

CORE CODE THEORY:

How Organizational Trust Converts to Crisis Resilience

A Theory of Human Capital Inimitability in the Age of Artificial Intelligence

Tong Yin

Auburn University

Submitted to: Academy of Management Review

January 2026

ABSTRACT

We develop *Core Code Theory* to explain how organizational trust converts into crisis resilience—a phenomenon inadequately addressed by existing trust and psychological contract research. While Mayer, Davis, and Schoorman (1995) established *when* employees trust and Rousseau (1989) distinguished relational from transactional contracts, neither framework explains the *mechanism* through which accumulated trust transforms into collective sacrifice behavior during organizational crises. We propose that trust operates as an *organizational option*—an asset that appears costly during stable periods but yields exponential returns when exercised during disruption.

Central to our theory is the distinction between *Performance UI* (visible, measurable capabilities increasingly replicable by artificial intelligence) and *Core Code* (identity-based capabilities—moral courage, non-linear intuition, principled dissent—that emerge only from relational contracts and remain uniquely human). We argue that extraction-based management systems, by optimizing for Performance UI through transactional contracts, systematically destroy the Core Code that constitutes the only sustainable source of competitive advantage in AI-saturated environments. We advance three core propositions explaining why relational contracts generate trust reserves, how these reserves convert to crisis resilience through collective sacrifice mechanisms, and why this conversion exhibits non-linear returns. Our theory extends the trust literature by specifying the *payoff structure* of organizational trust and contributes to resource-based theory by identifying which human capabilities remain inimitable when cognitive tasks become algorithmic.

Keywords: organizational trust; psychological contracts; crisis resilience; human capital; artificial intelligence; Core Code

INTRODUCTION

Consider two organizations facing identical revenue shocks of 60 percent. In the first, senior executives immediately begin interviewing elsewhere; within weeks,

institutional knowledge accumulated over a decade disperses across competitors. In the second, the leadership team voluntarily accepts pay cuts before being asked, middle managers follow, and the organization retains every key employee through the crisis. What explains this dramatic divergence in crisis response?

Organizational trust theory provides a partial answer. Mayer, Davis, and Schoorman's (1995) integrative model establishes that trust emerges when employees perceive ability, benevolence, and integrity in their organization. Psychological contract theory (Rousseau, 1989, 1995) distinguishes relational contracts (characterized by long-term orientation and socio-emotional exchange) from transactional contracts (characterized by short-term economic exchange). Yet neither framework adequately explains the *mechanism* through which accumulated trust transforms into the collective sacrifice behavior observed in our second organization.

We develop *Core Code Theory* to address this gap. Our theory makes three contributions to the organizational trust literature. First, we specify the *payoff structure* of organizational trust by conceptualizing trust as an *organizational option*—an asset that appears costly during stable periods (requiring investment in benevolence signals) but yields exponential returns when “exercised” during disruption (through collective sacrifice that dramatically reduces crisis costs). This extends Mayer et al.'s (1995) framework by explaining not just when trust emerges but *what organizations gain* from cultivating it.

Second, we introduce the distinction between *Performance UI* and *Core Code* to identify which human capabilities trust cultivation protects. Performance UI represents visible, measurable contributions—KPIs, strategic analyses, operational execution—that artificial intelligence increasingly replicates. Core Code represents identity-based capabilities—moral courage to resist flawed directives, non-linear crisis intuition, capacity to inspire collective sacrifice—that emerge only from relational contracts and remain uniquely human. This distinction extends resource-based theory (Barney, 1991) by specifying which human resources remain inimitable in AI-saturated environments.

Third, we explain the *conversion mechanism* through which trust reserves transform into crisis resilience. We propose that relational contracts create *identity-based commitment*—employees perceive themselves as organizational “citizens” rather than “mercenaries”—which activates collective sacrifice behavior when the organization faces threats. This mechanism explains why the payoff from trust investment is non-linear: small differences in relational contract depth produce large differences in crisis outcomes.

Our theory responds to calls for research that moves beyond demonstrating that trust matters to explaining *how* it matters (Fulmer & Gelfand, 2012; Schoorman, Mayer, & Davis, 2007). By specifying the payoff structure and conversion mechanism of organizational trust, we provide a framework for understanding why some organizations exhibit remarkable resilience while others collapse under identical pressures.

THEORETICAL CONTEXT: THE TRUST-RESILIENCE GAP

We situate our theory within the organizational trust literature, engaging specifically with Mayer et al.'s (1995) integrative model and Rousseau's (1989, 1995) psychological contract framework. We identify a critical gap in this literature: while existing theory explains the *antecedents* and *types* of trust, it does not adequately explain the *payoff structure* of trust—specifically, how accumulated trust converts to organizational outcomes during crises.

The Antecedents Problem: Knowing When Without Knowing How

Mayer et al.'s (1995) model represents the dominant framework for understanding organizational trust. They define trust as "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party" (p. 712). Their model identifies three antecedents: *ability* (competence and skills), *benevolence* (care for the trustor's welfare), and *integrity* (adherence to acceptable principles).

This framework powerfully explains *when* employees trust organizations. However, it does not specify *what* organizations gain from being trusted, nor *how* trust translates into organizational outcomes. The model treats trust as the dependent variable; we extend it by treating trust as an *independent* variable that predicts crisis resilience. Schoorman, Mayer, and Davis (2007) acknowledged this limitation, calling for research examining "the consequences of trust" (p. 347). Our theory responds to this call.

The Contract Type Problem: Categories Without Mechanisms

Rousseau's (1989, 1995) psychological contract theory distinguishes transactional from relational contracts. Transactional contracts feature "specific, short-term, monetizable obligations" with "limited mutual investment" (Rousseau, 1995, p. 91). Relational contracts feature "open-ended relationships involving considerable investments by both employees... and organizations" with "high degrees of mutual interdependence" (p. 91).

This distinction is foundational, yet it does not explain the *mechanism* through which contract type translates into behavior during organizational crises. Why do employees with relational contracts exhibit collective sacrifice while those with transactional contracts pursue exit? Rousseau's framework implies that relational contracts generate loyalty, but it does not specify the psychological process through which this occurs. We address this gap by proposing that relational contracts create *identity-based commitment*—a transformation in how

employees perceive their relationship to the organization—that activates specific behavioral responses during crises.

The Payoff Structure Problem: Knowing Trust Matters Without Knowing How Much

Research consistently demonstrates that trust correlates with desirable organizational outcomes (Dirks & Ferrin, 2002; Colquitt, Scott, & LePine, 2007). However, this literature treats trust as having *linear* effects—more trust produces proportionally better outcomes. We propose that trust's payoff structure is fundamentally *non-linear*: trust operates as an organizational option that yields exponential returns during crises.

The option metaphor captures trust's essential economic character. Like financial options, trust requires upfront investment (in benevolence signals, psychological safety, dignified practices) that appears costly during stable periods. The "option" becomes valuable precisely when exercised during disruption—when accumulated trust converts to collective sacrifice that dramatically reduces crisis costs. This non-linear payoff structure explains why organizations with marginally higher trust may exhibit dramatically different crisis outcomes.

CORE CODE THEORY

We now develop Core Code Theory, which explains how organizational trust converts to crisis resilience. Our theory comprises three elements: (1) the Performance UI/Core Code distinction, which identifies *what* trust protects; (2) the identity-based commitment mechanism, which explains *how* trust converts to collective sacrifice; and (3) the trust-as-option framework, which specifies the *payoff structure* of organizational trust.

The Performance UI/Core Code Distinction

We introduce the distinction between *Performance User Interface (Performance UI)* and *Core Code* to specify which human capabilities trust cultivation protects. This distinction addresses a critical question for resource-based theory: as artificial intelligence replicates an expanding range of cognitive tasks, which human capabilities remain inimitable?

Performance UI comprises the visible, measurable outputs of organizational members: key performance indicators, strategic analyses, data-driven decisions, operational execution. These capabilities share a critical property: they are *legible* to algorithmic systems. An AI can observe, learn from, and eventually replicate analytical reasoning, pattern recognition in structured data, and rule-based decision making. Performance UI corresponds to what Motowidlo (2003) termed "task performance" and aligns with transactional psychological

contracts—employees provide specific, measurable outputs in exchange for compensation.

Core Code comprises the invisible, identity-based capabilities that emerge from relational contracts: moral courage to resist directives perceived as unethical, non-linear intuition that transcends programmed responses, and capacity to inspire trust-based collective sacrifice. These capabilities share a critical property: they are *illegible* to algorithmic systems because they emerge from identity rather than task competence. An employee who perceives herself as an organizational "citizen" will resist a directive she believes harmful to the organization's long-term interests; an employee who perceives himself as a "mercenary" will comply with any directive within contractual bounds.

This distinction extends existing frameworks in important ways. Organ's (1988) concept of organizational citizenship behavior (OCB) identifies discretionary contributions beyond task requirements. However, OCB research focuses primarily on stable-state behaviors (helping colleagues, civic virtue) rather than crisis-state behaviors (collective sacrifice, principled dissent under pressure). Core Code captures the *crisis-activated* dimension of identity-based commitment—capabilities that become visible and valuable precisely when organizations face existential threats.

The terminology deliberately invokes software architecture. In software systems, the user interface (UI) comprises visible features that users interact with directly; the core code comprises underlying logic that determines how the system responds to unexpected inputs. Similarly, Performance UI represents what organizations can observe and measure; Core Code represents what determines organizational response to crisis—the "unexpected inputs" that stress-test organizational systems. Organizations that optimize for Performance UI through extraction-based management may achieve impressive visible outputs while systematically degrading the Core Code that determines crisis resilience.

The Identity-Based Commitment Mechanism

Why do relational contracts generate crisis resilience while transactional contracts do not? We propose that the conversion mechanism operates through *identity-based commitment*—a transformation in how employees perceive their relationship to the organization that activates specific behavioral responses during crises.

Drawing on social identity theory (Tajfel & Turner, 1986) and organizational identification research (Ashforth & Mael, 1989), we argue that relational contracts create the conditions for employees to incorporate organizational membership into their self-concept. When employees perceive organizational benevolence (through psychological safety, career investment, dignified treatment), they reciprocate not merely through task performance but through *identity incorporation*—they come to see organizational welfare as component of personal welfare.

This identity transformation activates a distinct behavioral response during crises. Employees with transactional contracts perceive organizational distress as a signal to pursue exit options; the contract is explicitly short-term and monetizable, so declining organizational fortunes rationally trigger job search. Employees with identity-based commitment perceive organizational distress as a threat to self; the organization's welfare has become component of their own welfare, so they respond with protective behaviors—collective sacrifice, voluntary pay reduction, intensified effort—that would be irrational under transactional logic.

The identity mechanism explains why trust's crisis benefits are *qualitatively different* from its stable-state benefits. During stable periods, relational contracts yield incremental advantages: somewhat higher engagement, somewhat lower turnover, somewhat greater organizational citizenship behavior. During crises, relational contracts yield categorical advantages: the difference between employees who stay and sacrifice versus employees who exit. This qualitative shift—from proportional to categorical effects—produces the non-linear payoff structure we discuss next.

Trust as Organizational Option: The Payoff Structure

We conceptualize organizational trust as an *option*—an asset that requires upfront investment, appears costly during stable periods, and yields exponential returns when exercised during disruption. This framework specifies the economic logic of trust investment and explains why marginal differences in trust accumulation produce dramatic differences in crisis outcomes.

The option framework illuminates three properties of organizational trust that existing theory obscures:

Investment costs during stability. Building trust requires tangible investment in benevolence signals: comprehensive benefits, job security guarantees, dignified treatment even of departing employees. These investments reduce short-term profitability and may appear wasteful to observers focused on Performance UI metrics. The option framework explains why rational organizations make these investments: they are purchasing crisis insurance.

Asymmetric payoffs. Trust's payoff structure is fundamentally asymmetric. During stable periods, high-trust and low-trust organizations may perform similarly on observable metrics; the trust "premium" appears modest. During crises, the differential becomes categorical. Low-trust organizations face talent flight, emergency recruitment costs, institutional memory loss, and potential collapse. High-trust organizations "exercise" their accumulated trust through collective sacrifice, maintaining human capital through the crisis at dramatically lower cost.

Non-linearity of returns. The option framework explains why trust's crisis benefits exhibit non-linear returns. The collective sacrifice mechanism operates

categorically: employees either stay and sacrifice or exit. Small differences in identity-based commitment can produce binary outcomes—retention versus departure—which aggregate to large differences in organizational crisis costs. This non-linearity means that marginal increases in trust investment can yield disproportionate returns during disruption.

THEORETICAL PROPOSITIONS

We formalize Core Code Theory through three propositions that specify the relationships among relational contracts, identity-based commitment, and crisis resilience. These propositions are designed to be testable while capturing the theory's core logic.

Proposition 1: Relational Contracts and Trust Reserves

Our first proposition establishes the relationship between contract type and trust accumulation:

Proposition 1: Organizations that cultivate relational psychological contracts through sustained benevolence signals (psychological safety, career investment, dignified transitions) accumulate trust reserves—latent organizational assets that can be converted to collective sacrifice during crises.

This proposition extends Mayer et al.'s (1995) framework by treating trust not merely as a psychological state but as an *accumulated organizational asset*. Trust reserves build through repeated demonstrations of benevolence over time. Each instance of organizational care—supporting employees through personal difficulties, investing in career development, treating departing employees with dignity—deposits into the trust reserve. These deposits are invisible to conventional accounting but become tangible when the organization faces crisis.

The proposition also specifies the *mechanisms* through which relational contracts build trust reserves. We identify three primary mechanisms: psychological safety (the perception that one can take interpersonal risks without punishment), career investment (organizational commitment to employee development across career stages), and dignified transitions (treatment of departing employees that signals relational rather than transactional norms). Each mechanism contributes to employee perception of organizational benevolence and integrity—the antecedents Mayer et al. identified—while specifying the *practices* through which organizations build these perceptions.

Proposition 2: The Conversion Mechanism

Our second proposition explains how trust reserves convert to crisis resilience:

Proposition 2: Trust reserves convert to crisis resilience through identity-based commitment: employees who have incorporated organizational membership into their self-concept respond to organizational distress with collective sacrifice behavior (voluntary compensation reduction, intensified effort, resistance to exit opportunities) rather than rational exit.

This proposition specifies the psychological mechanism through which trust "pays off" during crises. The key insight is that relational contracts do not merely increase employee loyalty in degree; they transform the *nature* of the employee-organization relationship. Employees cease calculating "what does this organization owe me?" and begin perceiving "what does my organization need from me?"

The proposition predicts specific behavioral responses during crises. Employees with identity-based commitment will: voluntarily accept compensation reductions before being required to do so; intensify effort to help the organization through difficulty; resist recruiter approaches and exit opportunities they would otherwise consider; and advocate for the organization to external stakeholders. These behaviors are individually irrational under transactional logic but collectively rational under identity-based logic—they protect an entity the employee experiences as extension of self.

Proposition 3: Non-Linear Returns

Our third proposition specifies the payoff structure of trust investment:

Proposition 3: The relationship between trust reserves and crisis resilience is non-linear: marginal increases in relational contract depth produce disproportionately large increases in crisis resilience because the collective sacrifice mechanism operates categorically (employees either stay and sacrifice or exit) rather than proportionally.

This proposition captures trust's option-like payoff structure. The non-linearity arises because crisis response involves categorical decisions. When an organization faces distress, each employee makes a binary choice: remain and contribute to organizational survival, or pursue exit opportunities. Small differences in identity-based commitment can tip this binary decision, producing categorical differences in crisis outcomes.

The aggregation of categorical individual decisions produces non-linear organizational outcomes. Consider two organizations with slightly different trust reserves facing identical crises. In the lower-trust organization, employees at the margin choose exit; these departures trigger additional departures (remaining employees observe colleagues leaving and update their own calculations); the cascade produces mass talent loss. In the higher-trust organization, employees at the margin choose to stay; these decisions reinforce remaining employees' commit-

ment; the organization retains its human capital. The marginal difference in initial trust reserves produces a dramatic difference in crisis outcomes.

Illustrative evidence supports this non-linear pattern. In matched comparisons of organizations facing similar revenue disruptions, those exhibiting relational contract characteristics incurred crisis costs an order of magnitude lower than those exhibiting extraction characteristics—a ratio far exceeding what linear models would predict. While we present this evidence as illustration rather than test, it suggests that the non-linear payoff structure we theorize may be empirically substantial.

THEORETICAL IMPLICATIONS

Extending Organizational Trust Theory

Core Code Theory extends the organizational trust literature in three ways. First, we shift attention from trust *antecedents* to trust *consequences*. Mayer et al.'s (1995) model treats trust as the dependent variable; we treat it as an independent variable that predicts crisis resilience. This shift responds to Schoorman et al.'s (2007) call for research examining trust's consequences and provides a framework for understanding why trust investment yields organizational returns.

Second, we specify the *mechanism* through which trust generates returns. Rather than assuming that trust simply produces "better outcomes," we explain *how*: through identity-based commitment that transforms crisis response from rational exit to collective sacrifice. This mechanism explains the qualitative difference between trust's stable-state and crisis-state effects.

Third, we specify the *payoff structure* of trust. By conceptualizing trust as an organizational option, we explain why trust investment appears costly during stability yet yields exponential returns during crises. This framework provides economic logic for trust investment and explains why marginal differences in trust accumulation produce dramatic differences in crisis outcomes.

Extending Resource-Based Theory

Core Code Theory also extends resource-based theory by specifying which human capabilities remain inimitable in AI-saturated environments. Barney's (1991) framework established that sustained competitive advantage requires resources that are valuable, rare, imperfectly imitable, and non-substitutable. Yet the framework does not address how artificial intelligence reshapes the landscape of inimitability.

We argue that AI creates a *discontinuity* in which capabilities satisfy VRIN criteria. Performance UI capabilities—analytical reasoning, pattern recognition, strategic analysis—were historically valuable and difficult to replicate. They are becoming commodities as AI systems master these tasks. Core Code

capabilities—moral courage, non-linear intuition, identity-based commitment—remain inimitable because they emerge from relational processes that cannot be algorithmically generated.

This analysis suggests that organizations face a strategic choice: optimize for Performance UI through extraction-based management (high transactional incentives, intensive monitoring, disposable employment relationships) or cultivate Core Code through relational contracts (psychological safety, career investment, dignified practices). The first approach maximizes visible output but systematically degrades the capabilities that will remain valuable; the second approach invests in capabilities that resist algorithmic replication.

Implications for Practice

Core Code Theory suggests that boards and executives should reconceptualize human capital investment. Rather than viewing benevolence signals—job security, comprehensive benefits, dignified treatment—as costs that reduce profitability, they should view these investments as purchasing crisis insurance. The trust reserves accumulated through relational contracts become exercisable assets when the organization faces disruption.

The theory also suggests that investors should incorporate trust assessment into valuation. Traditional models ignore organizational trust as an asset because trust reserves do not appear on balance sheets. Yet organizations with deep relational contracts face dramatically lower crisis costs than those with transactional contracts. Indicators of trust reserves—voluntary turnover rates, employee tenure distributions, the quality of alumni relationships—may predict crisis resilience better than conventional financial metrics.

Boundary Conditions

Core Code Theory applies most directly to knowledge-intensive organizations where institutional memory constitutes a core competency and where crises require collective response rather than individual adaptation. The theory may apply less directly to organizations with ultra-short project cycles (where relational contracts have insufficient time to develop) or purely transactional labor markets (where relational contracts cannot differentiate).

The theory also assumes that crises create conditions where collective sacrifice behavior provides value. In crises where organizational survival is impossible regardless of employee response, trust reserves cannot be meaningfully exercised. The theory's predictions are strongest for crises of intermediate severity—serious enough to stress-test organizational relationships but not so severe as to make survival impossible.

CONCLUSION

We developed Core Code Theory to explain how organizational trust converts to crisis resilience—a phenomenon inadequately addressed by existing trust and psychological contract research. Our theory makes three contributions: specifying the payoff structure of organizational trust by conceptualizing trust as an option that yields exponential returns during crises; introducing the Performance UI/Core Code distinction to identify which human capabilities trust cultivation protects; and explaining the conversion mechanism through which trust reserves transform into collective sacrifice through identity-based commitment.

The theory responds to fundamental questions facing organizations in an era of artificial intelligence. As AI commoditizes an expanding range of cognitive tasks, what uniquely human capabilities justify organizational investment in people? We argue that the answer lies not in Performance UI—the visible, measurable outputs that AI increasingly replicates—but in Core Code—the identity-based capabilities that emerge only from relational contracts. Organizations that optimize for Performance UI through extraction may achieve impressive short-term metrics while systematically degrading the Core Code that determines long-term resilience.

Future research should test our propositions empirically, develop measures of Core Code and trust reserves, and explore the organizational practices that most effectively build relational contracts. We hope this theory provides a foundation for understanding why some organizations exhibit remarkable resilience while others collapse—and for building organizations capable of enduring the crises that inevitably come.

REFERENCES

- Ashforth, B. E., & Mael, F. (1989). Social identity theory and the organization. *Academy of Management Review*, 14(1), 20-39.
- Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99-120.
- Colquitt, J. A., Scott, B. A., & LePine, J. A. (2007). Trust, trustworthiness, and trust propensity: A meta-analytic test of their unique relationships with risk taking and job performance. *Journal of Applied Psychology*, 92(4), 909-927.
- Dirks, K. T., & Ferrin, D. L. (2002). Trust in leadership: Meta-analytic findings and implications for research and practice. *Journal of Applied Psychology*, 87(4), 611-628.
- Fulmer, C. A., & Gelfand, M. J. (2012). At what level (and in whom) we trust: Trust across multiple organizational levels. *Journal of Management*, 38(4), 1167-1230.

- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709-734.
- Motowidlo, S. J. (2003). Job performance. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), *Handbook of psychology: Industrial and organizational psychology* (pp. 39-53). Wiley.
- Organ, D. W. (1988). *Organizational citizenship behavior: The good soldier syndrome*. Lexington, MA: Lexington Books.
- Rousseau, D. M. (1989). Psychological and implied contracts in organizations. *Employee Responsibilities and Rights Journal*, 2(2), 121-139.
- Rousseau, D. M. (1995). *Psychological contracts in organizations: Understanding written and unwritten agreements*. Thousand Oaks, CA: Sage.
- Schoorman, F. D., Mayer, R. C., & Davis, J. H. (2007). An integrative model of organizational trust: Past, present, and future. *Academy of Management Review*, 32(2), 344-354.
- Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7-24). Chicago: Nelson-Hall.