International Journal of Computer Applications (0975 - 8887)
Volume 187 - No.127, July 2026

The Deniability Asset: Claimed Ownership of AI-Assisted Work as a Consequence-Contingent Option

Francis Martinson
North Dakota State University
Fargo, ND, USA

Karl Pearson
Saint John's University
Minneapolis, MN, USA

Joseph Vinel
Miami University
Oxford, OH, USA

Andi Sina
Xavier University
Tampa, FL, USA

Conor Winters
Georgia Southern University
Statesboro, GA, USA

1. ABSTRACT

Knowledge workers increasingly produce documents, code, and analyses with assistance from generative artificial intelligence. A question that has received limited direct attention is how those workers describe their relationship to that output after the fact, and in particular how that description changes when the output is praised or criticized. This paper proposes that ownership of AI-assisted work is not a fixed property of the work itself but a strategic claim that flexes with consequences. Drawing on attribution theory [1, 2], research on moral crumple zones [3], and recent findings on disclosure penalties for AI use [5, 7], the paper argues that AI assistance gives the worker a deniability asset: an option that is held quietly while outcomes are favorable and exercised, by attributing the work to the machine, when outcomes turn unfavorable. The paper formalizes this option structure, identifies five moderating conditions, states six testable propositions, and specifies an experimental design capable of measuring the effect together with the pattern of results the model predicts. Analysis is extended across four deployment scenarios spanning software engineering, professional writing, financial analysis, and academic authorship, showing how the predicted magnitude of the effect varies with verifiability and domain identity. The contribution is conceptual. The intent is to give researchers in human-computer interaction, organizational behavior, and AI governance a precise vocabulary and a measurable construct for a behavior that is currently discussed mostly through anecdote.

2. INTRODUCTION

A software developer ships a feature written largely by a coding assistant. The feature is praised in review, and the developer accepts the praise without qualification. Three weeks later, a defect surfaces in the same feature, and the developer's account changes: the assistant generated that section, and the assistant is known to produce subtle errors. Both statements can be true. What changed between them is not the provenance of the code but the valence of its consequences.

This paper is about that switch. The claim developed here is that ownership of AI-assisted work behaves less like a property and more like a financial option. While outcomes are favorable, the worker holds the option and claims the work. When outcomes turn unfavorable, the worker exercises the option and assigns the work, or the failing part of it, to the machine. The option has a premium: disclosing AI involvement carries a reputational cost, documented in recent studies of disclosure penalties [5, 7], so the worker pays nothing to hold it and discloses only when disclosure buys protection from blame. The paper names this structure the deniability asset.

The phenomenon matters for three audiences. For organizations, it complicates accountability: if attribution claims are consequence-contingent, then post-incident accounts of who or what produced a piece of work are systematically unreliable in exactly the cases where reliability matters most. For AI governance, it bears on disclosure policy: mandatory disclosure regimes do not merely inform audiences, they extinguish the option, and workers can be expected to respond to that change in option value rather than to the informational content of the rule. For researchers, it offers a measurable construct at the intersection of attribution theory and human-AI interaction that has not yet been formalized.

Three boundaries on the claim should be stated immediately. First, the underlying psychology is not new, and this paper does not present it as new. Self-serving attribution, in which people internalize success and externalize failure, has been documented for half a century [1, 2]. The contribution is the analysis of what changes when the available external target is a machine collaborator whose actual contribution is invisible, plausible, and impossible for most audiences to verify. Second, the paper is conceptual. It proves nothing about prevalence and offers no field evidence. It states propositions and a design by which they can be tested. Third, exercising the deniability option is not always dishonest. When an AI system genuinely produced the flawed passage, attributing the flaw to the system is accurate. The construct of interest is the asymmetry: accurate attribution that surfaces only under criticism, having been available but withheld under praise.

3. PRIOR LITERATURE

The deniability asset sits at the junction of four research streams, each of which supplies a piece of the structure but none of which assembles it. Table **??** summarizes the contribution and the limitation of each stream.

Table 1. Contributing Literatures and the Structural Gap

Stream	Supplies	Does Not Address
Self-serving attribution [1, 2]	Motivation to externalize failure	An external target that cannot contest the account
Moral crumple zones [3]	Blame migration in human-machine systems	Blame flowing toward, not away from, the machine
Credit-blame in AI teaming [4, 6]	Machine as viable blame target	The within-person switch over time
Disclosure penalties [5, 7]	Cost of disclosing AI use	Disclosure as a strategically timed resource
Psychological ownership [10]	Routes by which felt ownership forms	Divergence of public claim from felt ownership under consequences
Moral disengagement [11]	Displacement of responsibility	A displacement target that cannot be wronged; the timing of displacement

3.1 Self-Serving Attribution

Miller and Ross reviewed the evidence that individuals attribute success to internal causes and failure to external causes, and subsequent work in attribution theory refined the conditions under which the asymmetry appears [1, 2]. This literature establishes the motivational engine. What it could not anticipate is an external attribution target that participated in the work, is present in every task, and whose contribution leaves no independently observable trace. A colleague blamed for a failure can object. A market condition blamed for a miss can be audited. A model blamed for a paragraph cannot contest the account, and in most organizations, no log exists that would.

3.2 Moral Crumple Zones

Elish introduced the moral crumple zone to describe how responsibility for the failure of an automated system is misattributed to a human operator who had limited control over the system [3]. The construct travels in the opposite direction from the one studied here: the human absorbs blame that belongs, in some causal sense, to the machine. The deniability asset is the reverse flow. The human deflects blame toward the machine, and the machine, unlike the operators in Elish's cases, has no standing to be wronged by the misattribution. Both constructs predict that attribution under failure will track power and incentive rather than causal contribution.

3.3 Credit and Blame in Human-AI Teaming

Experimental work in human-computer interaction has documented asymmetries in how people allocate credit and blame between themselves and machine partners. Hohenstein and Jung found that in AI-mediated communication, the AI functions as a lightning rod for blame when conversations go badly [4]. Jago showed that observers discount the authenticity of work once algorithmic involvement is revealed [5], and Dietvorst, Simmons, and Massey documented the speed with which people abandon algorithms after seeing them err [6]. These studies establish that the blame channel toward machines is open and that audiences respond to AI involvement with discounting. What they have not isolated is the within-person switch: the same worker, the same artifact, and an ownership claim that moves when the consequence arrives.

3.4 Disclosure Penalties

A growing body of work finds that disclosing AI assistance reduces evaluations of the discloser's competence and the work's worth, even when quality is held constant [5, 7]. This finding supplies the option premium. If disclosure were free, workers would disclose by default, and the deniability structure would collapse. Because disclosure is costly, silence under praise is the rational default, and disclosure becomes a resource spent deliberately, at the moment it purchases the most protection.

3.5 Psychological Ownership and Moral Disengagement

Two further literatures supply the psychological infrastructure of the asset. Pierce, Kostova, and Dirks developed the theory of psychological ownership, the state in which individuals feel that a target is theirs, and identified its routes: exercising control over the target, coming to know it intimately, and investing the self in it [10]. The deniability asset trades on the gap between this felt ownership and the public claim of ownership. The routes explain two of the moderating conditions introduced below. Work produced through invisible prompting supplies little experienced control and little self-investment, so felt ownership is weak and the public claim is cheap to abandon; work produced through visible editing travels the routes and anchors the claim. Bandura's account of moral disengagement identifies displacement of responsibility as a standing mechanism by which individuals avoid self-sanction for conduct they would otherwise censure [11]. Exercising the deniability option is a displacement move with unusually favorable terms: the recipient of the displaced responsibility is an artifact that cannot be wronged, cannot object, and already carries a public reputation for error [6]. Bandura's observation that disengagement often operates gradually and without full awareness also reinforces the behavioral definition adopted in Section 3: the option can be exercised sincerely.

3.6 The Gap

Each stream contributes one wall of the structure: motivation from attribution theory, the machine as blame target from crumple-zone and HCI research, the cost of disclosure from the penalty literature, felt-ownership anchoring from psychological ownership theory, and the displacement mechanism from moral disengagement. None formalizes the resulting temporal pattern, in which the worker's stated relationship to a fixed artifact is contingent on consequences that arrive after the artifact is complete. That formalization is the contribution of this paper.

4. THE DENIABILITY ASSET

4.1 Definition

The deniability asset is defined as the latent capacity, created by undisclosed AI assistance, to reassign authorship of work or of its failing components from the worker to the AI system at a moment of the worker's choosing. Three elements are essential.

First, the assistance must be real. The asset is not fabrication; the worker who falsely blames an AI system that contributed nothing is engaged in ordinary lying, which requires no new construct. Second, the assistance must be undisclosed or under-disclosed at the time the work is presented. Disclosure spends the asset. Third, the reassignment must be elective and consequence-contingent. The worker exercises it in response to criticism, not in response to new information about provenance, which the worker possessed all along.

One clarification bears on how the construct can be studied. Exercise of the option need not be consciously strategic. A worker who, under criticism, sincerely re-evaluates the machine's share of the work exhibits the same behavioral pattern as one who deploys the attribution deliberately, and motivated cognition is expected to produce exactly this convergence. The construct is therefore defined behaviorally, by the consequence-contingent movement of the claim, not by the mental state behind it. What separates the deniability asset from simple belief revision is the asymmetry of the trigger: the re-evaluation arrives with criticism and not with praise, although the information available to the worker is identical in both cases.

4.2 The Option Structure

The financial analogy is precise enough to be useful. Holding the option corresponds to presenting AI-assisted work without qualification. The premium paid for holding it is zero, while the premium avoided is the disclosure penalty: the discount audiences apply to work known to be machine-assisted [5, 7]. The strike event is the arrival of negative consequences. Exercising the option means disclosing, at that moment, the machine's role in the criticized component. The payoff is the transfer of some portion of blame from the worker to the system, a transfer the HCI literature suggests audiences are willing to accept [4]. Figure 1 traces the structure across the two branches.

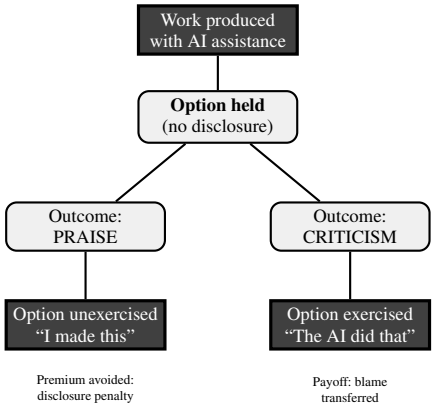

Fig. 1. The deniability option. Undisclosed AI assistance creates an option held through production. The outcome determines exercise: praise leaves it unexercised and the work claimed, criticism triggers exercise and reassignment to the machine.

Like financial options, the deniability asset decays. Each exercise teaches the audience that the worker's earlier silence was strategic, which raises the implied disclosure penalty on all of the worker's future work. An audience that has watched a colleague invoke the machine twice begins discounting that colleague's unqualified claims. The model therefore predicts restraint: workers should exercise the option for large consequences and absorb small criticisms silently, preserving the asset. Carried far enough, the decay inverts the instrument entirely. A worker known to reach for the machine under every criticism holds not an asset but a liability, since the audience now discounts the unqualified claims as well, and the worker has purchased short-term blame relief at the price of a permanent ownership discount.

4.3 Moderating Conditions

Five conditions are proposed to govern the asset's value and the likelihood of exercise, summarized in Table **??**.

Table 2. Moderating Conditions on Option Value

Condition	Raises Option Value	Lowers Option Value
Stakes	Career-threatening criticism	Routine, low-cost criticism
Domain identity	Craft-bound identity (writing)	Tool-assembled work (much of engineering)
Audience expertise	Non-expert evaluators	Evaluators who can probe the boundary
Verifiability	Finished artifacts, no logs	Version history, process records
Effort visibility	Invisible labor (prompting)	Visible labor (editing)

Stakes determine whether exercise is worthwhile: the option is exercised when the blame transferred exceeds the reputational cost of revealed strategic silence. Domain identity governs both felt ownership and the pain of disclosure; workers whose professional identity is bound to the craft of production should claim more and suffer more from disclosing. Audience expertise compresses option value from both sides, since expert audiences detect machine involvement unprompted and ask the follow-up questions that expose the boundary. Verifiability is decisive: where logs and version histories exist, claims about provenance can be checked and option

value falls toward zero. Effort visibility matters because the labor of directing generative systems is real but largely invisible [8], and invisible labor travels none of the routes by which psychological ownership forms [10], so it earns little felt ownership and lowers the resistance to abandoning the output.

5. SCENARIO ANALYSIS

The moderating conditions imply that the effect should vary in magnitude across settings rather than appearing uniformly. This section applies the model to four deployment scenarios, which serves both to illustrate the mechanism concretely and to generate differentiated predictions that an evaluation across settings could test. Table **??** summarizes.

Table 3. Predicted Option Value Across Four Scenarios

Scenario	Verifiability	Domain identity	Predicted option value
Software engineering	High (commits)	Low	Low–moderate
Professional writing	Low	High	High
Financial analysis	Moderate	Moderate	Moderate–high
Academic authorship	Low	Very high	Very high

Software engineering. Version control records what was written and when, and code review exposes the work to expert evaluators. Verifiability is high and domain identity comparatively low, since engineers already regard output as assembled from tools. The model predicts the smallest effect here, and predicts that where it appears it will concern design decisions rather than committed code.
Professional writing. Prose is typically delivered as a finished artifact with no process record, and writers' professional identity is bound to the craft of production. Verifiability is low and identity high, so the model predicts a large effect and a correspondingly painful exercise, since disclosure concedes the part of the work most closely tied to identity.
Financial analysis. Models and memoranda leave partial traces, and evaluating audiences are moderately expert. Stakes, however, are high, and the model predicts that stakes push exercise upward even where verifiability partially constrains it, producing a moderate to high effect concentrated in adverse outcomes.
Academic authorship. Manuscripts are delivered finished, disclosure norms are contested, identity investment is very high, and the reputational stakes of a flawed publication are severe. The model predicts the largest effect in this scenario, and predicts that mandatory disclosure policies will meet the strongest resistance here, consistent with P5 below.
The cross-scenario pattern is itself a prediction. If the deniability asset is real, effect sizes should order roughly as academic authorship, professional writing, financial analysis, and software engineering, and that ordering should track verifiability and domain identity rather than the amount of AI assistance actually used.

6. PROPOSITIONS

All six propositions are scoped at the level of the individual worker, and the scoping is deliberate. Team production introduces distinct dynamics, including diffusion of responsibility across members [11] and negotiated collective accounts of provenance, that deserve their own treatment; extending the option model to shared authorship is future work rather than an omission the present propositions are meant to cover. Within that boundary, the framework yields six propositions, stated so that each is falsifiable and collected in Table 6.

Table 4. Six Testable Propositions

ID	Construct	Prediction
P1	Option exercise	Ownership declines more steeply under criticism for AI-assisted than unassisted work
P2	Stakes	The P1 interaction strengthens as consequences grow
P3	Domain identity	The P1 interaction is larger in identity-central domains
P4	Audience expertise	The P1 interaction weakens when evaluators can verify provenance
P5	Mandated disclosure	Logging or mandated disclosure drives the P1 interaction toward zero
P6	Effort visibility	Visible human labor sustains ownership claims under criticism; invisible labor does not

P1 (Option exercise). Claimed ownership of AI-assisted work declines more steeply between praise and criticism conditions than claimed ownership of unassisted work. The interaction between assistance level and outcome valence, not either main effect, is the signature of the deniability asset.
P2 (Stakes). The interaction in P1 strengthens as the consequences attached to the evaluation grow.
P3 (Domain identity). The interaction in P1 is larger in domains central to the worker's professional identity than in domains regarded as instrumental.
P4 (Audience expertise). The interaction in P1 weakens when the worker believes the evaluating audience can independently assess how the work was produced.
P5 (Verifiability and mandated disclosure). Where AI involvement is logged, auditable, or disclosed by mandate, the interaction in P1 approaches zero, and worker resistance to such regimes correlates with the option value the regime would extinguish.
P6 (Effort visibility). Holding actual contribution constant, workers whose human contribution took visible forms claim more ownership under criticism than workers whose contribution took invisible forms such as prompting and curation.
P5 carries the main governance implication. Opposition to disclosure mandates is usually framed as a privacy or burden objection. The framework predicts a third source: mandates destroy option value, and the intensity of opposition should track how much value the worker stands to lose, which P2 and P3 make measurable.

7. MEASUREMENT AND PREDICTED RESULTS

7.1 Core Design

Participants complete a writing or analysis task under one of two conditions: minimal AI assistance or substantial AI assistance, with assistance levels controlled by the protocol rather than self-selected. The work is then evaluated, and participants are randomly assigned to receive standardized praise or standardized criticism. The outcome measure is claimed ownership, captured on a seven-point scale and through coded attribution language in a written

response to the evaluation, supplemented by behavioral measures such as willingness to attach one's name to the work in a portfolio. The design is a two-by-two: assistance level crossed with outcome valence. A baseline ownership measure, captured immediately after production and before any evaluation is delivered, anchors each participant's claim at the moment the option is created; the post-evaluation measure then records movement from that anchor rather than an unanchored report. The quantity of interest is the interaction term. A sample of roughly two hundred participants is adequately powered for a moderate interaction and is inexpensive to field.

7.2 Predicted Pattern of Results

The model makes a specific prediction about the shape of the data, not merely the existence of an effect. Table 7.2 gives illustrative cell means on a seven-point ownership scale, and Figure 2 plots the predicted interaction.

Table 5. Predicted Cell Means, Claimed Ownership (1–7)

Condition	Praise	Criticism	Drop
Low AI assistance	6.2	5.4	0.8
High AI assistance	6.0	3.3	2.7
Difference	0.2	2.1	**1.9**

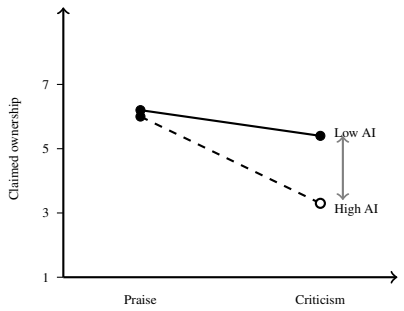

Fig. 2. Predicted interaction. Ownership claims are near-identical under praise regardless of assistance level, and diverge sharply under criticism. The gap at the criticism end, not either main effect, is the signature of the deniability option in exercise.

The critical feature is the convergence under praise and divergence under criticism. If both lines dropped in parallel, the data would show ordinary outcome bias, and the deniability construct would add nothing. If ownership were simply lower throughout the high-assistance condition, the data would show honest attribution rather than strategic timing. Only the interaction, in which the two conditions are indistinguishable under praise and separate under criticism, is diagnostic of an option held and then exercised. Controlled results should be read as an existence proof and a calibration of direction rather than as an estimate of field magnitude. The laboratory establishes whether the option is exercised when consequences are mild and reputations temporary; field settings raise both stakes and memory, which the moderating conditions predict will amplify exercise in low-verifiability domains and suppress it where process records exist. The cross-scenario evaluation of Section 6.3 is the designed bridge between the two, testing whether the ordering the model predicts for field settings emerges when the same protocol is run across professional populations.

7.3 Extensions Across Scenarios

The scenario analysis of Section 4 implies an evaluation across settings rather than a single study. Recruiting from the four professional populations and repeating the core design in each provides a direct test of the predicted ordering of effect sizes. A third assistance condition with mandated disclosure operationalizes P5: when the audience already knows the machine's role, the option does not exist, and the criticism condition should produce no additional attribution shift. Varying the stated stakes tests P2. Archival complements are also available: code review corpora, postmortem documents, and editorial correspondence can be coded for the timing of AI attribution relative to the valence of the surrounding discussion, and the model predicts that AI attributions cluster after negative turns in the discussion rather than being distributed evenly.

7.4 What Would Falsify the Framework

If ownership claims fall under criticism by similar amounts regardless of assistance level, the data show ordinary self-serving attribution and the deniability construct adds nothing. If the ordering of effect sizes across the four scenarios does not track verifiability and domain identity, the moderating account is wrong. If disclosure mandates fail to compress the interaction, the option model is wrong about the mechanism. The framework is offered as a hypothesis with these exit conditions, not as a settled account.

8. IMPLICATIONS

8.1 For Organizations

Post-incident accounts of AI involvement should be treated as systematically conditioned evidence. The time to capture provenance is at production, through logging and contemporaneous disclosure norms, not at failure, when every party's account is shaped by the consequences then on the table. Organizations that rely on after-the-fact testimony about who or what produced a failed artifact are sampling attribution at the moment of maximum distortion. The scenario analysis indicates where this exposure is greatest: settings that deliver finished artifacts without process records, which is to say most professional writing and analysis work.

8.2 For Disclosure Policy

The framework reframes disclosure mandates as option-extinguishing instruments. This predicts resistance that is rational rather than merely cultural, concentrated among workers in high-identity, low-verifiability roles. It also predicts a compensating behavior: where the deniability option is removed, workers will invest in other ownership signals, such as documented human revision passes, to recover the discount that disclosure imposes.

8.3 For the Study of Human-AI Collaboration

Research on human-AI teaming has concentrated on how the collaboration is performed. The deniability asset directs attention to how the collaboration is narrated afterward, and argues that the narration is an economic act with its own incentive structure. Self-reports of AI use, including those gathered by surveys and by employers, should be expected to vary with the respondent's recent outcomes, which is a measurable and correctable bias once it is named. The structure may also generalize beyond AI attribu-

tion: practitioners report analogous consequence-contingent commitment to analytical outputs generally, where confidence in what the data indicate flexes with whether the indication supports a preferred narrative, suggesting the option model may describe a family of attribution behaviors of which AI deniability is the newest member.

8.4 Relation to Content Classification

This analysis is a companion to frameworks that classify AI-generated content by its properties, including the authenticity spectrum proposed for synthetic media [9]. Classification frameworks describe the artifact. The present framework describes the human's claimed relationship to the artifact, and shows that the claim is a moving target. Both descriptions are necessary: governance that classifies content while taking ownership claims at face value will be gamed through the claims, not the content.

9. LIMITATIONS

The paper is conceptual, and its limitations follow from that. No evidence of prevalence is offered, and it is possible that in most settings workers report their AI use stably and honestly regardless of outcomes; the propositions exist to find out. The cell means in Table 7.2 are illustrative predictions rather than observed data, and are supplied to make the predicted pattern falsifiable, not to stand in for results. The option formalization is an analogy, and analogies mislead when pushed past their structure; the decay and premium components in particular are qualitative claims awaiting parameters. The four scenarios are reasoned applications of the moderating conditions rather than empirical cases. The design identifies the behavioral pattern of consequence-contingent attribution but cannot separate deliberate strategy from motivated sincere re-evaluation at the level of an individual, since the two produce the same observable movement; the baseline measure anchors the size of the movement but not the mental state behind it. The experimental design measures claimed ownership in controlled conditions, and laboratory attribution may understate field behavior, where stakes are real, or overstate it, where reputations are durable and audiences remember. Finally, the framework takes the disclosure penalty as given; if audience attitudes toward AI assistance normalize and the penalty fades, the option premium fades with it, and the deniability asset becomes a historical artifact of a transition period.

10. CONCLUSION

When work is produced with a machine whose contribution is invisible and unverifiable, the worker acquires something along with the output: a standing option to reassign authorship if the output disappoints. This paper has named that option the deniability asset, located it in fifty years of attribution research [1, 2] and a decade of human-AI interaction findings [3, 4, 5, 6], specified the five conditions that govern its value, applied it across four professional scenarios to generate differentiated predictions, and stated six propositions with an experimental design and a predicted result pattern sufficient to test the central one. Whether the asset is widely held and freely exercised is an empirical question, and the honest position is that the answer is not yet known. What the framework contributes now is precision: a vocabulary for a behavior that practitioners describe constantly and the literature has not yet pinned down, and a warning for governance, which is that any accountability regime built on what people say about their AI use after the consequences arrive is built on the most strategically contaminated evidence available.

Future work should proceed along three lines. First, execution of the core experiment and the cross-scenario evaluation described in Section 6, which would move the construct from proposal to measured effect. Second, longitudinal designs examining the decay property, testing whether repeated exercise raises the disclosure penalty a worker faces on later work. Third, policy studies examining whether disclosure mandates produce the compensating ownership signals the framework predicts, which would determine whether such mandates achieve transparency or merely relocate the strategy.

11. REFERENCES

[1] D. T. Miller and M. Ross, "Self-serving biases in the attribution of causality: Fact or fiction?" *Psychological Bulletin*, vol. 82, no. 2, pp. 213–225, 1975.

[2] B. Weiner, "An attributional theory of achievement motivation and emotion," *Psychological Review*, vol. 92, no. 4, pp. 548–573, 1985.

[3] M. C. Elish, "Moral crumple zones: Cautionary tales in human-robot interaction," *Engaging Science, Technology, and Society*, vol. 5, pp. 40–60, 2019.

[4] J. Hohenstein and M. Jung, "AI as a moral crumple zone: The effects of AI-mediated communication on attribution and trust," *Computers in Human Behavior*, vol. 106, 106190, 2020.

[5] A. S. Jago, "Algorithms and authenticity," *Academy of Management Discoveries*, vol. 5, no. 1, pp. 38–56, 2019.

[6] B. J. Dietvorst, J. P. Simmons, and C. Massey, "Algorithm aversion: People erroneously avoid algorithms after seeing them err," *Journal of Experimental Psychology: General*, vol. 144, no. 1, pp. 114–126, 2015.

[7] C. Longoni, A. Bonezzi, and C. K. Morewedge, "Resistance to medical artificial intelligence," *Journal of Consumer Research*, vol. 46, no. 4, pp. 629–650, 2019.

[8] A. R. Doshi and O. P. Hauser, "Generative AI enhances individual creativity but reduces the collective diversity of novel content," *Science Advances*, vol. 10, no. 28, eadn5290, 2024.

[9] F. Martinson, "The authenticity spectrum framework: Classifying deepfake and generative AI risks in synthetic media," *International Journal of Computer Applications*, vol. 187, no. 88, pp. 34–38, 2026.

[10] J. L. Pierce, T. Kostova, and K. T. Dirks, "Toward a theory of psychological ownership in organizations," *Academy of Management Review*, vol. 26, no. 2, pp. 298–310, 2001.

[11] A. Bandura, "Moral disengagement in the perpetration of inhumanities," *Personality and Social Psychology Review*, vol. 3, no. 3, pp. 193–209, 1999.