

Discredit by Saturation:

Topic Level Credibility Penalty as a Concealment Mechanism

Travis Gilly

ORCID: 0009-0007-2954-6313

Real Safety AI Foundation

Illinois, United States

♣ August 2026 (v0.2)

ABSTRACT

Some subjects acquire a reputation in public discourse that attaches to anyone who raises them. This paper describes a credibility penalty that operates at the level of topic rather than claim: a speaker who invokes a particular subject is discounted for having invoked it, before any assessment of what was said. The penalty is distinct from ordinary skepticism, which evaluates a claim and finds it wanting, and distinct from the canonical account of testimonial injustice, which locates the deficit in prejudice about the speaker's social identity. Here the sorting is done by subject matter, which means a documented conviction and an unfounded assertion drawn from the same subject receive the same reception. Where the subject has a substantial adjudicated core, the effect functions as concealment, because it suppresses the ordinary follow up that an adjudicated record normally invites, and it does so without requiring that anyone intend it. The paper works three cases: the federal proceedings arising from the conduct of Jeffrey Epstein and his associates, the MKULTRA program, and COINTELPRO. Each has a substantial adjudicated core, each is saturated to the point where invoking it carries a credibility cost, and in each the adjudicated record met a threshold ruling. Two of the three closed. In the third the door opened, and it opened through a doctrine that treats official concealment as a reason to extend a limitations period rather than as a reason to bar a claim. That difference is the paper's comparative finding: the mechanism is not inevitable, and what varies is whether the threshold doctrine encountered treats concealment as a bar or as a tolling event. The paper confines itself throughout to convictions, appellate holdings, and unsealed filings, specifies in advance what would show the effect does not exist, and declines to resolve whether the saturation is emergent or cultivated, because the mechanism operates either way.

Keywords: topic level credibility penalty; discredit by saturation; testimonial injustice; epistemic injustice; credibility deficit; persuasion knowledge; concealment without a concealer; child sexual abuse; adjudicated record; threshold foreclosure; institutional attention; falsifiability; chilling effects

I. INTRODUCTION

There is a familiar move in argument where a speaker raises a subject and the room does not evaluate the claim. It evaluates the speaker for having raised the subject. The claim may be true, false, documented, or invented. None of that is reached, because the sorting has already happened at the level of what was mentioned.

This paper is about that move, about the conditions under which a subject acquires the property that triggers it, and about what follows when the subject in question has a large adjudicated core.

The last condition is what makes the phenomenon worth analysis rather than complaint. Where a subject is wholly fabricated, a penalty attaching to anyone who raises it is a rough but serviceable heuristic. Where a subject has produced federal convictions affirmed on appeal, a penalty attaching to anyone who raises it suppresses inquiry into matters a court has already found to be true. The heuristic is then doing the opposite of its job, and it is doing it invisibly, because a filter that operates before evaluation leaves no trace of the evaluation it prevented.

Section II states how the cases were selected and what the paper excludes. Section III sets out the adjudicated record in the primary case. Section IV defines the penalty and distinguishes it from ordinary skepticism and from the canonical account of testimonial injustice. Section V describes the mechanism. Section VI argues that the penalty performs a concealment function without requiring a concealer, and compares the primary case against two others in which the same conditions held and the outcomes diverged. Section VII connects the argument to a companion analysis of the same credibility stock depleted from the opposite direction. Sections VIII and IX state the limitations and the empirical work that remains. Section X concludes.

A methodological commitment governs the whole. The paper does not resolve whether the saturation of its worked subject is emergent or cultivated. It does not need to, because the mechanism it describes produces the same effect under either hypothesis. A paper that reached for the cultivated answer would be performing the move it is trying to analyze.

II. METHODOLOGY

Three cases are used. A reader is entitled to ask how they were chosen and whether a different selection would have produced a different conclusion.

The criteria were fixed before the cases were, and there are three. The subject must have a substantial adjudicated core, so that the reception of well evidenced claims within it can be observed rather than assumed. The subject must be saturated in public discourse over a sustained period, to the point where invoking it carries a credibility cost independent of what is asserted. And the subject must have produced at least one threshold ruling, meaning a decision resolving whether a claim or a record could be reached at all, since a threshold ruling is where foreclosure becomes visible on the face of a published document rather than inferred from silence.

Three subjects satisfy all three criteria. The federal proceedings arising from the conduct of Jeffrey

Epstein and his associates. The MKULTRA program. And COINTELPRO. Each is treated below only through its adjudicated record.

The third criterion is what makes the set analytically useful rather than merely illustrative. Two of the three threshold rulings closed the door and one opened it, which means the set contains its own negative case. A comparison in which every instance confirms the thesis would be pattern matching. This one does not.

Candidate subjects that satisfy only two criteria were excluded and the exclusions are recorded. Institutional child abuse in religious organizations has the adjudicated core and the threshold rulings but is not saturated in the relevant sense, since raising it does not ordinarily discredit the speaker. Vaccine safety is saturated but lacks an adjudicated core on the contested claims. The Tuskegee study has the adjudicated core and produced a settlement, but invoking it confers credibility rather than costing it, which makes it a useful future control rather than a case.

Three cases are not a sample. The conclusions are stated as conditional on the criteria rather than as generalizations, and Section IX sets out the design that would test them at scale.

Two exclusions are absolute across all three cases. The paper makes no claim about any person who has not been adjudicated, and no claim about the contents of any record that has not been unsealed. This is not caution for its own sake. The argument in Section VI depends on the reader accepting that a substantial core of each subject is uncontested, and that acceptance is unavailable to a paper that has speculated past the record.

III. THE ADJUDICATED RECORD

This section states only what courts have adjudicated.

In 2007 the United States Attorney's Office for the Southern District of Florida entered a non prosecution agreement in the matter. The existence, provenance, and scope of that agreement were later established in litigation over whether it barred a prosecution in another district (*United States v. Maxwell*, 2024). Federal charges were never filed in the Southern District of Florida, and the Eleventh Circuit stated the point directly: because the government never filed charges, there was no preexisting proceeding in which a victim could have moved for relief (*In re Wild*, 2021).

Victims of the underlying conduct sought judicial relief in that district, and they won it. The district court granted partial summary judgment, holding that the government had violated the Crime Victims' Rights Act by negotiating the non prosecution agreement without conferring with them, that rights under the Act attach before formal charges are brought, and that the government must notify victims of an intention to enter such an agreement before binding itself. The court identified as particularly problematic the government's decision to conceal the existence of the agreement and to mislead the victims into believing that federal prosecution remained a possibility while telling them to be patient during a supposedly continuing investigation (*Doe 1 v. United States*, 2019). Sitting en banc, the Eleventh Circuit held that the Act does not authorize a victim to seek judicial enforcement of her rights in a freestanding civil action, and permits a motion only

within a preexisting proceeding. Because no charges had been filed, there was no proceeding, and therefore no forum. That holding abrogated the district court's ruling. The court reasoned further that implying a pre charge private right of action would seriously impair prosecutorial discretion contrary to the Act's own terms (*In re Wild*, 2021). A dissenting opinion read the same statute to grant two pre charge rights, the reasonable right to confer with the attorney for the government and the right to be treated with fairness, together with a private remedy where no prosecution is underway (*In re Wild*, 2021, Branch, J., dissenting).

In the Southern District of New York, Ghislaine Maxwell was convicted of conspiracy to transport minors with intent to engage in criminal sexual activity, of transportation of a minor with intent to engage in criminal sexual activity, and of sex trafficking of a minor. The Second Circuit affirmed, holding that the 2007 agreement did not bind prosecutors in a district that was not party to it, and that the extended limitations period for offenses involving the sexual abuse of a child applies to pre enactment conduct (*United States v. Maxwell*, 2024).

That is the floor. A jury convicted, a court of appeals affirmed, a district court found that the government had violated a victims' rights statute and had concealed the agreement while telling the victims to wait, and a court of appeals sitting en banc then held that no forum existed in which that finding could be made, because no one had charged the principal. Nothing in the remainder of this paper depends on any fact beyond those.

IV. THE PENALTY

A. Statement

Call the effect a topic level credibility penalty. It has three properties.

It is pre evaluative. It resolves before the content of the claim is assessed, and therefore does not depend on the content being weak.

It sorts by subject rather than by claim. Two assertions differing only in subject matter, matched for evidentiary quality, receive different reception. Correspondingly, two assertions within the same subject that differ sharply in evidentiary quality receive similar reception, since the sorting has already occurred at the level they share.

It attaches to the speaker by transfer. The penalty is incurred by raising the subject, which means it operates as a cost on speech acts rather than as a judgment about propositions, and a speaker anticipating it may decline to raise the subject at all. The anticipatory element is the standard signature of a chilling effect, on which there is a developed theoretical and empirical literature distinguishing a speech paradigm, a privacy and autonomy paradigm, and a collectivist paradigm (Penney, 2019).

The taxonomy this paper needs already exists in an unrelated field. The persuasion knowledge model distinguishes three knowledge structures whose interaction determines how a target receives a persuasion attempt: persuasion knowledge, agent knowledge, and topic knowledge (Friestad & Wright, 1994; Isaac & Calder, 2024). The claim here is that topic knowledge can come to dominate the other two, so that what an audience believes about a subject determines its response before anything about the speaker or the message is processed.

B. Distinguished from ordinary skepticism

Skepticism is claim level and post evaluative. A skeptic hears the claim, weighs the evidence, and discounts. That process functions correctly when weakly evidenced claims are discounted and well evidenced claims are not, and it is self correcting in the presence of better evidence.

The penalty described here is not self correcting, precisely because better evidence does not reach the stage at which it would operate. Empirical work applying the persuasion knowledge model finds that the dominant coping outcome is avoidance of the persuasion attempt rather than engagement with it, driven by beliefs about the tactic rather than by the content of the message (Mohr & Köhl, 2021). Avoidance is the behavioral form a pre evaluative filter takes.

This is why an adjudicated core does not dissolve the effect. Conviction and affirmance are evidence of the highest ordinary quality, and they are exactly the sort of evidence that cannot repair a sorting that happens before evidence is consulted.

C. Distinguished from testimonial injustice on the canonical account

Fricker's account of testimonial injustice locates the wrong in a credibility deficit caused by identity prejudice, and treats the systematicity of the deficit as following from the fact that identity prejudice tracks a person across contexts (Almagro, Osorio, & Villanueva Fernández, 2021).

The effect described here does not fit that account, and the misfit is informative rather than fatal. The deficit here is caused by the subject, not by the speaker's social identity, and it therefore does not track the speaker across contexts. The same speaker suffers no deficit on an unrelated subject. What is systematic is the topic, not the person.

Two departures from the canonical account make room for this. Voigt argues that we can fail in duties of testimonial justice in ways that do not wrong any particular individual (Voigt, 2017). Perrine argues that prejudice is not necessary for the harm, and proposes that testimonial injustice occurs where a speaker's capacity as a giver of knowledge is interfered with in important ways (Perrine, 2023). On Perrine's proposal the effect described here qualifies without modification, because interference with the capacity to give knowledge is exactly what a pre evaluative topic filter accomplishes.

A further complication is worth naming rather than avoiding. Almagro and colleagues describe cases in which public figures cultivate a persona for whom the conditions of testimonial injustice appear to apply, and exploit the appearance strategically (Almagro et al., 2021). A speaker who raises a saturated subject and is discounted can claim the mantle of the unjustly disbelieved, and some will do so in bad faith. That is a real hazard for this paper's argument. It is not a refutation, because the existence of parties who exploit an effect presupposes that the effect exists.

V. MECHANISM

How does a subject acquire the property?

The first stage is volume without resolution. A subject that generates sustained public attention and no terminal institutional answer accumulates commentary at a rate that outpaces adjudication. The ratio of assertion to adjudication rises. Since most of what circulates on any high volume subject is low quality, the observable average quality of speech about the subject falls, and falls further as the gap persists.

An adjacent literature on child maltreatment allegations describes the institutional consequence directly. Where a category of allegation acquires a reputation for tactical deployment, investigating bodies begin treating incoming reports in that category as presumptively suspect, and valid cases are discarded before they are assessed (Saini, Laajasalo, & Platt, 2020). That is the same sorting this paper describes, observed inside an institution rather than in public discourse, which establishes that the mechanism is not confined to lay audiences.

The second stage is the formation of a lay heuristic. Observers do not track quality claim by claim; they form a summary judgment about the class. The summary judgment is not irrational. It is a reasonable response to a genuinely degraded average, and it economizes on attention, which is the function heuristics serve.

The third stage is transfer. The summary judgment about the class becomes a judgment about anyone who invokes the class. This is the persuasion knowledge move: an audience deploys accumulated beliefs about an agent's goals and tactics to interpret and evaluate a persuasion attempt (Friestad & Wright, 1994), and evaluates the inferred motive rather than the message. The move is ordinarily protective, since it is what allows an audience to discount advertising. It becomes destructive when the inferred motive is derived from the subject rather than from anything the particular speaker has done.

The fourth stage is the trap that gives the paper its title. An adjudicated core keeps the subject in circulation. A wholly fabricated subject eventually exhausts itself, because nothing new attaches to it and attention migrates. A subject with real convictions, real victims, and real unresolved institutional questions cannot exhaust itself, because the underlying facts keep producing occasions to speak. So the subject stays saturated, the ratio of assertion to adjudication stays high, and the heuristic is continually reinforced by exactly the property that makes the heuristic wrong.

That is the structure. The factual core sustains the saturation, the saturation sustains the penalty, and the penalty suppresses inquiry into the factual core.

The loop distinguishes this effect from the canonical identity based case in one further respect. An identity prejudice can in principle be corrected by evidence about the group, and the remedy Fricker's account suggests is the cultivation of a virtuous hearer who corrects for prejudice in credibility judgments (Voigt, 2017). A topic filter of this kind is reinforced by evidence about the topic, because every new adjudicated fact produces new occasions to speak, most of which will be low quality for the same reasons the original ratio was unfavorable. Better evidence therefore strengthens the filter rather than weakening it. Correcting for a prejudice one holds about a group is tractable in a way that correcting for a base rate one has correctly observed about a topic is not.

VI. THE CONCEALMENT FUNCTION

A. *Concealment without a concealer*

A concealment mechanism, in the sense used here, is any regular feature of an information environment that reduces the probability of inquiry into a matter below what the matter's evidentiary standing would otherwise support. Nothing in the definition requires an agent, an intention, or a plan. A mechanism can be produced entirely by the aggregate of individually reasonable responses, and it will still have the effect.

This is where the refusal to adjudicate between the emergent and cultivated hypotheses does analytic work rather than merely avoiding trouble. If the saturation is emergent, the mechanism is a collective action failure. If it is cultivated, the mechanism is a strategy. The predicted observable is identical in both cases, which means the observable cannot discriminate between them, which means any paper claiming to have discriminated is claiming more than its evidence supports.

B. *The structural instance*

The strongest evidence for the concealment function is not anecdotal, and should not be sought in accounts of individual conversations, which are unfalsifiable in the way Section VIII warns against.

It is structural, and Section III already stated it. A district court found that the government had violated the victims' statutory rights and had concealed the agreement from them (*Doe I v. United States*, 2019). An en banc federal court of appeals then held that no judicial forum existed in which that question could be litigated at all, for the reason that the government had declined to charge the principal (*In re Wild*, 2021). The foreclosure therefore did not fall on an untested allegation. It fell on a finding a court had already made. The holding rested on a construction of a statute Congress wrote to protect crime victims, and it produced a result in which the protection is unavailable precisely where the executive has chosen not to proceed. A dissenting opinion read the same text the other way (*In re Wild*, 2021, Branch, J., dissenting).

That is an ordinary and legitimate outcome of statutory construction, and this paper takes no position on whether it was correctly decided. The observation is about what followed. A holding of that shape, on a statute of that purpose, with a dissent of that kind, is the sort of thing that ordinarily generates sustained legislative attention, a corrective amendment, or at minimum a durable scholarly literature on the gap it opened. The comparison the paper invites is between the institutional response this holding received and the response received by comparable gaps identified in statutes of comparable purpose in subjects that are not saturated. That comparison is proposed as a study in Section IX, and it is stated there as a test the thesis should be made to pass rather than as a result.

C. *Two further cases, and one that went the other way*

If the mechanism is real it should appear outside the primary case, and it should be possible to say what distinguishes an instance where it operated from one where it did not.

MKULTRA is the closest parallel. The program's existence, scope, and character are not contested: they were established by congressional investigation, and the Supreme Court itself recorded that the research

concerned the use of drugs and other biological, chemical, or physical agents in warfare or intelligence operations, and that the Agency had destroyed records. When a requester sought the names of the researchers and the institutions that housed them under the Freedom of Information Act, the Court held that section 102(d)(3) of the National Security Act qualifies as a withholding statute under Exemption 3, that the Director's authority extends to all sources of intelligence information rather than only those promised confidentiality, that MKULTRA researchers are intelligence sources within that definition, and that even superficially innocuous information such as institutional affiliation may be withheld where disclosure might permit an observer to deduce a source's identity (*Central Intelligence Agency v. Sims*, 1985). Justice Marshall, concurring, would have confined the term to sources providing information under an express or implied promise of confidentiality; the majority rejected that limitation (*Sims*, 1985, Marshall, J., concurring in the judgment). The Court expressly did not reach whether the destruction of the files had been lawful.

The structure matches the primary case at every joint. An adjudicated core. A saturated subject, to the point where the program's name now functions in ordinary speech as a marker of unseriousness. A threshold ruling that resolved access rather than merits. A separate opinion reading the same text narrowly. And a question the court noted it was not deciding, which no other body took up.

COINTELPRO is the negative case, and it is the more instructive of the two. The program satisfies the first two criteria as fully as either of the others. Its threshold ruling came out the other way. A district court held that the FBI's disruption operations against a lawful political party were patently unconstitutional and without statutory or regulatory authority, that surreptitious entries to obtain documents from the party's offices were obvious Fourth Amendment violations, and that the use of informants to gather private information about lawful political meetings was incompatible with the First Amendment, and it awarded damages under the Federal Tort Claims Act (*Socialist Workers Party v. Attorney General*, 1986).

The plaintiffs reached the merits because of how the court resolved a timeliness question. The two year administrative filing period did not begin until they had discovered, or with reasonable diligence should have discovered, the critical facts of both the injury and its cause, and where a defendant deliberately conceals material facts relating to its wrongdoing the clock does not start until the basis of the lawsuit is or should be known (*Socialist Workers Party*, 1986).

That is the comparative finding. In all three cases the underlying conduct was concealed, and in all three the concealment was the reason the claim or the request arrived when it did. In the primary case the concealment was not merely alleged but found, and found by the court whose ruling the en banc court then displaced (*Doe 1 v. United States*, 2019). In the first two, the threshold doctrine encountered treated the consequences of concealment as a bar: no charges meant no proceeding and so no forum, and the need to protect sources meant no names and so no record. In the third, the threshold doctrine treated concealment as a reason to extend the plaintiff's time rather than to end it.

The mechanism described in this paper is therefore not inevitable, and what varies is not the strength of the underlying record or the degree of public saturation. It is whether the doctrine standing at the threshold reads official concealment as something the claimant must bear or as something the concealing party must answer for. That is a property of the doctrine rather than of the subject, which is why saturation alone

does not determine the outcome and why the effect this paper describes operates on the margin rather than absolutely.

D. What the penalty forecloses

Four things, in rough order of tractability. Journalistic follow up, because an editor allocating scarce attention discounts a story by the reception its subject will receive. Discovery and unsealing motions, because the political cost of filing them is a function of how the filing will be characterized. Legislative inquiry, for the same reason at higher amplitude. And the ordinary willingness of an institution to answer a question, which is the least visible of the four and probably the most consequential, because an institution that anticipates the questioner will be discounted has little incentive to answer.

Each of the four operates through anticipation rather than through any completed act of suppression, which is why the effect leaves so little evidence of itself. The unwritten story, the unfiled motion, and the unasked question are not events, and an information environment cannot be audited for things that did not happen.

VII. THE SAME STOCK, DRAWN FROM TWO SIDES

A companion analysis by the present author describes a credibility stock attaching to allegations of child sexual exploitation, held in common by everyone who will ever need to make such an allegation truthfully, and depleted by official invocation of the allegation without a predicate. The mechanism there is overuse in pretext, and the injured population is diffuse, future, and unidentifiable, which is why no plaintiff can reach it. That pattern is documented independently: allegations concerning child pornography and child abuse have been invoked to justify expanded information gathering powers directed at particular communities, in circumstances where the allegation was doing work unrelated to the protection of any child (Mokrosinska, 2014).

The mechanism here is overuse in discourse. The two are not analogies. They are two withdrawals against one account.

The convergence matters for a reason beyond symmetry. Each argument, standing alone, is vulnerable to the objection that the credibility stock is a metaphor rather than a thing. Two independent depletion pathways producing the same predicted degradation, identified from unrelated starting points, is the ordinary form of evidence that the underlying quantity is real. Together they license a claim neither licenses alone: that the reception of child sexual abuse allegations is a shared resource with observable dynamics, and that it is currently being drawn down from both the official and the discursive side at once.

VIII. LIMITATIONS

Four results would show that the effect described here does not exist, or does not exist in the form claimed. They are stated as commitments rather than as concessions, because a paper of this kind is vulnerable to a specific failure: treating every objection as an instance of the phenomenon it describes, and thereby

becoming unfalsifiable.

If matched claims differing only in subject label receive statistically indistinguishable credibility ratings, the topic level sorting is not occurring and the paper is wrong.

If the discount tracks evidentiary quality within the saturated subject, so that adjudicated claims are received better than unadjudicated ones by roughly the margin observed in unsaturated subjects, then what is operating is ordinary skepticism with a low prior rather than a pre evaluative filter.

If the discount attaches to the speaker across unrelated subjects, then the mechanism is speaker level, the canonical account of testimonial injustice already covers it, and this paper adds nothing.

If institutional follow up on adjudicated findings within the saturated subject proceeds at rates comparable to follow up on adjudicated findings of similar gravity in unsaturated subjects, then the concealment function claimed in Section VI does not obtain, whatever is happening in individual conversations.

Four further limitations bear on scope.

Three cases are not a sample. They were selected against criteria stated in Section II, which is a defensible procedure and not a random one, and the set includes its own negative case, which is better than a set that does not. It remains too small to support a rate.

The comparison in Section VI.B holds subject matter and adjudicated standing roughly constant and lets the threshold doctrine vary, but the three cases also differ in era, in the identity of the concealing party, in the legal vehicle used, and in the remedy sought. Any of those could carry the difference the paper attributes to the doctrine. The design in Section IX is the way to separate them and this paper does not separate them.

The measurement problem is severe and is general to chilling effects research, where the difficulty of documenting the effect has sustained long standing skepticism among courts and scholars (Penney, 2019).

And the paper deliberately leaves the emergent and cultivated hypotheses unresolved, which means it cannot support any claim about responsibility for the condition it describes.

IX. FUTURE WORK

Each item here answers a limitation stated in Section VIII rather than extending the argument.

The small set and the confounds in the three case comparison are answered by the same study. Build a corpus of threshold rulings that resolved access to an adjudicated record or to a forum, code each for the gravity of the underlying finding, the purpose of the statute construed, the presence of a separate opinion reading the text the other way, and the era and identity of the party resisting disclosure. Code the subject matter for saturation on a measure independent of the ruling itself, such as the ratio of commentary to adjudication over a fixed window. Then compare rates of legislative follow up, corrective amendment, and scholarly treatment across the saturated and unsaturated groups. That design holds the confounds this paper cannot separate, and it measures events rather than absences, which is the response to the measurement problem.

The pre evaluative sorting claim is answered by an experiment. Present matched claims to matched

samples, varying only the subject label, and measure credibility ratings. This tests the first and second disconfirming conditions directly and is the cleanest available design.

The Tuskegee study, excluded in Section II because invoking it confers credibility rather than costing it, is the natural control for both designs. It has the adjudicated core and the settlement without the saturation, which is the contrast the argument needs and which no case in the present set supplies.

X. CONCLUSION

This is a diagnostic paper. It does not propose a reform, and it should be read with suspicion if it did, because the interventions that suggest themselves are worse than the condition.

Exhorting audiences to suspend the heuristic will not work, and would not be advisable if it did, since the heuristic exists because the average quality of speech on a saturated subject is in fact low. Elevating the credibility of speech within the saturated subject by fiat would import the failure mode the heuristic protects against. Neither correction survives contact with the reason the heuristic formed.

What the paper supports is narrower. Institutions can be evaluated on a criterion that does not require anyone to change their priors: whether follow up on an adjudicated finding proceeds at a rate explicable by the finding's evidentiary standing, or at a rate explicable by its subject. That is a measurable property of an institution, and it does not ask the public to be more credulous about anything. It also sidesteps the objection that the audience is behaving unreasonably, since it makes no claim about audiences at all. The objection has force: a summary judgment about a class whose observable average quality is genuinely low is not a prejudice in the sense the canonical account requires, and treating it as one would be a category error (Perrine, 2023).

The last observation is the one the paper exists to make. A subject can be simultaneously well documented and unspeakable, and the second condition can be produced by the first. Where that happens, the ordinary machinery by which a society converts findings into consequences stops turning, and it stops without anyone deciding to stop it. That is worth naming, and naming it is enough for one paper.

ACKNOWLEDGMENTS

The author discloses the use of speech-to-text software (Wispr Flow) and Anthropic’s Claude as assistive technology for a diagnosed neurodevelopmental condition. All intellectual contributions, research questions, theoretical arguments, methodological decisions, and conclusions are solely the author’s own.

Correspondence concerning this article should be addressed to Travis Gilly, Real Safety AI Foundation, Illinois, United States. Email: t.gilly@ai-literacy-labs.org.

REFERENCES

- Almagro, M., Osorio, J., & Villanueva Fernández, N. (2021). Weaponized testimonial injustice. *Las Torres de Lucca International Journal of Political Philosophy*, 10, 29–58. <https://doi.org/10.5209/1tdl.76461>
- Central Intelligence Agency v. Sims, 471 U.S. 159 (1985).
- Doe 1 v. United States, 359 F. Supp. 3d 1201 (S.D. Fla. 2019).
- Friestad, M., & Wright, P. (1994). The persuasion knowledge model: How people cope with persuasion attempts. *Journal of Consumer Research*, 21, 1–31. <https://doi.org/10.1086/209380>
- In re Wild, 994 F.3d 1244 (11th Cir. 2021) (en banc).
- Isaac, M. S., & Calder, B. J. (2024). Thirty years of persuasion knowledge research: From demonstrating effects to building theory to increasing applicability. *Consumer Psychology Review*, 8, 3–14. <https://doi.org/10.1002/arcp.1107>
- Mohr, S., & Kühn, R. (2021). Exploring persuasion knowledge in food advertising: An empirical analysis. *SN Business & Economics*, 1, 107. <https://doi.org/10.1007/s43546-021-00108-y>
- Mokrosinska, D. (2014). Privacy and the integrity of liberal politics: The case of governmental internet searches. *Journal of Social Philosophy*, 45, 369–389. <https://doi.org/10.1111/josp.12068>
- Penney, J. (2019). Chilling effects and transatlantic privacy. *European Law Journal*, 25, 122–139. <https://doi.org/10.1111/eulj.12315>
- Perrine, T. (2023). Prejudice, harming knowers, and testimonial injustice. *Logos & Episteme*, 14, 53–73. <https://doi.org/10.5840/logos-episteme20231413>
- Saini, M., Laajasalo, T., & Platt, S. (2020). Gatekeeping by allegations: An examination of verified, unfounded, and fabricated allegations of child maltreatment within the context of resist and refusal dynamics. *Family Court Review*, 58, 417–431. <https://doi.org/10.1111/fcre.12480>
- Socialist Workers Party v. Attorney General of the United States, 642 F. Supp. 1357 (S.D.N.Y. 1986).
- United States v. Maxwell, 118 F.4th 256 (2d Cir. 2024).
- Voigt, K. (2017). Testimonial injustice and speakers' duties. *Journal of Social Philosophy*, 48, 402–420. <https://doi.org/10.1111/josp.12210>