE CRIM. PSYCH. 300, 302 (2017).

[22] Lyn M. Van Swol, Michael T. Braun & Miranda R. Kolb, *Deception, Detection, Demeanor, and Truth Bias in Face-to-Face and Computer-Mediated Communication*, 42 COMMC'N RSCH. 1116, 1131 (2015).

[23] Timothy R. Levine et al., *Sender Demeanor: Individual Differences in Sender Believability Have a Powerful Impact on Deception Detection Judgements*, 37 HUM. COMMC'N RSCH. 377, 400 (2011).

[24] Unsurprisingly, lawyers and judges believe otherwise: "In a virtual jury trial, jurors lose the ability to lay eyes on witnesses in real time, and as a result, may miss nuances in behavior, speech patterns or other clues relevant to whether the witness is telling the truth." Paula Hinton & Tom Melsheimer, *The Remote Jury Trial Is a Bad Idea*, LAW360 (June 9, 2020), https://www.law360.com/articles/1279805. Hinton & Melsheimer offer no support for this assertion, other than (we imagine) their gut instincts based on their having "practic[ed] for over a combined 70 years." *Id.* The scientific consensus regarding deception detection has been reaffirmed for longer than seventy years. It involves dozens of studies pursued by scores of researchers. *See* sources cited *supra* note 17.

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press

example, a study from 2015 shows that people attempting to evaluate deception performed better when the interaction was a computer-mediated (text-based communication) rather than in-person communication.[25] At least one possible explanation for this finding is the unavailability of distracting and unreliable nonverbal cues.[26]

Despite popular belief in the efficacy of discerning people's honesty based on their demeanor, research shows that non-demeanor cues, meaning verbal cues, are more promising. A meta-analysis concluded that cues that showed promise at signaling deception tended to be verbal (content of what is said) and paraverbal (how it is spoken), not visual.[27] But verbal and paraverbal cues are just as observable from a video feed.

If we eliminate visual cues for fact-finders, and just give them audio feed, will that improve a jury's ability to detect deception? Unlikely. Audio-only detection accuracy does not differ significantly from audiovisual.[28] At this point, that should not be a surprise, considering the general low ceiling of deception detection accuracy – just above the fifty-fifty level. Only in high-pressure situations is it worthwhile (in a deception detection sense) to remove nonverbal cues.[29] To clarify: High-pressure situations likely make audio only better than audio plus visual, not the reverse. The problem for deception detection appears to be that, with respect to visual cues, the pressure turns the screws both on someone who is motivated to be believed but is actually lying and on someone who is being honest but feels as though they are not believed.

We should not think individual judges have any better ability to detect deception than a jury. Notwithstanding many judges' self-professed ability to detect lying, the science that humans are poor deception detectors has no caveat for the black robe. There is no evidence that any profession is better at deception detection, and a great deal of evidence to the contrary. For example, those whose professions ask them to detect lies (such as police officers) cite the same erroneous cues regarding deception.[30] More broadly, two meta-analyses from 2006 show that purported "experts" at deception detection are no better at lie detection than nonexperts.[31]

[25] Van Swol et al., *Deception, Detection, Demeanor, and Truth Bias*, at 1131.
[26] *Id.* at 1136.
[27] DePaulo et al., *Cues to Deception*, at 95.
[28] Bond & DePaulo, *Accuracy of Deception Judgements*, at 225.
[29] Saul M. Kassin, Christian A. Meissner & Rebecca J. Norwick, *"I'd Know a False Confession if I Saw One": A Comparative Study of College Students and Police Investigators*, 29 Law & Hum. Behav. 211, 222 (2005).
[30] Akehurst et al., *Lay Persons' and Police Officers' Beliefs*, at 461.
[31] Bond & DePaulo, *Accuracy of Deception Judgements*, at 229; Michael Aamodt & Heather Custer, *Who Can Best Catch a Liar? A Meta-Analysis of Individual Differences in Detecting Deception*, 15 Forensic Exam'r 6, 10 (2006).

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press

What about individuals versus groups? A 2015 study did find consistently that groups performed better at detecting lies,[32] a result the researchers attributed to group synergy – that is, that individuals were able to benefit from others' thoughts.[33] So, juries are better than judges at deception detection, right? Alas, almost certainly not. The problem is that only certain kinds of groups are better than individuals. In particular, groups of individuals who were familiar with one another before they were assigned a deception detection task outperformed both individuals and groups whose members had no preexisting connection.[34] Groups whose members had no preexisting connection were no better at detecting deception than individuals.[35] Juries are, by design, composed of a cross-section of the community, which almost always means that jurors are unfamiliar with one another before trial.[36]

There is more bad news. Bias and stereotypes affect our ability to flush out a lie. Females are labeled as liars significantly more than males even when both groups lie or tell the truth the same amount.[37] White respondents asked to detect lies were significantly faster to select the "liar" box for black speakers than white speakers.[38]

All of this is troubling, and likely challenges fundamental assumptions of our justice system. For the purposes of this chapter, however, it is enough to demonstrate that human inability to detect lying remains constant whether testimony is received in-person or remotely. Again, the science on this point goes back decades, and it is also recent. Studies conducted in 2014[39] and 2015[40] agreed that audiovisual and audio-only mediums were not different in accuracy detection. The science suggests

[32] Nadav Klein & Nicholas Epley, *Group Discussion Improves Lie Detection*, 112 Proc. Nat'l Acad. Scis. 7460, 7464 (2015). Groups consisted of three people, which is fewer than usually empaneled for any jury.

[33] *Id.* at 7463.

[34] Roger McHaney, Joey F. George & Manjul Gupta, *An Exploration of Deception Detection: Are Groups More Effective Than Individuals?* 45 Commc'n Rsch. 1111 (2018).

[35] *Id.*

[36] Meanwhile, there is some evidence that deliberation of the kind in which juries engage may make things worse. Holly K. Orcutt et al., *Detecting Deception in Children's Testimony: Factfinders' Abilities to Reach the Truth in Open Court and Closed-Circuit Trials*, 25 Law & Hum. Beh. 339 (2001).

[37] E. Paige Lloyd, Kevin M. Summers, Kurt Hugenburg & Allen R. McConnell, *Revisiting Perceiver and Target Gender Effects in Deception Detection*, 42 J. Nonverbal Behav. 427, 435 (2018).

[38] E. Paige Lloyd et al., *Black and White Lies: Race-Based Biases in Deception Judgements*, 28 Psych. Sci. 1134 (2017).

[39] Charlotte D. Sweeney & Stephen J. Ceci, *Deception Detection, Transmission, and Modality in Age and Sex*, 5 Frontiers Psych. 5 (2014).

[40] Scott E. Culhane, Andre Kehn, Jessica Hatz & Meagen M. Hildebrand, *Are Two Heads Better Than One? Assessing the Influence of Collaborative Judgements and Presentation Mode on Deception Detection for Real and Mock Transgressions*, 12 J. Investigative Psych. Offender Profiling 158, 165 (2015).

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press

that the medium of communication – in-person, video, or telephonic – has little if any relevant impact on the ability of judges or juries to tell truths from lies.[41]

The statements above regarding human (in)ability to detect deception apply equally to human (in)ability to detect mistakes, including the fact that scientists have long known that we are poor mistake detectors. Thirty years ago, Wellborn collected and summarized the then-available studies, most focusing on eyewitness testimony. Addressing jury ability to distinguish mistaken from accurate witness testimony, Wellborn concluded that "the capacity of triers [of fact] to appraise witness accuracy appears to be worse than their ability to discern dishonesty."[42] Particularly relevant for our purposes, Wellborn further concluded that "neither verbal nor nonverbal cues are effectively employed" to detect testimonial mistakes.[43] If neither verbal nor nonverbal cues matter in detecting mistakes, then there will likely be little lost by the online environment's suppression of nonverbal cues.

The research in the last thirty years reinforces Wellborn's conclusions. Human inability to detect mistaken testimony in real-world situations is such a settled principle that researchers no longer investigate it, focusing instead on investigating other matters, such as the potentially distorting effects of feedback given to eyewitnesses,[44] whether witness age affects likelihood of fact-

[41] For excellent summaries of the historical research and literature, including many of the studies referenced in note 17 comparing our ability to detect deception across mediums, see Jeremy A, Blumenthal, *A Wipe of the Hands, a Lick of the Lips: The Validity of Demeanor Evidence in Assessing Witness Credibility*, 72 Neb. L. Rev. 1157 (1993); Olin Guy Wellborn II, *Demeanor*, 76 Cornell L. Rev. 1075 (1990). Note that we did find the occasional study suggesting to the contrary. *See, e.g.*, Sara Landstrom, Par Anders Granhag & Maria Hartwig, *Children's Live and Videotaped Testimonies: How Presentation Mode Affects Observers' Perception, Assessment, and Memory*, 12 Leg. & Crim. Psych. 333 (2007), but the overwhelming weight of the scientific evidence is as summarized above.

[42] Wellborn, *Demeanor*, at 1088.

[43] *Id.* at 1091.

[44] *See, e.g.*, Laura Smalarz & Gary L. Wells, *Post-Identification Feedback to Eyewitnesses Impairs Evaluators' Abilities to Discriminate between Accurate and Mistaken Testimony*, 38 Law & Hum. Behav. 194 (2014); C. A. E. Luus & G. L. Wells, *The Malleability of Eyewitness Confidence: Co-witness and Perseverance Effects*, 79 J. Applied Psych. 714 (1994). The Smalarz and Wells experiment reveals the lengths to which researchers must go to produce an artificial situation in which human beings are able to distinguish between accurate and mistaken testimony (a situation that may be required to investigate something else). In this study, to assure a sufficient number of "mistaken" witnesses, investigators had to deceive witnesses intended to provide a mistaken eyewitness identification into believing that the correct perpetrator was present in a lineup. Smalarz & Wells, *Post-Identification Feedback*, at 196. Extraordinary measures to produce distinguishably accurate versus inaccurate hypothetical witnesses characterize other studies. One study resorted to cherry-picking from numerous mock witnesses those judged most and least accurate for later use. Michael R. Leippe, Andrew P. Manion & Ann Romanczyk, *Eyewitness Persuasion: How and How Well Do Fact Finders Judge the Accuracy of Adults' and Children's Memory Reports?* 63 J. Personality & Soc. Psych. 181 (1992). Another made no effort to provide fact-finders with witness testimony. Instead, fact-finders were provided with questionnaires that witnesses completed about their thought processes. David Dunning & Lisa Beth Stern, *Distinguishing Accurate from Inaccurate*

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press

finder belief,[45] and whether fact-finders understand the circumstances mitigating the level of unreliability of eyewitness testimony.[46] The most recent, comprehensive writing we could find on the subject was a 2007 chapter from Boyce, Beaudry, and Lindsay, which depressingly concluded (1) fact-finders believe eyewitnesses, (2) fact-finders are not able to distinguish between accurate and inaccurate eyewitnesses, and (3) fact-finders base their beliefs of witness accuracy on factors that have little relationship to accuracy.[47] This review led us to a 1998 study of child witnesses that found (again) no difference in a fact-finder's capacity to distinguish accurate from mistaken testimony as between video versus in-person interaction.[48]

In short, decades of research provide strong reason to question whether fact-finders can distinguish accurate from inaccurate testimony, but also strong reason to believe that no difference exists on this score between in-person versus online hearings, nor between judges and juries.

4.2 THE ABSENCE OF A DEHUMANIZATION EFFECT

Criminal defense attorneys have raised concerns of dehumanization of defendants, arguing that in remote trials, triers of fact will have less compassion for defendants and will be more willing to impose harsher punishments.[49] In civil trials, this concern could extend to either party. For example, in a personal injury case, the concern might be that fact-finders would be less willing to award damages because they are unable to connect with or relate to a plaintiff's injuries. Or, in less protracted but nevertheless high-stakes civil actions, such as landlord/tenant matters, a trier of fact (usually a judge) might feel less sympathy for a struggling tenant and therefore show greater willingness to evict rather that mediate a settlement.

Eyewitness Identifications via Inquiries about Decision Processes, 67 J. PERSONALITY & SOC. PSYCH. 818 (1994).

[45] Peter A. Newcombe & Jennifer Bransgrove, *Perceptions of Witness Credibility: Variations across Age*, 28 J. APPLIED DEV. PSYCH. 318 (2007); Leippe et al., *Eyewitness Persuasion*.

[46] Richard Schmechel, Timothy O'Toole, Catherine Easterly & Elizabeth Loftus, *Beyond the Ken? Testing Jurors' Understanding of Eyewitness Reliability Evidence*, 46 JURIMETRICS 177 (2006).

[47] Melissa Boyce, Jennifer Baeaudry & R. C. L. Lindsay, *Belief of Eyewitness Identification Evidence*, *in* 2 THE HANDBOOK OF EYEWITNESS PSYCHOLOGY 501 (R. C. L. Lindsay, D. F. Ross, J. D. Read & M. P. Toglia eds., 2007); *see also* Neil Brewer & Anne Burke, *Effects of Testimonial Inconsistencies and Eyewitness Credibility on Mock-Juror Judgments*, 26 LAW & HUM. BEHAV. 353 (2002) (finding that fact-finders favor witness confidence over testimonial consistency in their accuracy judgments); Steven Penrod & Brian Cutler, *Witness Confidence and Witness Accuracy: Assessing Their Forensic Relation*, 1 PSYCH. PUB. POL'Y & L. 817 (1995) (same); Siegfried L. Sporer, Steven Penrod, Don Read & Brian Cutler, *Choosing, Confidence, & Accuracy: A Meta-Analysis of the Confidence-Accuracy Relation in Eyewitness Identification Studies*, 118 PSYCH. BULL. 315 (1995) (same).

[48] Gail S. Goodman et al., *Face-to-Face Confrontation: Effects of Closed-Circuit Technology on Children's Eyewitness Testimony and Jurors' Decisions*, 22 LAW & HUM. BEHAV. 165 (1998).

[49] *See, e.g.*, BUNTON, NAPD STATEMENT.

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press

TABLE 4.1 *A hierarchy of study designs*

Tier	Randomized?	Setting	Example
1: RCTs	Yes	Field	Cuevas et al.
2: Lab Studies	Yes	Lab	Lee et al.
3: Observational Studies	No	Field	Walsh & Walsh

A review of relevant literature suggests that this concern is likely misplaced. While the number of studies directly investigating the possibility of online hearings is limited, analogous research from other fields is available. We focus on studies in which a decision-maker is called upon to render a judgment or decision that affects the livelihood of an individual after some interaction with that individual, much like a juror or judge is called upon to render a decision that affects the livelihood of a litigant. We highlight findings from a review of both legal and analogous nonlegal studies, and we emphasize study quality – that is, we prioritize randomized before nonrandomized trials, field over lab/simulated experiments, and studies involving actual decision-making over studies involving precursors to decisions (e.g., ratings or impressions).[50] With this ordering of pertinence, we systematically rank the literature into three tiers, from most to least robust: first, randomized field studies (involving decisions and precursors); second, randomized lab studies (involving decisions and precursors); and third, non-randomized studies. Table 4.1 provides a visual of our proposed hierarchy.

According to research in the first tier of randomized field studies – the most telling for probing the potential for online fact-finding based on testimonial evidence – proceeding via videoconference likely will not adversely affect the perceptions of triers of facts on the humanity of trial participants. We do take note of the

[50] *See* Joshua D. Angrist, *Instrumental Variables Methods in Experimental Criminological Research: What, Why and How*, 2 J. Exp. Criminol. 23, 24 (2006) (randomized studies are considered the gold standard for scientific evidence). The idea that randomized control trials are the gold standard in scientific research investigating causation is unfamiliar to some in the legal profession. Such studies are the gold standard because they create similarly situated comparison groups at the same point in time. Randomized studies create groups statistically identical to one another except that one is not exposed to the intervention or program (in the case of this section, video interaction), allowing us to know with as much certainty as science allows that the reason for any observed differences in outcome is the intervention or program. By contrast, a commonly used methodology that compares outcomes prior to an intervention's implementation to outcomes after the implementation could, for example, turn on factors that do not exist in both groups. If we think about a legal intervention, those factors could be a change in presiding judge, a new crop of lawyers working on these cases, a change in procedural or substantive law, a reduction in policing or change in arrest philosophy, implementation or abandonment of reliance on a risk-assessment tool, bail reform, etc. The randomized trial eliminates as far as possible the potentially influencing factors. For that reason, it is thought of as the gold standard.

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press

findings of studies beyond this first tier, which include some from the legal field. Findings in these less credible tiers are varied and inconclusive.

The research addressing dehumanization is less definitive than that addressing deception and mistake detection. So, while we suggest that jurisdictions consider proceeding with online trials and other innovative ways of addressing both the current crises of frozen court systems and the future crises of docket challenges, we recommend investigation and evaluation of such efforts through randomized control trials (RCTs).

4.2.1 *Who Would Be Dehumanized?*

At the outset, we note a problem common to all of the studies we found, in all tiers: none of the examined situations are structurally identical to a fact-finding based on an adversarial hearing. In legal fact-finding in an adversarial system, one or more theoretically disinterested observers make a consequential decision regarding the actions of someone with whom they may have no direct interaction and who, in fact, sometimes exercises a right not to speak during the proceeding. In all of the studies we were able to find, which concerned situations such as doctor-patient or employer-applicant, the decision-maker interacted directly with the subject of the interaction. It is thus questionable whether any study yet conducted provides ideal insight regarding the likely effects of online as opposed to in-person for, say, civil or criminal jury trials.

Put another way: If a jury were likely to dehumanize or discount someone, why should it be that it would dehumanize any specific party, as opposed to the individuals with whom the jury "interacts" ("listens to" would be a more accurate phrase), namely, witnesses and lawyers? With this in mind, it is not clear which way concerns of dehumanization cut. At present, one defensible view is that there is no evidence either way regarding dehumanization of parties in an online jury trial versus an in-person one, and that similar concerns might be present for some types of judicial hearings.

Some might respond that the gut instincts of some criminal defense attorneys and some judges should count as evidence.[51] We disagree that the gut instincts of any human beings, professionals or otherwise, constitute evidence in almost any setting. But we are especially suspicious of gut instincts in the fact-finding context. As we saw in the previous part, fact-finding based on testimonial hearings has given rise to some of the most stubbornly persistent, and farcically outlandish, myths to which lawyers and judges cling. The fact that lawyers and judges continue to espouse this kind of flat-eartherism suggests careful interrogation of professional gut instincts on the subject of dehumanization from an online environment.

[51] Bunton, NAPD Statement.

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press

4.2.2 *Promising Results from Randomized Field Studies*

Within our first-tier category of randomized field studies, the literature indicates that using videoconference in lieu of face-to-face interaction has an insignificant, or even a positive, effect on a decision-maker's disposition toward the person about whom a judgment or decision is made.[52] We were unable to find any randomized field studies concluding that videoconferencing, as compared to face-to-face communication, has an adverse or damaging effect on decision outcomes.

Two randomized field studies in telemedicine, conducted in 2000[53] and 2006,[54] both found that using videoconferencing rather than face-to-face communication had an insignificant effect on the outcomes of real telemedicine decisions. Medical decisions were equivalent[55] or identical.[56] It is no secret that medicine implemented tele-health well before the justice system implemented tele-justice.[57]

Similarly, a 2001 randomized field study of employment interviews conducted using videoconference versus in-person interaction resulted in videoconference applicants being rated higher than their in-person counterparts. Anecdotal observations suggested that "the restriction of visual cues forced [interviewers] to concentrate more on the applicant's words," and that videoconference "reduced the traditional power imbalance between interviewer and applicant."[58]

From our review of tier-one studies, then, we conclude that there is no evidence that the use of videoconferencing makes a difference on decision-making. At best, it may place a greater emphasis on a plaintiff's or defendant's words and reduce power imbalances, thus allowing plaintiffs and defendants to be perceived with greater humanity. At worst, videoconferencing makes no difference.

That said, we found only three tier-one studies. So, we turn our attention to studies with less strong designs.

[52] By way of explanation, the difference between field studies and lab studies is whether the experiment is conducted in a live setting or a contrived setting. While lab studies are common and valuable – especially when field experiments are difficult or, even more problematically, challenge ethics or decency – the scientific community generally places greater weight on field studies.

[53] Rod Eldford et al., *A Randomized Controlled Trial of Child Psychiatric Assessments Conducted Using Videoconferencing*, 6 J. Telemedicine & Telecare 73, 74–75 (2000).

[54] Carlos De Las Cuevas et al., *Randomized Clinical Trial of Telepsychiatry through Videoconference versus Face-to-Face Conventional Psychiatric Treatment* 12 Telemedicine & E-Health 341, 341 (2006).

[55] *Id.*

[56] *See* Angrist, *Instrumental Variables*.

[57] *See* Risto Roine, Arto Ohinmaa & David Hailey, *Assessing Telemedicine: A Systematic Review of the Literature*, 165 Canadian Med. Ass'n J. 765, 766 (2001) (noting the earliest reviews of telemedicine occurred in 1995).

[58] Derek S. Chapman & Patricia M. Rowe, *The Impact of Videoconference Technology, Interview Structure, and Interviewer Gender on Interviewer Evaluations in the Employment Interview: A Field Experiment*, 74 J. Occup. & Org. Psych. 279, 291 (2001).

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press

4.2.3 *Varied Findings from Studies with Less Strong Designs*

Randomized lab studies and non-randomized studies provide a less conclusive array of findings, causing us to recommend that use of remote trials be accompanied by careful study. Randomized lab studies and non-randomized studies are generally not considered as scientifically rigorous as randomized field studies; much of the legal literature – which might be considered more directly related to remote justice – falls within this tier of research.

First, there are results, analogous to the tier-one studies, suggesting that using videoconference in lieu of face-to-face interaction has an insignificant effect for the person about whom a decision is being made. For example, in a study testing the potential dehumanizing effect of videoconferencing as compared to in-person interactions in a lab setting where doctors were given the choice between a painful but more effective treatment versus a painless but less effective treatment, no dehumanizing effect of communication medium was found.[59] If the hypothesis that videoconferencing dehumanizes patients (or their pain) were true, in this setting, we might expect to see doctors prescribing a more painful but more effective treatment. No such difference emerged.

Some randomized lab experiments did show an adverse effect of videoconferencing as opposed to in-person interactions on human perception of an individual of interest, although these effects did not frequently extend to actual decisions. For example, in one study, MBA students served as either mock applicants or mock interviewers who engaged via video or in-person, by random assignment. Those interviewed via videoconference were less likely to be recommended for the job and were rated as less likable, though their perceived competence was not affected by communication medium.[60] Other lab experiments have also concluded that the videoconference medium negatively affects a person's likability compared with the in-person medium.

Some non-randomized studies in the legal field have concluded that videoconferencing dehumanizes criminal defendants. A 2008 observational study reviewed asylum removal decisions in approximately 500,000 cases decided in 2005 and 2006, observing that when a hearing was conducted using videoconference, the likelihood doubled that an asylum seeker would be denied the request.[61] In a Virtual Court pilot program conducted in the United Kingdom, evaluators found that the

[59] Min Kyung Lee, Nathaniel Fruchter & Laura Dabbish, *Making Decisions from a Distance: The Impact of Technological Mediation on Riskiness and Dehumanization*, 18 Proc. Ass'n Computing Mach. Conf. on Comput. Supported Coop. Work & Soc. Computing 1576, 1570 (2015).

[60] Fernando Robles et al., *A Comparative Assessment of Videoconference and Face-to-Face Employment Interviews*, 51 Mgmt. Decision 1733, 1740 (2013).

[61] We caution that besides stating that the study included "all asylum cases differentiated between hearings conducted via [videoconference], telephone, and in-person for FY 2005 and FY 2006," further selection details were not mentioned. Frank M. Walsh & Edward M. Walsh, *Effective*

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press

use of videoconferencing resulted in high rates of guilty pleas and a higher likelihood of a custodial sentence.[62] Finally, an observational study of bail decisions in Cook County, Illinois, found an increase in average bond amount for certain offenses after the implementation of CCTV bond hearings.[63] Again, however, these studies were not randomized, and well-understood selection or other biasing effects could explain all these results.

4.2.4 *Wrapping Up Dehumanization*

While the three studies first mentioned are perhaps the most analogous to the situation of a remote jury or bench hearing because they are analyzing the effects of remote legal proceedings, we cannot infer much about causation from them. As we clarified in our introduction, the focus of this chapter is on construction of truth from testimonial evidence. Some of the settings (e.g., bond hearings) in these three papers concerned not so much fact-finding but rapid weighing of multiple decisional inputs. In any event, the design weaknesses of these studies remain. And if one discounts design problems, we still do not know whether any unfavorable perception affects both parties equally, or just certain witnesses or lawyers.

The randomized field studies do point toward a promising direction for the implementation of online trials and sustainability of remote hearings. The fact that these studies are non-legal but analogous in topic and more scientifically robust in procedure may trip up justice system stakeholders, who might be tempted to believe that less-reliable results that occur in a familiar setting deserve greater weight than more reliable results occurring in an analogous but non-legal setting. As suggested above, such is not necessarily wise.

We found only three credible (randomized) studies. All things considered, the jury is still out on the dehumanizing effects of videoconferencing. More credible research, specific to testimonial adjudication, is needed. But for now, the credible

Processing or Assembly-Line Justice? The Use of Teleconferencing in Asylum Removal Hearings, 22 Geo. Immigr. L.J. 259, 259–71 (2008).

[62] The process was not random. The pilot took place in parts of London and Kent with two magistrates' courts participating (Camberwell Green and Medway) and sixteen police stations participating. Defendants in these participating courts had to give their consent before appearing in a Virtual Court (in December 2009, the consent requirement was removed). There was also a list of suitability criteria; if a defendant met any one of these criteria, the case was deemed unsuitable for videoconferencing. Matthew Terry, Steve Johnson & Peter Thompson, U.K. Ministry Just., Virtual Court Pilot Outcome Evaluation 1, 24–31 (2010).

[63] Dane Thorley & Joshua Mitts, *Trial by Skype: A Causality-Oriented Replication Exploring the Use of Remote Video Adjudication in Immigration Removal Proceedings*, 59 Int'l Rev. L. & Econ. 82 (2019); Ingrid V. Eagly, *Remote Adjudication in Immigration*, 109 Nw. U. L. Rev. 933 (2015); Shari Seidman Diamond, Locke E. Bowman, Manyee Wong & Matthew M. Patton, *Efficiency and Cost: The Impact of Videoconferenced Hearings on Bail Decisions*, 100 J. Crim. L. & Criminology 869, 897 (2010).

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press

research may relieve concerns about the dehumanizing effect of remote justice. Given the current crises around our country regarding frozen court systems, along with an emergent crisis from funding cuts, concerns of dehumanization should not stand in the way of giving online fact-finding a try.

4.3 A RESEARCH AGENDA

A strong research and evaluation program should accompany any move to online fact-finding.[64] The concerns are various, and some are context-dependent. Many are outside the focus of this chapter. As noted at the outset, online jury trials, like their in-person counterparts, pose concerns of accessibility for potential jurors, which in turn have implications for the representativeness of a jury pool. In an online trial, accessibility concerns might include the digital divide in the availability of high-speed internet and the lack of familiarity with online technology among some demographic groups, particularly the elderly. Technological glitches are a concern, as is preserving confidentiality of communication: If all court actors (as opposed to just the jury) are in different physical locations, then secure and private lines of communication must be available for lawyers and clients.[65] In addition, closer to the focus of this chapter, some in the Bench and the Bar might believe that in-person proceedings help focus jurors' attention while making witnesses less likely to deceive or to make mistakes; we remain skeptical of these assertions, particularly the latter, but they, too, deserve empirical investigation. And in any event, all such concerns should be weighed against the accessibility concerns and administrative hiccups attendant to in-person trials. Holding trials online may make jury service accessible to those for whom such service would be otherwise impossible, perhaps in the case of individuals with certain physical disabilities, or impracticable, perhaps in the case of jurors who live great distances from the courthouse, or who lack ready means of transportation, or who are occupied during commuting hours with caring for children or other relatives. Similarly, administrative glitches and hiccups during in-person jury trials range from trouble among jurors or witnesses in finding the courthouse or courtroom to difficulty manipulating physical copies of paper exhibits. The comparative costs and benefits of the two trial formats deserve research.

Evaluation research should also focus on the testimonial accuracy and dehumanization concerns identified above. As Sections 4.1 and 4.2 suggest, RCTs, in which hearings or trials are randomly allocated to an in-person or an online format, are

[64] Molly Treadway Johnson & Elizabeth C. Wiggins, *Videoconferencing in Criminal Proceedings: Legal and Empirical Issues and Directions for Research*, 28 Law & Pol'y 211 (2006).

[65] Eric T. Bellone, *Private Attorney-Client Communications and the Effect of Videoconferencing in the Courtroom*, 8 J. Int'l Com. L. & Tech. 24 (2013).

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press

necessary to produce credible evidence. In some jurisdictions, changes in law might be necessary.[66]

But none of these issues is conceptually difficult, and describing strong designs is easy. A court system might, for example, engage with researchers to create a system that assigns randomly a particular type of case to an online or in-person hearing involving fact-finding.[67] The case type could be anything: summary eviction, debt collection, government benefits, employment discrimination, suppression of evidence, and the like. The adjudicator could be a court or an agency. Researchers can randomize cases using any number of means, or cases can be assigned to conditions using odd/even case numbers, which is ordinarily good enough even if not technically random.

It is worth paying attention to some details. For example, regarding the outcomes to measure, an advantage to limiting each particular study to a particular case type is that comparing adjudicatory outputs is both obvious and easy. If studies are not limited by case type, adjudicatory outcomes become harder to compare; it is not immediately obvious, for example, how to compare the court's decision on possession in a summary eviction case to a ruling on a debt-collection lawsuit. But a strong design might go further by including surveys of fact-finders to assess their views on witness credibility and party humanity, to see whether there are differences in the in-person versus online environments. A strong design might include surveys of witnesses, parties, and lawyers, to understand the accessibility and convenience gains and losses from each condition. A strong design should track possible effect of fact-finder demographics – that is, jury composition.

Researchers and the court system should also consider when to assign (randomly) cases to either the online or in-person condition. Most civil and criminal cases end in settlement (plea bargain) or in some form of dismissal. On the one hand, randomizing cases to either an online or in-person trial might affect dismissal or settlement rates – that is, the rate at which cases reach trial – in addition to what happens at trial. Such would be good information to have. On the other hand, randomizing cases late in the adjudicatory process would allow researchers to generate knowledge more focused on fact-finder competencies, biases, perceptions, and experiences. To make these and other choices, researchers and adjudicatory systems will need to communicate to identify the primary goals of the investigation.

[66] *See, e.g.*, FED. R. CIV. P. 43 (contemplating "testimony . . . by contemporaneous transmission from a different location" only "[f]or good cause in compelling circumstances and with appropriate safeguards").

[67] We have focused this chapter on hearings involving fact-finding based on testimonial evidence. Hearings that do not involve fact-finding (status conferences, bond hearings, oral arguments, and so on) are now accepted, even routine. But the fact that something is accepted and routine does not make it advisable. For non-fact-finding hearings, concerns of truth detection might be lessened, but dehumanization and other concerns remain extant. We recommend strong evaluation of online versus in-person hearings that do not involve fact-finding as well.

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press

Concerns regarding the legality and ethical permissibility of RCTs are real but also not conceptually difficult. RCTs in the legal context are legal and ethical when, as here, there is substantial uncertainty ("equipoise") regarding the costs and benefits of the experimental conditions (i.e., online versus in-person trials).[68] This kind of uncertainty/equipoise is the ethical foundation for the numerous RCTs completed each year in medicine.[69] Lest we think the consequences of legal adjudications too high to permit the randomization needed to generate credible knowledge, medicine crossed this bridge decades ago.[70] Many medical studies measure death as a primary outcome. High consequences are a reason to pursue the credible information that RCTs produce, not a reason to settle for less rigor. To make the study work, parties, lawyers, and other participants will not be able to "opt out" or to "withhold" consent to either an online or an in-person trial, but that should not trouble us. Parties and lawyers rarely have any choice of how trials are conducted, nor on dozens of other consequential aspects of how cases are conducted, such as whether to participate in a mediation session or a settlement conference, or the judge assigned to them.[71]

Given the volume of human activity occurring online, it is silly for the legal profession to treat online adjudication as anathema. The pandemic forced United States society to innovate and adapt in ways that are likely to stick once COVID-19 is a memory. Courts should not think that they are immune from this trend. Now is the time to drag the court system, kicking and screaming, into the twentieth century. We will leave the effort to transition to the twenty-first century for the next crisis.

[68] Holly Fernandez-Lynch, D. James Greiner & I. Glenn Cohen, *Overcoming Obstacles to Experiments in Legal Practice*, 367 Science 1078 (2020); *see also* Michael Abramowicz, Ian Ayres & Yair Listokin, *Randomizing Law*, 159 U. Pa. L. Rev. 929 (2011).

[69] Charles Fried, Medical Experimentation: Personal Integrity and Social Policy (2nd ed. 2016).

[70] Harry Marks, The Progress of Experiment: Science and Therapeutic Reform in the United States, 1900–1990 (2000).

[71] Studies in which courts have randomized compulsory mediation sessions or settlement conferences are collected in D. James Greiner & Andrea Matthews, *Randomized Control Trials in the United States Legal Profession*, 12 Ann. Rev. L. & Soc. Sci. 295 (2016). Regarding randomization of assignments of judges, see the excellent Adam M. Samaha, *Randomization in Adjudication*, 51 Wm. & Mary L. Rev. 1 (2009).

https://doi.org/10.1017/9781009255301.007 Published online by Cambridge University Press