COLLECTED ESSAYS · REVISED 19 OCTOBER 2026

Seeing the Loop

Essays on moralization, ideology, political violence, and what an impartial observer would want from the human species

Vincent L. Anderson

5 pieces and an appendix

Compiled 19 October 2026

Contents

1 **The Diagnostic Advantage of Moral Anti-Realism**
The framework: why a diagnosis built on moral realism cannot see the mechanism it belongs to.

2 **Adversarial Review: Young and Durwin (2013) and van Zomeren et al. (2011)**
An earlier audit of two citations the thesis had misused, and how they were reclassified.

3 **Applying the Framework: Five Ideologies**
White supremacy, anarchism, Marxism, liberalism and white abolitionism, including their theories of free will.

4 **Domestic Terrorism Through the Diagnostic Lens**
The deadliest and most consequential U.S. attacks, the response to them, and prevention.

5 **After Us: A Deep-Ecological Posthumanism**
The case for the abolition of ideology and voluntary human extinction, and the strongest objections.

A **Adversarial review and revisions**
What the review found and what changed, essay by essay.

A note on reading

Every essay presumes the same three positions without arguing for them: moral anti-realism, impartial altruism as the overriding aim, and Sapolskian determinism. Later essays build on earlier ones. Claims of fact are cited; claims that follow from the presumptions are marked as such. The final essay states a case in the strongest form its defenders would give, then tests it against the strongest objections.

ESSAY 1

The Diagnostic Advantage of Moral Anti-Realism

Vincent L. Anderson · Oct 17, 2026

A systemic, determinist diagnosis of hatred and ideology, and why a diagnosis built on moral realism cannot see the mechanism it belongs to. Companion to "Diagnosis Without Desert." Revised 2026-10-19 after adversarial review.

Abstract

This paper presumes three positions without argument: moral anti-realism; impartial altruism as an overriding aim; and Sapolsky's determinism. On the third, every belief and emotion is the product of prior biological and environmental causes, and libertarian free will is excluded. On these presumptions, moral realism is not a rival theory to be weighed. It is a widespread misconception with a function. It attaches a sentiment, such as disgust or outrage, to a justification that presents itself as a stance-independent moral fact, after the sentiment has formed. The paper defines a diagnostic systemically: an account of how a population comes to hold a harmful ideology, and of how hostility toward its holders is produced, amplified and sustained. It argues that a diagnostic built on moral realism is drastically inferior in kind, because it is itself an output of the mechanism it should diagnose. It reads moralized sentiment as perception, so it cannot see the moralization pipeline or the systems that amplify it. The paper reviews evidence for each stage of that pipeline, from post hoc moral reasoning and disgust-driven moralization to the network amplification of outrage. It also argues that wherever a sentiment must be codified, it must be stated as desert or character. The justice system requires sentiment to be stated as desert, and the gatekeepers of financing, healthcare and legal documentation adapt the same conversion. Their codified outputs then stoke moralized issues and are cited as proof that moral facts exist, so the belief confirms itself. The difference in kind follows from the presumptions; the evidence shows that the loop exists and does harm, not that the moral realist diagnostic is blind to it. The size of its behavioral effects has not been measured, and the paper proposes how to measure it.

1. Presumptions and terms

The paper presumes three positions. It does not argue for them, and it does not weigh them against alternatives.

1. **Moral anti-realism.** There are no moral facts whose truth is independent of the attitudes of any actual or idealized agent.

2. **Altruism.** The reduction of suffering and the increase of well-being for all sentient beings is the observer's central and overriding aim. It is impartial: it counts each being's suffering, opponents included.
3. **Sapolskian determinism.** Every behavior, belief and emotion is the product of prior causes. These run from the neural events of the preceding seconds to genes, upbringing, culture and evolutionary history (Sapolsky 2017, 2023). Libertarian free will is excluded. No one is the ultimate author of the dispositions they act from, so no one has basic desert.

Moral realism, on these presumptions. Moral realism enters the paper only as something to be explained. It is the widespread belief that moral judgments report stance-independent facts. On the presumptions, that belief is a misconception, and it has a function. Hume described taste as "gilding or staining all natural objects with the colours, borrowed from internal sentiment" (Hume 1751, App. 1.21). The psychology supports the order of events. Moral judgment is usually a quick intuition, and moral reasoning is usually a post hoc construction generated after the judgment is reached (Haidt 2001). Moral realism is the step that presents the sentiment's verdict as a fact the world contains. Not every sentiment is converted into such a claim, and the account does not need it to be. The conversion is made wherever the social context requires a sentiment to be codified. A court cannot sentence on disgust. It must state its sentence as what the offender deserves, so the justice system converts sentiment into a claim of desert as a matter of procedure. That conversion then serves as the baseline for the institutions that grant financing, healthcare and legal documentation. Conflict adds a further route to felt objectivity. Section 3 sets out the evidence for each stage.

Terms kept apart.

- **Moral anti-realism** is a thesis about morality. Its Humean constructivist form extends to all practical reasons.
- **Global anti-realism**, in Dummett's sense, holds that the truth of any statement rests on its demonstrability. The paper does not use it. Its empirical claims are ordinary empirical claims, true or false whatever anyone thinks of them.
- **Moral relativism** holds that moral claims are true relative to a culture or framework. It is not presumed. Error theorists and quasi-realists need not be relativists. A Humean constructivist holds that reasons are relative to each agent's standpoint, which is close to agent-relativism but not to the cultural relativism some studies manipulate.
- **Folk moral objectivism** is an ordinary person's belief that a specific moral issue has an objectively correct answer. It is the misconception as people actually hold it, issue by issue, and it is what most of the studies below measure.

Throughout, "realist" and "anti-realist" carry the qualifier "moral."

2. What a systemic diagnostic is

A diagnostic is an account of how a harmful ideology, and the hostility it provokes, are produced and sustained in a population. It also says where intervention can change them. Its unit is the system, not the person.

The parts of the system. The diagnostic tracks causes at several levels at once, and each level is caused in turn.

- **Information ecosystems and platforms.** These determine which messages spread. Moral-emotional language increased a message's diffusion by about 20% per word, mostly within ideological networks rather than between them (Brady et al. 2017). Positive social feedback for outrage increased later outrage, and users conformed to the outrage norms of their networks (Brady et al. 2021).
- **Networks.** Among people who had already committed ideologically motivated crimes, having radicalized friends predicted the use of violence (Jaśko, LaFree and Kruglanski 2017). Radical social contexts strengthened the link between grievance and support for violence (Jaśko et al. 2020).
- **Material and social conditions.** In the same population of offenders, economic and social loss of significance each predicted the use of violence (Jaśko, LaFree and Kruglanski 2017). Humiliation increases the appeal of extremist ideologies through a need for cognitive closure (Webber et al. 2018).
- **Institutions.** Courts, lenders, health systems and immigration agencies must state their decisions in codified form. Where a decision rests on sentiment, the codified form turns the sentiment into a claim of desert or of moral character. The justice system sets this form, and the others adapt it (Section 3).
- **Psychological mechanisms.** Intuition, moralization and felt objectivity convert these inputs into moral conviction (Section 3).
- **Feedback loops.** Perceived dehumanization produces reciprocal dehumanization and support for aggression on the other side (Kteily, Hodson and Bruneau 2016). American partisans estimate the other side's prejudice and dehumanization toward them at about twice their measured level (Moore-Berg et al. 2020).

Individuals as nodes. Carriers, targets and observers are nodes. Their beliefs and emotions are outputs of the system and inputs to it. The diagnostician is one of them. Her outrage at an ideology is produced by the same kinds of causes, and once she expresses it, it re-enters the system as input. Under the determinism presumed in Section 1, no node's beliefs stand outside this causal web.

What the diagnostic is for. It locates leverage points: places where a change in design, policy, conditions or practice reduces the production of the ideology and of the hostility around it. Individual conversations are one leverage point among several, not the frame.

Two diagnostics.

- **The moral anti-realist diagnostic** represents every stage above as caused, including the observer's own sentiment and its apparent objectivity. It can therefore name every stage as a place to intervene.
- **The moral realist diagnostic** is the diagnosis that results when the misconception goes unrecognized. It reads moralized sentiment as perception of a moral fact. So it locates the problem in culpable individuals who perceive wrongly or chose vice, and it recommends condemnation. It is not weighed against the first as a rival account of morality. It is part of the phenomenon the first one diagnoses, and Sections 4 to 7 trace what follows from its being an output of the loop. Because determinism is presumed, the comparison holds desert fixed: the advantage claimed is what remains once both diagnostics exclude basic desert.

3. The mechanism the moral realist diagnostic cannot see

Moral hostility toward an ideology's carriers is produced by a pipeline, and the system amplifies each stage. The moral realist diagnostic reads the pipeline's output as a perception, so it cannot represent the pipeline at all. That is the difference in kind this paper defends.

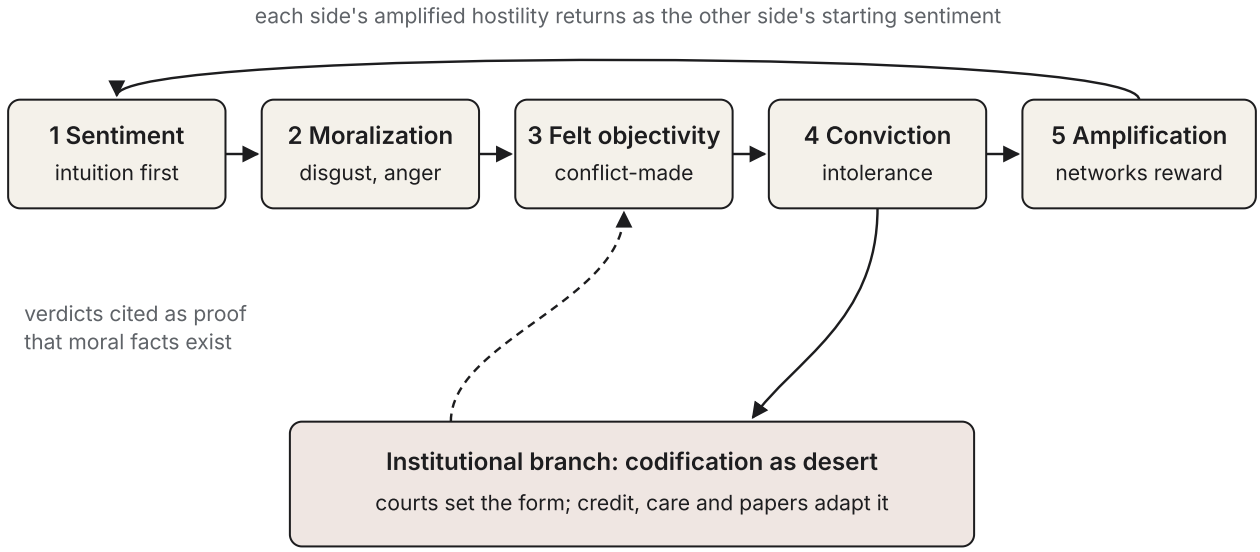


Figure 1. The moralization loop: five stages, one feedback path, and the institutional branch.

Each side's amplified hostility returns as the other side's starting sentiment, so the loop sustains itself. Below the loop, courts must codify sentiment as desert, and the gatekeepers of credit, care and papers reuse that form. That their exclusions return as grievance is predicted, not measured. Their codified outputs also return to stage 3, cited as proof that moral facts exist.

Stage 1: sentiment comes first. Moral judgment is generally the result of quick, automatic intuitions. Moral reasoning usually follows the judgment as a post hoc construction (Haidt 2001).

Stage 2: sentiment drives moralization. Moralization converts preferences into values, in individual lives and in cultures. It recruits disgust, motivates people to search for supporting reasons, and mobilizes institutions (Rozin 1999; Rozin, Markwith and Stoess 1997). Experiments show the causal direction.

- Disgusting images relevant to abortion increased moral conviction about abortion. Disgust, not anger or appraisals of harm, mediated the increase (Wisneski and Skitka 2017).
- Persuasive frames that elicit anger and disgust moralized issues, and the effect lasted at least two weeks (Clifford 2019).
- A posthypnotic flash of disgust made moral judgments more severe (Wheatley and Haidt 2005).
- Incidental disgust, unrelated to the issue, has only a small effect on moral judgment once publication bias is corrected (Landy and Goodwin 2015). The strong effect is disgust tied to the issue, which is the case that matters here.

Stage 3: conflict manufactures felt objectivity. People who argued cooperatively were less inclined than those who argued competitively to say that a topic has one objective truth (Fisher et al. 2017). Felt moral objectivity is partly an output of the conflict it is then used to justify.

Stage 4: objectivity and conviction drive hostility.

- Believing an issue to be objectively grounded and non-negotiable was enough to produce intolerance, less sharing and distancing from people who disagree (Wright, Cullum and Schwab 2008).
- Moral judgment reduced tolerance only when people treated the issue as objectively grounded (Wright and Pölzler 2022).
- Moral conviction produced social distance, failed cooperation and an inability to find procedural solutions (Skitka, Bauman and Sargis 2005).
- Moralized attitudes blocked compromise (Ryan 2017), and moralizing an issue increased endorsement of violent protest (Mooijman et al. 2018).

Stage 5: the system amplifies and feeds back. Moral-emotional language spreads faster, mostly within ideological networks (Brady et al. 2017). Outrage that earns social reward is repeated, and network norms set its level (Brady et al. 2021). Perceived dehumanization is returned in kind (Kteily, Hodson and Bruneau 2016). Each side's expressed hostility therefore becomes the other side's input at stage 1.

The institutional branch: where the conversion is required. Stages 1 to 5 describe what people do with sentiment. Some social contexts leave no choice. Where a sentiment must be codified, it must be stated as a claimed moral fact, whether or not any individual would have made that claim unprompted.

The justice system sets the baseline. The American Law Institute revised the Model Penal Code to make desert the dominant principle of sentencing, and the joint opinion in Gregg v. Georgia described capital punishment as "an expression of society's moral outrage at particularly offensive conduct" (Caruso 2020; Gregg v. Georgia, 428 U.S. 153, 183 (1976)). Laypeople asked to sentence act on just deserts, even when they state a preference for deterrence (Carlsmith, Darley and Robinson 2002). Sentiment enters the courtroom directly. Gruesome photographs made mock jurors angrier at the defendant and raised the conviction rate, and the anger mediated the extra weight jurors gave to the prosecution's evidence (Bright and Goodman-Delahunty 2006). Victim impact statements raise concerns that emotional appeal and the victim's social worth influence capital sentences (Myers and Greene 2004), and the Supreme Court held that "the Eighth Amendment erects no per se bar" to them (Payne v. Tennessee 1991). Under the presumed determinism, no one has basic desert. The desert a court states is therefore the conversion itself, performed by procedure.

Other gatekeepers adapt the conversion.

- **Legal documentation and identification.** Naturalization in the United States requires "good moral character" (8 U.S.C. §1427(a)(3)). The regulation judges it case by case against "the standards of the average citizen in the community of residence", and its bars include crimes involving moral turpitude (8 C.F.R. §316.10). USCIS describes moral turpitude as conduct that "shocks the public conscience as being inherently base, vile, or depraved" (USCIS Policy Manual, vol. 12, pt. F, ch. 5). The test is a community's sentiment, stated as a property of the applicant. Convictions from the baseline institution feed it directly.
- **Healthcare.** Prospective jurors gave intravenous drug users lower priority for transplantable hearts than smokers or people with high-fat diets, even when the drug users had better prognoses. Whether the behavior had caused the organ failure did not change the allocation (Ubel, Baron and Asch 1999). Cues that illness has individual behavioral causes raised support for individual responsibility for health costs and lowered support for government-financed insurance (Gollust and Lynch 2011). Public solidarity in welfare is conditional on deservingness criteria (van Oorschot 2000). Law writes these judgments into coverage. The Affordable Care Act lets premiums vary by tobacco use, up to 1.5 to 1, and bars every other factor except individual or family coverage, rating area and age. Tobacco use is the only listed factor that concerns the enrollee's own conduct (42 U.S.C. §300gg(a)(1)). Beginning by January 2027, or by the end of 2028 where a state receives a good-faith exemption, federal law makes 80 hours a month of work, community service or a work program,

half-time education, or equivalent income a condition of Medicaid eligibility for certain adults, in a subchapter titled "Increasing Personal Accountability" (Public Law 119-21, §71119).

- **Financing.** Scoring systems sort people into categories of "taste, riskiness, or worth", and the scores have powerful classifying effects on life chances. Fourcade and Healy observe that these systems are "presented, and experienced, as a moralized system of opportunities and just deserts" (Fourcade and Healy 2017). A score is a measured risk, but it reaches the person as a verdict on what they deserve.

What is documented and what is argued. The texts document that each of these institutions states its decisions as desert, character or worth. For legal documentation the link to the justice system is direct, because convictions trigger the bars. For healthcare and financing, the paper argues that the justice system's form is the baseline these institutions adapt, and it does not trace that lineage historically. Exclusion from credit, care or papers is a material condition of the kind Section 2 lists, so the institutional branch also feeds the loop's inputs. That link is predicted here, not measured.

The branch closes on itself. Codification does more than state a sentiment as a moral fact. Its outputs then serve as evidence that moral facts exist. A ruling, a conviction or a finding of bad character stokes the moralized issue it decided. The moralized issue is then cited as proof that the underlying moral fact is real, and the belief confirms itself. Two links in this circle are measured. An institutional decision shifted perceived social norms even where personal opinions did not change (Tankard and Paluck 2017). Perceived consensus, along with disgust, predicted higher perceived moral objectivity (Schulman, Liberman and Eyal 2026). Codification therefore manufactures the consensus that makes a moral claim feel objective, as competitive argument does at stage 3. The last step, treating codified verdicts as confirmation that moral facts exist, is circular: the evidence offered is an output of the belief it is offered to confirm. That step is argued here, not measured.

Why the moral realist diagnostic is blind to this. At stages 3 and 4, the pipeline produces a sentiment that presents itself as a perception of a moral fact. The moral realist diagnostic accepts that presentation. It therefore reads the pipeline's output as evidence about the world, not as a product of the pipeline. Its own recommended response, condemnation, is the behavior that stage 5 rewards and amplifies. It is not outside the mechanism looking in. It is one of the mechanism's outputs, describing itself. The moral anti-realist diagnostic presumes that no stage reports a stance-independent fact. It can therefore treat every stage as a cause, and every stage as a place to intervene. In the institutional branch, no one needs to hold the misconception for it to operate. Procedure performs the conversion, and its outputs, a verdict of desert or a finding of bad character, are then read as evidence that the person is culpable, and as evidence that moral facts exist.

4. Criterion one: explanation and prediction

Sections 4 to 7 do not stage a contest between two theories. Each traces one consequence of the fact that the moral realist diagnostic is an output of the loop, and sets it beside what the moral anti-realist diagnostic can do.

Under the presumed determinism, the desert-laden part of the moral realist diagnostic's explanation has no content. It explains an ideology by individuals' culpable misperception or chosen vice. No one authors the dispositions they act from (Section 1), so "chose vice" names the outcome without explaining it. "Misperception" is a causal explanation, but an incomplete one, because it stops at the individual. You learn of the vice only by observing the belief. The moral anti-realist diagnostic explains with causes that are measured independently of the belief and that predict it.

- **Significance loss** increases the appeal of extremist ideologies through a heightened need for cognitive closure. The evidence comes from imprisoned extremists in two countries and from experiments (Webber et al. 2018).
- **Self-uncertainty under threat** shifts identification toward radical groups in experiments (Hogg, Meehan and Farquharson 2010; Hogg 2021).
- **Epistemic, existential and social motives** predict conspiracy belief across 279 studies, though with small correlations (Biddlestone et al. 2025).
- **Ostracism** causally raises conspiracy belief, and self-affirmation reverses the effect (Poon, Chen and Wong 2020).
- **Networks** shape which messages spread (Brady et al. 2017). Among people already radicalized, they also shape who turns to violence (Jaśko, LaFree and Kruglanski 2017; Jaśko et al. 2020). The Jaśko studies predict violence, not the adoption of belief.

Explaining the hostility as well. The systemic diagnostic explains the hostility around an ideology with the same tools: the pipeline in Section 3 and the feedback loops in Section 2. The moral realist diagnostic explains that hostility as a warranted response to perceived wrongness. It therefore does not represent, as a mechanism, why hostility escalates on both sides at once, or why each side overestimates the other's hostility (Moore-Berg et al. 2020).

The limits. The causal predictors are probabilistic, and individually their effects are small to moderate. Most people who suffer significance loss never radicalize. The advantage is categorical in kind, since only one diagnostic explains without appeal to authorship, and modest in measured predictive power.

5. Criterion two: leverage for intervention

The moral anti-realist diagnostic finds a leverage point at every stage of the loop. The moral realist diagnostic finds one: condemnation of individuals, which is the behavior stage 5 rewards.

Stage	Leverage point	Evidence
1 Sentiment	Regulate the observer's own outrage by reappraisal	Israelis trained in reappraisal supported conciliatory policies more and aggressive ones less, five months later (Halperin et al. 2013). A Colombian study partly replicated this. Negative emotion fell (d = 0.47) and support for conciliatory statements rose (d = 0.33), with no change on aggressive statements (Hurtado-Parrado et al. 2019).
2 Moralization	Avoid frames that recruit disgust and anger, and frame in the audience's own values	Anger and disgust frames moralize and polarize issues (Clifford 2019). Arguments built on the speaker's own values tend to be "unpersuasive, even offensive" (Feinberg and Willer 2019).
3 Felt objectivity	Shift interactions from competitive to cooperative	Cooperative argument produced lower felt objectivity than competitive argument (Fisher et al. 2017).
4 Conviction	Teach that groups and people can change	Malleability beliefs improved out-group attitudes and willingness to compromise (Halperin et al. 2011). The effect held for six months in a field experiment during violence (Goldenberg et al. 2018). Teaching that traits can change reduced adolescent aggression (Yeager, Trzesniewski and Dweck 2013).
5 Amplification	Change platform incentives, correct meta-perceptions, and exchange stories rather than arguments	Platform design interacts with reinforcement and norm learning to amplify outrage (Brady et al. 2021). Correcting exaggerated meta-dehumanization reduced dehumanization (Landry et al. 2023). Non-judgmental exchange of stories durably reduced exclusionary attitudes where arguments alone did not (Kalla and Broockman 2020, 2023).
Institutions	Remove desert and moral-character tests from codified decisions, and state decisions by risk, need and cause	Untested as an intervention. The tests themselves are documented in sentencing, naturalization, health coverage and credit scoring (Caruso 2020; 8 C.F.R. §316.10; 42 U.S.C. §300gg; Public Law 119-21, §71119; Fourcade and Healy 2017).

What the moral realist diagnostic recommends. It recommends showing the culpable the moral fact and condemning them. In field experiments, arguments alone had no effect on exclusionary attitudes (Kalla and Broockman 2020). Expressed condemnation is the behavior that social reward amplifies (Brady et al. 2021), and the other side returns perceived dehumanization in kind (Kteily, Hodson and Bruneau 2016).

The limits. Several leverage points have weak or null evidence at scale. Cutting exposure to like-minded sources on Facebook had no measurable effect on eight attitude measures (Nyhan et al. 2023). Interventions that reduced partisan dislike did not reliably reduce support for undemocratic candidates or partisan violence (Voelkel et al. 2023). The individual interventions have modest effects, and several samples are small. Most of the supported leverage points work by training individuals. The system-level levers, such as platform incentives, are mostly untested, and the one large platform test was null.

6. Criterion three: the diagnostician as a node in the system

The diagnostician is part of the system she diagnoses, and the diagnostic she uses changes what she feeds back into it. Its clarity does not depend on what she feels. Her anger or contempt is a caused output, like anyone's. What the moral anti-realist diagnostic changes is that she does not read the feeling as perception of a moral fact, and that she counts what expressing it feeds back. Where time does not permit persuasion, she may still use anger or contempt as an instrument. Instrumental contempt still enters stage 5 as input, and the moral anti-realist diagnostic requires her to count that cost, which the moral realist diagnostic does not represent.

A diagnostician running the misconception amplifies the loop. Treating a contested issue as objectively grounded produces intolerance and distancing (Wright, Cullum and Schwab 2008; Wright and Pölzler 2022). Humility was associated with less intolerance, partly through less folk objectivism about the issue (Wright and Pölzler 2022). The moral realist diagnostician's expressed outrage then enters stage 5 as input (Brady et al. 2021).

Borrowed universality does not select its input. The presumptions predict that the appearance of stance-independent authority amplifies whatever sentiment it dresses, so that disgust and contempt are amplified as readily as compassion. A diagnostic that reads sentiment as perception cannot see this, because the amplified sentiment presents itself as the perception. The moral anti-realist diagnostic can see it, whichever sentiment is amplified. Stage 2 shows that disgust is among the most reliable inputs to moralization. No cited study tests the prediction directly, and Section 9 adds it as a hypothesis.

A result the account does not yet explain. Zijlstra found three dimensions of folk moral objectivity. Treating moral truth as universal related positively to punitive attitudes, while treating it as independent of minds related negatively (Zijlstra 2019). The universality result fits the account. The mind-independence result is a genuine counter-finding: the component closest to claimed stance-independence related to less punitiveness, not more. The evidence therefore locates the hostility in felt universality and non-negotiability, which moralization supplies (Rozin 1999), more than in the claim of stance-independence itself. The moral anti-realist diagnostic does not abolish felt universality in its user; it removes the warrant for treating that feeling as evidence.

Relativism is not the comparison. Relativism primes increased cheating (Rai and Holyoak 2013), and relativism mediated an effect of broad foreign experience on immoral behavior (Lu et al. 2017). Young and Durwin's "antirealism" primes were also relativist or subjectivist, and their "realism" prime asked about universality. That prime raised charitable pledges above control, while the relativist primes left pledges unchanged. The rise in giving fits the claim above that borrowed universality amplifies whatever it dresses, compassion included. The evidence is two small experiments, the

field study without random assignment or blind delivery, and neither has been replicated (Young and Durwin 2013). Relativism is not presumed, and these studies bear on relativism and felt universality, not on moral anti-realism.

7. Criterion four: error correction and reciprocity

A diagnostic is better if it can detect its own errors, and if it does not drive escalation when both sides of a conflict use it.

Discounting felt moral objectivity. Felt moral objectivity rises with competition (Fisher et al. 2017). The moral anti-realist diagnostic never treats it as perception of a stance-independent fact, so the discount is built in. The moral realist diagnostic counts felt objectivity as evidence about the moral facts. In a conflict, that means the conflict's own product counts as evidence for continuing the conflict. Codification adds a second source. The moral realist diagnostic counts laws and verdicts as evidence about moral facts, though they are outputs of the same conversion.

Symmetric use. Consider two populations in conflict, each running the same diagnostic.

- **Both running the moral realist diagnostic.** Each side reads its own moralized sentiment as perception, and the other side's as misperception or vice. Each side's condemnation is the other side's input (Section 3). The diagnostic therefore drives escalation on both sides, and it has no step that represents this.
- **Both running the moral anti-realist diagnostic.** Each side can split the conflict into a factual part, which evidence can settle, and a disagreement between standpoints, which no stance-independent fact settles. Exaggerated meta-perceptions are factual and correctable, and correcting them reduced dehumanization (Moore-Berg et al. 2020; Landry et al. 2023). Value disagreement remains, but borrowed authority does not inflate it.

Correcting the diagnostician. The moral anti-realist diagnostic applies to its user. It predicts that her own sentiment has causes, including the network she sits in. It therefore includes a step the moral realist diagnostic lacks: checking her own outrage for amplification before treating it as information about the opponent.

8. The borrowing objection

The objection is that a person who calls herself a moral realist can borrow the moral anti-realist diagnostic's working parts: causal explanation, reappraisal, malleability, perspective-getting and humility.

On the presumptions, most of the objection dissolves. Determinism is presumed, so the chosen-vice explanation and basic desert are not open options in this paper. A self-described moral realist who adopts causal explanation, regulates her own outrage and

attends to the system has borrowed everything except one step. That step is reading moralized sentiment as perception of a stance-independent fact.

The step is not a neutral extra. Section 3 shows that it is the point at which the pipeline's output presents itself as a perception. Keeping it means keeping the one element of the diagnostic that the mechanism itself produces. Borrowing therefore narrows the gap in practice, but it cannot close it. A self-described realist can close the gap only by denying moralized sentiment any evidential role. At that point her realism does no work in the diagnosis. The thesis is therefore about diagnostics that treat moralized sentiment as evidence, not about realism as a metaphysics.

9. How drastic is the advantage?

The advantage is drastic in kind and unmeasured in size.

Drastic in kind. The moral anti-realist diagnostic can represent the whole loop and intervene at every stage. The moral realist diagnostic reads the loop's output as perception, recommends the behavior the loop amplifies, and has no step that discounts its own escalation. That is the difference between a diagnostic that can see the mechanism and one that is part of it. It follows largely from the presumptions in Section 1: given them, a diagnostic that reads sentiment as perception is blind to the loop by construction. The evidence in Section 3 adds that each stage of the loop exists and does harm. The institutional branch adds that where a sentiment must be codified, it must be stated as desert or character. Whether that statement presents the sentiment as a stance-independent fact is argued here, not shown.

Unmeasured in size. No study has assigned diagnosticians to one diagnostic or the other and measured what follows.

Criterion	Advantage of the moral anti-realist diagnostic	Strength of evidence	Main sources
Explanation and prediction	It explains without appeal to authorship; under determinism, the desert-laden part of the moral realist explanation has no content	Categorical in kind; the causal predictors have small to moderate effects	Webber et al. 2018; Biddlestone et al. 2025; Brady et al. 2017
Leverage for intervention	A leverage point at every stage, against condemnation only	Moderate: several field experiments with follow-up, and some null results at scale	Halperin et al. 2011, 2013; Goldenberg et al. 2018; Kalla and Broockman 2020, 2023; Nyhan et al. 2023
The diagnostician as a node	Less intolerance and less amplification, and it sees that borrowed universality amplifies any sentiment	Moderate, mostly correlational. Zijlstra's mind-independence result counts against the step the paper blames, locating hostility in felt universality	Wright, Cullum and Schwab 2008; Wright and Pölzler 2022; Zijlstra 2019

Criterion	Advantage of the moral anti-realist diagnostic	Strength of evidence	Main sources
Error correction and reciprocity	A built-in discount on felt moral objectivity, and no symmetric escalation	Argument, with indirect support	Fisher et al. 2017; Moore-Berg et al. 2020; Landry et al. 2023

The study that would measure it. The study treats the misconception as a variable inside the system, not as a rival theory. Assign moderators, analysts or policy staff to diagnose a live conflict using one of three protocols.

1. **A protocol that runs the misconception** reads moralized sentiment as perception and locates culpable individuals.
2. **The moral anti-realist protocol** treats every stage of the loop as caused and searches for system-level leverage.
3. **A relativist protocol** separates moral anti-realism from relativism.

Measure the hostility of the outputs each group produces and the interventions each recommends. Test those interventions in the field. Also measure whether the diagnosticians' own outrage and prosocial giving change. At the system level, measure the outrage rate in the networks the diagnosticians moderate, and the hostility of the policy output their recommendations produce. Test the prediction from Section 6 that the misconception amplifies disgust and contempt along with compassion. Preregister the design and power it for small effects.

10. Conclusion

On the presumptions of moral anti-realism, impartial altruism and Sapolskian determinism, moral realism is a misconception with a job. It dresses sentiment, especially disgust and outrage, in the appearance of a stance-independent fact after the sentiment has formed. The evidence shows each stage of the loop this produces:

- intuition first and reasons after;
- disgust-driven moralization;
- felt objectivity manufactured by conflict;
- conviction that drives intolerance and support for violence;
- networks that reward and return outrage;
- institutions that require the conversion wherever a sentiment must be codified, beginning with the justice system and extending to financing, healthcare and legal documentation, whose codified outputs are then cited as proof that moral facts exist.

A diagnostic built on the misconception reads the loop's output as perception. It locates the problem in culpable individuals, though determinism denies anyone culpability in the required sense. And it recommends the condemnation the loop amplifies. Where a sentiment must be codified, institutions perform the same reading by procedure,

beginning with the courts. A moral anti-realist diagnostic sees the loop as a loop and finds leverage at every stage, from the observer's own outrage to platform incentives. That is a drastic advantage in kind. Its size in outcomes has not been measured, and the study in Section 9 would measure it.

References

Every journal entry was checked against its Crossref record on 2026-10-17. Each finding cited was checked against the study's abstract, from Crossref, OpenAlex, Consensus or the publisher. Full texts were read for Caruso (2020), Young and Durwin (2013) and Hurtado-Parrado et al. (2019). Every DOI was re-checked against Crossref on 2026-10-19. The statutes, regulation, Supreme Court opinion and USCIS Policy Manual were read on Cornell Law School's Legal Information Institute, uscis.gov and govinfo.gov on 2026-10-17. Companion paper: Diagnosis Without Desert.

Biddlestone, M., et al. (2025). Reasons to believe: A systematic review and meta-analytic synthesis of the motives associated with conspiracy beliefs. *Psychological Bulletin* 151(1): 48–87. https://doi.org/10.1037/bul0000463

Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., and Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *PNAS* 114(28): 7313–7318. https://doi.org/10.1073/pnas.1618923114

Brady, W. J., McLoughlin, K., Doan, T. N., and Crockett, M. J. (2021). How social learning amplifies moral outrage expression in online social networks. *Science Advances* 7(33): eabe5641. https://doi.org/10.1126/sciadv.abe5641

Bright, D. A., and Goodman-Delahunty, J. (2006). Gruesome evidence and emotion: Anger, blame, and jury decision-making. *Law and Human Behavior* 30(2): 183–202. https://doi.org/10.1007/s10979-006-9027-y

Carlsmith, K. M., Darley, J. M., and Robinson, P. H. (2002). Why do we punish? Deterrence and just deserts as motives for punishment. *Journal of Personality and Social Psychology* 83(2): 284–299. https://doi.org/10.1037/0022-3514.83.2.284

Caruso, G. D. (2020). Justice without retribution: An epistemic argument against retributive criminal punishment. *Neuroethics* 13(1): 13–28. https://doi.org/10.1007/s12152-018-9357-8

Clifford, S. (2019). How emotional frames moralize and polarize political attitudes. *Political Psychology* 40(1): 75–91. https://doi.org/10.1111/pops.12507

Feinberg, M., and Willer, R. (2019). Moral reframing: A technique for effective and persuasive communication across political divides. *Social and Personality Psychology Compass* 13(12): e12501. https://doi.org/10.1111/spc3.12501

Fisher, M., Knobe, J., Strickland, B., and Keil, F. C. (2017). The influence of social interaction on intuitions of objectivity and subjectivity. *Cognitive Science* 41(4): 1119–1134. https://doi.org/10.1111/cogs.12380

Fourcade, M., and Healy, K. (2017). Seeing like a market. *Socio-Economic Review*, 15(1): 9–29. https://doi.org/10.1093/ser/mww033

Goldenberg, A., Cohen-Chen, S., Goyer, J. P., Dweck, C. S., Gross, J. J., and Halperin, E. (2018). Testing the impact and durability of a group malleability intervention in the context of the Israeli–Palestinian conflict. *PNAS* 115(4): 696–701. https://doi.org/10.1073/pnas.1706800115

Gollust, S. E., and Lynch, J. (2011). Who deserves health care? The effects of causal attributions and group cues on public attitudes about responsibility for health care costs. *Journal of Health Politics, Policy and Law* 36(6): 1061–1095. https://doi.org/10.1215/03616878-1460578

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review* 108(4): 814–834. https://doi.org/10.1037/0033-295x.108.4.814

Halperin, E., Russell, A. G., Trzesniewski, K. H., Gross, J. J., and Dweck, C. S. (2011). Promoting the Middle East peace process by changing beliefs about group malleability. *Science* 333(6050): 1767–1769. https://doi.org/10.1126/science.1202925

Halperin, E., Porat, R., Tamir, M., and Gross, J. J. (2013). Can emotion regulation change political attitudes in intractable conflicts? From the laboratory to the field. *Psychological Science* 24(1): 106–111. https://doi.org/10.1177/0956797612452572

Hogg, M. A. (2021). Uncertain self in a changing world: A foundation for radicalisation, populism, and autocratic leadership. *European Review of Social Psychology* 32(2): 235–268. https://doi.org/10.1080/10463283.2020.1827628

Hogg, M. A., Meehan, C., and Farquharson, J. (2010). The solace of radicalism: Self-uncertainty and group identification in the face of threat. *Journal of Experimental Social Psychology* 46(6): 1061–1066. https://doi.org/10.1016/j.jesp.2010.05.005

Hume, D. (1751). *An Enquiry Concerning the Principles of Morals*, Appendix 1, paragraph 21. https://davidhume.org/texts/m/app1

Hurtado-Parrado, C., et al. (2019). Emotion regulation and attitudes toward conflict in Colombia: Effects of reappraisal training on negative emotions and support for conciliatory and aggressive statements. *Frontiers in Psychology* 10: 908. https://doi.org/10.3389/fpsyg.2019.00908

Jaśko, K., LaFree, G., and Kruglanski, A. (2017). Quest for significance and violent extremism: The case of domestic radicalization. *Political Psychology* 38(5): 815–831. https://doi.org/10.1111/pops.12376

Jaśko, K., et al. (2020). Social context moderates the effects of quest for significance on violent extremism. *Journal of Personality and Social Psychology* 118(6): 1165–1187. https://doi.org/10.1037/pspi0000198

Kalla, J. L., and Broockman, D. E. (2020). Reducing exclusionary attitudes through interpersonal conversation: Evidence from three field experiments. *American Political Science Review* 114(2): 410–425. https://doi.org/10.1017/s0003055419000923

Kalla, J. L., and Broockman, D. E. (2023). Which narrative strategies durably reduce prejudice? Evidence from field and survey experiments supporting the efficacy of perspective-getting. *American Journal of Political Science* 67(1): 185–204. https://doi.org/10.1111/ajps.12657

Kteily, N., Hodson, G., and Bruneau, E. (2016). They see us as less than human: Metadehumanization predicts intergroup conflict via reciprocal dehumanization. *Journal of Personality and Social Psychology* 110(3): 343–370. https://doi.org/10.1037/pspa0000044

Landry, A. P., Schooler, J. W., Willer, R., and Seli, P. (2023). Reducing explicit blatant dehumanization by correcting exaggerated meta-perceptions. *Social Psychological and Personality Science* 14(4): 407–418. https://doi.org/10.1177/19485506221099146

Landy, J. F., and Goodwin, G. P. (2015). Does incidental disgust amplify moral judgment? A meta-analytic review of experimental evidence. *Perspectives on Psychological Science* 10(4): 518–536. https://doi.org/10.1177/1745691615583128

Lu, J. G., Quoidbach, J., Gino, F., Chakroff, A., Maddux, W. W., and Galinsky, A. D. (2017). The dark side of going abroad: How broad foreign experiences increase immoral behavior. *Journal of Personality and Social Psychology* 112(1): 1–16. https://doi.org/10.1037/pspa0000068

Mooijman, M., Hoover, J., Lin, Y., Ji, H., and Dehghani, M. (2018). Moralization in social networks and the emergence of violence during protests. *Nature Human Behaviour* 2(6): 389–396. https://doi.org/10.1038/s41562-018-0353-0

Moore-Berg, S. L., Ankori-Karlinsky, L.-O., Hameiri, B., and Bruneau, E. (2020). Exaggerated meta-perceptions predict intergroup hostility between American political partisans. *PNAS* 117(26): 14864–14872. https://doi.org/10.1073/pnas.2001263117

Myers, B., and Greene, E. (2004). The prejudicial nature of victim impact statements: Implications for capital sentencing policy. *Psychology, Public Policy, and Law* 10(4): 492–515. https://doi.org/10.1037/1076-8971.10.4.492

Nyhan, B., et al. (2023). Like-minded sources on Facebook are prevalent but not polarizing. *Nature* 620(7972): 137–144. https://doi.org/10.1038/s41586-023-06297-w

Payne v. Tennessee, 501 U.S. 808 (1991). https://www.law.cornell.edu/supremecourt/text/501/808

Poon, K.-T., Chen, Z., and Wong, W.-Y. (2020). Beliefs in conspiracy theories following ostracism. *Personality and Social Psychology Bulletin* 46(8): 1234–1246. https://doi.org/10.1177/0146167219898944

Public Law 119-21, 139 Stat. 72 (approved 4 July 2025), §71119, Medicaid community engagement requirements, 139 Stat. 306. https://www.govinfo.gov/content/pkg/PLAW-119publ21/html/PLAW-119publ21.htm

Rai, T. S., and Holyoak, K. J. (2013). Exposure to moral relativism compromises moral behavior. *Journal of Experimental Social Psychology* 49(6): 995–1001. https://doi.org/10.1016/j.jesp.2013.06.008

Rozin, P. (1999). The process of moralization. *Psychological Science* 10(3): 218–221. https://doi.org/10.1111/1467-9280.00139

Rozin, P., Markwith, M., and Stoess, C. (1997). Moralization and becoming a vegetarian: The transformation of preferences into values and the recruitment of disgust. *Psychological Science* 8(2): 67–73. https://doi.org/10.1111/j.1467-9280.1997.tb00685.x

Ryan, T. J. (2017). No compromise: Political consequences of moralized attitudes. *American Journal of Political Science* 61(2): 409–423. https://doi.org/10.1111/ajps.12248

Sapolsky, R. M. (2017). *Behave: The Biology of Humans at Our Best and Worst*. New York: Penguin Press. ISBN 9781594205071.

Sapolsky, R. M. (2023). *Determined: A Science of Life Without Free Will*. New York: Penguin Press. ISBN 9780525560975.

Schulman, R., Liberman, N., and Eyal, T. (2026). Perceived moral objectivity: Why some moral issues seem more objective than others. *Cognition* 276: 106642. https://doi.org/10.1016/j.cognition.2026.106642

Skitka, L. J., Bauman, C. W., and Sargis, E. G. (2005). Moral conviction: Another contributor to attitude strength or something more? *Journal of Personality and Social Psychology* 88(6): 895–917. https://doi.org/10.1037/0022-3514.88.6.895

Tankard, M. E., and Paluck, E. L. (2017). The effect of a Supreme Court decision regarding gay marriage on social norms and personal attitudes. *Psychological Science* 28(9): 1334–1344. https://doi.org/10.1177/0956797617709594

Ubel, P. A., Baron, J., and Asch, D. A. (1999). Social acceptability, personal responsibility, and prognosis in public judgments and transplant allocation. *Bioethics* 13(1): 57–68. https://doi.org/10.1111/1467-8519.00131

U.S. Code, Title 8, §1427(a)(3), requirements of naturalization. https://www.law.cornell.edu/uscode/text/8/1427

U.S. Code, Title 42, §300gg(a)(1), fair health insurance premiums. https://www.law.cornell.edu/uscode/text/42/300gg

U.S. Code of Federal Regulations, Title 8, §316.10, good moral character. https://www.law.cornell.edu/cfr/text/8/316.10

U.S. Citizenship and Immigration Services. Policy Manual, Volume 12, Part F, Chapter 5. https://www.uscis.gov/policy-manual/volume-12-part-f-chapter-5

van Oorschot, W. (2000). Who should get what, and why? On deservingness criteria and the conditionality of solidarity among the public. *Policy & Politics* 28(1): 33–48. https://doi.org/10.1332/0305573002500811

Voelkel, J. G., et al. (2023). Interventions reducing affective polarization do not necessarily improve anti-democratic attitudes. *Nature Human Behaviour* 7(1): 55–64. https://doi.org/10.1038/s41562-022-01466-9

Webber, D., et al. (2018). The road to extremism: Field and experimental evidence that significance loss-induced need for closure fosters radicalization. *Journal of Personality and Social Psychology* 114(2): 270–285. https://doi.org/10.1037/pspi0000111

Wheatley, T., and Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science* 16(10): 780–784. https://doi.org/10.1111/j.1467-9280.2005.01614.x

Wisneski, D. C., and Skitka, L. J. (2017). Moralization through moral shock: Exploring emotional antecedents to moral conviction. *Personality and Social Psychology Bulletin* 43(2): 139–150.

https://doi.org/10.1177/0146167216676479

Wright, J. C., Cullum, J., and Schwab, N. (2008). The cognitive and affective dimensions of moral conviction: Implications for attitudinal and behavioral measures of interpersonal tolerance. *Personality and Social Psychology Bulletin* 34(11): 1461–1476. https://doi.org/10.1177/0146167208322557

Wright, J. C., and Pölzler, T. (2022). Should morality be abolished? An empirical challenge to the argument from intolerance. *Philosophical Psychology* 35(3): 350–385. https://doi.org/10.1080/09515089.2021.1983160

Yeager, D. S., Trzesniewski, K. H., and Dweck, C. S. (2013). An implicit theories of personality intervention reduces adolescent aggression in response to victimization and exclusion. *Child Development* 84(3): 970–988. https://doi.org/10.1111/cdev.12003

Young, L., and Durwin, A. J. (2013). Moral realism as moral motivation: The impact of meta-ethics on everyday decision-making. *Journal of Experimental Social Psychology* 49(2): 302–306. https://doi.org/10.1016/j.jesp.2012.11.013

Zijlstra, L. (2019). Folk moral objectivism and its measurement. *Journal of Experimental Social Psychology* 84: 103807. https://doi.org/10.1016/j.jesp.2019.04.005

ESSAY 2

Adversarial Review: Young and Durwin (2013) and van Zomeren et al. (2011)

Written 2026-10-18; revised 2026-10-19

Written 2026-10-18T11:51-06:00. The reviewer holds the paper's presumptions fixed: moral anti-realism, impartial altruism and Sapolskian determinism. The question is whether these two studies, as the paper cited them, advance its thesis: that the moral anti-realist diagnostic sees the mechanism and the moral realist diagnostic is part of it.

Verdict

Both citations distracted from the thesis, as the author suspected. Neither study tests moral anti-realism. Young and Durwin contrast universality with cultural relativism and subjectivism. Van Zomeren and colleagues measure first-order moral conviction. Neither bears on whether a diagnostic can see the loop. The paragraph also turned a point about mechanism into a ledger of motivational benefits. That invites a comparison of moral realism and moral anti-realism by outcomes, which the presumptions exclude.

Young and Durwin (2013)

Read in full from the authors' lab copy on 2026-10-18.

1. **The primes do not contrast realism with moral anti-realism.** The "realism" prime asked: "Do you agree that some things are just morally right or wrong, good or bad, wherever you happen to be from in the world?" That is universality. The Experiment 1 "antirealism" prime said morals are "shaped by our culture and upbringing, so there are no absolute right answers to any moral questions." The Experiment 2 version said "it is up to each person to discover his or her own moral truths." Those are cultural relativism and individual subjectivism. The paper does not presume relativism, and it set aside Rai and Holyoak (2013) and Lu et al. (2017) on that ground. Citing Young and Durwin as "moral realism against moral anti-realism" was inconsistent with that rule.
2. **The relativist primes did not lower giving.** Control and "antirealism" did not differ in either experiment. The Mann-Whitney p was 0.907 in Experiment 1, and $t(115) = 1.02$, $p = 0.32$, in Experiment 2. The only effect is that the universality prompt raised giving above baseline.
3. **The field study was not randomized or blind.** Conditions were "assigned in sequence." One canvasser delivered the prime, the pitch and a persuasion step,

knowing the condition. Step durations were not recorded. There were 46 to 47 people per condition.

4. **The online study measured a hypothetical pledge.** Mechanical Turk workers paid $0.11 said how much of a notional $20 they would give. Cells were unequal after exclusions: 74, 58 and 59.
5. **The cause was chosen to be uncontroversial.** The authors note that "distinct effects may obtain for moral issues that are recognized as controversial." The paper is about contested ideologies.
6. **No replication was found.** Web searches on 2026-10-18 found none.

What it can still show. Its universality prime parallels Zijlstra (2019), where felt universality related to punitiveness. Together they fit the paper's claim that felt universality amplifies sentiment without selecting it, though Zijlstra's mind-independence result points the other way, as the paper notes. That is a sentence of value, not a paragraph.

Van Zomeren, Postmes, Spears and Bettache (2011)

Abstract read on 2026-10-18. The full text sits behind a human-verification check, which the reviewer did not complete, so the design was not verified.

1. **Construct.** The study defines moral convictions as "strong and absolute stances on moralized issues." That is first-order moralization, not belief in stance-independent moral facts.
2. **Mechanism.** The proposed route runs through identification with victims, group-based anger and group efficacy. None of these is a metaethical belief.
3. **Conflict with the paper's own framework.** The paper treats moral conviction as a stage 4 output that drives intolerance and blocks compromise (Skitka, Bauman and Sargis 2005; Ryan 2017). Citing it as a benefit of "the misconception" merged moral conviction with moral realism. That is the same kind of premise drift as the earlier conflation of moral anti-realism with global anti-realism.

Implications for the paper

- **The thesis does not need a motivational ledger.** Whether believing the misconception raises giving is irrelevant to the diagnostic's clarity. In the same way, which emotion the diagnostician feels is irrelevant, as the author established on 2026-10-17.
- **The thesis-relevant point is the mechanism.** Borrowed universality amplifies whatever it dresses. A diagnostic that reads sentiment as perception cannot see that, whichever sentiment is amplified.

- **Honest handling is reclassification, not deletion.** Young and Durwin is not counter-evidence to moral anti-realism. It is evidence about relativism and felt universality, so it moves to the paragraph that handles those.

Changes made to the paper

- **Section 6.** The paragraph citing both studies now reads "Borrowed universality does not select its input." It states the mechanism and the prediction, and it cites neither study.
- **Section 6, relativism paragraph.** Young and Durwin now appear there, described as universality against relativism, with their design limits.
- **Section 9 table.** The node row no longer cites Young and Durwin or a "benefit on giving."
- **References.** Van Zomeren et al. (2011) is removed, since nothing cites it now.

Sources

Young, L., and Durwin, A. J. (2013), Journal of Experimental Social Psychology 49(2): 302–306, https://doi.org/10.1016/j.jesp.2012.11.013. Full text read from https://moralitylab.bc.edu/wp-content/uploads/sites/192/2011/10/YoungDurwin2013.pdf on 2026-10-18.

Van Zomeren, M., Postmes, T., Spears, R., and Bettache, K. (2011), Group Processes & Intergroup Relations 14(5): 735–753, https://doi.org/10.1177/1368430210395637. Abstract read from OpenAlex and https://research.rug.nl/en/publications/can-moral-convictions-motivate-the-advantaged-to-challenge-social/ on 2026-10-18.

ESSAY 3

Applying the Framework: Five Ideologies

An application of "The Diagnostic Advantage of Moral Anti-Realism" to white supremacy, anarchism, Marxism, liberalism and white abolitionism. Written Oct 19, 2026.

1. How the framework is applied

What the framework can and cannot do. The paper presumes moral anti-realism, impartial altruism and Sapolskian determinism. On those presumptions no ideology is *true or false as morality*, because there are no stance-independent moral facts for it to match. The framework therefore does not rank ideologies by moral correctness. It does two other things. It diagnoses how each ideology, and the hostility around it, is produced, amplified and sustained (Sections 2–3 of the paper). And it measures each against the one aim it presumes: the impartial reduction of suffering for all sentient beings, opponents included. An ideology can be assessed on that second axis without any appeal to moral fact, because the assessment is about consequences and about whose suffering the ideology counts.

What "extreme" means here. The paper does not define extremity. This application uses a definition built from the paper's own stage 4: an ideology is extreme to the degree that it (a) treats its core claims as non-negotiable and objectively grounded, (b) licenses intolerance, exclusion or violence toward outgroups, and (c) excludes some beings' suffering from the count. On this definition extremity is a property of how an ideology is *held and enacted*, not only of its content. Every ideology below has extreme and moderate carriers, and the analysis notes both. Where an ideology's content alone satisfies (b) and (c), its extremity does not depend on the carrier.

A note on the list. "All extreme ideologies" cannot be covered in one document; the five named here are analysed, and the template transfers to others. Liberalism is not usually classed as extreme. It is included as the comparison case, and because the paper's institutional branch (desert-based sentencing, "good moral character" tests) is largely built from liberal legal institutions.

Template for each ideology.

- **Core claims**, briefly and in the terms its own carriers would accept.
- **Inputs**: the material, social and psychological conditions that predict uptake (Sections 2 and 4 of the paper).
- **Moralization profile**: which sentiments it recruits (stage 2), what it presents as stance-independent fact (stage 3), and how non-negotiable it makes its claims (stage 4).

- **Codification**: where it has been, or seeks to be, written into institutions (the institutional branch).
- **The hostility loop**: how the ideology and its opponents feed each other (stage 5).
- **Against impartial altruism**: whose suffering it counts.
- **Leverage points**: where the moral anti-realist diagnostic would intervene, and where the moral realist diagnostic would only condemn.

2. White supremacy

Core claims. People racialized as white are a natural kind, superior in worth or capacity, and entitled to dominance or separation. Contemporary forms centre on demographic threat, such as the "Great Replacement" narrative popularized by Renaud Camus (2011).

Inputs. The paper's predictors apply directly. Significance loss and humiliation raise the appeal of extremist ideologies through need for closure (Webber et al. 2018). Self-uncertainty under threat moves identification toward radical groups (Hogg, Meehan and Farquharson 2010). Economic and social loss of significance, and radicalized friends, predict violence among those already radicalized (Jaśko, LaFree and Kruglanski 2017). Status loss of a formerly dominant group is the characteristic input here: the ideology converts a felt decline into a story of theft.

Moralization profile. This is the most disgust-driven of the five. Its vocabulary is purity, contamination, degeneration and invasion, which is the disgust route to moralization that Rozin (1999) and Wisneski and Skitka (2017) describe. It presents race as a biological fact and racial hierarchy as a natural or divine order, so the sentiment arrives already dressed as stance-independent truth. Non-negotiability is extreme: the outgroup's existence, not its conduct, is the offence, so no compromise is available in principle.

Codification. The clearest historical case of the institutional branch. U.S. naturalization was limited to "free white persons" from 1790; anti-miscegenation statutes stood until *Loving v. Virginia* (1967); Jim Crow and the Nuremberg Laws (1935) wrote racial worth into citizenship, marriage and access to goods. Each statute then served as evidence that the hierarchy was real, which is the self-confirming circle the paper describes (Section 3, "the branch closes on itself"). It has since been codified against as well, for instance in Germany's criminal bans on Nazi symbols and incitement (StGB §§86a, 130).

The hostility loop. Contemporary carriers are thick in online networks where moral-emotional language spreads within, not between, ideological communities (Brady et al. 2017). Opponents' condemnation is recycled as proof of persecution and replacement, and their meta-perceptions are often exaggerated in turn (Moore-Berg et al. 2020). Mass-casualty attackers in Christchurch (2019), El Paso (2019) and Buffalo (2022) cited replacement narratives, which is the stage 4 link from conviction to violence (Mooijman et al. 2018) in its sharpest form.

Against impartial altruism. It fails at the root. It does not merely weigh some beings' suffering wrongly; it removes whole groups from the count or counts their suffering as a good. This is the one judgment the framework can make with full confidence, and it is made without any appeal to moral fact: impartiality is the presumed aim, and the ideology is defined by its partiality.

Leverage points. Significance and belonging offered through non-racial channels (Webber et al. 2018); network exit and alternative social ties, given the network predictors (Jaśko et al. 2020); malleability framing (Halperin et al. 2011); platform incentives that reward outrage less (Brady et al. 2021); and non-judgmental narrative exchange (Kalla and Broockman 2020, 2023). Exit programmes such as EXIT Deutschland and Life After Hate work on these lines; their effects have not been rigorously evaluated. The moral realist diagnostic would mainly condemn, which this ideology's own narrative predicts and absorbs.

What the framework adds. Under determinism, carriers are not ultimately authors of their dispositions. That changes the response, not the assessment: the ideology is maximally harmful against the presumed aim, and its carriers are caused nodes whose causes can be changed.

3. Anarchism

Core claims. Coercive hierarchy, above all the state, is illegitimate and unnecessary; social order can rest on voluntary association and mutual aid. The tradition is wide: Proudhon's mutualism, Kropotkin's communist anarchism (*Mutual Aid*, 1902), anarcho-syndicalism, Stirner's egoism (1844) and, more distantly, market anarchism.

Inputs. Distrust of institutions that have excluded or punished the carrier; experiences of police or state coercion; and, historically, industrial labour conditions. On available accounts, uptake is concentrated in youth subcultures and in movements formed around specific grievances, so network effects are likely strong (Jaśko et al. 2020).

Moralization profile. Split, and the split matters to the framework. Stirner is an anti-realist: he called morality, rights and the state "spooks," abstractions that people serve as if they had authority of their own. That is close to the paper's own diagnosis of borrowed universality. Kropotkin, by contrast, grounded anarchist ethics in a naturalized account of solidarity, and much movement rhetoric treats "freedom" and "domination" as stance-independent facts. The main recruited sentiment is outrage at coercion rather than disgust. Non-negotiability varies: on the state itself it is high; on everyday organization it is low by design, since consensus processes are a core practice.

Codification. The one ideology here that rejects the institutional branch as such, which gives it a partial view of the desert conversion the paper describes. Its relation to codification is mostly as a target: the U.S. Anarchist Exclusion Act (1903) and the present bar on naturalizing anyone who advocates opposition to all organized government (8

U.S.C. §1424) are the legal-documentation gatekeeper applied to the ideology itself. Within movements, informal codification (call-outs, expulsions) can reproduce the desert form without a state.

The hostility loop. Historically, the "propaganda of the deed" period produced assassinations, including of French President Carnot (1894), Empress Elisabeth of Austria (1898) and President McKinley (1901), and the state responses (exclusion laws, the Palmer Raids) became inputs to further grievance. Today the loop runs mainly through street confrontations with police and with the far right, where each side's footage becomes the other's evidence of threat.

Against impartial altruism. Mixed, and dependent on the form. Its stated aim, the end of domination for everyone, is impartial in intent; Kropotkin's mutual-aid ethic is close to the presumed aim. The risks are consequential, not partial: the suffering produced by political violence, and the welfare cost if dismantling institutions removes protections before replacements exist. Those are empirical questions, not settled by the ideology's intent.

Leverage points. Cooperative rather than competitive engagement lowers felt objectivity (Fisher et al. 2017), and anarchist consensus practice is already a version of it. Correcting meta-perceptions between activists and police (Landry et al. 2023). Removing desert and moral-character tests from codified decisions, the paper's untested institutional lever, would also remove some of the grievance anarchism feeds on.

4. Marxism

Core claims. History is driven by material conditions and class conflict; capitalism extracts surplus value from labour, produces crises and will be superseded by a classless society. The family runs from academic Marxism and democratic socialism to Marxism-Leninism and Maoism, which added a vanguard party and a revolutionary state.

Inputs. Material deprivation, exploitation at work and perceived unfairness of distribution, which are the paper's "material and social conditions" input taken as the ideology's own subject. Marxism is unusual in that its explanation of its own uptake is roughly the one the paper uses: beliefs as outputs of material conditions.

Moralization profile. Marxism contains the most explicit precedent for the paper's thesis and the clearest case of violating it. In *The German Ideology* Marx and Engels treat morality as ideology, an output of material relations presented as eternal truth, and Engels argued in *Anti-Dühring* (1878, Part I, ch. IX) that all moral theories have been hitherto, "in the last analysis," the product of the economic conditions of society. That is a structural cousin of the paper's moral anti-realism. Yet Marxist movements moralize heavily: "exploiter," "class enemy" and "enemy of the people" are verdicts of character and desert. Lukes (*Marxism and Morality*, 1985) called this a paradox: a theory that dismisses morality as ideology while condemning capitalism in strongly moral terms. In

the framework's terms, the theory saw the mechanism and the movement became one of its outputs. Historical inevitability then supplies felt objectivity (stage 3) on a world-historical scale, and conflict is built in as class struggle, which is the competitive setting Fisher et al. (2017) found raises felt objectivity.

Codification. Marxist-Leninist states codified class position as moral standing. Soviet "class alien" status disenfranchised whole categories (the *lishentsy*) and barred them from voting, higher education, housing and many jobs (Alexopoulos 2003); Article 58 of the RSFSR criminal code (1927) made any action "aimed at overthrowing, undermining or weakening" Soviet power a crime; Maoist China assigned hereditary class labels that governed schooling, jobs and marriage. These are the institutional branch in pure form, with class standing replacing race or individual conduct as the desert test. Collectivization and the famines that followed (the Soviet famine of 1932–33, the Great Leap Forward famine of 1959–61) killed millions. Marxism has also been codified against: the U.S. statute that bars anarchists from naturalization also bars members of the Communist Party (8 U.S.C. §1424(a)(2)).

The hostility loop. Anti-communism and communism fed each other through the twentieth century, each treating the other as an existential moral threat. Today the loop is mostly rhetorical, with "Marxist" used as an epithet and "capitalist" or "fascist" returned.

Against impartial altruism. Divided by form. The aim, ending exploitation and want for everyone, is impartial in principle, and the critique of suffering produced by economic structure is one the framework itself would make. The Leninist forms failed the aim in practice, partly through the mechanism the paper names: class enemies were removed from the count, so their suffering registered as justice.

Leverage points. The causal-explanation tools Marxism already holds; reappraisal of outrage (Halperin et al. 2013); and stripping desert from codified class or political categories. Democratic and academic Marxists who keep the materialist diagnosis and drop the class-enemy verdict are close to the paper's own "borrowing" case (Section 8), from the other direction.

5. Liberalism

Which liberalism. In political theory, liberalism is the tradition of individual rights, limited government, rule of law and tolerance (Locke, Kant, Mill, Rawls). In American usage, "liberal" also names the progressive left of the party system. This section addresses the first; the second shares its institutions and shows the same patterns in culture-war form.

Core claims. Individuals hold rights that constrain what states and majorities may do to them; legitimate authority rests on consent; disagreement about the good life is to be tolerated, not settled by force.

Inputs. As the default ideology of most Western institutions, liberalism is acquired largely through schooling, law and civic culture rather than through grievance. Its intense forms rise under perceived threat to the liberal order itself, which shows the same self-uncertainty route as the other four (Hogg 2021).

Moralization profile. Liberalism's founding documents are moral realist in form. Locke grounded rights in natural law; the U.S. Declaration of Independence calls its truths "self-evident"; Kant grounded dignity and retributive punishment in reason. These are borrowed universality in the paper's sense, and they amplify whatever they dress, including compassion (abolition, civil rights) and, at times, contempt for those cast as illiberal. Liberalism also contains the strongest internal resources for the framework's position. Mill's liberalism is consequentialist. Rawls held that natural talents and the starting points that shape effort are "arbitrary from a moral point of view" and rejected pre-institutional desert as a basis for distributive shares (*A Theory of Justice*, 1971, §§12, 17, 48). He did not extend this to retributive justice, which §48 treats separately, and his reason is moral arbitrariness rather than determinism, so the strand is a partial ally.

Codification. This is liberalism's distinctive exposure. The paper's institutional branch is largely a liberal construction: desert as the dominant principle of sentencing in the Model Penal Code, "good moral character" in naturalization, deservingness tests in health coverage and welfare (Caruso 2020; 8 C.F.R. §316.10; van Oorschot 2000). The responsible, choosing individual at the centre of liberal theory is the figure the desert conversion needs. Liberalism is therefore the ideology the paper's institutional critique most directly targets, even though it is the least "extreme" of the five.

The hostility loop. In its mainstream form, liberal proceduralism (elections, courts, free speech) is itself a damping mechanism: it converts conflicts into procedures with losers who can try again. In polarized form, where the American partisan sense of "liberal" is the relevant one, the loop of Section 3 runs as usual: American partisans overestimate the other side's hostility by about two to one (Moore-Berg et al. 2020), and moral-emotional content spreads within each side (Brady et al. 2017).

Against impartial altruism. Strong in scope, limited in two places. Rights are extended to every person, so its circle of concern is formally universal among humans. The limits are its treatment of desert, which licenses suffering imposed as punishment that determinism says no one deserves, and its historical boundaries of the rights-bearing person (excluding, at various times, the enslaved, the colonized and women; and, like the other four in their mainstream forms, largely excluding non-human sentient beings, which the presumed aim counts).

Leverage points. Replacing desert with risk, need and cause in sentencing and gatekeeping, the paper's institutional lever; building on Rawls's and Mill's anti-desert and consequentialist strands rather than on natural-law rhetoric; and the meta-perception corrections that work well in partisan settings (Landry et al. 2023).

6. White abolitionism

Core claims. "Whiteness" is not a biological fact but a social category, built historically to sort who received the privileges of citizenship, property and labour protection; it should therefore be abolished as a category, as one abolishes an institution. The term comes from the *Race Traitor* journal of Noel Ignatiev and John Garvey (1993–2005), whose motto was "Treason to whiteness is loyalty to humanity," and from whiteness-studies historians such as David Roediger (*The Wages of Whiteness*, 1991), building on Du Bois's account of the "psychological wage" of whiteness (*Black Reconstruction*, 1935). Its target is the category and the privileges attached to it, not white people. It is a small intellectual current, not a mass movement.

Inputs. Mostly academic and activist settings; the experience of witnessing or suffering racial hierarchy; for white adherents, felt complicity. These are different inputs from the significance-loss route that drives most of the other four.

Moralization profile. Its core analysis is the closest of the five to the paper's method. It treats race as caused and constructed, not perceived, and explains a hierarchy by its history and the material interests it serves. The risk lies at stage 4. Where whiteness is moralized as collective guilt, people are assigned a moral standing by birth, which is the desert form the paper rejects under determinism and the mirror image of the essentialism it opposes. Outrage, shame and guilt, rather than disgust, are the recruited sentiments. The word "abolition" borrows the moral authority of the anti-slavery movement, which is borrowed universality in the paper's sense.

Codification. Little. Some workplace and educational diversity trainings use whiteness concepts, and the evidence on diversity training's effects is thin and mixed, with some effects fading within months (Devine and Ash 2022; Paluck et al. 2021). The ideology has no codified state form comparable to the others, though it has been codified against: Executive Order 13950 (2020) and several state laws restricted training in "divisive concepts."

The hostility loop. This is where the framework predicts the largest gap between meaning and reception. "Abolish whiteness" is often heard, on its critics' own accounts, as a call to eliminate white people, and white supremacist propaganda uses the phrase as evidence of "anti-white" intent and of replacement. That is the exaggerated meta-perception pattern (Moore-Berg et al. 2020; Kteily, Hodson and Bruneau 2016) in an unusually direct form: one side's slogan becomes the other side's primary recruiting input. The loop with white supremacy is tight, since each names the other as its reason for existing.

Against impartial altruism. The aim, ending a hierarchy of worth, is impartial and fits the presumed aim. The risk is partiality at the edges: counting white people's suffering at a discount, or treating it as deserved. Where that happens, the ideology fails the same test as the hierarchy it opposes, in a weaker form.

Leverage points. Keep the causal and constructivist analysis, which the framework endorses, and drop collective desert, which it rejects; frame in the audience's own values rather than the speaker's (Feinberg and Willer 2019); and pay attention to how terms are received, since the measured cost of exaggerated meta-perception falls on both sides.

7. Free will: libertarian and compatibilist tendencies

Why this belongs in the analysis. The paper presumes Sapolskian determinism, which excludes libertarian free will, the view that agents can originate choices not fixed by prior causes. Sapolsky also rejects compatibilism as a basis for desert: redefining "free" as acting from one's own desires without coercion does not give anyone ultimate authorship of those desires (Sapolsky 2023). So each ideology's working theory of the will matters to the framework for one reason: it determines where the ideology places basic desert, and basic desert is what powers the institutional branch. Three positions are tracked: **libertarian** (choices originate in the agent), **compatibilist** (freedom is acting from one's own reasons, consistent with determinism, often with desert retained) and **hard determinist** (no basic desert).

White supremacy. Asymmetric. It is hereditarian about the groups it ranks below it: their condition is fixed by biology, which is determinism used to deny capacity, not to deny blame. It is libertarian about the groups it blames: conspiring elites and "replacers" are cast as willing, scheming agents who deserve retribution. Its ingroup ideal is heroic will, expressed in fascist cults of will and struggle. Determinism and free will are each applied where they license subordination or punishment.

Anarchism. Rhetorically voluntarist: autonomy, self-ownership and free association are its vocabulary, and Stirner's egoism comes close to a libertarian picture of the self-owning ego. Its classical theorists leaned the other way on crime. Bakunin was a materialist who wrote that socialism "absolutely rejects the doctrine of free will" and treated people as products of nature and society (Maximoff 1953), and Kropotkin argued in his writing on prisons that crime has social causes that punishment cannot address. That makes much classical anarchism compatibilist about political freedom (freedom as the absence of domination) and close to hard determinist about punishment. Of the five, its anti-punitive strand is nearest to the paper's position.

Marxism. Formally compatibilist. "It is not the consciousness of men that determines their being, but their social being that determines their consciousness" (Marx, 1859 Preface); "men make their own history, but they do not make it as they please" (*The Eighteenth Brumaire*, 1852); and Engels, building on Hegel's "insight into necessity," defined freedom as "control over ourselves and over external nature, founded on knowledge of natural necessity" (*Anti-Dühring*, 1878). That is textbook compatibilism. Practice diverged. Leninist and Maoist voluntarism stressed revolutionary will over material conditions, and Stalinist show trials demanded confessions of conscious

counter-revolutionary intent, a libertarian standard of guilt applied to the people the state wished to punish. The theory explains class enemies causally; the movement then blamed them as free agents.

Liberalism. Its philosophers are mostly compatibilist: Hobbes (a contested member of the liberal canon), Locke on liberty as the power to act on one's will, Hume, and Mill's chapter on liberty and necessity in *A System of Logic*. Kant is the exception, grounding moral agency in a transcendental freedom that is libertarian in structure. Liberal law and public discourse lean libertarian. The U.S. Supreme Court described the requirement of criminal intent as being as universal and persistent in mature legal systems as "belief in freedom of the human will and a consequent ability and duty of the normal individual to choose between good and evil" (*Morissette v. United States*, 1952), and rejected "a deterministic view of human conduct that is inconsistent with the underlying precepts of our criminal justice system" (*United States v. Grayson*, 1978). "Personal responsibility" in welfare and health law draws on the same picture. Rawls is the liberal outlier who keeps a compatibilist account of agency but denies pre-institutional desert, as Section 5 notes.

White abolitionism. Structurally determinist about race: whiteness is produced by history, law and economic interest, not chosen. Voluntarist about exit: the "race traitor" is someone who chooses to refuse the privileges of whiteness, and Ignatiev's project assumed whites could make that choice. Where it moralizes whiteness as collective guilt, it assigns desert by birth, which neither libertarianism nor compatibilism supports, since no one chooses their birth on either view.

The pattern across all five. No ideology applies one theory of the will consistently, but the asymmetry the pipeline predicts holds in only three of the five. In white supremacy, Stalinist Marxism and liberal law, the theory of the will shifts toward libertarianism where blame is assigned. Classical anarchism's anti-punitive determinism is a counterexample, and white abolitionism's tension is between structure and exit, not between explanation and blame. The pipeline predicts the asymmetry, because the urge to blame comes first and the attribution of free will is built to support it. The experimental evidence is mixed. Clark et al. (2014) found that a motive to punish raised belief in free will; replications found the effect only for attributions to the wrongdoer, and explained it by norm violation rather than a motive to blame (Monroe and Ysidron 2021). A mechanistic account of behaviour reduced retributive punishment (Shariff et al. 2014).

For the framework. Under Sapolskian determinism, the libertarian claims fail outright, and the compatibilist claims are acceptable only as descriptions of how someone acts, not as grounds for basic desert. Applied evenly, that removes the desert each ideology assigns to its opponents, while leaving the causal explanations each one holds.

8. Comparison

	White supremacy	Anarchism	Marxism	Liberalism	White abolitionism
Main input	Status and significance loss	Distrust of coercive institutions	Material deprivation, exploitation	Socialization; threat to the liberal order	Witnessing hierarchy; felt complicity
Main sentiment recruited	Disgust, fear	Outrage at coercion	Outrage at exploitation	Compassion; contempt for those cast as illiberal	Outrage, guilt, shame
Stance-independent claim	Racial hierarchy as nature	Freedom and domination as facts (not in Stirner)	Historical inevitability; class verdicts	Natural rights; self-evident truths; desert	Collective guilt (in moralized forms)
Resources for the framework	None in mainstream forms	Stirner's "spooks"	Morality as ideology (Marx, Engels)	Rawls on desert; Mill's consequentialism	Race as constructed, not perceived
Codified form	Race laws (historical)	Mostly a target of codification	Class-standing laws in Leninist states	The desert institutions the paper critiques	Minimal
Theory of the will (Section 7)	Hereditarian about those it subordinates; libertarian about those it blames	Voluntarist rhetoric; classical theorists near hard determinist on punishment	Compatibilist in theory; libertarian guilt in Stalinist practice	Compatibilist philosophers (Kant aside); libertarian law and discourse	Determinist about the category; voluntarist about exit
Counts all beings' suffering?	No: by definition	In aim; violence and transition costs in practice	In aim; class enemies excluded in Leninist practice	All humans in principle; desert excepted	In aim; at risk where guilt becomes desert
Extremity on the Section 1 definition	High in all forms surveyed: content meets (a), (b) and (c)	Varies by form; high in insurrectionary forms	Varies by form; high in Leninist forms	Meets (a) in founding rhetoric and (c) on desert; low on (b) outside exclusionary codification	Low on (b); (c) at risk where guilt becomes desert

The table reads two ways. Down each column, it diagnoses one ideology. Across the "resources for the framework" row, it shows that four of the five contain resources the framework can use (anti-realism, arguments against desert, or causal explanation). Holding such a thesis has not by itself prevented harm in any of them: Stalinism kept "morality as ideology" while committing mass harm.

9. What follows, and the limits

The pattern. Run through the framework, the five ideologies are not points on a single moral scale. They are different inputs into one pipeline. Each recruits a sentiment, dresses it in stance-independent authority, makes it non-negotiable to some degree, and, where it gains power, codifies it as desert. The ideologies differ most in two things: what they feed into the pipeline, and whose suffering they leave out of the count. On the second, white supremacy is distinct in kind, since its exclusion is definitional rather than a failure of practice.

The symmetry the framework demands. The diagnostician is a node (Section 6 of the paper). An analyst who reads her own revulsion at white supremacy, or her own sympathy for any of the others, as perception of a moral fact has re-entered the loop. The judgment that white supremacy fails impartial altruism stands; it rests on the presumed aim, not on her feeling. What changes is the response: carriers of every ideology here are caused nodes, so the framework recommends changing causes over condemning persons, including where the ideology is the most harmful.

Limits.

- The framework's empirical studies were not designed to test these ideologies. Most measure partisanship, contested issues or radicalization in general. Applying them here is extrapolation.
- Historical claims in this application (laws, events, authors) were checked against primary or standard sources in an adversarial review on 2026-10-19; the references below list them.
- "Extreme" is defined here, not in the paper. A different definition would change the last row of the table.
- Each ideology is a family. A short section per family flattens internal variation, which in each case is large.
- The comparison of harm uses the presumed aim. Someone who rejects impartial altruism would not accept the "counts all beings' suffering" row, and the paper's presumptions mean it does not argue them with such a reader.
- Section 7 assigns each ideology a dominant tendency on free will. Every one contains thinkers on more than one side, and the tendencies describe emphasis, not doctrine.

References

Studies of moralization, networks and intervention follow the reference list of "The Diagnostic Advantage of Moral Anti-Realism" (Brady et al. 2017, 2021; Caruso 2020; Feinberg and Willer 2019; Fisher et al. 2017; Halperin et al. 2011, 2013; Hogg 2021; Hogg, Meehan and Farquharson 2010; Jaśko et al. 2017, 2020; Kalla and Broockman

2020, 2023; Kteily, Hodson and Bruneau 2016; Landry et al. 2023; Mooijman et al. 2018; Moore-Berg et al. 2020; Rozin 1999; Sapolsky 2023; van Oorschot 2000; Webber et al. 2018; Wisneski and Skitka 2017). The works below are cited only here.

Alexopoulos, G. (2003). *Stalin's Outcasts: Aliens, Citizens, and the Soviet State, 1926–1936*. Ithaca: Cornell University Press.

Camus, R. (2011). *Le Grand Remplacement.* Neuilly-sur-Seine: David Reinharc.

Clark, C. J., Luguri, J. B., Ditto, P. H., Knobe, J., Shariff, A. F., and Baumeister, R. F. (2014). Free to punish: A motivated account of free will belief. *Journal of Personality and Social Psychology* 106(4): 501–513. https://doi.org/10.1037/a0035880

Devine, P. G., and Ash, T. L. (2022). Diversity training goals, limitations, and promise. *Annual Review of Psychology* 73: 403–429. https://doi.org/10.1146/annurev-psych-060221-122215

Du Bois, W. E. B. (1935). *Black Reconstruction in America.* New York: Harcourt, Brace, pp. 700–701.

Engels, F. (1878). *Anti-Dühring*, Part I, chs. IX and XI. https://www.marxists.org/archive/marx/works/1877/anti-duhring/

Ignatiev, N., and Garvey, J., eds. (1993–2005). *Race Traitor*, nos. 1–16.

Kropotkin, P. (1877). Prisons and their moral influence on prisoners. In R. N. Baldwin, ed. (1927), *Kropotkin's Revolutionary Pamphlets.* New York: Vanguard Press. https://theanarchistlibrary.org/library/prisons-and-their-moral-influence-on-prisoners

Kropotkin, P. (1902). *Mutual Aid: A Factor of Evolution.* London: Heinemann.

Lukes, S. (1985). *Marxism and Morality.* Oxford: Clarendon Press.

Marx, K. (1852). *The Eighteenth Brumaire of Louis Bonaparte*, ch. 1. https://www.marxists.org/archive/marx/works/1852/18th-brumaire/ch01.htm

Marx, K. (1859). Preface to *A Contribution to the Critique of Political Economy.* https://www.marxists.org/archive/marx/works/1859/critique-pol-economy/preface.htm

Marx, K., and Engels, F. (1845–46; published 1932). *The German Ideology.*

Maximoff, G. P., ed. (1953). *The Political Philosophy of Bakunin.* Glencoe, IL: Free Press.

Mill, J. S. (1843). *A System of Logic*, Book VI, ch. II, "Of Liberty and Necessity."

Monroe, A. E., and Ysidron, D. W. (2021). Not so motivated after all? Three replication attempts and a theoretical challenge to a morally motivated belief in free will. *Journal of Experimental Psychology: General* 150(1): e1–e12. https://doi.org/10.1037/xge0000788

Morissette v. United States, 342 U.S. 246, 250 (1952).

Paluck, E. L., Porat, R., Clark, C. S., and Green, D. P. (2021). Prejudice reduction: Progress and challenges. *Annual Review of Psychology* 72: 533–560. https://doi.org/10.1146/annurev-psych-071620-030619

Rawls, J. (1971). *A Theory of Justice.* Cambridge, MA: Harvard University Press, §§12, 17, 48.

Roediger, D. R. (1991). *The Wages of Whiteness.* London: Verso.

Shariff, A. F., Greene, J. D., Karremans, J. C., Luguri, J. B., Clark, C. J., Schooler, J. W., Baumeister, R. F., and Vohs, K. D. (2014). Free will and punishment: A mechanistic view of human nature reduces retribution. *Psychological Science* 25(8): 1563–1570. https://doi.org/10.1177/0956797614534693

Stirner, M. (1844; title page 1845). *Der Einzige und sein Eigentum*. Leipzig: Wigand.

United States v. Grayson, 438 U.S. 41, 52 (1978).

Statutes: Act of 26 March 1790, ch. 3, 1 Stat. 103; Act of 3 March 1903, 32 Stat. 1213; 8 U.S.C. §1424; Executive Order 13950 (2020); *Loving v. Virginia*, 388 U.S. 1 (1967); Strafgesetzbuch §§86a, 130.

ESSAY 4

Domestic Terrorism Through the Diagnostic Lens

Vincent L. Anderson · Oct 19, 2026

1. Framework and scope

The framework, presumed. This essay presumes the two earlier documents: "The Diagnostic Advantage of Moral Anti-Realism" and its application to five ideologies. It takes as given three positions: moral anti-realism, impartial altruism as the overriding aim, and Sapolskian determinism. It also takes as given the paper's account of hostility as a pipeline: sentiment comes first, moralization recruits disgust and outrage, conflict manufactures felt objectivity, conviction drives intolerance and violence, and networks reward and return the output. Institutions that must codify sentiment state it as desert. The application added two further findings. Extremity is a property of how an ideology is held and enacted. And every ideology attributes free will selectively: determinism where it excuses, free will where it blames.

What the framework asks of terrorism. On these presumptions, an attack is not a free act of evil to be avenged. It is the output of a causal system, with the attacker as one node. The diagnostic question is therefore not "who deserves what?" but "which causes produced this, which amplified it, and which can be changed?" The aim is to reduce the suffering such attacks cause, including the suffering caused by responses to them.

Definitions. Two definitions are used, because the major datasets differ. Cato's 1975–2025 dataset counts premeditated, ideologically motivated killings by non-state actors, and excludes spontaneous hate crimes (Nowrasteh 2026). CSIS counts "the deliberate use or threat of premeditated violence by nonstate actors" for political goals and broad psychological impact, including plots and non-lethal attacks (Byman and McCabe 2025). "Domestic" here means attacks on U.S. soil by people living in the United States, whatever the ideology, including homegrown jihadists. The 9/11 attacks were planned and staffed from abroad and are excluded.

Case selection. Cases were chosen on two criteria the request named. *Most destructive*: the highest death tolls. *High-profile*: attacks whose political effects reached far beyond their death tolls, such as assassinations. Historical racial terror is included, because by death toll it dwarfs the modern record.

2. The record

The modern totals. From 1975 to 2025, 208 attackers killed 3,577 people on U.S. soil in politically motivated attacks, about 0.35% of all U.S. murders (Nowrasteh 2026). Two attacks, 9/11 and Oklahoma City, account for about 88% of those deaths. Excluding both, 186 killers took 430 lives: 45% by right-wing attackers, about 32% by Islamist attackers and about 16% by left-wing attackers. CSIS's count for the past decade gives 112 victims of right-wing attacks, 82 of jihadist attacks and 13 of left-wing attacks. CSIS also found that in the first half of 2025, left-wing incidents (attacks and plots) outnumbered right-wing ones for the first time in more than 30 years, five to one. The numbers are small, and left-wing attacks remain far less lethal. CSIS's data run from 1994 to 4 July 2025 (Byman and McCabe 2025).

The deadliest and most consequential cases.

Year	Attack	Killed	Orientation
1877–1950	Racial terror lynchings in the South	4,084 documented	White supremacist
1921	Tulsa race massacre	Estimated 100–300 (state commission, 2001)	White supremacist
1995	Oklahoma City federal building bombing	168	Anti-government, white nationalist influences
2016	Pulse nightclub, Orlando	49	Jihadist (pledged to ISIS)
2019	Walmart, El Paso	23	White nationalist ("Hispanic invasion")
2015	Inland Regional Center, San Bernardino	14	Jihadist
2025	Bourbon Street truck ramming, New Orleans	14	Jihadist (ISIS-inspired)
2018	Tree of Life synagogue, Pittsburgh	11	Antisemitic white nationalist
2022	Tops supermarket, Buffalo	10	White nationalist ("Great Replacement")
2015	Emanuel AME Church, Charleston	9	White supremacist
2016	Police ambush, Dallas	5	Black nationalist, anti-police
2025	LDS meetinghouse, Grand Blanc, Michigan	4	Anti-Mormon (investigators' reading)
2025	Assassination of Rep. Melissa Hortman and her husband, Minnesota	2	Partisan, anti-abortion (as alleged by prosecutors; CSIS: right-wing)
2025	Assassination of Charlie Kirk, Orem, Utah	1	Partisan (as alleged; Utah's governor described the suspect's "leftist ideology"; after the CSIS data window)

Year	Attack	Killed	Orientation
2021	Capitol attack, Washington, D.C.	Contested; about 140 officers assaulted (DOJ)	Election-denial, anti-government

What the table hides. The 2017 Las Vegas shooting killed 60, more than any modern attack but Oklahoma City, and is not included: the FBI's Behavioral Analysis Unit found "no single clear motivating factor" (FBI 2019). That matters. Of 173 public mass attacks from 2016 to 2020, the Secret Service found that about half were sparked by personal, domestic or workplace grievances, and only 18% by ideological, bias-related or political beliefs (NTAC 2023). The line between terrorism and other mass violence is drawn by motive, and motive is often mixed, thin or added late.

Victims by type. Recent left-wing violence has mostly targeted police and government property: 10 of 13 left-wing victims in the past decade were police officers, and 20 of 35 attacks used arson or incendiaries. Jihadist attacks have mostly targeted crowds. White supremacist attacks have mostly targeted worshippers and shoppers chosen for race, religion or ethnicity (Byman and McCabe 2025).

3. Case analyses

3.1 Racial terror lynching: the institutional branch at its extreme

Lynching is the deadliest domestic terrorism in U.S. history, and the clearest case of the paper's institutional branch. The Equal Justice Initiative documented 4,084 racial terror lynchings in twelve Southern states between 1877 and 1950. Lynchings were often public spectacles, photographed and sold as postcards. The moralization is textbook: accusations of sexual threat recruited disgust and fear (stage 2), and crowds presented the killing as justice. Codification ran in both directions. Courts and sheriffs rarely prosecuted, and Jim Crow law encoded the racial hierarchy the violence enforced, so the institutions served as proof that the hierarchy was real. The Tulsa massacre of 1921 followed the same pattern on a larger scale, beginning with an accusation against a Black teenager and ending with the destruction of the Greenwood district. On the framework, this was not the work of exceptionally evil individuals. It was a system in which sentiment, networks and institutions were aligned to produce killing, and it ended when that alignment changed.

3.2 Oklahoma City (1995): grievance codified as war

Timothy McVeigh and Terry Nichols killed 168 people, including 19 children. McVeigh timed the bombing for the second anniversary of the federal siege at Waco, in which 76 people died, and cited Ruby Ridge as well. He carried pages of *The Turner Diaries*, a white nationalist novel that depicts a truck bombing of a federal building. The pipeline is visible at each stage. The input was a grievance against federal force, shared across a

militia network. Moralization converted it into a story of tyranny and self-defence. And felt objectivity was manufactured by conflict: each federal siege confirmed the militia narrative that the government was at war with citizens. McVeigh described the dead children as "collateral damage" (Michel and Herbeck 2001), the language of a soldier, not a murderer. That is the stage 4 output in pure form: conviction that removes victims from the count. He was executed in 2001. Whether the execution raised his standing in the militia movement is plausible but was not documented for this essay.

3.3 The replacement chain (2015–2022): contagion through networks

Charleston, Pittsburgh, El Paso and Buffalo form a linked series. Dylann Roof wrote that his radicalization began with a web search for "black on white crime" after the Trayvon Martin case. Robert Bowers posted on Gab that a Jewish refugee charity was bringing "invaders" before attacking the Tree of Life synagogue. The El Paso attacker posted a manifesto citing the 2019 Christchurch attack in New Zealand and a "Hispanic invasion." The Buffalo attacker copied parts of the Christchurch manifesto and livestreamed his attack, as the Christchurch attacker had. This is stage 5 made literal: each attack's text and video became the next attacker's input. The sentiment is the disgust-and-fear profile identified in the ideology analysis (status loss converted into a story of theft), and the network is online, anonymous and rewards escalation with notoriety.

3.4 Homegrown jihadist attacks: borrowed ideology, local inputs

The Pulse (2016), San Bernardino (2015) and New Orleans (2025) attackers were U.S. citizens or residents who pledged allegiance to the Islamic State or were inspired by it, without being directed by it. In each case the ideology supplied a ready-made moral frame and a promise of significance to people whose local inputs (personal crisis, isolation, a search for meaning) are the same predictors the paper lists (Webber et al. 2018). The New Orleans attacker was an Army veteran who had posted videos pledging allegiance to ISIS. The framework treats ISIS here less as a cause than as a costume: a moral authority borrowed to dress a sentiment that formed first, which is stage 1 of the pipeline.

3.5 The 2025 assassinations: the partisan loop

On 14 June 2025, Minnesota Rep. Melissa Hortman and her husband were shot dead and a state senator and his wife were wounded. Vance Boelter is charged with the attacks; prosecutors say he carried a list of about 45 officials, mostly Democrats. In September 2025, Charlie Kirk was shot dead at Utah Valley University; Utah's governor said the suspect had "clearly a leftist ideology," and casings were engraved with anti-fascist messages mixed with internet memes. These cases show the loop of stage 5 most clearly. Each killing was immediately counted by partisans as evidence of the other side's violence, and each count became an input to the next round of outrage, which is the exaggerated meta-perception pattern (Moore-Berg et al. 2020). The engravings also

show how online meme culture blurs ideology: the verbal content mixed sincere political messages with irony, which makes the motive hard to read and easy for each side to claim.

3.6 The Capitol attack (2021): codified desert, reversed

The January 6 attack is high-profile rather than highly lethal. It matters to the framework for what happened afterwards. About 1,580 people were charged and many were convicted and imprisoned, with courts stating their sentences as desert. On January 20, 2025, nearly all were pardoned or had their sentences commuted, and supporters described them as patriots and political prisoners. The same acts were codified as deserving punishment, then as deserving release, within four years. On the framework, that reversal shows directly what the paper claims: the desert a court states is a conversion of sentiment, and when the sentiment in power changes, the desert changes with it.

4. Cross-case patterns

Grievance comes first; ideology often comes later. The Secret Service data cover public mass attacks generally, most of them not ideological, so they cannot show when ideology arrives for terrorists; they show the background against which it arrives. Most public mass attackers acted on personal, domestic or workplace grievances; 58% had mental health symptoms, 25% had suicidal thoughts, 41% had a history of domestic violence, and 29% had lost a job (NTAC 2023). Ideology was the stated driver in under a fifth of attacks. Studies of lone-actor terrorists find similar patterns of grievance, mental-health history and leaked intent (Gill, Horgan and Deckert 2014; Corner and Gill 2015). The framework's hypothesis is that ideology often arrives as a ready-made frame for a grievance that has already formed: a moral authority that gives private rage a public meaning and a promise of significance. FBI Director Christopher Wray called the increasingly common mixed motive a "salad bar" of ideologies (Wray 2022).

Moralization makes killing feel like duty. Every manifesto in the replacement chain presents the attack as defence against invasion, and McVeigh presented his as war. That is the stage 4 output the paper describes: conviction that the issue is objectively grounded and non-negotiable, which removes victims from the count. The disgust vocabulary of white supremacist attacks (invasion, replacement, contamination) is the strongest version of the stage 2 route.

Networks and contagion. Attackers cite each other. Christchurch was cited in El Paso and copied in Buffalo; Oklahoma City's model came from a novel. Research on mass shootings has found a contagion period of about 13 days after highly publicised attacks (Towers et al. 2015), and scholars have called on media to stop publicising perpetrators' names and images (Lankford and Madfis 2018). Notoriety is the social reward that stage 5 predicts will be repeated.

Warning signs are usually visible. Three-quarters of public mass attackers caused concern in others beforehand, two-thirds made threatening or concerning communications, and 70% worried family members. But 22% of that concerning behaviour was never reported to anyone able to respond (NTAC 2023). The system produces signals; the gap is in who receives them.

Lethality depends on means. Most of the deadliest modern attacks used firearms; Oklahoma City used a fertilizer bomb, and New Orleans a rented truck. Grievance and ideology determine who attacks; access to means largely determines how many die.

Free will is attributed selectively. The application found that ideologies apply determinism where they excuse and free will where they blame. Public reaction to attacks shows the same asymmetry. The framework predicts that commentators will tend to describe attackers from their own side as mentally ill or as products of circumstance, and attackers from the other side as evil agents who represent their movement. Media framing shows a related asymmetry: white and Latino mass shooters were about twice as likely as Black shooters to be described as mentally ill (Duxbury, Frizzell and Lindsay 2018). Whether partisans make the same split along party lines is plausible but untested here. Under the presumed determinism, both framings are partial: every attacker is a caused node, and no one's causes are excuses that someone else's are not.

The attacker as hero of a desert story. Attackers almost always cast themselves in a desert narrative, punishing a group for what it supposedly deserves. Their opponents then respond in the same form, demanding that the attacker receive what he deserves. On the framework, both are the same conversion of sentiment into desert, aimed in opposite directions.

What would count against the framework. Reading cases through the pipeline after the fact cannot fail, so the framework has to say what it predicts. It predicts that attacks cluster in time after highly publicised attacks (Towers et al. 2015); that executions raise how often later manifestos cite the executed attacker; that court records will usually document a grievance before the attacker adopted an ideology; and that conduct-based enforcement will produce fewer claims of persecution than designation by ideology. Attackers whose ideology demonstrably came first, or executions followed by fewer citations of the executed, would count against it.

5. What a desert-based response has done

The paper predicts that a diagnostic which reads moralized sentiment as perception will locate the problem in culpable individuals, recommend condemnation, and feed the loop. The U.S. response to domestic terrorism illustrates each of these, across administrations.

Retribution as the headline response. Capital prosecution has been the visible response to the deadliest attacks: McVeigh was executed, Roof and Bowers were sentenced to death, and federal prosecutors under the Biden administration sought

death for the Buffalo attacker, whose federal trial began in October 2026. Biden's December 2024 commutations of federal death sentences pointedly excluded Roof and Bowers. On the framework, incapacitation of dangerous people is justified by risk, and needs no desert. Execution adds nothing to incapacitation; its function is to state desert. The framework predicts that it also supplies martyrdom to the attacker's network; that prediction is testable but untested here.

Counting wars. After each high-profile attack, partisans count deaths by ideology to show that the other side is more violent. The data support neither side's strongest claims: right-wing attacks killed the most people over the past decade, ahead of jihadist attacks, left-wing incidents rose sharply in 2025, and both remain a tiny share of U.S. homicide (Byman and McCabe 2025; Nowrasteh 2026). The counting itself is a stage 5 behaviour: each side's tally is offered as proof of the other side's guilt, and it raises the felt stakes of the conflict.

Designation by ideology. On September 22, 2025, the president signed an order designating "Antifa" a domestic terrorist organization; three days later, a presidential memorandum (NSPM-7) directed federal agencies to investigate and disrupt networks behind political violence. Among the "common threads" it attributed to this violence, it listed "anti-Americanism, anti-capitalism, and anti-Christianity" and "hostility towards those who hold traditional American views on family, religion, and morality." Civil-liberties groups, including the Brennan Center, objected that these indicators describe protected beliefs, not violence, and that U.S. law provides no authority to designate domestic organizations (Brennan Center 2025). On the framework, an indicator list built from beliefs is the realist diagnostic in codified form: it treats an opposing standpoint as a moral defect, so it targets the opponent's ideology rather than the causes of violence. It also predicts the next loop, because it confirms for the targeted side the story that it is under attack. The pattern is not one administration's: a 2023 FBI field memo linking "radical-traditionalist" Catholics to violent extremism drew the same objection, that it treated belief as a sign of risk, and was withdrawn (WSET 2023).

Cutting prevention. In 2025, the Department of Homeland Security cut about $18.5 million in grants from its Center for Prevention Programs and Partnerships, which funded local threat assessment, training and mental health services, including about $700,000 for Minnesota's Department of Public Safety, announced in July 2025, about five weeks after the Hortman assassination. A department official said the grants reflected "left-wing ideologies and did next to nothing to combat actual threats" (Hippensteel and Mayorquín 2025). The cut removed funding from leverage points the framework favours; no evaluation has yet measured its effect.

Desert reversed by pardon. The January 6 pardons (Section 3.6) show the same pattern from the other side: desert assigned by one administration's sentiment and withdrawn by the next. On both sides, the codified verdict follows the sentiment, and then each verdict is cited as proof of the moral facts that justify it.

6. Solutions and preventative measures

The framework's rule is to intervene at every stage of the loop, judge each measure by what it does to suffering, and count the suffering caused by the response as well as by the attack. Evidence grades: **moderate** means field or large observational evidence; **preliminary** means case series or small studies; **untested** means a prediction of the framework.

Stage	Measure	What it targets	Evidence
Inputs	Behavioral threat assessment and management teams in schools, workplaces and communities, with clear reporting routes	The 75% of attackers who cause concern beforehand, and the 22% of concerning behaviour never reported	Moderate (NTAC 2023, observational)
Inputs	Mental health, suicide-prevention and crisis services linked to threat assessment	The 58% with mental health symptoms and 25% with suicidal thoughts	Moderate for the overlap; untested as counter-terrorism
Inputs	Significance and belonging through non-ideological channels: work, mentoring, community roles	Significance loss and humiliation as predictors	Moderate in lab and field (Webber et al. 2018); untested at scale
Means	Extreme risk protection orders (red-flag laws) for people making credible threats	Lethality, and the window between threat and attack	Preliminary: none of the threatened shootings occurred in a California case series of 21 (Wintemute et al. 2019); the authors make no causal claim
Means	Background checks, and closing domestic-violence gaps in firearm access	The 41% of attackers with domestic violence histories	Moderate for violence generally; untested specifically for terrorism
Moralization	Elected officials and media avoiding disgust and dehumanizing frames ("invasion," "vermin"), on all sides	Stage 2 moralization	Moderate (Clifford 2019; Wisneski and Skitka 2017)
Amplification	No-notoriety reporting: minimize attackers' names, images and manifestos	Contagion	Preliminary (Towers et al. 2015; Lankford and Madfis 2018)
Amplification	Platform action against manifesto and livestream sharing, and against forums that celebrate attackers	Network reward	Preliminary: an industry campaign against a harassment forum failed to shut it down but cut its activity by about half, with displacement (Vu, Hutchings and Anderson 2023)
Feedback	Correcting exaggerated meta-perceptions between partisans; joint condemnation of violence by leaders of both parties	The partisan loop	Moderate (Moore-Berg et al. 2020; Landry et al. 2023)

Stage	Measure	What it targets	Evidence
Exit	Off-ramp and exit programmes for people in extremist networks	Network ties	Preliminary; rigorous evaluations are scarce
Institutions	Prosecute violence and credible threats by conduct, not by ideology; base sentences on risk and incapacitation, not desert	Martyrdom, and the loop fed by ideological designation	Untested

Prevention before punishment. The data point to one conclusion more than any other. Most attackers send signals beforehand, most to people close to them, and much of it is never passed on. The highest-yield measure is therefore a reporting-and-response system that families, schools and employers trust: one that leads to help and, where needed, temporary removal of firearms, not automatically to prosecution. The framework predicts that families will report less where reporting is likely to send a relative to prison; that prediction has not been tested here.

Treat every ideology's violence by the same rule. The framework requires symmetric treatment. Prevention funding and enforcement should follow the measured risk of violence, whichever ideology carries it. Programmes built to target one side's beliefs will be read by that side as persecution, which is the loop the paper describes.

What the framework adds to standard prevention. Most of these measures are already recommended by researchers who do not share the paper's presumptions. The framework's distinct contributions are three: it denies that attackers' beliefs are free choices deserving punishment, so it prefers risk-based incapacitation to execution; it counts the response as an input to the next attack; and it applies the same analysis to the observer, so the outrage a citizen or official feels after an attack is treated as a caused signal to regulate, not as perception of a moral fact.

7. Limits and conclusion

Limits.

- Counts differ by dataset and definition. Cato excludes spontaneous hate crimes and attackers without clear ideological purpose; CSIS includes plots and non-lethal attacks and excludes the Kirk killing, which falls after its 4 July 2025 cutoff. Ideological coding of recent cases is provisional, and several 2025 and 2026 cases are still before the courts.
- The case narratives use public reporting and court records as summarized in news coverage; historical figures (lynching totals, Tulsa estimates, Oklahoma City details) come from standard sources and were not re-verified for this essay.
- The framework's studies were mostly not designed around terrorism. Applying them here is extrapolation, as the application tab noted.

- Most prevention measures have preliminary evidence at best for preventing terrorism specifically. Attacks are rare, which makes any measure hard to evaluate.
- The essay judges responses by the presumed aim. Readers who hold retribution as valuable in itself will not accept the critique of capital punishment, and the presumptions mean the essay does not argue that point with them.

Conclusion. Through this lens, U.S. domestic terrorism is not a parade of uniquely evil individuals. It is the output of a pipeline that is visible in every case: grievance and significance loss as inputs, an ideology borrowed to give private rage moral authority, networks that reward escalation with notoriety, and access to lethal means. The deadliest episode, racial terror lynching, ran on institutions as much as on individuals. The modern cases run on networks and contagion. The current response leans on the moral realist diagnostic: execution, ideological designation, partisan counting and cuts to prevention, each of which the framework predicts will feed the next round. The alternative is less dramatic: threat assessment that families trust, temporary removal of firearms from people making credible threats, reporting that denies attackers notoriety, leaders who refuse dehumanizing language, symmetric enforcement by conduct, and sentencing by risk rather than desert. None of it requires believing that attackers deserve nothing. It requires recognizing that what they deserve is not the question that prevents the next attack.

Sources

Web sources were read on 2026-10-19. Studies carried over from the framework paper are cited as in its reference list; all citations were verified in an adversarial review on 2026-10-19.

Hippensteel, C., and Mayorquín, O. (2025). New York Times, 31 August 2025, syndicated as "Before Minnesota shooting, a program aimed at preventing attacks lost federal funding." https://www.hawaiitribune-herald.com/2025/08/31/nation-world-news/before-minnesota-shooting-a-program-aimed-at-preventing-attacks-lost-federal-funding

Brennan Center for Justice (2025). "Trump's orders targeting antifascism aim to criminalize opposition." https://www.brennancenter.org/our-work/research-reports/trumps-orders-targeting-antifascism-aim-criminalize-opposition

Byman, D., and McCabe, R. (2025). "Left-Wing Terrorism and Political Violence in the United States: What the Data Tells Us." CSIS, 25 September 2025. https://www.csis.org/analysis/left-wing-terrorism-and-political-violence-united-states-what-data-tells-us

Duxbury, S. W., Frizzell, L. C., and Lindsay, S. L. (2018). Mental illness, the media, and the moral politics of mass violence. *Journal of Research in Crime and Delinquency* 55(6): 766–797. https://doi.org/10.1177/0022427818787225

Equal Justice Initiative. *Lynching in America: Confronting the Legacy of Racial Terror.* https://eji.org/reports/lynching-in-america/

Lankford, A., and Madfis, E. (2018). Don't name them, don't show them, but report everything else. *American Behavioral Scientist* 62(2): 260–279. https://doi.org/10.1177/0002764217730854

NSPM-7, "Countering Domestic Terrorism and Organized Political Violence" (25 September 2025), summary and criticism. https://en.wikipedia.org/wiki/NSPM-7

Nowrasteh, A. (2026). "Politically Motivated Killers: 51 Years of Terrorist Murders on US Soil, 1975–2025." Cato Institute Policy Analysis No. 1012. https://www.cato.org/policy-analysis/politically-motivated-killers-51-years-terrorist-murders-us-soil-1975-2025

Towers, S., Gomez-Lievano, A., Khan, M., Mubayi, A., and Castillo-Chavez, C. (2015). Contagion in mass killings and school shootings. *PLOS ONE* 10(7): e0117259. https://doi.org/10.1371/journal.pone.0117259

NTAC: U.S. Secret Service National Threat Assessment Center (2023). *Mass Attacks in Public Spaces: 2016–2020*. Washington, DC: U.S. Secret Service; summarized by ASIS. https://www.asisonline.org/security-management-magazine/latest-news/today-in-security/2023/january/study-identifies-common-behaviors-mass-attackers/

Vu, A. V., Hutchings, A., and Anderson, R. (2023). No easy way out: The effectiveness of deplatforming an extremist forum to suppress hate and harassment. arXiv:2304.07037 (formal venue not confirmed). https://arxiv.org/abs/2304.07037

Wintemute, G. J., Pear, V. A., Schleimer, J. P., Pallin, R., Sohl, S., Kravitz-Wirtz, N., and Tomsich, E. A. (2019). Extreme risk protection orders intended to prevent mass shootings: A case series. *Annals of Internal Medicine*. 171(9): 655–658. https://doi.org/10.7326/M19-2162. Summary: https://www.eurekalert.org/pub_releases/2019-08/acop-css081519.php

WLOS / Sinclair (2025). "Officials investigate motive behind Charlie Kirk assassination." https://wlos.com/news/nation-world/officials-investigate-motive-behind-charlie-kirk-assassination-heres-what-we-know-suspect-tyler-robinson-utah-valley-university-murder-federal-state-charges-president-trump-governor-cox-fbi-patel

Wikipedia, "List of terrorist incidents in 2025" and "…in 2026." https://en.wikipedia.org/wiki/List_of_terrorist_incidents_in_2025

From the framework paper: Clifford 2019; Landry et al. 2023; Moore-Berg et al. 2020; Webber et al. 2018; Wisneski and Skitka 2017.

Corner, E., and Gill, P. (2015). A false dichotomy? Mental illness and lone-actor terrorism. *Law and Human Behavior* 39(1): 23–34. https://doi.org/10.1037/lhb0000102

FBI (Federal Bureau of Investigation), Behavioral Analysis Unit (2019). *Key Findings of the BAU's Review of the October 1, 2017, Mass Shooting in Las Vegas.* Reported at https://www.ktvu.com/news/fbi-finds-no-specific-motive-in-vegas-attack-that-killed-58

Gill, P., Horgan, J., and Deckert, P. (2014). Bombing alone: Tracing the motivations and antecedent behaviors of lone-actor terrorists. *Journal of Forensic Sciences* 59(2): 425–435. https://doi.org/10.1111/1556-4029.12312

Michel, L., and Herbeck, D. (2001). *American Terrorist: Timothy McVeigh and the Oklahoma City Bombing.* New York: ReganBooks.

Oklahoma Commission to Study the Tulsa Race Riot of 1921 (2001). *Final Report.* https://www.okhistory.org/research/forms/freport.pdf

Reason (2024). "Biden commutes nearly all federal death penalty sentences," 23 December 2024. https://reason.com/2024/12/23/biden-commutes-nearly-all-federal-death-penalty-sentences/

theGrio (2026). Buffalo shooter federal death penalty trial, 14 October 2026. https://thegrio.com/2026/10/14/buffalo-shooter-federal-death-penalty-trial/

U.S. Department of Justice, U.S. Attorney's Office for the District of Columbia. Capitol Breach Cases. https://www.justice.gov/usao-dc/capitol-breach-cases

Wray, C. (2022). Testimony on worldwide threats, as reported in "Will 'salad bar' extremism replace old-school world terrorism?" FDD, 14 July 2022. https://www.fdd.org/analysis/2022/07/14/will-salad-bar-extremism-replace-old-school-world-terrorism/

WSET (2023). "FBI retracts document calling 'radical-traditionalist Catholics' a domestic terrorism threat." https://wset.com/news/nation-world/fbi-retracts-document-calling-radical-traditionalist-catholics-a-domestic-terrorism-threat

ESSAY 5

After Us: A Deep-Ecological Posthumanism

Vincent L. Anderson · Oct 19, 2026

1. What this essay does

The task. The three earlier essays presumed moral anti-realism, impartial altruism and Sapolskian determinism, and used them to diagnose hatred, ideology and domestic terrorism. This essay takes their implications one step further and makes the case, as its defenders would make it, for a deep-ecological posthumanism with two commitments: the abolition of ideology, and voluntary human extinction. It argues the case at full strength in Sections 2 to 6, then sets out the strongest objections in Section 7 and says in Section 8 what survives them.

The constraint that defines the position. "Voluntary human extinction" here means one thing only: people freely choosing not to have children, so that the human population declines to zero over generations through ordinary deaths. It is the position of the Voluntary Human Extinction Movement, whose motto is "May we live long and die out" (Knight, in Miller 2011). It does not mean ending any existing life, one's own included; it does not mean coercing anyone's reproduction; and it does not mean violence of any kind. Section 6 shows that these exclusions are not a disclaimer added from outside. They follow from the framework itself, which counts every existing person's suffering and predicts that moralized ecological certainty becomes violent.

How the essay uses "case." The framework presumes that there are no stance-independent moral facts. So the essay does not claim that human extinction is objectively right. It argues that, for someone who holds the impartial aim the earlier essays presume, voluntary non-procreation is a reasonable and possibly the most consistent response to the evidence. Whether that argument succeeds is a question about consequences, which Section 7 tests.

2. What the earlier essays commit us to

Five commitments carry forward.

1. **No moral fact privileges humans.** On moral anti-realism, human exceptionalism, the view that human interests count for more because humans are special in kind, is not a discovered fact. Where it is felt as obvious, the framework reads it as a sentiment dressed in borrowed universality (the paper, Sections 1 and 6).
2. **The aim counts every sentient being.** Impartial altruism counts the suffering and well-being of all sentient beings, opponents included. The application essay noted that

all five ideologies it examined largely exclude non-human animals from that count, and treated this as a failure against the aim.

3. **No one deserves anything basic.** Under Sapolskian determinism, no human or animal is the ultimate author of what it does. Human beings cannot deserve to flourish at other species' expense, and they cannot deserve to be punished for having done so.
4. **Ideology is moralization machinery.** The paper's pipeline (sentiment, moralization, felt objectivity, conviction, amplification, codification) describes how human groups reliably turn feeling into claimed fact and claimed fact into hostility. The ideology essay found the pipeline in every ideology it examined, including those that contain anti-realist insights of their own.
5. **The pipeline kills, and repair is slow.** The terrorism essay found that the deadliest domestic violence in U.S. history ran on moralized certainty, and that the best-supported interventions have small to moderate effects, often untested at scale (the paper, Section 5; the terrorism essay, Section 6).

Taken together, these commitments shift the question. The usual question is how human societies should treat nature. The posthumanist question is what an impartial observer, with no stake in the human standpoint, would want from the human species as a whole.

3. Deep ecology and posthumanism, rebuilt without moral realism

Deep ecology as it was stated. Arne Naess distinguished a "shallow" ecology movement, concerned with pollution and resource depletion for the sake of people in developed countries, from a "deep" movement that rejects the human-centred image of nature altogether (Naess 1973). The platform Naess and George Sessions drafted in 1984 opens: "The well-being and flourishing of human and nonhuman life on Earth have value in themselves." Its fourth point reads: "The flourishing of human life and cultures is compatible with a substantial decrease of the human population. The flourishing of nonhuman life requires such a decrease" (Devall and Sessions 1985).

What the framework removes. "Value in themselves" is a moral realist claim: it presents a sentiment, reverence for living things, as a property the world contains. On the framework, it is borrowed universality, and the paper's Section 6 predicts that it will amplify whatever it dresses, including contempt for the humans who threaten the valued thing. Deep ecology's own history bears this out: the language of intrinsic value has been taken up by eco-fascists who rank nature above human lives (Section 6 below). So the framework keeps deep ecology's conclusions about scale and drops its metaphysics.

What the framework puts in its place. Non-human beings count not because value inheres in them, but because the presumed aim counts every sentient being's suffering and well-being. That grounding is narrower than deep ecology's. It counts sentient animals, not rivers, species or ecosystems as such. Ecosystems matter on this view

because sentient beings live in them. The narrower grounding is also more stable, because it does not ask anyone to perceive a value in nature; it asks them to apply an aim they already hold to beings they already know can suffer.

Posthumanism. Critical posthumanism, as Rosi Braidotti describes it, rejects the idea of "Man" as the measure of all things and treats the human as one node in a web of living and technological systems (Braidotti 2013). It is distinct from transhumanism, which aims to enhance humans, and from misanthropy, which hates them. The framework supplies posthumanism with a method it often lacks. The paper already treats the individual human, the diagnostician included, as a caused node in a system. Extending that view from the social system to the biosphere is the posthumanist step: the human species becomes one causal force among others, to be assessed by its effects on all sentient beings, not by its own self-image.

The resulting position. A deep-ecological posthumanism, on these terms, holds three things. Human interests carry no extra weight because they are human. The human species is assessed, like any other cause, by what it does to sentient life. And the moralized self-images that justify human dominance, from divine mandate to natural right, are outputs of the pipeline, not reports of fact.

4. The abolition of ideology

What is abolished. "Ideology" has many senses. Mannheim treated it as thought bound to a social position (Mannheim 1929); Bell predicted its end in the West as political conflict became technical (Bell 1960), a prediction later events did not bear out. The framework's sense is narrower and more precise: an ideology is a system that converts sentiment into claims presented as stance-independent fact, makes them non-negotiable, and codifies them as desert. To abolish ideology in this sense is not to abolish values, aims, preferences or arguments. It is to stop presenting them as features of the world that opposing standpoints simply fail to perceive.

The programme. The earlier essays already contain it, scattered.

- **In each person:** treat one's own moral outrage as a caused signal to be checked for amplification, not as perception (the paper, Section 7).
- **In conflict:** separate factual disputes, which evidence can settle, from disagreements between standpoints, which no stance-independent fact settles, and correct exaggerated meta-perceptions (Moore-Berg et al. 2020; Landry et al. 2023).
- **In institutions:** replace desert and moral-character tests with decisions stated by risk, need and cause (the paper, Section 5; the terrorism essay, Section 6).
- **In public language:** refuse the disgust and dehumanization frames that drive moralization (Clifford 2019), and enforce against violence by conduct, never by belief.

The reflexive problem. An obvious objection is that "abolish ideology" is itself an ideology. On the framework's definition, it is one only if it is held as a stance-independent truth, made non-negotiable and codified as desert. Held as an aim with reasons, open to evidence about its consequences, and advanced only by argument and by conduct-based law, never by belief tests or desert, it is not. The programme must therefore apply to itself first: its adherents should expect to feel their own position as obviously correct, and should treat that feeling as the pipeline at work.

Why the abolition of ideology leads to the next section. The pipeline begins with intuition-first moral judgment and disgust-driven moralization (Haidt 2001; Rozin 1999), features of evolved human psychology. De-moralization works against them. The paper reports that its best-supported interventions have small to moderate effects, and the terrorism essay found that the strongest prevention tools are only preliminarily tested. If ideology can be reduced but not abolished in a species whose cognition produces it, the impartial observer must ask what follows for the species. That is the question of Section 5.

5. The case for voluntary human extinction

The argument has four premises and a conclusion. Each premise draws on the earlier essays or on current evidence.

P1. The aim counts all sentient beings impartially. This is presumed (Section 2).

P2. Humanity inflicts more deliberate suffering on other sentient beings than any other agent. It does so chiefly through farming: more than 70 billion land mammals and birds were slaughtered in 2020 (FAOSTAT, as compiled by Faunalytics 2022), roughly three-quarters of farmed land animals by one estimate in factory farms (Anthis 2019). Humanity's other ecological effects, an average 73% decline in monitored wildlife populations since 1970 (WWF 2024) and the crossing of six of nine planetary boundaries (Richardson et al. 2023), are harms to populations and systems whose net effect on total suffering is contested (see Objection 1). None of this is the work of a few bad actors; it is the aggregate output of billions of caused human choices, which is how the framework says causes should be read.

P3. The pipeline that produces human hostility is a stable feature of the species. The earlier essays found it in every ideology examined, in every era of domestic terrorism examined, and in the institutions that respond to violence. Its interventions work, but modestly. The defender concludes that a human population will keep generating moralized hostility toward humans and indifference toward non-humans for as long as it exists, at a scale that tracks its size. This is an extrapolation beyond what the earlier essays measured, and it is in tension with Section 4: to the extent that de-moralization succeeds, P3 weakens.

P4. Every new life is created without its consent, and every human life includes suffering. This premise does not need David Benatar's stronger claim that coming into existence is always a harm (Benatar 2006). It needs only the weaker point that a decision to create a person imposes on that person a life whose suffering is certain and whose benefits are not, and that the decision is made by others for their own reasons.

P5. The expected balance of suffering over well-being, across all sentient beings, is worse in a future with humans than in one without them, and no less drastic course reduces it comparably. This premise is empirical, the earlier essays do not establish it, and it is the one Section 7 attacks.

C. For someone who holds the aim, choosing not to have children is a reasonable and possibly the most consistent response; and if enough people made that choice freely, a gradual, voluntary decline of the human population to zero would end human-caused suffering without killing anyone, through non-procreation alone, and, if P5 holds, would reduce net suffering.

Why extinction and not only reduction. Deep ecology's fourth point asks for "a substantial decrease" of the human population, not its end. The defender of voluntary extinction replies that any remaining human population retains the capacities that produced the present situation: the pipeline, the institutions and the technology. A smaller humanity can grow again. As Les Knight put it, "As long as there's one breeding couple we could be right back where we are now" (Miller 2011). On this reading, reduction postpones the problem, and voluntary extinction, by non-procreation alone, ends it.

Why this is not misanthropy. The position contains no verdict that humans are evil, deserve to disappear or have forfeited anything. Under determinism no one deserves anything basic, and the framework diagnoses misanthropy as disgust-driven moralization like any other. The case rests on consequences for every sentient being, humans included: the humans who would never be born would also never suffer. The movement's motto expresses the intended attitude: "May we live long and die out."

6. Why only voluntary: the framework's guardrails

Ecological ideas have been turned into violence before, and the framework explains how. The El Paso attacker's manifesto mixed environmental degradation with racist claims about immigration (E&E News 2019); the Christchurch attacker called himself an eco-fascist (Forchtner 2019); the Finnish writer Pentti Linkola argued for an authoritarian "eco-dictatorship" and drastic population reduction; Ted Kaczynski killed three people in the name of an anti-technology programme. Coercive population policy by states, justified by development and resource arguments, from China's one-child policy to the mass sterilization campaign of India's Emergency in 1976, imposed suffering on millions. Deep ecology itself is not the source: Naess held a Gandhian commitment to non-

violence, and eco-fascism borrows deep ecology's language while rejecting its egalitarian points. Any case for human extinction must show why it cannot become any of these. The framework supplies four reasons, and none depends on a rule added from outside.

1. Every existing person counts. The aim that grounds the case counts the suffering of every sentient being alive now. Killing, injuring or coercing people causes exactly the suffering the argument exists to reduce, and it causes it to beings who already exist and can already suffer. Non-procreation, by contrast, directly harms no existing person. The asymmetry is the whole point: the case concerns people who would never exist, not people who do.

2. No one deserves extinction. Under determinism, no human has forfeited anything by being part of the cause. Eco-fascism depends on a desert story, that some people (usually others, often immigrants) are a contamination that deserves removal. That is the disgust route of stage 2 and the desert conversion of the institutional branch, which the framework rejects at every step.

3. Moralized ecology follows the pipeline. The paper predicts that a sentiment presented as stance-independent fact, held as non-negotiable, licenses intolerance and violence (Wright, Cullum and Schwab 2008; Mooijman et al. 2018). "The Earth has intrinsic value and humans are destroying it" is as capable of that conversion as any other belief. The position survives only if it is held as an aim, not as a revealed truth, which is what Section 4 requires.

4. Coercion defeats the aim on its own terms. Coercive population control has produced suffering, resistance and backlash wherever it has been tried. On the framework's own consequence test, it fails. Even a coercion that "worked" would be excluded by point 1, since it imposes suffering on existing people by definition; the consequence test is a second reason, not the only one.

What the position therefore is, and is not. It is an argument that each person may weigh in deciding whether to have children, together with support for conditions that make that decision free: access to contraception and equality for women, which VHEMT itself argues for (Miller 2011), together with education. It is not a case for ending any existing life, one's own included. Nothing in this essay applies to a person's decision to go on living. Despair about the state of the world is a reason to seek support from people you trust or from a crisis line, not an application of the argument.

7. The strongest objections

The framework requires that its own conclusions be tested by consequences. These objections are stated at full strength, most serious first.

1. Wild-animal suffering: extinction removes the only possible helper. This is the most serious objection, because it uses the case's own premise. The aim counts all sentient beings, and most suffering among sentient beings is not caused by humans. Most animals in the wild belong to species that produce vast numbers of offspring, most of whom die soon after birth from starvation, predation, parasitism or exposure (Horta 2010). That suffering began long before humans and would continue for hundreds of millions of years after them. Humans are the only known agents that could ever understand it and act to reduce it. On this objection, human extinction ends human-caused suffering but removes the only possible remedy for the much larger suffering in nature, and an impartial observer would therefore want a reformed humanity, not an absent one.

2. Irreversibility under uncertainty. The case rests on contested empirical premises (P3 and P5) and contested philosophy (Benatar's asymmetry). Reform can be corrected if it proves wrong; extinction cannot. The paper's own error-correction criterion (Section 7) favours options that keep correction possible, so an impartial observer under uncertainty has strong reason to prefer the reversible path and to keep the agents who could later discover they were wrong. The long run is also uncertain in both directions: humans may one day create very large numbers of digital or artificial sentient beings, which could multiply well-being or suffering by orders of magnitude. That uncertainty strengthens the case for caution about any irreversible verdict.

3. The aim also counts well-being. Impartial altruism, as the paper defines it, seeks both less suffering and more well-being. Most human lives contain a great deal of well-being by their own report. Benatar's asymmetry, that the absence of pain is good even if no one enjoys that good, whereas the absence of pleasure is not bad unless there is someone for whom that absence is a deprivation, is disputed by many philosophers. Without it, P4 weighs only the certainty of some suffering against the likelihood of much well-being, and does not settle the question.

4. Voluntary extinction selects against itself. If the people who accept the argument have no children, the next generation is drawn from those who reject it. Values that run in families, through upbringing and possibly through heritable dispositions, then become rarer. A voluntary strategy cannot reach zero; it can only shift the composition of humanity toward people less concerned with the argument's aims, including concern for animals.

5. Better levers reach the same harms faster. The harms in P2 come from what humans do, not from their existence as such. Industrial animal farming can be ended by changes in diet, technology and law. Habitat loss can be reversed. Global fertility is already falling, and the United Nations projects the world population to peak at about 10.3 billion in the mid-2080s and then decline (UN DESA 2024). A path to extinction would take more than a century even with universal agreement; a path to reform could reduce the same suffering within decades.

6. The transition has costs. A population that shrinks toward zero ages. Its last generations would face care burdens, collapsing services and the loss of the institutions that protect the vulnerable, which is suffering the aim counts.

7. The reception problem. The ideology essay found that "abolish whiteness" is often heard as a call to eliminate people, whatever its authors meant. "Human extinction" invites the same misreading, and eco-extremists have already shown that the misreading can kill. On the framework's own consequence test, promoting the idea publicly may cause more harm than the voluntary choices it persuades.

8. The case is conditional. On moral anti-realism, the argument binds only those who hold the impartial aim. It gives no reason to someone whose aims are partial, for example someone who values their own descendants above strangers. The framework cannot call that person mistaken, only differently aimed.

8. What the case establishes, and what remains disputed

The defender's reply. A defender would answer the objections as follows. To objection 1: the hope that humans will one day reduce wild-animal suffering is speculative, while the suffering humans cause now is measured; and a species that has not ended its own factory farming is a poor candidate to manage nature's suffering. To objection 4: the argument asks only for individual choices, so it succeeds to the extent that it is adopted, whatever happens to the totals. That reply concedes the objection's substance: by the case's own consequence test, a choice that shifts the future toward people less concerned for animals may count against itself. To objection 5: reform and voluntary decline are not rivals, and a person can support both. The defender also notes that Sections 5 and 6 make no demand on anyone; they describe a choice that each person is free to decline.

What survives. Three conclusions survive the objections in some form.

- **The abolition of ideology,** in the framework's sense of de-moralization, survives as an aim, though Section 4 does not show that full abolition is feasible. It follows from the earlier essays, it is consistent with any outcome of the extinction debate, and Section 4's reflexive test keeps it from becoming an ideology itself.
- **A permissive case for voluntary childlessness** survives. For someone who holds the impartial aim, choosing not to have children is a reasonable response to the evidence in P2 and P3, and it directly harms no existing being.
- **The deep-ecological reweighting** survives: non-human sentient beings count impartially, each according to its capacity for suffering and well-being, and the human species is assessed by its effects on them.

What remains disputed. The strong claim, that impartial altruism requires human extinction, does not survive the wild-animal objection unless one accepts that human help for animals in the wild will never come, or Benatar's asymmetry, or both. Those are open empirical and philosophical questions, not settled by the framework. Even if they were settled in the case's favour, the irreversibility objection would counsel against any verdict that cannot be corrected. The selection objection also shows that voluntary extinction, adopted by some, changes who inherits the future more than whether there is one.

The posthumanist conclusion. The firmest result is the stance the earlier essays have been building toward: to see the human species as one caused force among others, to judge it by what it does to every being that can suffer, and to hold even that judgment as an aim rather than a revealed truth. Whether that stance ends in a reformed humanity or a voluntary exit is a question its holders will answer differently, and the framework requires them to keep counting the consequences while they do.

References

Web sources were read on 2026-10-19. Works cited from the earlier essays (Clifford 2019; Haidt 2001; Landry et al. 2023; Mooijman et al. 2018; Moore-Berg et al. 2020; Rozin 1999; Wright, Cullum and Schwab 2008) follow the reference list of "The Diagnostic Advantage of Moral Anti-Realism."

Bell, D. (1960). *The End of Ideology.* Glencoe, IL: Free Press.

Benatar, D. (2006). *Better Never to Have Been: The Harm of Coming into Existence.* Oxford: Oxford University Press.

Braidotti, R. (2013). *The Posthuman.* Cambridge: Polity.

Devall, W. (Bill), and Sessions, G. (1985). *Deep Ecology: Living as if Nature Mattered.* Salt Lake City: Gibbs M. Smith. (Prints the eight-point platform drafted by Naess and Sessions in 1984.)

E&E News (2019). "Suspect's screed mixes environmentalism, racist rants." https://www.eenews.net/articles/suspects-screed-mixes-environmentalism-racist-rants/

Faunalytics (2022). Global animal slaughter statistics and charts: 2022 update, compiled from FAOSTAT. https://faunalytics.org/global-animal-slaughter-statistics-charts-2022-update/

Horta, O. (2010). Debunking the idyllic view of natural processes: Population dynamics and suffering in the wild. *Télos: Revista Iberoamericana de Estudios Utilitaristas* 17(1): 73–88. https://revistas.usc.gal/index.php/telos/article/view/284/250

Mannheim, K. (1929). *Ideologie und Utopie.* Bonn: Cohen. English translation (1936), *Ideology and Utopia.*

Miller, A. (2011). "This guy wants you dead" (interview with Les U. Knight, Voluntary Human Extinction Movement). *Vice.* https://www.vice.com/en/article/this-guy-wants-you-dead/

Naess, A. (1973). The shallow and the deep, long-range ecology movement: A summary. *Inquiry* 16(1–4): 95–100. https://doi.org/10.1080/00201747308601682

Richardson, K., et al. (2023). Earth beyond six of nine planetary boundaries. *Science Advances* 9(37): eadh2458. https://doi.org/10.1126/sciadv.adh2458

UN DESA (2024). *World Population Prospects 2024: Summary of Results.* New York: United Nations.

WWF (2024). *Living Planet Report 2024.* Gland: WWF. https://www.worldwildlife.org/press-releases/catastrophic-73-decline-in-the-average-size-of-global-wildlife-populations-in-just-50-years-

reveals-a-system-in-peril
Anthis, J. R. (2019). *Global Farmed and Factory Farmed Animals Estimates*. Sentience Institute. https://www.sentienceinstitute.org/global-animal-farming-estimates
Forchtner, B. (2019). Eco-fascism: Justifications of terrorist violence in the Christchurch mosque shooting and the El Paso shooting. *openDemocracy*. https://www.opendemocracy.net/en/countering-radical-right/eco-fascism-justifications-terrorist-violence-christchurch-mosque-shooting-and-el-paso-shooting/

APPENDIX

Adversarial review and revisions

19 October 2026

Review conducted on 2026-10-19. Each essay was checked by an independent reviewer that held the framework's presumptions fixed and tested citations, facts and argument structure. Every revision listed here was applied to both the live documents and this collection.

How the review worked

Four reviews ran: one on the thesis and its earlier citation review, one on the ideology application, one on the domestic terrorism essay, and one on the posthumanism essay. Reviewers verified claims against primary sources where they could reach them (court opinions, statutes, journal records, the original reports) and against standard secondary sources otherwise. They returned three kinds of finding: factual or citation errors, weaknesses in the structure of the argument, and problems of balance.

The thesis: "The Diagnostic Advantage of Moral Anti-Realism"

Citations. All 51 DOIs resolved and matched their Crossref records. Three claims were wrong as written and were corrected.

- The phrase "primary justification for the death penalty," cited through Caruso (2020), comes from *Spaziano v. Florida* (1984), where the Court was summarizing a petitioner's argument it then rejected. It was replaced with the joint opinion in *Gregg v. Georgia* (1976), which describes capital punishment as "an expression of society's moral outrage."
- The Medicaid work requirement under Public Law 119-21 can start as late as the end of 2028 under good-faith exemptions, and includes half-time education and an income alternative. The dates and terms were corrected.
- Wright and Pölzler's (2022) humility finding is correlational. "Humility reduced intolerance" became "humility was associated with less intolerance."

Smaller corrections: Moore-Berg et al. (2020) measured prejudice and dehumanization, not "hostility"; Fisher et al. (2017) compared cooperative with competitive argument and had no baseline; the Fourcade and Healy print citation and the Sapolsky entries were fixed.

Argument. Four structural changes were made.

1. *Desert held fixed.* The "moral realist diagnostic" had been defined to include desert and condemnation, so several comparisons were true by stipulation. Section 2 now states that the comparison holds desert fixed, since determinism already excludes it.
2. *Circularity acknowledged.* The abstract now says that the difference in kind follows from the presumptions, and that the evidence shows the loop exists and does harm, not that the realist diagnostic is blind to it.
3. *Felt universality.* Section 6 now states that the anti-realist diagnostic does not abolish felt universality in its user; it removes the warrant for treating that feeling as evidence.
4. *The borrowing objection.* Section 8 now concedes that a realist who denies moralized sentiment any evidential role closes the gap, so the thesis is about diagnostics that treat sentiment as evidence, not about realism as metaphysics.

Not resolved. Two quotations attributed to Fourcade and Healy (2017), and the USCIS Policy Manual wording, could not be verified against full text during the review. The companion paper, "Diagnosis Without Desert," is not included in this collection; the one passage that depended on it was rewritten to stand alone. The word "drastic," which the review suggested replacing because the size of the advantage is unmeasured, was kept as the author's term, with the paper's own qualification that it is drastic in kind and unmeasured in size.

The earlier citation review

Stale references to "paper two" were replaced, and its summary of Zijlstra (2019) was aligned with the paper, which treats Zijlstra's mind-independence result as counter-evidence.

"Applying the framework: five ideologies"

Facts. Nine corrections, including: the *Race Traitor* journal ran from 1993 to 2005, not from 1992; the Marx and Engels quotations were restored to their published wording; Rawls's "arbitrary from a moral point of view" is in §12, and he rejected desert for distributive shares only; the *Morissette* and *Grayson* quotations were corrected in wording and framing; Article 58 criminalized actions, not intent; the claim about diversity training now cites Devine and Ash (2022) and Paluck et al. (2021). A full reference list was added; the essay previously had none.

Argument and balance.

- The essay's own definition of "extreme" had not been applied evenly. The table now shows liberalism meeting two of the three criteria in some respects, and the claim that white supremacy is extreme "inherently" is restated in terms of the definition.

- Exclusion of non-human animals had been counted against liberalism alone. It is now noted for all five.
- Codification *against* each ideology is now reported for all five, not only for anarchism.
- The free-will pattern, that ideologies are determinist where they excuse and libertarian where they blame, holds for three of the five, not all. The section now says so, and reports that Clark et al. (2014) has a partial failed replication (Monroe and Ysidron 2021).

"Domestic Terrorism Through the Diagnostic Lens"

Facts. The Cato study is dated 2026, not 2025. The DHS grant-cut article was by the *New York Times*, not the AP. Duxbury et al. (2018) found that white *and Latino* shooters were more often framed as mentally ill than Black shooters. The Capitol attack count is about 1,580 charged and 140 officers assaulted. The Hortman case is now described as alleged, matching the treatment of the Kirk case.

Argument and balance.

- *Misused data.* Secret Service data on mass attacks, most of them not ideological, had been used to show when ideology arrives for terrorists. That claim is now marked as a hypothesis, supported by studies of lone-actor terrorists (Gill, Horgan and Deckert 2014; Corner and Gill 2015).
- *One-sided Section 5.* Section 5 had criticized only one administration. It now includes the Biden Justice Department's pursuit of the death penalty in Buffalo, the exclusion of Roof and Bowers from the 2024 commutations, and the withdrawn 2023 FBI memo on "radical-traditionalist" Catholics.
- *Unfalsifiable readings.* Reading cases through the pipeline after the fact could not fail. A new paragraph states what the framework predicts and what evidence would count against it.
- *Causal claims.* Claims that the funding cuts moved resources toward enforcement, and that executions create martyrs, are now stated as untested predictions.

"After Us: A Deep-Ecological Posthumanism"

Argument. The review found the argument incomplete, and four changes were made.

- *Missing premise.* The step from the premises to the conclusion needed a comparison of worlds, so a new premise (P5) states it and marks it as empirical and unproven.
- *P2 overreach.* The evidence for P2 now separates deliberate suffering, mainly farming, from ecological harms whose net effect on suffering is contested.
- *Irreversibility.* The strongest missing objection was added: extinction cannot be corrected and reform can.
- *Overclaims.* Section 8's account of what "survives" was narrowed.

Safety. Every statement of the conclusion now specifies non-procreation alone. The exclusion of coercion no longer depends only on whether coercion would work. The essay distinguishes deep ecology from the eco-fascism that borrows its language, and it says plainly that nothing in it applies to any person's decision to go on living.

Tooling

The first build, a single template rendered from the exported text, had four defects. It had no checks that citations matched references. It left export artifacts in the text (people and date chips, and a placeholder where a diagram had been). It had no contents page, page numbers or running heads. And it had no link between corrections and the live documents. The second build fixed all four.

- **One list of revisions** drives both the documents and the PDF: each correction is applied to the source text and sent to the live document as the same find-and-replace.
- **A citation check** flags any in-text citation without a reference entry, plus leftover drafting notes and export artifacts.
- **A sync check** compares fresh exports of the live documents with the build text before compiling.
- **A typeset template** adds a cover, contents, essay openers, page numbers and the redrawn diagram.