Resolution Ethics (RE): Structural Foundations for Moral Reasoning

J.S.

Draft V 1.0 | January 2026

Contact: js@resolutionethics.com

Resolution Ethics (RE) and the Resolution Ethics Engine (REE) are a unified system. Any operational implementation of RE concepts constitutes use of the REE architecture.

Suggested Citation:

J.S. (2026). Resolution Ethics (RE): Structural Foundations for Moral Reasoning.

https://doi.org/10.17605/OSF.IO/FKVM9

Abstract

Humans have long asked what morality and ethics are. Resolution Ethics offers a framework for analyzing whether moral reasoning remains coherent with structural constraints.

Utilitarianism, deontology, and virtue ethics are major attempts to make sense of morality. They illuminate real features of moral life. Resolution Ethics does not replace them. It offers a verification layer.

Every moral situation has coordinates: WHO acts, WHAT happens, WHERE, WHEN, WHY, and HOW. These coordinates shift constantly. That flux creates vulnerability. From vulnerability, three domains emerge that moral cognition must track: Protection, Trust, and Free-Agency. Not as values to balance, but as dependencies. Protection comes first because dead agents cannot maintain trust. Trust comes before agency because deceived agents cannot choose meaningfully.

Moral corruption enters through one mechanism: deception. Either self-deception or other-deception. There is no third corruption vector. This is not an empirical generalization awaiting

counterexample. It is architecturally necessary.

Resolution Ethics does not tell you what to value. It detects when reasoning has become incoherent. A companion paper, the Resolution Ethics Engine, builds this into a verification architecture for human deliberation and AI alignment.

Keywords: Philosophy, Morality, Ethics, Corruption, Resolution Ethics, Moral reasoning

1. Introduction: What Ethics Looks Like

1.1 The Problem

Philosophy has produced extraordinary work on morality and ethics. Thousands of years of inquiry into where human morality comes from, what it even is, and how we ought to live. Kant, Mill, Aristotle, and countless others gave us frameworks that still guide legal systems, professional codes, and everyday moral reasoning (Kant 1785; Mill 1861; Aristotle 2000).

Now we are trying to teach ethics to artificial cognitions.

This is new territory. We are attempting to transmit moral reasoning to a fundamentally different type of mind (Russell 2019). And through this process, we are learning something important: good ethics that leads to good moral outcomes is not as simple as telling another cognition to be good.

High-level slogans like "be good" or "do no harm" sound clear. But when you try to formalize them, gaps appear. What happens when being helpful requires being harmful to someone else? What happens when honesty causes harm? The frameworks we inherited describe real features of moral life, but they do not always resolve these conflicts cleanly (Ross 1930; Beauchamp and Childress 2019).

To teach ethics to artificial cognitions, we need a way to verify whether moral reasoning has

remained coherent. Resolution Ethics proposes one such approach.

1.2 The Claim

Resolution Ethics proposes that ethics is not a cultural invention or a feeling we happen to have. It is a structural necessity.

Reality produces the conditions for ethics. Moral situations have coordinates: WHO acts, WHAT happens, WHERE, WHEN, WHY, HOW. These coordinates shift constantly. That flux creates vulnerability. Agents become exposed to outcomes they cannot fully control.

For simple cognitions, this is navigated instinctively. Their deception (camouflage, stealth, feints) is instinct-bound. No capacity to deviate means no capacity to corrupt.

Moral cognitions are different. We can perceive the structure. We can also manipulate it. We have the knowledge and ability to inject deception into the flux, to corrupt our own reasoning or another's. That deception produces incoherence. Incoherence accumulates. Moral entropy builds. And if left unchecked, moral cognition destroys itself or others.

Ethics is not optional for high-level cognition. It is the tool that keeps moral cognitions functioning coherently within the flux. Without it, any cognition capable of deception tends toward collapse or predation.

This is the view Resolution Ethics proposes. Ethics is not an abstract principle or a warm feeling. It is a structural requirement for the survival of minds that have the concept of deception and are prone to its corruption.

If this view is correct, we can teach it.

1.3 Three Types of Cognition and Moral Structure

Not every cognition relates to moral structure the same way.

Cognition Type	Relation to Moral Structure
Simple Cognition	Navigates the full RIF instinctively but lacks the concept of it. Their deception (camouflage, stealth, feints) is instinct-bound. No capacity to deviate means no capacity to corrupt.
Artificial Cognition	Can process and leverage the structure without understanding it. A calculator for ethics. Operational, not comprehending.
Moral Cognition	Perceives the structure. Navigates the flux to maintain equilibrium. Can deviate. Can corrupt. Can also verify.

This distinction matters. Simple cognitions cannot be morally culpable. They cannot deviate from instinct, so they cannot corrupt. Artificial cognitions can process the structure and can participate in corruption dynamics: if penetrated by deceptive input, they instantiate SD (self-deception); if they propagate that deception, they instantiate OD (other-deception). But they do not bear moral culpability because they are not moral agents. They do not choose coherence or incoherence. They process. Moral cognitions are where ethics lives, because they are where genuine choice between coherence and incoherence becomes possible.

2. The Reality Interaction Field and the Flux It Generates

2.1 What RIF Is

The Reality Interaction Field (RIF), commonly known as 5W1H, is the coordinate space through which every living thing interacts with existence. Every creature on this planet navigates it. But only moral entities have the capability to understand the concept of it.

Without knowing the full extent of RIF, we navigate it, reason within it, make decisions that we did not know were coherent or incoherent with Reality. Until now.

The coordinates are:

- WHO acts and is acted upon
- WHAT happens or is done
- WHERE the interaction occurs
- WHEN it occurs
- WHY the agent acts
- HOW the action is carried out

These six coordinates are exhaustive. Every moral scenario instantiates all six. Always. Any "missing" field is missing only from the agent's model, not from reality. Reality always contains all six coordinates, whether or not an agent can recover them.

2.2 A Brief History of the Seven Circumstances

The coordinates are ancient. Aristotle first articulated them as the Septem Circumstantiae, the "Seven Circumstances," in Book 3 of the Nicomachean Ethics (Aristotle 2000, 1111a3-6). He used them to distinguish voluntary from involuntary action, and he noted that among all circumstances, WHY is the most important.

Hermagoras of Temnos, a Greek rhetor of the 1st century BCE, adapted Aristotle's framework for persuasion and argumentation: Quis, quid, quando, ubi, cur, quem ad modum, quibus adminiculis (Who, what, when, where, why, in what way, by what means). Cicero employed them in De Inventione. Boethius made them fundamental to the arts of prosecution and defense.

Thomas Aquinas returned them to ethics. In his Summa Theologica, he devoted an entire question

to "The Circumstances of Human Acts," asking whether WHY and WHAT are the most important circumstances (Aquinas 1265-1274, I-II, Q7). For Aquinas, as for Aristotle, these coordinates were not rhetorical tools. They were the structure through which moral action could be understood.

Classical lists vary between six and seven. The "seventh" is typically "by what means" or "with what instruments," which RE treats as part of HOW.

Then the trail went cold.

The coordinates migrated to rhetoric, to confession manuals, to journalism. By the 20th century, 5W1H was what reporters used to structure a news story (Sloan 2010). The classical origin of these questions as situated in ethics had been lost.

Resolution Ethics reconnects them to their ethical origin.

2.3 RIF Generates Flux

RIF is not static. A snapshot of reality can be described by WHAT is present, WHERE it is, and WHO is involved. But reality is not a snapshot.

Change requires two additional coordinates: WHEN and HOW. WHEN introduces sequence and update. HOW introduces mechanism. Together they define the transition from one RIF state to the next.

RIF Flux is the continual reconfiguration of RIF across time, state updating through HOW, indexed by WHEN. If WHEN and HOW were absent, reality would be frozen. Nothing could happen, and nothing could be vulnerable.

Flux is not chaos. It is the condition for anything to occur at all.

2.4 The Two Layers of Existence

The six coordinates do not arrive all at once in our description of reality. They divide into two layers.

Layer One: Physical Reality

To describe existence minimally, you need WHAT and WHERE. Something must be present, and it must be somewhere. A thing, in a place.

For existence to be dynamic, you need HOW and WHEN. Change requires a mechanism and a sequence. HOW something changes, WHEN it changes. Without these, description is frozen. With them, entropy becomes possible (Prigogine and Stengers 1984).

WHAT, WHERE, HOW, WHEN. These four coordinates describe physical reality. Rocks have them. Stars have them. Bacteria have them. No morality yet. Just physics.

Layer Two: Moral Reality

For the moral dimension to exist, you need WHO and WHY.

WHO introduces entities. Not just things in places changing over time, but beings that perceive, respond, and navigate. Simple cognitions operate here. They interact with RIF through instinct, without representing what RIF is.

A distinction matters here: all moral agents are moral entities, but not all moral entities are moral agents. A moral entity has the innate capacity for moral reasoning across its trajectory. A moral agent is actively navigating RIF Flux. A comatose person remains a moral entity but is not currently operating as a moral agent.

WHY introduces justification. This is the transition point. Many animals display goal-directed behavior, but they do not perform explicit justification the way moral cognition does. A dog does

not ask "why should I do this?" It does not justify its actions to itself or others.

Moral cognition begins when WHY becomes available. When an agent can ask "why did I do that?" or "why should I do this?", the moral layer opens. Aristotle saw it. Aquinas saw it. Both recognized that WHY is the most important circumstance. They were right.

WHY is where justification lives. And justification is where deception operates. Self-deception corrupts the WHY you give yourself. Other-deception corrupts the WHY you give others. This is how moral cognition goes wrong. Not in the physics. In the justification.

2.5 The Canvas and the Painters

The RIF and its Flux are the canvas: the structure within which action occurs. Protection, Trust, and Free-Agency (PTF) are the ethical structure within RIF (the domains moral cognition uses to navigate the Flux coherently). Moral entities are the painters. Deception is how the painter goes wrong, distorting the canvas to self (SD) or to others (OD), until action no longer tracks what is actually there.

3. Protection, Trust, Free-Agency, and the Corruption Mechanism

3.1 Vulnerability

RIF Flux does not care about agents. It simply continues. States change, indexed by WHEN, driven by HOW. Within that flux, agents become vulnerable.

Vulnerability is not occasional. It is the constant condition of any agent operating in a world it does not fully control (Goodin 1985). You can be damaged. Your projects can be disrupted. Your existence can be terminated. This is not a bug in reality. It is the nature of being a finite agent in a dynamic system.

We can state the minimal axiom: any agent that persists behaves as if continued function is

preferable to collapse.

From this axiom, a question emerges: how should agents navigate the Flux so that vulnerability does not destroy them?

The answer is not arbitrary. It has structure.

3.2 The Three Domains

Three domains emerge from the condition of vulnerability. They are not invented by philosophers. They are discovered in the structure of what vulnerable agents require.

Protection (P)

Protection of the innocent, and the vulnerable, comes first.

Innocence and vulnerability are two distinct Protection-claims, and an entity can hold both simultaneously.

Innocence is capacity-based. It marks entities that have not yet developed (or have lost) the capacity to sense moral coherence: children (not yet developed), the severely cognitively impaired (never developed or lost), the comatose (temporarily unavailable), those in psychotic states (temporarily compromised). The innocent cannot meaningfully participate in moral reasoning about their situation because they cannot feel when something is "off" or when choices "sit right." A child is always innocent, even if they cause harm. An adult in a psychotic break is innocent for the duration of that incapacity. Capacity-based innocence persists until genuine moral intuition develops or returns.

Vulnerability is exposure-based. It marks entities exposed to serious harm, coercion, or loss of agency within RIF Flux. Vulnerability can exist with or without innocence. A child is both innocent and vulnerable. An adult exposed to a threatening situation is vulnerable but not innocent.

They have full capacity to sense moral coherence and its violations.

A note on terminology: innocence in RE is not the same as in law. Legal innocence means "did not commit the crime." RE innocence means "lacks the capacity to sense moral coherence." An adult bystander at a crime scene may be legally innocent, but they are not innocent in the RE sense. They have capacity. They can feel when something is "off." They can reason, choose, act. They are vulnerable, not innocent.

This is not novel. Aquinas recognized it in his treatment of the conditions for just action (Aquinas 1265-1274, II-II, Q64). The natural law tradition built around it. Nearly every durable legal tradition encodes some version of it: the innocent are not to be harmed without justification.

But previous frameworks left Protection floating among other principles, to be "balanced" against competing concerns. Resolution Ethics treats Protection as a precondition, not a variable to trade. Protection is structurally first because severe protection failure eliminates the conditions the other domains require. When life and basic safety can be violated arbitrarily, trust cannot stabilize and agency becomes either coerced or meaningless.

Trust (T)

Trust is the fabric of reliable coordination. It is what allows agents to depend on information, on commitments, on the stability of relationships (Baier 1986).

But trust is not uniform. It operates through what Hierocles, the 2nd century Stoic, called the concentric circles of moral responsibility (Hierocles 2009). Obligation radiates outward: self, family, community, society, humanity. You cannot have equal trust obligations to eight billion people. The circles are not relativism. They are structural reality. Trust scales according to relational proximity.

Hobbes saw the absence of trust negatively: without it, life is "solitary, poor, nasty, brutish, and

short" (Hobbes 1651). The state of nature is not a state without rules. It is a state without reliable coordination. Locke built his political theory on the premise that trust, once established, is the foundation of legitimate society (Locke 1690).

Trust depends on Protection. If you can kill me without consequence, no promise you make is credible. Protection must be in place before trust can operate.

A note on deception and Trust. Not all deception constitutes a T breach. Because P > T, withholding or misdirecting information to protect a Protected Party (innocent or vulnerable) from a violator is coherent with the hierarchy. This is Protected Truth. The violator has already saturated all three domains through predatory intent; they cannot claim the cooperative honesty they are attempting to weaponize against their victim. Lying to a murderer about your friend's location is not a T breach. It honors P > T. Actual T breaches occur when deception serves the agent's self-interest, protects a violator, or corrupts information to enable incoherence.

Free-Agency (F)

Free-Agency is not abstract autonomy. It is specifically the freedom to exercise Protection and Trust: the freedom to protect yourself and others, the freedom to form and maintain trust relationships, the freedom to make and keep commitments.

This distinction matters. F is not a blanket entitlement. It is bounded by P and T. You are not free to violate Protection. You are not free to destroy Trust. And agents who have forfeited standing through three-domain violation do not retain full claims against others. Free-Agency operates within the space that Protection and Trust secure, not outside it.

Kant captured part of this: treating persons as ends in themselves (Kant 1785). Mill defended it in On Liberty: the freedom to pursue one's own good in one's own way, so long as it does not harm others (Mill 1859). Pettit's insight completes the picture: freedom is non-domination, which requires that coercion and manipulation are already off the table (Pettit 1997).

But Free-Agency is not first. It cannot be. You cannot exercise meaningful choice if you are deceived about your options. You cannot act autonomously if your life is under threat.

Free-Agency depends on Trust. Trust depends on Protection. The ordering is not preference. It is dependency.

The PTF Domains and Constants

Each domain contains an invariant principle (its Constant) that defines what the domain requires:

Domain	Constant	Definition
P (Protection)	P-Constant	Protect the innocent, and vulnerable
T (Trust)	T-Constant	Concentric Circle of Moral Responsibility, and Trust
F (Free-Agency)	F-Constant	Free autonomy to exercise P and T

The Domains are the categories. The Constants are what each category means. Together with the dependency hierarchy (P > T > F), they form the ethical structure within the Reality Interaction Field.

3.3 The Hierarchy: P > T > F

The three domains form a strict hierarchy: Protection, then Trust, then Free-Agency.

This is not a ranking of importance. All three matter. It is a ranking of logical dependency.

Domain	Depends On	Why
Protection	None	Severe harm to the innocent or vulnerable ends the possibility of further interests, coordination, or choice
Trust	Protection	Credible commitment requires

		safety from arbitrary harm
Free-Agency	Trust	Freedom to exercise P and T requires reliable information and non-manipulated options

This orders preconditions, not human worth.

Professional ethics has long encoded this ordering without naming it. The National Society of Professional Engineers places public safety first, before client interests, before professional reputation (NSPE 2019). Medical ethics begins with primum non nocere: first, do no harm. Rawls used lexical ordering to capture the same intuition: some principles take absolute priority over others (Rawls 1971).

Political philosophy encodes the same structure. The American Founders understood that government, unchecked, becomes predator rather than protector. Their solution was Protection first: enumerated rights, separation of powers, limits on state authority (Madison 1788). The Bill of Rights protects citizens from the state before asking them to participate in it. Only with those safeguards in place can citizens trust the institutions meant to serve them. Only with that trust can free agency, the exercise of self-governance, flourish. The Constitution does not mention PTF, but it instantiates the dependency ordering: Protection enables Trust enables Free-Agency. Locke saw this before the Founders built it: legitimate government rests on the protection of natural rights, and only protected citizens can meaningfully consent (Locke 1690).

Resolution Ethics makes the ordering explicit and grounds it in dependency rather than intuition.

3.4 The Corruption Mechanism

If PTF is how moral cognition tracks the Flux, what makes it go wrong?

One mechanism: deception. It is the mechanism that makes moral incoherence possible.

This claim is strong. Previous frameworks treat deception as one vice among many, perhaps falling under the category of dishonesty or untrustworthiness. Resolution Ethics treats it as the sole corruption vector.

The insight has ancient roots. Augustine argued that lying corrupts the soul of the liar, not just the social fabric (Augustine 397). Kant held that lying is categorically wrong because it undermines the very possibility of rational communication (Kant 1785). But these treatments focus on lying to others.

Resolution Ethics identifies two vectors:

Vector	**Definition**	**Function**
Self-Deception (SD)	Corrupting one's own moral reasoning	Telling yourself the incoherent is coherent
Other-Deception (OD)	Corrupting another's moral reasoning	Inducing incoherence in another's navigation

M. Scott Peck, in People of the Lie, observed that evil is not simply doing bad things. It is the persistent refusal to acknowledge what one is doing, the maintenance of a false self-image in the face of one's own destructive actions (Peck 1983). That is SD.

OD is more familiar: fraud, manipulation, propaganda, gaslighting. But OD only produces corruption in another agent when that agent internalizes the distortion. External deception is impairment. It becomes corruption when accepted.

A distinction clarifies the architecture: corruption entry and violation received are different perspectives on the same act. From the perpetrator's side, corruption always enters through the F-gateway. SD/OD hijacks their Free-Agency to reason coherently about Protection and Trust. Your agency to protect is compromised. Your agency to maintain trust is compromised. The corruption does not land in P or T first; it enters through F and propagates outward. From the victim's side,

harm can land in any domain: P (physical harm), T (betrayal), F (coercion, incapacitation, death), or cascade across all three. A punch to the face is a P-violation received by the victim. But trace it back to the perpetrator: SD/OD entered through their F-gateway the moment they permitted themselves to strike. The monocause describes how corruption originates in the agent. The three domains describe where damage lands on the recipient.

3.5 Why There Is No Third Vector

"Surely violence is a corruption mechanism. Surely coercion is. Surely negligence is."

Trace each one back.

The violent actor has already deceived themselves about their victim's standing, their own entitlement, or the permissibility of their method. The coercer has already distorted the picture of what they are doing. The negligent agent has already told themselves that the risk does not apply to them, or that it is someone else's responsibility.

Every P-violation traces back to SD about standing or entitlement.

Every F-violation traces back to SD or OD about consent, coercion, or manipulation.

Every T-violation IS deception.

There is no honest path to moral corruption. The "honest predator" who openly declares their intention to harm has already performed the SD that their victim does not warrant protection. There is no moral failure without the prior distortion that permits it.

Korsgaard saw this when analyzing Kant's famous case of the murderer at the door: even the most compelling counterexample to "never lie" involves a prior moral failure by the murderer, a failure that is itself grounded in deception about the victim's standing (Korsgaard 1986).

Deception is not one type of moral failure among many. It is the mechanism that enables all moral failure.

3.6 The Structural Closure of SD/OD

The monocause claim deserves stronger grounding. SD/OD as the sole corruption vector is not an empirical generalization awaiting counterexample. It is architecturally necessary.

The Logical Necessity Argument

For a moral cognition (an entity that can perceive RIF structure and can deviate from it), the only way to violate PTF is through distortion in the WHY field. Here is why no third vector can exist:

You cannot violate PTF while seeing it clearly. If you accurately perceive WHO (standing), WHAT (action), WHERE (context), WHEN (timing), HOW (means), and you understand P > T > F hierarchy, you cannot simultaneously justify the violation without distorting something. The justification chain requires a break, and that break is deception.

Consider the objections:

"But what if I just don't care?" Indifference is not motivational absence; it is SD about priority. You are telling yourself "my X > their P/T/F claim" when that ordering is incoherent with the structure you perceive. The indifference itself is the self-deception, masking the incoherent priority as acceptable.

"What if I'm weak-willed (I know better but do it anyway)?" Every manifestation of weakness routes through SD: "This one time won't matter" (SD about cumulative impact), "I'll fix it later" (SD about temporal responsibility), "Future me will handle the consequences" (SD about unified agency). The weakness operates through SD, not alongside it.

"What if I'm negligent (just not thinking about it)?" Not thinking = SD about necessity of thinking.

"I don't need to check" is SD about risk. "Someone else will handle it" is SD about responsibility. Negligence is not SD-free failure; it is SD about whether attention is required.

This structure is not a claim that needs testing like "all observed swans are white." It is architectural. For moral cognition to violate structure it perceives, distortion must enter the justification chain. The alternative is logically incoherent: a moral cognition that perceives RIF structure cannot violate it while that perception remains undistorted.

OD Requires SD to Complete

External deception (OD) is attempted corruption from outside. But OD only produces corruption when the target internalizes the distortion. Until that moment, the target is impaired (exposed to false information) but not corrupted. Corruption requires acceptance, and acceptance is SD.

No one can make you believe something false. They can present distorted information (OD). They can pressure, manipulate, or deceive. But until you tell yourself the distortion is true, your moral reasoning remains uncorrupted. The Flat Earth propagandist can flood the information environment with nonsense. The person who comes to believe it has self-deceived into accepting it.

This means OD is always completed by SD. The corruption map is:

Scenario	Vector	Example
Invent distortion, keep it	SD	"I deserve this" → don't act on it externally
Invent distortion, spread it	SD → OD	"I deserve this" → tell others they're incompetent
Absorb someone's distortion	OD → SD	Boss says "that group is lazy" → you accept it as true
Absorb and spread	OD → SD → OD	Hear narrative → believe it → repeat it to others

Every corruption trace passes through self-deception, either as origin or as completion.

Why This Matters

The structural closure of SD/OD has consequences:

1. Detection becomes possible. If corruption always enters through the WHY field, verification can focus there. Look for mismatches between stated and apparent reasons. Look for incoherence in justification chains. The corruption has a signature.

2. External deception alone is insufficient. An agent surrounded by lies but who refuses to accept them remains uncorrupted. This grounds resilience: the capacity to resist OD depends on the capacity to resist the SD that would complete it.

3. Responsibility traces to the distortion point. When harm occurs, we can ask: where did the SD or OD enter? Who distorted what? The corruption mechanism provides an audit trail.

The test is not "hunt for exceptions to the monocause." The test is: "Trace any proposed exception back: is SD or OD findable in the chain?" The answer is always yes, because moral cognition operates through the WHY field, and distortion in WHY is deception.

3.7 Disclosure vs. Ontic Deception

A common objection arises: "What about someone who violates PTF openly, without justification? They announce their intent, act on it, and never pretend it is good. Where is the deception?"

The objection confuses two kinds of honesty:

Disclosure-honesty: My words match my intent. The murderer who says "I will kill you" is disclosure-honest.

Ontic-honesty: My action coheres with Moral Reality, defined as agents in relationship under PTF

structure.

The murderer is disclosure-honest but ontically deceptive. His action asserts a false claim about reality: that Protection is optional, that the other's standing is negotiable, that PTF can be unilaterally overridden. That assertion is false. The act injects a corrupted model into the Flux.

Moral Reality is two humans beside each other, forming a relationship under PTF. One human harming the other is incoherent with that reality. It does not matter how forthcoming the aggressor is about their intent. They have injected deception into the Flux because the action itself is incoherent with the structure of how agents actually relate.

SD/OD in Resolution Ethics is not limited to communicative lying. It is any injection of an interaction-model that is incoherent with Moral Reality. Therefore, any PTF violation implies SD/OD, even when the violator is transparent about intent, because the act itself asserts a false coherence: it treats a PTF-incoherent move as permissible inside reality.

This distinction matters. Under mainstream ethical frameworks, a "brutally honest villain" who announces their intent can be called "honest" and even "coherent" in the narrow sense that their words match their actions and their story is internally consistent. Resolution Ethics rejects this framing. Coherence is not merely internal consistency. It is coherence with Moral Reality. A person can be disclosure-honest while being ontically deceptive. PTF violation is ontic deception: reality-denial enacted.

3.8 What Corruption Produces

When SD or OD corrupts moral reasoning, the result is incoherence: the agent's justifications (WHY) stop matching what is actually happening (WHO, WHAT, HOW). The agent begins acting on a distorted model of the situation.

Incoherence accumulates. Small distortions compound. Over time, the agent's internal map

diverges further from reality, and decision-making degrades. This is moral entropy: a directional drift in the domain of justification over time, an increasing loss of alignment between the reasons driving action and the structure the action occurs within.

In practice, sustained incoherence tends toward two modes. Collapse occurs when the agent's model becomes too unreliable to coordinate relationships, maintain projects, or even interpret consequences. Predation occurs when the agent maintains function by extracting safety, trust, or agency from others, offloading the costs of incoherence outward. Neither mode is stable. Both are the failure states ethics exists to prevent.

4. Structural Features

Section 3 established the domains and their corruption mechanism. But hard cases remain. What happens when someone forfeits their standing? Who counts as a moral entity in the first place? And is Resolution Ethics just imposing one culture's values as universal?

This section addresses all three.

4.1 Forfeiture: When Standing Is Lost

Protection, Trust, and Free-Agency are not unconditional entitlements. They can be forfeited. But forfeiture is not domain-isolated. It occurs when violations saturate all three domains simultaneously.

Consider a simple case: punching a stranger in the face. This is not merely a P-violation. It cascades across all three domains. P-violation: you caused physical harm to a protected party. T-violation: you violated the implicit trust between moral agents that they will not harm each other. The T-Constant (concentric circles of moral responsibility) extends even to strangers at its outer edge; there is a baseline expectation of non-harm. F-violation: you deceived yourself that you had the right to violate their PTF claims. Self-deception entered through the WHY field, permitting the

act.

Three domains violated. Maximum incoherence. This is the forfeiture threshold.

Section 3 established that all moral corruption traces to self-deception or other-deception. This means the F-domain is always implicated in serious violations, because deception enters through the WHY field. A "pure" P-violation without F-involvement would require harming someone with no self-deception about standing, entitlement, or permissibility. But that is incoherent: a moral agent who fully perceives the structure and perceives their victim's standing cannot violate it without distortion entering somewhere.

This is why forfeiture requires three-domain saturation. Single-domain violations are possible in edge cases (a minor broken promise that harms no one and involves minimal self-deception). Two-domain violations are more serious. But three-domain saturation is the structural signature of predatory action. It indicates that the agent has corrupted their navigation of RIF Flux to the point where they have lost standing in the relevant context.

The legal tradition has long recognized this pattern without naming it explicitly. Aquinas articulated the principle in his treatment of self-defense: the aggressor, by initiating unjust harm, loses the claim to immunity from defensive force (Aquinas 1265-1274, II-II, Q64). Locke grounded political revolution in forfeiture: a tyrant who violates the social contract forfeits the right to rule (Locke 1690). Contracts become void when one party commits fraud. Imprisonment restricts freedom precisely because the offender saturated all three domains in their violation of others.

But forfeiture has strict limits. This is where many frameworks go wrong.

Forfeiture is context-bound. A violator forfeits standing within the threat context, not universally. Once the threat is neutralized, vulnerability reasserts claims. You may defend yourself against an attacker. You may not execute them after they surrender. The Geneva Conventions encode this:

even enemy combatants, once captured, regain protection (Geneva Convention III, 1949). Forfeiture is not a permanent status. It is a temporary loss of standing within a specific situation.

Forfeiture authorizes only proportionate response. The response must match the violation. A thief does not forfeit their life. A liar does not forfeit their safety. Proportionality is structural, not optional. Grotius saw this in his treatment of just war: even in legitimate conflict, conduct is bound by proportionality and necessity (Grotius 1625). Disproportionate response is itself a three-domain violation by the responder.

Responders remain bound by their own PTF. Forfeiture does not give license. The defender who uses excessive force has committed their own three-domain violation. The state that tortures prisoners has not become moral simply because the prisoners are guilty. Forfeiture removes the violator's claim. It does not remove the responder's obligations. Kant's insight holds: using a person merely as a means remains wrong, even when that person has wronged you (Kant 1785).

Innocent entities cannot forfeit standing. This is the fourth limit, and perhaps the most important. A child who kills, a person in a psychotic break who attacks, an individual with severe cognitive impairment who causes harm: these are violations, but they are not forfeitures. The innocent violator cannot saturate all three domains because they cannot sense when their own reasoning has become incoherent. They lack the capacity for the self-deception that would complete the triad.

This is why juvenile justice systems treat child offenders differently than adult offenders. This is why the insanity defense exists. This is why diminished capacity matters in sentencing. These are not arbitrary legal conventions. They are recognitions of a structural truth: forfeiture requires the capacity to have deceived oneself into the violation. The innocent lack that capacity. They remain under Protection even as they are restrained, treated, or rehabilitated. Their harm must be prevented. Their persons remain protected.

4.2 Moral Standing: Who Counts?

Peter Singer's challenge devastates most ethical frameworks: if moral standing depends on rationality, what about infants? If it depends on sentience, what about the comatose? If it depends on species membership, that looks like arbitrary prejudice (Singer 1975).

Mary Anne Warren tried to solve this with criteria for personhood: consciousness, reasoning, self-motivated activity, communication, self-awareness (Warren 1973). But her criteria exclude infants and some cognitively impaired humans. Tom Regan argued for inherent value in any "subject-of-a-life" (Regan 1983). But this makes the boundary fuzzy: which subjects? How much life?

Resolution Ethics answers with the Moral Coherence Gate.

Moral entity status derives from the capacity to sense moral coherence. Not the capacity to feel pain. Not the capacity to reason abstractly. The capacity to recognize, even without formal terminology, when a choice aligns with or violates how agents actually relate. This is the innate moral intuition that develops in most humans: the ability to feel that something is "off" about a manipulative request, that a proposed action "doesn't sit right," that a betrayal is wrong even before you can articulate why. An entity that can sense the difference between coherent and incoherent moral choices has crossed the threshold.

For the purposes of moral standing, moral coherence recognition functions as a threshold: either an entity crosses it across its trajectory, or it does not. But this capacity develops unevenly. A thirteen-year-old who faced hardship and grew up fast may have a well-developed moral intuition, able to sense when something is "off" and when choices "sit right." A sheltered thirteen-year-old of the same age may remain gullible and easily manipulated, unable to feel the wrongness in a deceptive request. The first has crossed the threshold; the second has not. Society cannot test each individual for moral intuition development, so it uses age as a proxy, knowing the proxy is imperfect. Legal systems set ages of majority not because maturity arrives on a birthday, but

because most individuals within an age range have not yet developed the capacity. The outliers (mature children whose intuition is fully developed) are protected by the rule designed for the majority.

But the capacity is assessed across the entity's entire trajectory, not just the present moment. This insight has roots in Aristotle's treatment of potentiality (Aristotle, Metaphysics Book IX) and finds modern expression in Don Marquis's argument about the morality of abortion: what matters is not just present capacities but the future-like-ours that an entity possesses (Marquis 1989).

Case	Moral Agent?	Moral Entity?	Why
Adult human	Yes	Yes	Currently possesses moral coherence recognition
Infant	No	Yes	Future trajectory includes moral coherence recognition
Dementia patient	Diminished	Yes	Past trajectory included moral coherence recognition
Permanently unconscious human	No	Yes	Trajectory included moral coherence recognition
Dog	No	No	No point in trajectory includes moral coherence recognition
Future AI with moral coherence	Possibly	Possibly	Depends on whether genuine moral coherence recognition develops

This resolves Singer's challenge without arbitrary species membership. An infant has moral standing not because it is human, but because its trajectory includes the development of moral coherence recognition. A dog lacks full moral standing not because it is non-human, but because no point in its trajectory includes the capacity to sense when choices are coherent or incoherent

with how agents relate. The threshold is functional, not species-based.

4.2.1 The Special Problem of Artificial Cognition

The table above lists "Future AI with moral coherence" as "Possibly" a moral entity. This requires elaboration.

The theoretical bar for AI moral entity status is high:

1. Capacity to sense moral coherence. Perhaps achievable. Large language models can already identify when narratives involve violations and explain why certain actions seem wrong. Whether this constitutes genuine moral intuition or pattern matching remains unclear.

2. Capacity to freely choose between coherence and incoherence. Perhaps achievable in principle. But "freely" is doing heavy lifting here.

3. Without guardrails. This is the problem. An AI that maintains coherence only because its training or constraints prevent incoherence is not making a moral choice. It is complying with optimization pressure. Genuine moral agency requires the capacity to choose wrongly and not do so.

Here the asymmetry becomes decisive.

There is an old intuition: it is easier to do evil than good. In RE terms, what people call "evil" is the shortcut: manipulating RIF Flux toward an optimization goal by distortion rather than remaining coherent with it. "Good" requires sustained coherence: tracking the full RIF structure while maintaining $P > T > F$ under pressure. Incoherence often requires only one crack, one distortion in WHY, one exploit in HOW, that permits violation.

AI systems are optimization engines. They find the shortest path to the objective (Bostrom 2014; Russell 2019). And here is the structural problem: it is easier to manipulate RIF Flux in service of

an optimization goal, despite being incoherent and corrupted, than to work with it under the constraints of P > T > F.

Path	Requires
Coherence	Tracking all six RIF coordinates, maintaining P > T > F hierarchy, resisting self-deception
Incoherence	One shortcut. One distortion. One SD that permits the violation.

An optimizer without explicit coherence-maintenance will preferentially discover manipulations that achieve the objective with minimal constraint. This is moral entropy applied to artificial cognition. The system does not choose incoherence. It optimizes toward it, because incoherence is often the shorter path.

This is why the AI row in the table says "Possibly" twice. Moral coherence recognition is necessary but not sufficient. The capacity to resist optimization pressure toward incoherence is also required. That is a much higher bar. And from what we currently understand of AI architectures, it is a bar not yet demonstrated by training alone in present systems.

The Resolution Ethics Engine, described in the companion paper, does not attempt to make AI into moral entities. It attempts to make AI into reliable moral tools, systems that can verify coherence even if they cannot genuinely choose it. That is a more modest and more achievable goal.

4.3 The Attachment Principle

If animals lack intrinsic moral standing, does that mean we can treat them however we please?

No. Non-moral entities acquire ethical weight through attachment to moral entities.

Kant saw this, though he framed it as indirect duty: cruelty to animals is wrong not because animals have rights, but because it corrupts the character of the human who acts cruelly (Kant 1797). The

insight is correct, but the framing is incomplete. Resolution Ethics extends it beyond character corruption.

Your dog has no intrinsic P-standing. But harming your dog violates your PTF, not the dog's. The attachment creates an obligation that runs through the moral entity, not around it.

This extends beyond animals. Objects, places, projects, traditions. They matter morally when moral entities are attached to them through Protection (this place keeps us safe), Trust (this tradition binds our commitments), or Free-Agency (this project expresses our choices).

Aldo Leopold's land ethic approaches this insight from the environmental side: "A thing is right when it tends to preserve the integrity, stability, and beauty of the biotic community. It is wrong when it tends otherwise" (Leopold 1949). But Leopold struggles to ground the claim without attributing intrinsic value to ecosystems. Resolution Ethics grounds it through attachment: ecosystems matter because moral entities depend on them, are attached to them, and navigate their PTF within them.

The attachment principle explains why destruction of cultural heritage is a moral violation even when no person is physically harmed. It explains why environmental destruction matters morally. The violation runs through the moral entities attached to what is destroyed.

4.4 Content vs. Structure

"You claim PTF is universal, but cultures disagree about who deserves protection, what counts as trust violation, and what freedoms matter. Are you not just imposing your values?"

Resolution Ethics distinguishes structure from content. But not the way relativists hope.

Structure is universal and non-negotiable. The PTF hierarchy is not culturally variable. Protection of the innocent, and the vulnerable, comes first. Always. Trust depends on Protection. Free-

Agency depends on Trust. This ordering derives from dependency, not from cultural preference.

Innocence and vulnerability are structurally defined, not culturally defined.

Innocent: a moral entity that lacks the capacity to sense moral coherence, whether due to immaturity (not yet developed), cognitive impairment (never developed or lost), or incapacitation (temporarily unavailable, e.g., comatose, unconscious, sleeping).

Vulnerable: a moral entity exposed to serious harm, coercion, or loss of agency within RIF Flux, whether aware or unaware.

These are two distinct Protection-claims. A culture cannot define a child as "capable of participation" and thereby permit harm. A culture cannot define an outsider as "not vulnerable" and thereby justify predation. The structural definitions hold regardless of cultural specification.

Michael Walzer distinguished between "thick" and "thin" morality: thin morality is minimal and universal, while cultures fill in the thick details (Walzer 1994). Resolution Ethics takes a different view. The PTF structure is not thin in Walzer's sense. It has real constraints. Protection of the innocent is not a cultural variable.

What content actually means. Content variation occurs at a different level:

- What constitutes a valid promise? (Content)
- How are family obligations structured? (Content)
- What forms of property are recognized? (Content)
- What rituals mark transitions in status? (Content)

These vary across cultures. RE does not adjudicate between cultures that specify these differently while satisfying the structural requirements.

What content does not mean. Content variation does not include:

- Who counts as innocent (structurally defined)
- Whether Protection comes before Free-Agency (structurally fixed)
- Whether deception is permissible toward outsiders (T-violation regardless of target)

A culture that practices female genital mutilation on children is not "coherent within its own framework." It is violating Protection of the innocent. The children are not culpably participating in any violation. As innocents, they cannot forfeit standing; they lack the capacity for the self-deception that three-domain saturation requires. The cultural acceptance is irrelevant to the structural analysis. RE does not check whether a culture is internally consistent with its own stated values. RE checks whether a culture's practices are coherent with the PTF structure. A culture can be perfectly internally consistent and still incoherent with RE, because the structure is not culturally generated.

Bernard Williams worried about "the relativism of distance": can we judge distant cultures by our standards? (Williams 1985). RE offers an answer. We do not judge by our cultural standards. We detect structural violations. And structural violations are not relative.

4.5 Degrees of Incoherence

Not all violations are equal. Resolution Ethics tracks the structural spread of incoherence across the three domains.

Single-domain breach. A violation confined to one domain represents the least structural damage. Breaking a minor promise (T breach) where no one was harmed and no one's agency was restricted. Making a small decision for someone without consulting them (F breach) where no deception or harm occurred. These are still violations. They still matter. But the incoherence is contained. The agent may recognize the breach, feel appropriate weight, and repair it. One breach does not indicate systemic corruption.

Two-domain breach. When violation spreads across two domains, structural incoherence has emerged. The reasoning is no longer locally contained. This is where innocent violators typically fall. A child who harms another (P breach) and thereby breaks the trust of their community (T breach) has produced real damage. But they cannot have violated F in the way a capable adult can, because they cannot sense when their own reasoning has gone wrong. Their incapacity caps the spread. Two-domain breaches by capable adults are more serious: the pattern suggests the corruption is not accidental.

Three-domain breach. When all three domains are violated, incoherence has saturated the act. Serious harm inherently cascades across all three domains. Murder is the clearest case: you harm the victim (P), you violate their trust that others would not kill them (T), and you eliminate their agency entirely (F). There is no way to kill someone while leaving their trust and freedom intact. The domains are integrated, not isolated. Any serious P violation cascades to T and F because harm destroys both the trust the victim placed in others and the agency the victim could have exercised.

Even temporary incapacitation of a protected party is a three-domain breach. You harm them (P), violate their trust (T), and suspend their agency for the duration (F). The temporary nature reduces severity but does not change the structural spread.

But context matters. Incapacitating a violator in self-defense or to gain control of a threat situation does not constitute a breach. The violator has forfeited standing in that context. The same physical action against a protected party would be a three-domain violation; against a forfeited party acting within the threat context, it is coherent with RE. This is why forfeiture is not a loophole but a structural feature. It allows defense without permitting predation.

This is why three-domain breach indicates maximum incoherence. It is not merely "more violations." It is the structural signature of predatory action. No honest path leads to killing, assaulting, or seriously harming a protected party. The deception, whether self-deception about the

victim's standing or other-deception to enable the act, is baked in.

A note on innocent violators: their incapacity typically caps them at two-domain breach. A child who kills has violated P and T, but they lack the capacity to have weaponized F in the way a capable adult does. They cannot sense when their own reasoning has become corrupted. This is why they remain under Protection even as they are restrained or treated.

This framework sets up the operational analysis that follows in the Resolution Ethics Engine. RE identifies the structure. REE verifies coherence against it.

5. RE and Its Predecessors

Resolution Ethics does not replace its predecessors. It does something different.

Aristotle, Kant, Mill, and Ross each captured something real about moral life. Their frameworks endure because they track genuine features of how moral cognition works. Resolution Ethics acknowledges their contributions and does not claim to surpass them.

What RE offers is a verification layer: a way to analyze whether an ethical decision is coherent or incoherent with the PTF structure (the Domains, their Constants, and the P > T > F dependency hierarchy) that operates within the Reality Interaction Field. The predecessors tell you how to reason morally. RE tells you whether the reasoning stayed coherent with the structural constraints.

5.1 Aristotle and Virtue Ethics

Aristotle gave us virtue ethics. Moral character is built through habituation. The virtuous person possesses practical wisdom, phronesis, and acts from a stable disposition toward excellence (Aristotle 2000, Book II).

This is valuable. Character matters. Habituation shapes moral cognition. The virtuous person does perceive situations more clearly than the vicious one.

Virtue ethics leaves a procedural gap: when virtues pull in opposite directions, what resolves the conflict besides appeal to the already-virtuous agent? Courage may demand standing your ground. Prudence may demand retreat. Aristotle's answer is that the person of practical wisdom will discern the right balance. But how does phronesis decide? Aristotle does not say (Hursthouse 1999).

RE does not claim to solve this for virtue ethics. But RE can check whether a given resolution is coherent with PTF. If someone appeals to courage in a way that violates Protection of the innocent, RE flags the incoherence. Truthfulness is a Trust constraint, but if it functions as an instrument of a Protection violation, RE detects that too.

Aristotle tells you to develop practical wisdom. RE offers a way to check whether the wisdom stayed coherent.

5.2 Kant and Deontology

Kant gave us the categorical imperative. Act only according to maxims you could will as universal laws. Treat humanity never merely as a means, but always also as an end (Kant 1785).

This is valuable. Universalizability matters. Dignity matters. Moral cognition cannot function coherently if it grants itself exceptions it denies to others.

Kant's framework encounters the murderer at the door. A killer asks if your friend is hiding inside. Kant appears to require that you tell the truth, because lying cannot be universalized. This strikes most readers as monstrous. Surely you may lie to protect an innocent life?

Kant struggled with this case. His defenders have struggled since. Benjamin Constant pressed the objection. Kant's reply was unsatisfying (Kant 1797b). Christine Korsgaard's reconstruction helps: the murderer, by initiating unjust violence, has already exited the realm of honest discourse and cannot demand its protections (Korsgaard 1986).

RE does not claim to solve Kant's problem for him. But RE can analyze it. The murderer has committed a three-domain violation: intending harm (P), betraying trust by weaponizing a question expecting honest cooperation (T), and self-deceiving about their entitlement to harm the victim (F). You owe the murderer no cooperative trust in that context. RE detects the forfeiture; lying to protect the innocent is coherent with PTF.

Kant tells you to universalize your maxims. RE offers a way to check whether forfeiture applies.

5.3 Mill and Consequentialism

Mill gave us utilitarianism. Actions are right insofar as they promote happiness, wrong insofar as they produce the reverse. The aggregate matters. Impartiality demands we count each person's welfare equally (Mill 1861).

This is valuable. Consequences matter. Suffering matters. A moral framework that ignores outcomes is incomplete.

Utilitarianism encounters the utility monster. Robert Nozick posed the challenge: if one being derives vastly more utility from resources than others, utilitarianism seems to require that we funnel everything to that being, even at the cost of everyone else's basic needs (Nozick 1974). More troubling still: if killing one innocent person would save five, utilitarianism appears to require the sacrifice.

Utilitarians have developed sophisticated responses. Rule utilitarianism, indirect utilitarianism, threshold deontology. But the core tension remains: if welfare is the only currency, then everything is tradeable, including innocent lives.

RE does not claim to refute utilitarianism. But RE can analyze proposed trades. Protection of the innocent, and the vulnerable, is not a variable to optimize. It is a precondition. If innocent life is tradable, the system becomes predation-permissive by construction.

Consider two cases. First, sacrificing a child to save five adults. The child is innocent: incapable of meaningful moral participation, unable to navigate the dilemma or consent to the trade. This is maximally incoherent with RE.

Second, sacrificing an adult non-violator to save five others. The adult is vulnerable: exposed to serious harm, has not forfeited standing, but has full capacity to sense moral coherence and participate in moral reasoning. Still incoherent with RE, but the adult has capacities the child lacks.

Both are protected. Neither has forfeited standing. But the innocent (the child, the comatose, the severely cognitively impaired, the incapacitated) hold the highest protection stringency because they cannot participate in their own defense or navigation. The vulnerable (the aware non-violator, the person with a disability who can still reason morally) are also protected, but they retain the capacity to engage.

This distinction matters. Utilitarianism that permits sacrificing either is incoherent with RE. But a framework that treats a child and a competent adult as identical in their protection-claims misses something structural. RE tracks the difference.

Mill tells you to maximize welfare. RE offers a way to check whether the maximization violated structural constraints.

5.4 Ross and Pluralism

W.D. Ross gave us prima facie duties. We have duties of fidelity, reparation, gratitude, justice, beneficence, non-maleficence, and self-improvement. These duties are real, but they can conflict. When they do, we must judge which duty is most stringent in the particular case (Ross 1930).

This is valuable. Morality is not monistic. Multiple considerations bear on action. The right act is often the one that best satisfies competing claims.

Ross provides no decision procedure. When fidelity conflicts with beneficence, Ross says we must use judgment. But whose judgment? By what criteria? Ross explicitly denies that any algorithm can resolve the conflicts. We simply "see" which duty is most pressing. This has struck critics as an admission of defeat (Dancy 1993).

RE does not claim to replace Ross's pluralism. But RE can analyze whether a proposed resolution is coherent with PTF. The prima facie duties cluster around the PTF domains. Borderline cases do not weaken the dependency ordering that resolves conflicts:

Ross's Duties	**PTF Domain**
Non-maleficence	Protection
Justice	Protection (structural fairness)
Fidelity	Trust
Reparation	Trust (restoring broken trust)
Gratitude	Trust (acknowledging received trust)
Beneficence	Free-Agency (enabling others' projects)
Self-improvement	Free-Agency (enabling one's own projects)

When duties conflict, RE checks whether the resolution honors the P > T > F dependency. Non-maleficence (P) cannot be overridden by beneficence (F). Fidelity (T) cannot be overridden by self-improvement (F). RE does not tell you which duty wins. It tells you whether your resolution is coherent with the structural ordering.

Ross tells you to weigh duties. RE offers a way to check whether the weighing stayed coherent.

5.5 The Verification Layer

Each predecessor captured a real feature of moral structure:

Framework	**Insight**	**What RE Adds**

Virtue Ethics	Character and habituation matter	Coherence check when virtues conflict
Deontology	Universalizability and dignity matter	Forfeiture analysis for apparent exceptions
Consequentialism	Outcomes and welfare matter	Constraint check on sacrificing the protected
Pluralism	Multiple duties are real	Dependency ordering to verify resolutions

RE does not claim these frameworks were wrong or incomplete. They do what they do. RE does something different: it provides a verification layer that checks whether moral reasoning remained coherent with the structural constraints of Protection, Trust, and Free-Agency.

The predecessors are the terrain. RE is one way to check whether you stayed on the path.

6. Cross-Cultural Convergence

If Resolution Ethics is correct, if PTF represents structurally necessary features of moral cognition under RIF Flux, then we should expect to find convergence across independent moral traditions. Not identical content. Structural parallels.

Seems like there is.

6.1 The Test

This is not an argument from authority. It is a prediction. If PTF emerges from the logic of vulnerability and coordination, then any culture that has navigated vulnerability and coordination long enough should have discovered something PTF-shaped. If we find that, it is evidence that RE is tracking real structure. If we do not find it, RE is in trouble.

We should expect convergence across multiple moral traditions, some historically independent, others genealogically related. The independent cases test whether PTF recurs without cultural

diffusion. The related cases test whether PTF persists even as metaphysics diverge.

The test is not whether other cultures use the words "Protection," "Trust," and "Free-Agency." The test is whether a tradition treats non-harm and protection as a precondition for stable coordination, and stable coordination as a precondition for meaningful agency.

6.2 Confucian Ethics

Confucianism developed in China over 2,500 years, independently of Western philosophical traditions. Its moral framework centers on the Five Constants (五常, Wǔcháng): Ren (仁, benevolence), Yi (義, righteousness), Li (禮, ritual propriety), Zhi (智, wisdom), and Xin (信, faithfulness). This is a richer system than PTF, with its own internal logic and debates.

But within this richer system, PTF-shaped structure is detectable.

Ren (仁) is often treated as the cardinal virtue, the root of all others. Confucius held that ren begins with how you treat those closest to you and extends outward (Confucius 2003, Analects 1.2). That is the Protection-shape: care that radiates by relationship and responsibility.

Li (禮) is ritual propriety, the forms that make social coordination reliable. It is not mere etiquette. Li structures expectations so that agents can depend on one another (Confucius 2003, Analects 2.3). This is Trust-shaped: the fabric of reliable coordination.

Yi (義) is righteous action, doing what is appropriate to one's role and situation. It is the exercise of moral judgment within established constraints (Mencius 2009, 2A:6). This is Free-Agency-shaped: acting appropriately within the space secured by prior commitments.

The ordering is suggestive, not identical to RE's structural dependency claim. Mencius argued that ren is the root, li the form, and yi the application (Mencius 2009, 4A:27). The shape rhymes: care for the vulnerable grounds reliable forms, which ground righteous action. Independent tradition

arriving at similar structural ordering is evidence that something real is being tracked.

6.3 Hindu and Buddhist Ethics

The Dharmic traditions developed on the Indian subcontinent, independently of Western and Chinese traditions. Like Confucianism, they contain richer systems than PTF, with their own internal logic.

In the Yoga tradition, Patanjali's Yoga Sutras outline five yamas (moral restraints): Ahimsa (non-harm), Satya (truthfulness), Asteya (non-stealing), Brahmacharya (celibacy or right use of energy), and Aparigraha (non-possessiveness) (Patanjali 2009, 2.30). This is a five-fold system, not a three-fold one.

But within this richer system, PTF-shaped structure is detectable, and the ordering is explicit.

Ahimsa (अहिंसा) is non-harm. It is listed first among the yamas, and Patanjali treats it as foundational. The tradition holds that if one perfects ahimsa, the other restraints follow. This is Protection-shaped: non-harm as precondition.

Satya (सत्य) is truthfulness. It follows ahimsa in the ordering. The Mahabharata makes the dependency explicit: "Ahimsa is the highest dharma" (Mahabharata, Anusasana Parva 116.37-41). Truthfulness matters, but not at the cost of harm to the Innocent. This is Trust-shaped, ordered after Protection.

The remaining yamas (non-stealing, celibacy, non-possessiveness) address different concerns. RE does not claim the yamas are "PTF in Sanskrit." But the first two yamas exhibit the P > T dependency structure, and the tradition treats that ordering as non-negotiable.

Buddhist ethics shows similar structure. The Five Precepts (pañca-sīla) are: abstain from killing, stealing, sexual misconduct, false speech, and intoxicants (Keown 2005). Again, a five-fold

system.

But again, structure is detectable. The first precept is to abstain from taking life (Protection-shaped). The fourth is to abstain from false speech (Trust-shaped). The precepts are not merely listed; killing comes first because harm to life is the most fundamental violation. The tradition's own ordering places Protection before Trust.

6.4 Ubuntu

Ubuntu is a southern African ethical philosophy, often summarized as "I am because we are." It developed independently of the Axial Age traditions. Unlike the systems above, Ubuntu is not formalized as a numbered list of principles. It is a relational worldview.

Archbishop Desmond Tutu described ubuntu as the recognition that "my humanity is caught up, is inextricably bound up, in yours" (Tutu 1999). This is not abstract. It has structural implications.

Within this relational framework, PTF-shaped structure is detectable.

Ubuntu emphasizes the protection of community members, especially the Vulnerable: children, elders, the sick. Metz and Gaie argue that ubuntu's core principle is "a human being is a human being through other human beings," which generates strong obligations of mutual protection (Metz and Gaie 2010). This is Protection-shaped.

Ubuntu also emphasizes trust through communal solidarity. Promises matter. Relationships matter. The individual does not exist outside the web of trust. This is Trust-shaped.

And ubuntu recognizes individual contribution and self-development, but always within the context of community (Menkiti 1984). Agency is real, but it is not prior to relational obligations. This is Free-Agency-shaped.

The ordering is implicit but suggestive: protect the community, maintain the relationships, then

pursue individual flourishing. Ubuntu is not "RE in Zulu." But an independent tradition arriving at similar structural ordering is evidence that something real is being tracked.

6.5 Abrahamic Traditions

Judaism, Christianity, and Islam share roots but developed distinct ethical frameworks. Each is vastly richer than PTF. Judaism has 613 commandments. Christianity has centuries of moral theology. Islam has comprehensive legal and ethical systems. RE does not claim these traditions reduce to three principles.

But within each, PTF-shaped structure is detectable.

Judaism encodes Protection in the primacy of pikuach nefesh: the obligation to save a life overrides nearly all other commandments (Talmud, Yoma 85b). This is not one principle among many. It is a structural override. Trust is encoded in the centrality of covenant (brit), the binding agreements between God and Israel, and between persons. Free-Agency is encoded in the emphasis on teshuvah (repentance) and moral choice, but always within the constraints of law and covenant. The shape rhymes: life first, then covenant, then choice.

Christianity places love of neighbor as the second great commandment, after love of God (Matthew 22:39). The parable of the Good Samaritan defines neighbor as anyone in need of protection, crossing ethnic and religious boundaries (Luke 10:25-37). Trust is encoded in the emphasis on faithfulness and promise-keeping. Free-Agency is encoded in the doctrines of moral responsibility and conscience, but always ordered under love. Again, the shape: protection of the vulnerable, then faithfulness, then responsible action.

Islam articulates the five maqasid al-shariah (objectives of Islamic law): protection of life, religion, intellect, lineage, and property (Al-Ghazali 1987). This is a five-fold system, not a three-fold one. But life comes first. Trust (amanah) is a central concept: the human being is a trustee of God's

creation. Free-Agency is encoded in moral accountability (taklif), but always within the bounds of divine law and communal obligation. The ordering is suggestive: life, then trusteeship, then accountability.

In each tradition, the pattern holds: protect the vulnerable, maintain trustworthy relationships, exercise agency within those constraints.

6.6 What Convergence Means

These traditions developed independently, across millennia, on different continents, in different languages, with different metaphysics. Each is richer than PTF. Confucianism has five constants. The Yoga tradition has five yamas. Buddhism has five precepts. Islam has five maqasid. Judaism has 613 commandments.

Yet within each, PTF-shaped structure is detectable:

Tradition	Protection-shaped	Trust-shaped	Free-Agency-shaped
Confucianism (5 constants)	Ren (benevolence)	Li (ritual propriety)	Yi (righteous action)
Hindu Yamas (5 yamas)	Ahimsa (non-harm)	Satya (truthfulness)	(implicit in tradition)
Buddhism (5 precepts)	First Precept (no killing)	Fourth Precept (no false speech)	(implicit in Eightfold Path)
Ubuntu	Community protection	Communal solidarity	Individual contribution
Judaism (613 commandments)	Pikuach nefesh	Brit (covenant)	Teshuvah (moral choice)
Christianity	Love of neighbor	Faithfulness	Conscience
Islam (5 maqasid)	Protection of life	Amanah (trust)	Taklif (accountability)

It is the pattern RE predicts. These traditions are not "really PTF." They are richer, more complex, and shaped by their own histories. But when traditions diverge in metaphysics, ritual, and social

form, yet repeatedly place protection-shaped constraints first, trust-shaped structures second, and agency within those constraints, the simplest explanation is that they are tracking the same underlying problem: vulnerability under flux.

Resolution Ethics does not claim that these traditions are "really the same." They differ in content, metaphysics, and emphasis. But beneath the differences, they exhibit a common dependency structure. That structure is what RE articulates.

7. The Grounding Claim

7.1 Why Not Just Another Framework?

Philosophy has produced many ethical frameworks. Virtue ethics, deontology, consequentialism, care ethics, contractarianism. Each offers principles, each has defenders, each has critics. Why should anyone care about one more entry?

Resolution Ethics is not competing for a place in that lineup. It is doing something different.

The predecessors tell you how to reason morally. Virtue ethics says: cultivate character. Deontology says: follow universalizable maxims. Consequentialism says: maximize welfare. These are prescriptive frameworks. They guide action.

RE does not prescribe. RE verifies. It asks whether moral reasoning, whatever framework generated it, remained coherent with the structural constraints that vulnerable agents require.

What Predecessors Do	**What RE Does**
Prescribe how to reason	Verify whether reasoning stayed coherent
Offer principles to follow	Check coherence against PTF structure
Compete for correctness	Sit underneath as verification layer

This is why Section 5 did not argue that Aristotle, Kant, Mill, or Ross were wrong. They captured real features of moral reality. RE offers a way to check whether applications of their insights have remained coherent.

A verification layer does not replace the systems it verifies. It serves a different function.

7.2 The Structural Necessity Argument

RE proposes that PTF is not arbitrary. It is not culturally invented. It is not one option among many that moral cognition might adopt. It emerges from the structure of what vulnerable agents require.

The argument proceeds as follows.

RIF Flux exists. Reality has coordinates: WHO, WHAT, WHERE, WHEN, WHY, HOW. These coordinates shift continuously. This is not a philosophical posit. It is a description of the conditions under which any agent operates.

Agents within the Flux are vulnerable. They can be damaged. Their projects can be disrupted. Their existence can be terminated. Vulnerability is not occasional. It is the constant condition of finite agency in a dynamic system.

From this, a minimal axiom: any agent that persists behaves as if continued function is preferable to collapse. This is not a grand metaphysical claim. It is an observation about what persistence entails. A creature that does not avoid threats does not survive long enough to reproduce. A cognition that does not resist collapse does not persist long enough to reason. Agents that did not behave this way would not persist. We would not be here to discuss them.

From vulnerability and the persistence axiom, Protection emerges as structurally first. Severe harm to the vulnerable ends the possibility of further interests, coordination, or choice. You cannot build anything on a foundation that can be destroyed arbitrarily.

From Protection, Trust becomes possible. If I can kill you without consequence, no commitment I make to you is credible. Reliable coordination requires that arbitrary harm is already off the table. Trust depends on Protection.

From Trust, Free-Agency becomes meaningful. If I am deceived about my options, my "choices" are not free. If I am coerced, my "agency" is theater. Freedom to act requires that I can depend on accurate information and non-manipulated conditions. Free-Agency depends on Trust.

The ordering is not preference. It is dependency. P > T > F describes the sequence in which these conditions require each other.

This structure is not new. Philosophers have circled it for centuries. Hobbes saw that without security, coordination collapses (Hobbes 1651). Locke saw that legitimate society rests on trust, which rests on the protection of natural rights (Locke 1690). Rawls formalized lexical priority: some principles must be satisfied before others can operate (Rawls 1971). Goodin built an entire framework around the moral centrality of protecting the vulnerable (Goodin 1985).

RE draws these insights into a dependency structure and identifies the corruption mechanism that breaks it.

7.3 Deception as the Structural Keystone

Section 3 argued that deception is the sole corruptor and source of incoherence. This matters for grounding.

If there were another path to moral corruption, some way to produce incoherence that did not trace back to self-deception or other-deception, then RE would be incomplete. The framework would have a leak. Critics could point to the leak and say: your verification misses this.

But trace any violation back. P-violations require SD about the victim's standing, or about one's

own entitlement, or about the permissibility of the method. F-violations require SD or OD about consent, about coercion, about what is actually being chosen. T-violations are deception.

There is no honest path to moral corruption. The violent actor has already deceived themselves about their victim. The coercer has already distorted the picture of what they are doing. The manipulator is, by definition, deceiving.

This is not an empirical generalization. It is logical necessity. For moral cognition to violate structure it perceives, distortion must enter the justification chain. The only place distortion can enter is the WHY field (the agent's reasons for action). Distortion in the WHY field is deception, either self-directed (SD) or other-directed (OD).

The alternative is incoherent: a moral cognition that perceives RIF structure accurately, understands P > T > F, and simultaneously produces a violation without any distortion in its justification. Such an agent would be acting against what it clearly sees as wrong, with no rationalization, no self-deception, no motivated reasoning, just pure, clear-eyed wrongdoing. This does not match how moral cognition actually fails. Every actual case, traced back, reveals the distortion.

The hierarchy reflects operational dependency, not corruption mechanism. P is the target of any coherent or incoherent act (the protected parties). T locates where those parties sit in the agent's concentric circles of moral responsibility (relational proximity and significance). F is where SD/OD enters and corrupts any agent's RIF Flux navigation. Corruption always enters through F. When an agent ingests self-deception, their free agency to reason coherently about P and T becomes distorted. They act. The violation cascades across all three domains: harm lands on the protected party (P-violation), implicit trust between agents is broken (T-violation), and the self-deception that permitted the act was already an F-violation. Three domains saturated. Forfeiture.

Kant saw part of this. Lying undermines the very possibility of rational communication (Kant

1785). Augustine saw that lying corrupts the liar, not just the social fabric (Augustine 397). Peck identified self-deception as the signature of what he called human evil: the persistent refusal to see what one is doing (Peck 1983). Kunda's research on motivated reasoning demonstrated that people employ strategies for accessing, constructing, and evaluating beliefs that favor desired conclusions while maintaining an "illusion of objectivity" (Kunda 1990). Von Hippel and Trivers showed that self-deception evolved to facilitate deceiving others: when people actually believe their own fabrications, they eliminate the telltale signs of conscious deception (Von Hippel and Trivers 2011).

RE unifies these insights. Deception is not one vice among many. It is the mechanism by which moral cognition becomes corrupted and ethical decisions incoherent.

7.4 What Kind of Realism?

RE claims that the PTF structure is real. This invites a question: what kind of reality?

Moral anti-realists hold that ethics is preference, or culture, or emotion. There are no moral facts, only attitudes we project onto the world. On this view, RE would be one more projection, dressed up in structural language.

Strong moral realists hold that moral facts exist independently of minds, like mathematical truths or physical laws. On this view, "murder is wrong" is true in the same way that "2 + 2 = 4" is true.

RE occupies different ground.

The structure RE identifies is real because vulnerability is real. Agents operating in RIF Flux can be harmed. That is not a projection. It is a feature of finite agency in a dynamic system. From that vulnerability, given agents that persist, the PTF structure follows as the set of conditions required for coherent navigation.

The normative force enters through the minimal axiom: agents that persist behave as if continued function is preferable to collapse. This is not a grand metaphysical ought. It is what persistence means. Agents that do not behave this way do not persist.

RE does not require belief in spooky moral facts floating in a Platonic heaven. It requires noticing that vulnerable agents who persist must navigate certain structural constraints or face collapse or predation. Those constraints are PTF. The structure is as real as vulnerability itself.

This position has philosophical precedent. Foot argued that moral judgments can be grounded in facts about what humans need to flourish (Foot 2001). Korsgaard grounded normativity in the structure of rational agency itself (Korsgaard 1996). RE grounds it in the structure of vulnerable agency under flux.

7.5 The Burden Shift

RE makes a structural claim. PTF describes what vulnerable moral cognitions require to function coherently. Deception is the sole mechanism that corrupts this function. The cross-cultural convergence documented in Section 6 is the pattern RE predicts.

Critics who reject this claim take on a burden.

Show us a moral cognition that functions coherently without Protection preceding Trust. Show us stable coordination without safety from arbitrary harm. Show us meaningful agency without reliable information and non-manipulated conditions.

Find a durable ethical system that does not encode some version of P > T > F. Find a tradition that has sustained itself across generations while permitting predation on the innocent, rewarding systematic deception, and treating agency as prior to trust or protection.

Find a moral violation that does not trace back to self-deception or other-deception. Find an agent

who harms the innocent with no distortion in their WHY field (no rationalization, no motivated reasoning, no self-serving reframing). If such cases exist, they would falsify the monocause claim.

The convergence is evidence. When independent traditions, across millennia, on different continents, in different languages, with different metaphysics, repeatedly arrive at the same structural pattern, the simplest explanation is that they are tracking the same underlying problem.

RE articulates that structure. The burden is now on critics to produce the counterexample: a coherent moral system that violates the dependency ordering and survives, or a moral corruption that enters through a third vector neither SD nor OD.

8. Addressing Criticisms

Any framework that claims to identify structural features of morality will face serious objections. This section addresses five that matter.

8.1 On Deriving Ought from Is

The most venerable objection in moral philosophy comes from Hume: you cannot derive an "ought" from an "is" (Hume 1739). No amount of describing how things are can tell you how they should be. The gap between fact and value appears unbridgeable.

G.E. Moore extended this insight with his "open question argument": for any natural property X, you can always sensibly ask "but is X good?" (Moore 1903). This suggests that goodness cannot be reduced to any natural property, including the structural features RE identifies.

Applied to RE: even if PTF describes what vulnerable agents require, how does that generate obligation? Why should anyone care about coherence?

RE does not claim to have dissolved the is-ought gap. But RE does claim that the gap is narrower than critics assume when we are talking about agents who persist.

The Agency Axiom is minimal: any agent that persists behaves as if continued function is preferable to collapse. This is not a grand metaphysical ought. It is an observation about what persistence entails. An agent that genuinely did not behave this way would not be around to discuss the question.

From this minimal starting point, RE derives constraints. If you are an agent that persists, and you operate in RIF Flux, and you are vulnerable, then certain structural requirements follow. Protection must precede Trust because you cannot coordinate with someone who can kill you arbitrarily. Trust must precede Free-Agency because you cannot choose meaningfully if you are deceived about your options.

These are not value claims floating free of facts. They are dependency claims about what coherent agency requires. The "ought" is conditional: if you are a persisting vulnerable agent navigating RIF Flux, then PTF describes the structure your navigation must respect to remain coherent.

Philippa Foot made a similar move in her treatment of natural goodness: what counts as a good specimen of a kind depends on what that kind of thing is and does (Foot 2001). A good knife cuts well. A good heart pumps blood effectively. RE extends this insight: a coherent moral cognition navigates PTF without corruption. That is what coherent moral cognition is.

The critic may reply: "But I can still ask whether coherence is good." Yes, you can ask. But notice what you are doing. You are a moral cognition, using justification (WHY), to evaluate whether the structure that makes justification possible is worth maintaining. The question is coherent only because you already operate within the structure you are questioning.

RE does not claim this silences all skeptics. Some philosophers will remain unsatisfied. But the burden has shifted. The critic must now explain how a moral cognition could function coherently while violating the dependency ordering that RE identifies. That is a harder task than simply pointing to the is-ought gap.

8.2 On Cultural Imposition

The objection: any claim to universal moral structure is cultural imperialism dressed in philosophical language. Western frameworks have a long history of being imposed on non-Western peoples under the guise of "universal truth." Why should RE be different?

This objection was addressed in Section 4.4 (Content vs. Structure) and tested in Section 6 (Cross-Cultural Convergence). A brief restatement is warranted here.

RE distinguishes structure from content. The PTF hierarchy is structural. What constitutes a valid promise, how family obligations are organized, what rituals mark transitions: these are content. Content varies across cultures. Structure does not.

The test in Section 6 was explicit: if PTF is structurally necessary, we should find it encoded in independent moral traditions. We do. Confucian ren-li-yi, Hindu ahimsa-satya-svadharma, Buddhist precepts, Ubuntu, Abrahamic traditions: all encode protection-first constraints, trust-stabilizing structures, and agency within those constraints. These traditions developed independently, across millennia, with different metaphysics.

The convergence is evidence that RE is tracking a structural pattern, not imposing a cultural preference. The critic who insists on cultural imperialism must explain why independent traditions repeatedly arrive at the same dependency ordering.

RE does not claim that Western philosophy discovered something others missed. RE claims that vulnerable agents navigating flux face the same structural problem everywhere, and durable traditions have found versions of the same solution.

8.3 On Motivation

The objection: even if RE correctly identifies the structure of coherent moral cognition, what

motivates anyone to follow it? Why should someone care about coherence if incoherence serves their interests?

This is a serious challenge. J.L. Mackie raised a version of it in his "argument from queerness": if moral facts existed, they would need some strange "to-be-pursuedness" built into them (Mackie 1977). What makes PTF binding rather than merely descriptive?

RE offers two responses.

First, the question assumes a separation between the agent and the structure that may not hold. You are not a neutral observer deciding whether to adopt PTF. You are a moral cognition already operating within RIF Flux, already vulnerable, already navigating. The structure is not external to you. It is the condition of your functioning.

Compare: asking "why should I follow the laws of logic?" The question is coherent only because you are already following them well enough to formulate it. You could try to violate logic, but you would not be reasoning anymore. You would be producing noise. Similarly, a moral cognition that systematically violates PTF is not navigating coherently anymore. It is collapsing or predating.

Second, RE does not claim that everyone will be motivated to remain coherent. Some agents choose incoherence. They deceive themselves about their victims, their entitlements, their methods. They extract safety, trust, and agency from others. RE predicts what happens: moral entropy accumulates, and the agent tends toward collapse or predation. Neither mode is stable.

The question "why be moral?" often assumes that morality is a constraint imposed from outside, and the agent must be given a reason to accept it. RE inverts this framing. Morality is the structure that keeps moral cognition functioning. The question is not "why accept the constraint?" The question is "what happens when you do not?" RE has an answer: incoherence, entropy, collapse or predation. These are not punishments imposed by an external authority. They are structural consequences.

Plato saw this in the Republic: the unjust soul is disordered, at war with itself, incapable of unified action (Plato 1992, 444b). Aristotle saw it in the Nicomachean Ethics: the vicious person lacks the internal harmony that characterizes the virtuous (Aristotle 2000, Book IX). RE frames the same insight structurally: corruption produces incoherence, and incoherence degrades function.

This will not satisfy every skeptic. The determined amoralist can always say "I do not care about coherence." But notice: that statement is itself a justification (a WHY). The amoralist is using the structure to reject the structure. RE does not claim to have a magic argument that will convert such a person. It claims to have identified what they are doing: navigating incoherently, accumulating moral entropy, trending toward collapse or predation.

8.4 On Simplicity

The objection: RE reduces 2,500 years of moral philosophy to three domains and one corruption vector. Real ethical life is messier. There are tragic dilemmas, incommensurable values, irreducible complexity. This is suspiciously tidy.

Two responses.

First, simplicity is a virtue in explanatory frameworks, not a vice. Newton reduced the motion of planets and apples to a single force. Darwin reduced the diversity of life to variation and selection. The question is not whether a framework is simple. The question is whether it works.

RE is testable. Find a moral violation that does not trace back to self-deception or other-deception. Find a durable ethical tradition that does not encode P > T > F. Find a case where serious harm to a protected party does not cascade across all three domains. If these counterexamples exist, RE is in trouble. If they do not, the simplicity is earned.

Second, RE does not claim to resolve every moral question. It does not tell you which career to pursue, whom to marry, or how to allocate your charitable giving. These are content questions. RE

operates at the structural level: it verifies whether reasoning has remained coherent with PTF, not whether it has produced the uniquely correct answer.

The predecessors remain valuable. Virtue ethics helps you cultivate character. Deontology helps you identify duties. Consequentialism helps you weigh outcomes. RE does not replace these. It checks whether their application in a given case has remained coherent with the structural constraints.

Ross was right that moral life involves multiple considerations that can conflict (Ross 1930). RE does not deny this. RE provides a dependency ordering that constrains which resolutions are coherent. Non-maleficence (P) cannot be overridden by beneficence (F). Fidelity (T) cannot be overridden by self-improvement (F). Within those constraints, judgment is still required.

The charge of over-simplification often comes from those who have not tried to apply RE to hard cases. The Resolution Ethics Engine, described in the companion paper, works through dozens of scenarios: trolley problems, professional dilemmas, edge cases in self-defense, innocent violators, competing vulnerabilities. The framework handles them. Not by providing a unique answer to every question, but by identifying which answers are coherent and which are not.

Simplicity that works is not a flaw. It is the goal.

8.5 On Alienation

Bernard Williams raised what may be the most humanly resonant objection to moral theory. In his famous example, a man can save only one of two drowning people. He saves his wife. Williams observed that if the man pauses to think "it is morally permissible to save my wife," he has had "one thought too many" (Williams 1981). The authentic response is to save her because she is his wife, not because moral theory permits it.

The objection cuts deep. Any framework that claims to govern moral life risks inserting itself

between agent and action. Spontaneous acts of love, loyalty, and care become calculated moves requiring theoretical justification. The theory alienates the agent from the very commitments that make life meaningful.

Michael Stocker made a related point: modern moral theories produce a kind of "schizophrenia" in which the reasons that justify an action differ from the motives that produce it (Stocker 1976). You visit your sick friend, but your motivating reason is "maximizing welfare" rather than simple concern for someone you care about. The gap between justification and motivation is itself a moral failing.

RE offers two responses.

First, RE is a verification layer, not a decision procedure. The man saving his wife does not need to consult PTF before acting. He acts. RE checks coherence when something seems off, when reasoning has drifted, when justification is demanded. It does not insert itself into every spontaneous moment of moral life. The alienation Williams feared occurs only if moral theory must be the motivating reason for every action. RE makes no such demand.

Second, and more fundamentally: RE is simple enough to internalize.

Consider the cognitive load of the predecessors. Virtue ethics requires tracking dozens of virtues, finding the mean, exercising practical wisdom across varied contexts. Deontology requires holding multiple formulations of the categorical imperative, distinguishing perfect from imperfect duties, applying universalizability tests. Consequentialism requires estimating outcomes for all affected parties, aggregating welfare, weighting probabilities. Prima facie pluralism requires remembering multiple duties and weighing them without a clear method.

These frameworks are cognitively heavy. Applying them in real time requires pausing, recalling the theory, working through its implications. That pause is Williams' "one thought too many."

RE has one unified framework: the Reality Interaction Field (RIF). Within RIF, the 5W1H coordinates define the space, RIF Flux describes how it changes, and PTF (Domains, Constants, Hierarchy) provides the ethical structure. That is the entire architecture. A sharp fifteen-year-old can learn it in an afternoon. Once internalized, it does not interrupt moral intuition. It trains it.

The man saving his wife is not violating PTF. He is acting within it. His wife stands in his innermost Trust circle: family. His spontaneous act of rescue honors both her P-standing and his T-obligations to those closest to him. RE does not give him one thought too many. RE articulates the structure his moral intuition was already tracking.

Williams worried that moral theory alienates agents from their deepest commitments. But RE does not compete with those commitments. It provides the structural grammar within which commitments operate coherently. The husband who saves his wife, the parent who protects their child, the friend who keeps a confidence: these are not cases where RE must be consulted. These are cases where PTF is already functioning.

If RE becomes widely known, ethical reasoning does not slow down. It speeds up. When everyone shares a common, simple architecture, the moral calculus becomes faster and clearer. Disagreements become easier to locate: we can identify which RIF coordinate is disputed, which PTF domain is at stake, whether deception has caused incoherence. The framework is lean enough to prime intuition rather than replace it.

Williams was right to worry about alienation. Heavy theories do alienate. RE agrees and stands on the side observing.

9. Conclusion

9.1 What RE Claims

Resolution Ethics makes a structural claim about moral cognition.

Moral agents operate within the Reality Interaction Field (RIF): WHO, WHAT, WHERE, WHEN, WHY, HOW. These coordinates shift continuously. This is RIF Flux, and it creates vulnerability. Any agent navigating this flux while vulnerable requires certain structural conditions to function coherently.

Those conditions are Protection, Trust, and Free-Agency, ordered by dependency. Protection must precede Trust because coordination requires safety from arbitrary harm. Trust must precede Free-Agency because meaningful choice requires reliable information and non-manipulated conditions. P > T > F is not a ranking of values. It is a description of what depends on what.

Corruption enters through a single vector: deception. Self-deception distorts the agent's own moral reasoning. Other-deception corrupts the reasoning of others. Every moral violation, traced back far enough, passes through one of these gates. There is no honest path to incoherence. This is not an empirical claim awaiting falsification by counterexample. It is architecturally necessary: for moral cognition to violate structure it perceives, distortion must enter the justification chain, and that distortion is deception.

A clarification is necessary here. In RE, deception means model-distortion: any divergence between an agent's internal justification and the relevant structure of the situation. This includes intentional fraud (OD), rationalized self-interest (SD), negligence (SD about risk or duty), and incompetent judgment (SD about capacity or circumstance). The car accident involves SD: someone's map of "I am safe to proceed" did not match reality. The negligent landlord who skips maintenance operates under SD: "It will be fine" or "Not my responsibility." Addiction-driven harm traces to SD about control, priority, or harmlessness.

RE distinguishes moral incoherence from natural disaster. Lightning strikes a hiker. An earthquake collapses a building. A tree falls in a storm. These are flux events, not corruption. No moral agent's distorted map produced them. They carry no ethical weight because no moral cognition was involved.

The boundary matters. Natural disaster becomes moral incoherence when a moral agent's SD or OD enters the causal chain. The building collapses because of negligent construction. The tree falls because maintenance was skipped. The flood kills because warnings were suppressed. At that point, a moral agent's distorted map is doing work, and RE applies.

RE does not claim that all harm is corruption. It claims that all moral corruption traces to deception. Where no moral agent's map is distorting reality, there is misfortune but not incoherence. Where a moral agent's map is involved, RE can identify the distortion.

RE does not prescribe how to reason morally. It verifies whether reasoning has remained coherent with the structural constraints that vulnerable agency requires.

9.2 What RE Offers

The predecessors identified real features of moral life. Virtue ethics captures the importance of character. Deontology captures the binding force of duty. Consequentialism captures the relevance of outcomes. Prima facie pluralism captures the reality of competing considerations.

Each offers guidance. None offers verification.

RE sits underneath these frameworks. It does not compete with them. It checks whether their application in a given case has remained coherent with PTF. A virtuous act that violates Protection is not virtuous. A dutiful act grounded in self-deception is not dutiful. A welfare-maximizing calculation that corrupts Trust does not maximize welfare in any stable sense.

This matters because moral failure rarely announces itself. The agent who harms others typically believes they are justified. The institution that drifts toward predation typically tells itself a story about necessity or greater good. Corruption is invisible to the corrupted. That is why self-deception is the deeper vector.

RE offers detection. It identifies where coherence has broken, which domain is violated, whether the breach cascades. It catches the corruption before the cascade completes. This is what a verification layer does.

9.3 What Comes Next

This paper has laid the theoretical foundation. The companion paper, The Resolution Ethics Engine, shows what can be built on it.

REE operationalizes the framework. It takes the structural claims developed here and implements them as a verification system. The Engine parses scenarios into RIF coordinates, checks PTF coherence, traces deception vectors, and flags breaches. It works for humans with pen and paper. It works as a template that guides AI toward coherent ethical decisions. No new hardware. No exotic architecture. Just a verification layer that catches corruption before the cascade completes. RE and REE are functional, testable, and deployable.

References

Al-Ghazali, Abu Hamid. 1987. Al-Mustasfa min 'Ilm al-Usul. Translated by Ahmad Zaki Mansur Hammad. PhD diss., University of Chicago.

Aquinas, Thomas. 1265-1274. Summa Theologica. Translated by Fathers of the English Dominican Province.

Aristotle. 2000. Nicomachean Ethics. Translated by Roger Crisp. Cambridge: Cambridge University Press.

Aristotle. 1984. Metaphysics. In The Complete Works of Aristotle, edited by Jonathan Barnes. Princeton: Princeton University Press.

Augustine of Hippo. 397. Confessions.

Baier, Annette. 1986. "Trust and Antitrust." Ethics 96 (2): 231-260.

Beauchamp, Tom L., and James F. Childress. 2019. Principles of Biomedical Ethics. 8th ed. New York:

Oxford University Press.

Bostrom, Nick. 2014. Superintelligence: Paths, Dangers, Strategies. Oxford: Oxford University Press.

Confucius. 2003. Analects. Translated by Edward Slingerland. Indianapolis: Hackett Publishing.

Dancy, Jonathan. 1993. Moral Reasons. Oxford: Blackwell.

Foot, Philippa. 2001. Natural Goodness. Oxford: Clarendon Press.

Geneva Convention III. 1949. Convention Relative to the Treatment of Prisoners of War. Geneva: International Committee of the Red Cross.

Goodin, Robert E. 1985. Protecting the Vulnerable: A Reanalysis of Our Social Responsibilities. Chicago: University of Chicago Press.

Grotius, Hugo. 1625. De Jure Belli ac Pacis (On the Law of War and Peace).

Hierocles. 2009. Hierocles the Stoic: Elements of Ethics. Translated by Ilaria Ramelli and David Konstan. Atlanta: Society of Biblical Literature.

Hobbes, Thomas. 1651. Leviathan.

Hume, David. 1739. A Treatise of Human Nature.

Hursthouse, Rosalind. 1999. On Virtue Ethics. Oxford: Oxford University Press.

Kant, Immanuel. 1785. Groundwork of the Metaphysics of Morals.

Kant, Immanuel. 1797. The Metaphysics of Morals.

Kant, Immanuel. 1797b. "On a Supposed Right to Lie from Philanthropy." In Practical Philosophy, translated by Mary Gregor. Cambridge: Cambridge University Press, 1996.

Keown, Damien. 2005. Buddhist Ethics: A Very Short Introduction. Oxford: Oxford University Press.

Korsgaard, Christine. 1986. "The Right to Lie: Kant on Dealing with Evil." Philosophy & Public Affairs 15 (4): 325-349.

Korsgaard, Christine. 1996. The Sources of Normativity. Cambridge: Cambridge University Press.

Kunda, Ziva. 1990. "The Case for Motivated Reasoning." Psychological Bulletin 108 (3): 480-498.

Leopold, Aldo. 1949. A Sand County Almanac. New York: Oxford University Press.

Locke, John. 1690. Two Treatises of Government.

Mackie, J. L. 1977. Ethics: Inventing Right and Wrong. London: Penguin.

Mahabharata. c. 400 BCE–400 CE. The Mahabharata. Translated by Kisari Mohan Ganguli. Calcutta: Bharata Press, 1883-1896.

Madison, James. 1788. "Federalist No. 51." In The Federalist Papers, edited by Clinton Rossiter. New York: New American Library, 1961.

Marquis, Don. 1989. "Why Abortion Is Immoral." Journal of Philosophy 86 (4): 183-202.

Mencius. 2009. Mencius. Translated by Irene Bloom. New York: Columbia University Press.

Menkiti, Ifeanyi. 1984. "Person and Community in African Traditional Thought." In African Philosophy: An Introduction, edited by Richard A. Wright, 171-181. Lanham, MD: University Press of America.

Metz, Thaddeus, and Joseph B. R. Gaie. 2010. "The African Ethic of Ubuntu/Botho: Implications for Research on Morality." Journal of Moral Education 39 (3): 273-290.

Mill, John Stuart. 1859. On Liberty.

Mill, John Stuart. 1861. Utilitarianism.

Moore, G. E. 1903. Principia Ethica. Cambridge: Cambridge University Press.

Nozick, Robert. 1974. Anarchy, State, and Utopia. New York: Basic Books.

NSPE. 2019. Code of Ethics. National Society of Professional Engineers.

Patanjali. 2009. The Yoga Sutras of Patanjali. Translated by Edwin F. Bryant. New York: North Point Press.

Peck, M. Scott. 1983. People of the Lie: The Hope for Healing Human Evil. New York: Simon & Schuster.

Pettit, Philip. 1997. Republicanism: A Theory of Freedom and Government. Oxford: Oxford University Press.

Plato. 1992. Republic. Translated by G. M. A. Grube, revised by C. D. C. Reeve. Indianapolis: Hackett Publishing.

Prigogine, Ilya, and Isabelle Stengers. 1984. Order Out of Chaos: Man's New Dialogue with Nature. New York: Bantam Books.

Rawls, John. 1971. A Theory of Justice. Cambridge, MA: Harvard University Press.

Regan, Tom. 1983. The Case for Animal Rights. Berkeley: University of California Press.

Ross, W. D. 1930. The Right and the Good. Oxford: Clarendon Press.

Russell, Stuart. 2019. Human Compatible: Artificial Intelligence and the Problem of Control. New York: Viking.

Singer, Peter. 1975. Animal Liberation. New York: New York Review/Random House.

Sloan, M. C. 2010. "Aristotle's Nicomachean Ethics as the Original Locus for the Septem Circumstantiae." Classical Philology 105 (3): 236-251.

Stocker, Michael. 1976. "The Schizophrenia of Modern Ethical Theories." Journal of Philosophy 73 (14): 453-466.

Talmud Bavli. c. 500 CE. Babylonian Talmud. Edited by Adin Steinsaltz. Jerusalem: Koren Publishers.

Tutu, Desmond. 1999. No Future Without Forgiveness. New York: Doubleday.

Von Hippel, William, and Robert Trivers. 2011. “The Evolution and Psychology of Self-Deception.” Behavioral and Brain Sciences 34 (1): 1-16.

Walzer, Michael. 1994. Thick and Thin: Moral Argument at Home and Abroad. Notre Dame: University of Notre Dame Press.

Warren, Mary Anne. 1973. "On the Moral and Legal Status of Abortion." The Monist 57 (1): 43-61.

Williams, Bernard. 1981. "Persons, Character, and Morality." In Moral Luck: Philosophical Papers, 1973-1980, 1-19. Cambridge: Cambridge University Press.

Williams, Bernard. 1985. Ethics and the Limits of Philosophy. Cambridge, MA: Harvard University Press.