

When the Truth Becomes Fiction, Part III: Nobody Has Won With It Yet

The fear was that a fake would be believed. The sharper risk is that something real stops counting as proof. American courts have now heard that argument several times. It has lost every time, and the reason it lost is not reassuring.

Jennifer McKinney · August 16, 2026 · 9 min read

Gary Schildhorn told a Senate committee that it is fundamental, when we are harmed by somebody, that there be a remedy. Through the legal system, or through law enforcement. He said it after being told nobody could help him, because the synthetic voice that nearly took his money had not quite managed to take it.

He had spent a career inside that machinery. He found the place it stops.

This is about a different edge of the same problem. Not a remedy that does not reach. Evidence that stops settling anything.

The liar's dividend, introduced in court

In 2018 Walter Huang died when his Tesla, with Autopilot engaged, struck a highway barrier. His family sued. The case ran in the Superior Court of California, Santa Clara County, as number 19CV346663.

The plaintiffs wanted to use public statements Elon Musk had made in 2016 about what Autopilot could do. Recorded, at public events, years before the crash.

Tesla's response was that the recordings may or may not be authentic, because Musk, like many public figures, is the subject of deepfake videos and audio purporting to show him saying things he never said.

Not that these particular recordings were fake. That they were the sort of thing that can be fake, and so the company could not be asked to admit he said it.

Judge Evette Pennypacker issued a tentative ruling on 27 April 2023 ([tentative ruling](#)).

Mr. Musk, and others in his position, can simply say whatever they like in the public domain, then hide behind the potential for their recorded statements being a deep fake to avoid taking ownership of what they did actually say and do. The Court is unwilling to set such a precedent by condoning Tesla's approach here.

She ordered Musk deposed.

The argument has a name. Robert Chesney and Danielle Citron called it the liar's dividend in the *California Law Review* in 2019, before the tools were good. Their formulation was that deepfakes make it easier for liars to avoid accountability for things that are in fact true ([full text](#)).

The dividend is not paid for making a convincing fake. It is paid for the fact that convincing fakes exist.

The scoreboard so far

The question is how often it has worked. Not how often commentators warn about it. How often a court has excluded, discounted or disbelieved genuine evidence because someone raised the possibility of AI.

The record is small and it is searchable ([CourtListener](#)).

Case	Outcome	What was argued and what happened
Huang v. Tesla	Failed	Superior Court of California, Santa Clara County, no. 19CV346663. Tesla answered a request to admit Musk's recorded 2016 statements by arguing they may or may not be authentic, since he is the subject of many deepfake videos. Judge Pennypacker rejected it and ordered him deposed.

United States v. Reffitt	Failed	US District Court, District of Columbia, no. 1:21-cr-00032. At the first January 6 trial, defense counsel raised on cross-examination whether prosecution video could have been AI-manipulated, without offering any factual basis. Convicted on all five counts in March 2022.
Wisconsin v. Rittenhouse	Failed	Kenosha County Circuit Court. The defense objected to pinch-to-zoom on already-admitted drone footage, arguing Apple's software uses AI to construct what it thinks is there. The judge put the burden on the prosecution, which could not produce an expert on short notice. An objection to a display function, not to a deepfake.
United States v. Khalilian	Failed	US District Court, District of Nevada, no. 2:23-cr-00222. The defense moved to exclude a voice recording on the ground it could be a deepfake. The court indicated a witness familiar with the voice would likely suffice to authenticate. Not excluded.

Valenti v. Dfinity USA Research	Backfired	US District Court, Northern District of California, no. 3:21-cv-06118. Facing a disqualification motion built on video, plaintiff's counsel offered an expert report alleging the videos were deepfaked. The court read the allegation as evidence the firm was protecting its own interests over the class's.
Washington v. Puloka	Excluded	King County Superior Court, no. 21-1-04851-2. The defense sought to admit an AI-upscaled enhancement of genuine bystander video. The court excluded the enhancement as opaque, not peer reviewed, and not reproducible. AI skepticism protected the record rather than eroding it.
Mendones v. Cushman & Wakefield	Succeeded	Superior Court of California, Alameda County, no. 23CV028772. The inverse case. Self-represented plaintiffs submitted video exhibits with a summary judgment motion. Judge Kolakowski found them AI-generated, called them deepfakes, and dismissed with prejudice in September 2025.

Compiled from opinion and docket search across federal and state courts, government publishing, and the federal evidence rulemaking record. Trial-level records are unevenly digitised, so this is a floor rather than a certified ceiling.

The pattern is consistent. In every case where someone pointed at authentic material and said it might be synthetic, they lost, or the claim rebounded on them.

The one case where an AI argument decided something went the other way entirely. Judge Victoria Kolakowski found the exhibits were generated, called them deepfakes in the order, and dismissed the case with prejudice.

So the courtroom scoreboard, so far, reads: fabricated evidence submitted and caught, once. Real evidence successfully dismissed as fabricated, never.

Why the defense keeps failing, and why that may not last

There is a reason, it is measurable, and it complicates the argument I am making.

In February 2025 the *American Political Science Review* published five preregistered experiments by Kaylyn Jackson Schiff, Daniel Schiff and Natália Bueno, run on more than fifteen thousand American adults. The question was whether politicians can escape accountability by claiming that evidence of a scandal is misinformation ([open access](#)).

Against text, it works. Against video, it mostly does not.

Denial target	Effect size (standard deviations)
Text report, claiming uncertainty	0.17
Text report, rallying supporters	0.21
Video evidence	0

The authors' funding statement credits Emory, Georgia Tech, Yale and Purdue, and they declare no conflicts of interest. One of four video conditions produced a small non-null effect. The rest did not.

People still believe video. The liar's dividend is real and it is currently cashing mostly against transcripts and documents, not against footage.

The finding cuts against the alarm and it has an expiry date nobody can read. The experiments measured how audiences responded in 2022 and 2023, when most people had not been fooled by a synthetic video of anyone they knew. The result describes a public that still has the reflex to believe its eyes.

A judge refusing to let the possibility of fakery excuse the actual record. A public that has not yet learned to doubt what it watches. Both are holding. Neither is guaranteed.

What the rulemakers are doing about it

The federal judiciary is not waiting for the case law to arrive.

Stage	Action	Detail
November 2024	Advisory Committee considers a deepfake-specific authentication rule and declines to advance it	The proposal would have shifted the burden to the proponent once an opponent showed a reasonable person could find the item fabricated
May 2025	Committee votes 8 to 1 to publish proposed Rule 707 for comment	The dissent came from the Department of Justice
June 2025	Standing Committee approves publication	

February 2026	Public comment period closes	Submissions from the Department of Justice, trial lawyers, bar associations, civil liberties groups, academics, and at least one company selling verification infrastructure
Pending	No transmission to the Judicial Conference or the Supreme Court	Earliest plausible effective date under the standard cycle is December 2027

Sources are the Advisory Committee on Evidence Rules agenda books for [November 2025](#) and [May 2026](#), and the public comment docket. Proposed Rule 707 would apply the reliability standard already used for expert testimony to machine-generated evidence offered without an expert witness.

Rule 707 addresses evidence everyone agrees came from a machine. If a machine produces an expert-like conclusion, the conclusion faces the same reliability test a human expert would face. Sensible, and overdue.

It does not touch the problem in this piece. A deepfake does not announce itself as machine output. It arrives claiming to be a recording of something that happened. The rule that would have reached it is the authentication amendment the Committee set aside in November 2024.

Rule 707	Reached	Scope
What it covers	Yes	Evidence openly produced by a machine. Algorithmic output offered without an expert witness. Expert-like conclusions from software.
What it does not cover	No	A deepfake presented as a genuine recording. A challenge to real footage as synthetic. Authentication of contested audiovisual evidence.

The gap on the right is what the shelved November 2024 authentication amendment would have addressed.

Louisiana moved first among the states. An act effective 1 August 2025 requires attorneys to exercise reasonable diligence in verifying evidence before offering it, and creates a pretrial process for raising authenticity concerns about AI-altered material.

A reprieve, not an outcome

The deepfake defense has not worked. Judges have rejected it. Audiences still trust video. The scoreboard is clean.

That is a reprieve, not an outcome.

The reason the defense fails is that a claim of fakery still sounds like an excuse. It sounds like an excuse because most people have never been fooled by a synthetic recording of someone they know. Schildhorn has. He was persuaded by his own son's cadence, and once you have had that experience the sentence "that video could be fake" stops sounding like a dodge and starts sounding like a description of the world.

Courts are made of people who watch the same footage as everyone else. The doubt does not need to be planted case by case. It arrives on its own, from ordinary life, and it arrives at the

same rate the technology reaches ordinary people.

Nobody has won with the argument yet. That is a fact about the present tense.

Sources

- [Tentative ruling of Judge Evette D. Pennypacker, Huang v. Tesla, 27 April 2023](#)
- [Chesney and Citron, Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, 107 California Law Review 1753 \(2019\)](#)
- [Schiff, Schiff and Bueno, The Liar's Dividend, American Political Science Review 119\(1\), February 2025](#)
- [Advisory Committee on Evidence Rules agenda book, November 2025](#)
- [Advisory Committee on Evidence Rules agenda book, May 2026](#)
- [CourtListener opinion and docket search](#)