Here is a **JHSS-ready rewrite of Section 3 (Framework + Experimental Design + Testable Predictions)** that fully incorporates the missing rigor, removes circularity, and aligns with the stronger ChatGPT framework.

---

3. Toward a Mechanism-Specific, Regime-Based Framework of Altruism

3.1 From Traits to Mechanisms: A Regime-Based Model

A central limitation of existing research is the treatment of altruism as a single latent trait inferred from observed helping behavior. This approach is inherently ambiguous, as similar behaviors can arise from fundamentally different underlying processes. For example, helping a sibling, reciprocating a favor, responding to distress, and enforcing fairness norms may all appear behaviorally similar but are driven by distinct cognitive and motivational mechanisms.

To address this issue, we propose a **regime-based framework** in which altruistic behavior is generated by **distinct mechanisms that are selectively activated depending on context**, rather than by a single trait or a weighted combination of factors. Specifically, we identify four primary regimes:

- **Kin-based regime (K):** behavior driven by genetic relatedness and familial attachment
- **Reciprocity-based regime (R):** behavior driven by expectations of future return
- **Empathy-based regime (E):** behavior driven by affective responses to others' distress
- **Norm-based regime (N):** behavior driven by internalized rules and fairness principles

In this framework, behavior at any given moment is generated by the **dominant active mechanism**, not by additive contributions from multiple mechanisms. This resolves a key source of circularity in prior models, where variables are both used to define and explain altruism.

3.2 Non-Circular Identification of Mechanisms

A critical challenge for any mechanism-based model is determining which mechanism is active in a given situation without inferring it directly from the behavior being explained. Inferring "empathy" from helping behavior, for example, is circular if helping is also the outcome to be explained.

To avoid this, we treat mechanism identification as a **separate inference problem** based on **independent evidence**, rather than behavioral outcomes alone. Specifically, mechanisms are identified using three complementary approaches:

(1) Contextual Manipulation (Causal Induction)
Experimental conditions are designed such that only one mechanism is plausibly operative. For example:

- Kin condition: recipient is a close family member; no repetition or norm cues
- Reciprocity condition: repeated interaction with future payoff
- Empathy condition: visible distress; no future interaction
- Norm condition: rule violation with opportunity for costly punishment

This approach induces mechanisms rather than inferring them post hoc.

(2) Process-Level Signatures (Mechanistic Evidence)
Each mechanism is associated with distinct cognitive and behavioral signatures:

- Empathy: rapid decisions, attention to distress cues, elevated emotional arousal
- Reciprocity: slower, deliberative decisions sensitive to future payoff
- Norm enforcement: sensitivity to rule violations, willingness to punish at a cost
- Kin: preference based on relational identity, relatively insensitive to norms

These signatures can be measured using reaction time, attention patterns, and self-reported motivations.

(3) Perturbation (Causal Necessity)
Selective disruption of mechanisms provides additional evidence:

- Cognitive load reduces reciprocity (strategic reasoning)
- Emotion suppression reduces empathy-driven responses
- Removal of norm cues reduces norm-based punishment

If removing a mechanism eliminates the behavior, this supports its causal role.

Together, these approaches allow for **non-circular, causally grounded identification** of underlying mechanisms.

3.3 A Unified Experimental Framework

To operationalize this model, we propose a single experimental paradigm that systematically varies context while holding the decision structure constant.

Core Task:
Participants repeatedly decide whether to incur a cost (e.g., money, time, or effort) to benefit another individual.

Conditions (within-subject design):

- **Kin (K):** recipient is a family member; one-shot interaction
- **Reciprocity (R):** repeated interaction with same partner; potential future return
- **Empathy (E):** recipient presented in distress; no future interaction
- **Norm (N):** third-party rule violation; participant can pay to enforce fairness

Process Measures:

- Reaction time (fast vs deliberative decisions)
- Attention (e.g., focus on faces vs payoff information)
- Post-decision motivation reports (e.g., “felt bad,” “expected return,” “rule violation”)

This design enables direct comparison of behavior across contexts within the same individual, allowing for the identification of **mechanism switching** rather than stable trait expression.

3.4 Testable Predictions and Falsifiability

The regime-based framework generates clear, falsifiable predictions:

(1) Dissociation Across Conditions
Each mechanism should dominate behavior only in its corresponding context:

- Kin preference in K condition
- Strategic cooperation in R condition

- Emotion-driven helping in E condition
- Costly punishment in N condition

(2) Process-Level Differentiation
Distinct mechanisms should exhibit different cognitive signatures:

- Empathy → fast, affect-driven responses
- Reciprocity → slower, deliberative responses
- Norm enforcement → rule-sensitive decisions

(3) Variable Irrelevance (Critical Test)
Variables relevant to one mechanism should be irrelevant in others:

- Future payoff should not affect empathy-driven behavior
- Emotional distress should not affect reciprocity-driven decisions
- Norm framing should have minimal effect in kin-based contexts

Failure of this prediction would directly falsify the model.

(4) Within-Subject Regime Switching
The same individual should exhibit different behaviors across conditions, demonstrating that altruism is not a fixed trait but a **context-dependent process**.

3.5 Implications for the Nature–Nurture Debate

This framework reframes the nature-versus-nurture debate by shifting the focus from whether altruism is innate or learned to **how different mechanisms are activated under different conditions**. Genetic and environmental factors do not independently determine altruism; instead, they influence the **probability of entering specific regimes**.

For example:

- Genetic factors may modulate sensitivity to social cues (e.g., empathy)
- Environmental factors may shape norm internalization or expectations of reciprocity

Thus, altruistic behavior emerges from the interaction between predispositions and context, rather than from a single underlying source.

Why this version is now JHSS-ready

This rewrite:

- Eliminates circular reasoning
- Introduces causal identification logic
- Provides a unified, testable experimental design
- Generates falsifiable predictions
- Moves from descriptive review → mechanistic theory