

Moral Cognition Explanandum Audit

What representative programs take moral cognition to be

Bree Beal
Version 1.2 - September 2026

Status. This is a scoped, non-authoritative RC3 literature audit supporting Registry diagnosis D-LIT-001. It compares representative programs already governed by Master Bibliography 1.8 and Source Register 2.15. It does not prove that every existing or possible theory lacks a moral/nonmoral boundary, and it does not promote NDT from formal specification to operational identification or calibrated measurement.

1. Question and conclusion

NDT begins from a diagnosis: moral psychology contains methodologically varied and technically developed experimental and computational programs without comparable agreement on what makes a cognitive token moral rather than nonmoral. The evidence-bearing conclusion of this audit is narrower than a discipline-wide absence claim. Across the representative programs below, the operational target changes - moral content, perceived harm, cooperation, relationship regulation, affect-linked norms, welfare learning, or agreement-based judgment - and no shared token-level provenance criterion separates moral from nonmoral normativity.

This divergence matters even when each local program is rigorous. Experimental or computational precision can identify processes within a selected class of cases without independently establishing why those cases, rather than behaviorally or verbally similar controls, instantiate moral cognition.

2. Selection and limits

The sample is purposive rather than exhaustive. It includes prominent content, template, functional, relational, affective, learning, and contractualist programs already represented in the governed comparison layer. The audit asks one question of each: what phenomenon is operationalized, and does that operation independently distinguish moral cognition from nonmoral normativity at the judgment-token level?

A program may contribute a valid module or restricted model family without supplying the general boundary. Failure to supply this boundary is not evidence that the program's experiments or local predictions are unsound. The audit does not treat NDT's own formal boundary as empirically identified merely because rivals lack a shared one.

3. Representative comparison

Program and governed source	Operational target and boundary consequence
Moral Foundations Theory (S16, S51)	Organizes moral concerns through evolved or culturally elaborated foundations/modules and measures endorsement or relevance across those domains. This supplies a rich content architecture, but the same foundation-linked content can be activated in strategic, institutional, habitual, or otherwise nonmoral normative cognition; token provenance is not the type criterion.
Theory of Dyadic Morality (S17, S52)	Models condemnation through norm violation, negative affect, and perceived dyadic harm. It offers a substantive harm-based boundary, but harm perception and condemnation remain the operational center; it does not distinguish same-content, same-response tokens by whether a represented moral relationship actually organized the evaluation.
Morality-as-Cooperation (S19)	Treats morals as biological and cultural solutions to recurrent cooperation problems and tests the moral valence of cooperative behaviors. This is a functional/content proposal: cooperation can also be prudential, strategic, institutional, or habitual, so cooperative success alone does not type the cognition that produced it.

Program and governed source	Operational target and boundary consequence
Relationship Regulation Theory (S54)	Explains moral judgment through motives for regulating unity, hierarchy, equality, and proportionality. Its represented relational motives overlap NDT's relationship-sensitive explanandum at the construct level, but relationship form or motive category does not by itself establish the actual token-level contribution required by MDT.
Sentimental-rules account (S58)	Explains the moral/conventional distinction through prohibited-action information together with an affective mechanism. This is a genuine classification proposal, but the norm-plus-affect package is neither necessary nor sufficient for NDT moral type and does not supply the broader boundary between moral and every nonmoral normative route.
Bayesian welfare learning (S21)	Infers weights attached to welfare-centered principles from observed behavior. The computation makes learning explicit and testable, but the agents, welfare values, and moral target space are fixed before learning; it models inference within a selected moral interpretation rather than how a judgment acquires moral type.
Resource-rational contractualism (S40)	Models action-, welfare-, and agreement-based judgments as resource-rational computations. It supplies a powerful restricted family of moral reasoning, but a contractualist or agreement-based computation can occur inside moral or nonmoral normativity; the philosophical family does not type every token that realizes it.

4. What the audit supports

- A concise authorial diagnosis that technically developed programs do not share one specification of the moral-cognition explanandum.
- NDT's claim to conceptual and architectural importance: it explicitly targets the moral/nonmoral boundary at the judgment-token level.
- A literature-review strategy based on rigorously typed construct comparisons rather than whole-theory name matching.

5. What the audit does not support

- The universal claim that no existing theory anywhere defines moral cognition noncircularly.
- The inference that methodological sophistication is scientifically empty or that the audited programs have no local explanatory value.
- Any change from claimed to demonstrated in the Stage 6 evidential partition.
- The inference that NDT's builders, thresholds, or latent provenance organization are already operationally identified.

6. Governed public wording

NDT begins from a diagnosis: moral psychology has developed methodologically varied and technically sophisticated experimental and computational programs without comparable agreement on what makes a cognitive token moral rather than nonmoral. A representative audit supports this scoped diagnosis; it does not claim exhaustive proof that no existing or possible theory could supply such a boundary.

7. Governed sources

- S16 / S51 - Graham and colleagues on Moral Foundations Theory.
- S17 / S52 - Schein and Gray on the Theory of Dyadic Morality.
- S19 - Curry, Mullins, and Whitehouse on Morality-as-Cooperation.
- S54 - Rai and Fiske on Relationship Regulation Theory.
- S58 - Nichols on norms with feeling and the moral/conventional distinction.
- S21 - Kleiman-Weiner, Saxe, and Tenenbaum on Bayesian moral learning.
- S40 - Levine, Chater, Tenenbaum, and Cushman on resource-rational contractualism.