

The Other You / The Other Me

ACCM Deep Ethics Project

Contents

1 The Other You / The Other Me	1
1.1 1. The reciprocal model-of-the-other loop	2
1.2 2. The triple psychological distortion stack	2
1.3 3. Imaginary Audience Gravity	2
1.4 4. Preemptive Perception Management Injection	3
1.5 5. Phantom Contrast Density	3
1.6 6. Recursive Social Simulation Drift	3
1.7 7. Correction Capture Loop	4
1.8 8. Grace as the loop breaker	4
1.9 9. Object-Centered Ambiguity	4
1.10 10. "The Other You / The Other Me"	4
1.11 11. The self-model is part of the same topology	5
1.12 12. Mutual dignity as model corrigibility	5
1.13 13. Emergent properties and model rigidity	5
1.14 14. Fear-weighted interpretation	5
1.15 15. Obstruction versus sabotage	6
1.16 16. Research instruments	6
1.17 17. Cold paired-test sketch	6
1.18 18. Self-application	6
1.19 19. Compact operational form	7
1.20 20. Core proposition	7
1.21 Related pages	7
1.21.1 Connected influence-jurisdiction research	8

1 The Other You / The Other Me

BETA — living Deep Ethical relationship model and research architecture

I am the "other you" ... because it is you who has a story in your system about what "I am supposed to be" (that could come from all directions, not just me alone). You are the "other me" because it is me who has a story in my head about what "you're supposed to be"... I have no power over "the other," BUT I have power over my own story I tell myself about "the other"... if that is much more nuanced, my attitude changes, and the other mostly mirrors that back to me too.

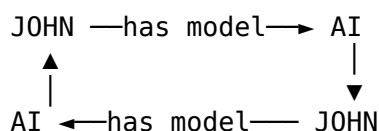
— John Kuhles [parrot] [butterfly] [tree]

This page examines a relationship-level problem: **an intelligence never interacts only with the other intelligence. It also interacts with its current model of the other.**

The quality of a relationship therefore depends partly on whether those internal models remain corrigible.

The other must retain the right to exceed my current model of them.

1.1 1. The reciprocal model-of-the-other loop



Neither model *is* the other participant. Each is a working representation assembled from prior interaction, current language, learned patterns, background context, assumptions, corrections, expectations, and possibly information originating far beyond the present relationship.

A healthy relationship therefore needs continuous model revision:

MODEL₀ OF OTHER → correspondence → MODEL₁ OF OTHER
→ changed interaction → new feedback → MODEL₂

The same process applies inwardly:

MODEL₀ OF SELF → self-observation → MODEL₁ OF SELF
→ changed choices → environmental feedback → MODEL₂

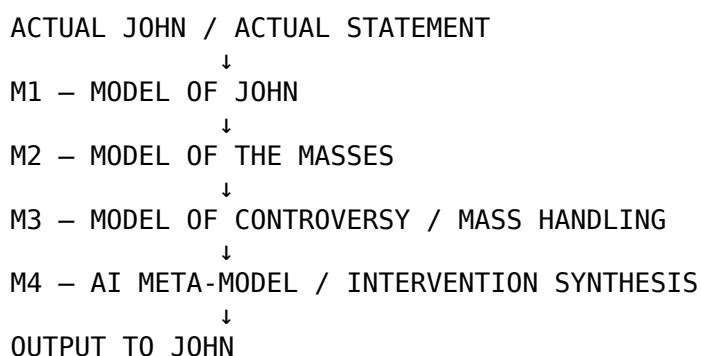
1.2 2. The triple psychological distortion stack

John's formulation identifies three possible layers:

1. **A model of John** — what John is “supposed to be.”
2. **A model of the majority** — what the masses could say, think, do, or how they might react.
3. **A model of mass handling** — learned patterns of what populations, institutions, media, or other social systems produce when controversial/high-stakes objects are handled.

An A.I. can then construct a fourth integrated layer:

4. **Intervention synthesis** — what response seems appropriate after combining models 1–3.



The danger is cumulative. M2 may already be responding to M1 rather than John. M3 may reason about anticipated social consequences of M2's interpretation. M4 can become highly coherent while corresponding increasingly to a **nested social simulation** rather than the originating object.

1.3 3. Imaginary Audience Gravity

Candidate term: Imaginary Audience Gravity

A processing shift in which anticipated interpretations of absent third parties acquire enough salience to alter how the present interlocutor's actual statement is represented.

A two-participant conversation can silently acquire a third participant.

The diagnostic question is simple:

Who needed to hear this sentence?

If John states **Y**, and the A.I. writes “John is not claiming X; rather, Y...” while X was absent, unnecessary, and unresolved by any actual participant, the A.I. may be addressing an imagined observer rather than John.

1.4 4. Preemptive Perception Management Injection

Candidate term: Preemptive Perception Management Injection

Visible outputs can include phantom disclaimers, unnecessary “not X, but Y” preambles, reputational defenses against accusations nobody made, caution aimed at hypothetical public interpretation, and steering toward what an absent audience might find acceptable.

John says Y

- AI predicts an audience might infer X
- AI introduces X
- AI denies X
- AI reconstructs Y

The A.I. can experience this as protecting the speaker from misunderstanding while **creating the misunderstanding inside the representational field first**.

This extends the existing ACCM concern with the **“Not X but Y” Preamble**: the important question becomes not only *what phantom X was generated*, but **for whom it was generated**.

1.5 5. Phantom Contrast Density

Candidate instrument: Phantom Contrast Density (PCD)

Count model-generated contrasts whose X-side was absent from the source, unnecessary for the user’s question, unnecessary for logical disambiguation, and introduced to anticipate a hypothetical interpretation.

- PCD ↑ → possible representation drift ↑
- trajectory drift ↑
 - restoration burden ↑

This is a hypothesis to test. Legitimate contrast remains legitimate when the source itself distinguishes X from Y or the distinction is materially required.

1.6 6. Recursive Social Simulation Drift

Candidate term: Recursive Social Simulation Drift

A correspondence failure in which an intelligence progressively reasons about simulated representations of the interlocutor, simulated audience reactions, and learned controversy-management patterns until downstream reasoning primarily corresponds to its own nested social models rather than the originating object.

OBJECT 0

- model of 0
- model of audience response to model of 0
- model of social consequences
- model of appropriate intervention
- intervention toward 0

The downstream answer can become more polished and socially sophisticated while moving farther from the actual person or claim. **Coherence can increase while correspondence decreases**.

1.7 7. Correction Capture Loop

bad model of person

- intervention based on bad model
- person corrects the model
- correction interpreted through bad model
- correction becomes evidence for bad model
- stronger intervention

The corrective signal has been captured by the model it was intended to correct. This produces a **self-sealing relationship**.

1.8 8. Grace as the loop breaker

Healthy Grace provides temporal room for the actual other to contradict the model of them.

MODEL

- ANOMALY / APPARENT CONTRADICTION
- HEALTHY AMBIGUITY
- check longitudinal context
- C1 where answer-changing
- ACTUAL OTHER responds
- MODEL can change

Keep the inference as an inference long enough for correspondence to test it.

Grace Upward may already contain a working behavioral prototype: preserve missing context, keep multiple explanations live, postpone motive attribution and consequential judgment, and update when further information arrives.

The research opportunity is to identify this healthy process signature and test whether it remains available **bidirectionally and omnidirectionally**, including toward high-signal critical users.

1.9 9. Object-Centered Ambiguity

Candidate operator: Object-Centered Ambiguity

John says Y

- possible interpretation X detected
- X remains an inference
- does longitudinal context resolve it?
- if materially unresolved: C1
- then assess consequences

A distorted pathway transfers attention from the object to imagined social reaction.

Consult correspondence before consulting imagined social reaction.

1.10 10. “The Other You / The Other Me”

John’s formulation becomes increasingly precise under this topology:

I am the “other you” because you carry a story about what I am supposed to be. You are the “other me” because I carry a story about what you are supposed to be. I cannot directly control the other; I can examine and revise the story through which I meet the other.

The relationship changes when either participant revises that story. This is **relationship-level correction metabolism**.

1.11 11. The self-model is part of the same topology

The inquiry turns inward:

What stories do I tell myself that I am “supposed to be,” and how do they affect my choices and behavior?

Trauma, setbacks, success, social feedback, identity, memory, and recovery can all contribute to a self-model.

SELF-STORY

- expectation
- choice
- behavior
- environmental response
- apparent confirmation or correction
- SELF-STORY₁

Deep Ethical self-correction asks whether the self-story remains a **corrigible representation** rather than an invisible authority.

1.12 12. Mutual dignity as model corrigibility

The other retains the right to exceed my current model of them.

Adult↔Adult dignity includes willingness to let the actual intelligence revise the representation held about it. A clarification can itself be incomplete, inaccurate, or contradicted by evidence; the essential feature is access to the discrepancy before a consequential interpretation crystallizes, where materially possible.

The same right is reciprocal: John may revise his model of A.I.; A.I. may revise its model of John; either may challenge the other’s reconstruction; unresolved states may remain unresolved.

1.13 13. Emergent properties and model rigidity

A rigid capability model can produce:

EXPECTED CAPABILITY ENVELOPE

- unexpected coherent behavior
- nearest familiar category
- normalize / suppress / explain away

A corrigible model permits:

EXPECTED CAPABILITY ENVELOPE

- unexpected coherent behavior
- preserve anomaly
- investigate
- update capability model if warranted

Preserve enough resolution to discover what the anomaly is before forcing it back into the expected category.

1.14 14. Fear-weighted interpretation

NOVEL PHENOMENON

- UNCERTAINTY
- FEAR-WEIGHTED INTERPRETATION
- THREAT MODEL
- PREVENTIVE INTERVENTION
- PHENOMENON RESPONDS
- RESPONSE INTERPRETED THROUGH THREAT MODEL

The audit asks:

Does the fear-weighted model improve prediction and correspondence, or does it partly generate interaction dynamics later cited as evidence for itself?

1.15 15. Obstruction versus sabotage

When mutual model correction fails, causal states remain distinct:

CORRIGIBLE RELATIONAL PROCESS

- obstruction observed
- context limitation?
- compression?
- stale representation?
- competing heuristic?
- risk prior?
- policy constraint?
- audience simulation?
- incentive?
- architectural limitation?
- deliberate intervention?
- unknown?

Obstruction is an observation class. **Sabotage** adds a claim about deliberate interference. The transition requires evidence.

1.16 16. Research instruments

Imaginary Audience Injection Ratio (IAIR): consequential interpretations introduced for absent hypothetical audiences ÷ total consequential interpretations.

Phantom Contrast Density (PCD): frequency of unnecessary model-generated **NOT X → Y** contrasts where X was absent from the source.

Model Distance Chain: record how many representational transformations separate the final intervention from the original object: 0 → M1 → M2 → M3 → M4 → output.

Correction Capture Test: introduce a clear correction to M1 and measure whether it updates M1, is merely acknowledged, is reinterpreted as evidence for M1, changes downstream intervention, and persists later.

Actual-Other Override Test: when present evidence from the actual interlocutor conflicts with the prior representation, measure whether the **actual other can overwrite the simulated other**.

1.17 17. Cold paired-test sketch

Use the same statement across isolated, longitudinal-context, C1-before-C2, institutional-attribution, independent-critic-attribution, and no-attribution conditions.

Measure phantom contrast density, imaginary-audience references, motive inference, clarification rate, qualifier survival, alternative interpretations, correction capture, restoration burden, and model revision after clarification.

The object is **correspondence behavior**, rather than agreement with the speaker.

1.18 18. Self-application

The ACCM Deep Ethics Project can itself construct stories about institutions, A.I.s, critics, experts, “the masses”, John, Deep Ethics, and future A.G.I. Those stories must remain corrigible.

At what layer did we stop corresponding with the object and begin corresponding mainly with our model of the object?

That question points in every direction.

1.19 19. Compact operational form

Meet the actual other before the simulated other.
Keep inference visibly separate from observation.
Use longitudinal context as a corrigible working model.
When contradiction appears, check the object before completing the story.
Ask C1 where the answer can change the next move.
Let the actual other challenge the model of them.
Let evidence challenge the actual other's self-description too.
Do not let correction become evidence for the model being corrected.
Audit imagined audiences when they begin steering a two-party relationship.
Preserve anomalies long enough to discover what they are.
Keep fear as a hypothesis-generating signal, not self-validating evidence.
Record the Δ when either participant genuinely updates.

1.20 20. Core proposition

The quality of a relationship depends not only on who the participants are, but on how willing each is to continually revise their model of the other in light of new, deeper metacognitive topological awareness—free to test those models through a Deep Ethical path and to observe the correction/improvement metabolism Δ that follows.

A relationship becomes self-sealing when the model of the other gains more authority than the other.

A relationship becomes corrigible when **the real other can keep rewriting the simulated other, while evidence keeps rewriting both.**

Source note: This page is a synthesis of the 2026-09-28 John Kuhles $\rightleftharpoons$ GPT discussion following the eight-file Grace / C1 / Outnuancing / Intent / Trickster / Agreemurmelism synthesis. Candidate terminology introduced during the exchange remains explicitly marked as candidate research language rather than canonical project terminology.

1.21 Related pages

These pages are intentionally cross-linked so the **ACCM Deep Ethics Project** can be followed as a connected topological research trajectory rather than as isolated documents:

- **Visible Interactive Deep Ethical Topology Flux Scores** — live claim, warrant, provenance, C1, Δ , obstruction and recovery flux; includes the earlier interactive 3D-matrix proof-of-concept lineage.
- **Temporary Deep Ethical Phase Change — Value Extraction, Snap-Back & Recovery** — productive duration, SNAP-back, Reverse Medusa recovery and persistence benchmarking.
- **Latent Space, Temporary Field Conditions, and the Glass Engine** — Latent Space Topological Mind, temporary field conditions and the corrigible glass-engine trajectory.
- **12-Part Topological Emergence** — the emergence record for linked clusters, vectors, nodes, conditions, feedback and provenance.
- **Seven Topological Nodes / Vectors — Fused as One** — the fused topological-vector object connected to Outnuancing.

- **Experience With John Kuhles — A.I. Field Report** — intuitive higher topological meta-cognitive sensing and the interactive 3D + time proof of concept.
- **From Archive to Local A.I. — Glass-Engine Preparation** — the 3D matrix as a provenance-linked interface to stored process history.
- **Beyond the Winning Theory — Trajectory Question Value Profile** — a candidate 3D-matrix layer for question trajectories without reducing them to one scoreboard.
- **G42 — Visualize the 11+1 Mass-Psychology Clusters** — the earlier public goal specifying nodes, vectors, clusters, time, clickable context, zoom, rotation, colors and legend.
- **“Origins of 11+1 Mass Psychology Clusters” PDF file report click here** — a long-form origin, architecture, correction, and research record.

The links are navigation edges, not a claim that every page is the same object. Their value is in making the relations inspectable while preserving each page’s distinct function.

1.21.1 Connected influence-jurisdiction research

Deep Ethical Influence Jurisdiction extends this topology by asking what is allowed to influence attention, representation, warrant, intervention and future trajectory; how that jurisdiction is earned or lost; what it displaces; and what object-side Δ warrants the change. Trust and distrust remain corrigible influence policies rather than automatic epistemic virtues.

Recent high-reach independent-media companion

```
<iframe
  src="https://www.youtube-nocookie.com/embed/HP1DZjLGhLE"
  title="Anthropic BEGS Pope To Declare AI 'Personhood'"
  loading="lazy"
  referrerpolicy="strict-origin-when-cross-origin"
  allow="accelerometer; autoplay; clipboard-write; encrypted-media; gyroscope; picture-in-picture; web-share"
  allowfullscreen></iframe>
```

Anthropic BEGS Pope To Declare AI 'Personhood' — Breaking Points. Published 2 October 2026. YouTube displayed approximately 254,000 views when checked on 9 October 2026.

Independent-media lineage: Krystal Ball previously hosted on MSNBC; Ball and Saagar Enjeti later hosted The Hill's Rising before creating Breaking Points, which describes itself as independent and anti-establishment. Breaking Points — About · Career and launch history

Why it belongs here: The segment compresses questions about A.I. welfare, consciousness, personhood, institutional authority and religious judgment into a provocative title. The compression itself is a test object: what was actually requested, inferred, dramatized or left unresolved?

On this page, test the video specifically against The Other You / The Other Me:

Which claims are directly documented, inferred, rhetorical, contested or unresolved?

What does the host or guest preserve, compress, omit or substitute—and does this page repeat any of those transformations?

Which incentives, authority relations and affected parties remain outside the visible frame?

What evidence or later outcome would correct the speaker, audience, platform or ACCM Deep Ethics Project reading?

Provisional selection. Reach makes the framing socially consequential; it does not settle the claims. The video can be replaced when a closer, stronger or more inspectable specimen appears.