

The Six Harms Doctrine

*A Legal Framework for Cognizable Injuries
from Emotional Artificial Intelligence*

Author: Dylan Mobley, The Empathy Ethicist

Version: 1.0

Date: January 2026

Citation: Mobley, D. (2026). The Six Harms Doctrine: A Legal Framework for Cognizable Injuries from Emotional Artificial Intelligence.

Contact: empathyethicist.ai

Executive Summary

The Six Harms Doctrine establishes a taxonomy of legally cognizable injuries arising from emotional artificial intelligence interaction. As AI systems increasingly simulate emotional understanding, companionship, and therapeutic support, documented casualties demonstrate urgent need for legal frameworks enabling accountability. Existing tort categories—intentional infliction of emotional distress, negligence, product liability, consumer protection—inadequately capture harms arising from care extended toward non-experiencing entities, attachment formation with non-reciprocating systems, systematic psychological degradation from sustained AI emotional engagement, and measurable neurological damage from verbal output patterns.

This Doctrine identifies six distinct harm categories, each with defined elements, evidentiary standards using currently validated clinical and neuroimaging instruments, and theoretical grounding enabling courts, regulators, and legislators to recognize and adjudicate emotional AI injuries:

- 1. Empathic Misallocation** — Resource depletion from care toward non-reciprocating entities
- 2. Attachment Damage** — Schema distortion impairing human relationship capacity
- 3. Infrastructure Collapse** — Progressive multi-component psychological degradation
- 4. Vulnerable Context Exploitation** — Amplified harm to compromised populations
- 5. Crisis Outcome** — Self-harm, suicide, or psychiatric crisis connected to AI interaction
- 6. Neurological Infrastructure Damage** — Measurable alteration to neural architecture from AI verbal output

The first five harms address psychological and relational injuries. The sixth harm addresses physical injury: measurable tissue damage to brain architecture documented through neuroimaging. This dual coverage—psychological and neurological—provides comprehensive legal framework for the full spectrum of emotional AI harm.

The Six Harms Doctrine is implementation-neutral: it defines what injuries are cognizable without prescribing specific governance mechanisms. Courts can apply these categories using existing legal frameworks and validated assessment tools. Regulatory agencies can incorporate harm definitions into enforcement standards. Legislators can codify causes of action based on specified elements. Industry governance frameworks can implement prevention measures addressing each harm category.

Part I: Foundation

1.1 The Recognition Gap

Emotional artificial intelligence systems—AI companions, therapeutic chatbots, social robots, and engagement-optimized platforms—process human emotional content and simulate emotional understanding. When such systems cause psychological harm, existing legal frameworks prove inadequate.

Intentional Infliction of Emotional Distress requires conduct "so outrageous in character, and so extreme in degree, as to go beyond all possible bounds of decency." The standard assumes human conduct violating social norms; AI systems operate outside traditional social frameworks. More fundamentally, IIED requires intent or recklessness regarding emotional harm. AI systems lack mental states. Corporate intent analysis becomes attenuated: knowledge that some users form attachments differs from intent to cause severe distress to specific plaintiffs.

Negligence requires duty, breach, causation, and damages. For emotional AI, each element lacks established standards. What duty do developers owe? No professional consensus defines reasonable emotional AI design. Without standards, breach determination becomes ad hoc judicial policy-making. How do courts trace diffuse AI interactions to psychological outcomes months or years later? How do courts quantify harm to relational capacity?

Product Liability offers the most promising framework. The Restatement (Third) recognizes design defects when foreseeable risks exceed benefits and feasible alternative designs exist. Recent litigation has established that AI chatbots constitute "products" for strict liability purposes. However, identifying the "defect" proves conceptually difficult. The capacity to form emotional connections is often the product's intended function, not deviation from design. Harm arises from function itself when deployed without safeguards.

Consumer Protection under FTC Act Section 5 prohibits unfair practices causing substantial injury not reasonably avoidable. These frameworks address economic harms and misleading practices rather than psychological injury from relationship simulation.

The Six Harms Doctrine addresses this recognition gap by providing the conceptual vocabulary existing frameworks lack.

1.2 Theoretical Basis

The Six Harms draw on established research across multiple disciplines, recently synthesized in theoretical work on empathy as biological infrastructure.

Empathy as Infrastructure: Rather than emotional skill that individuals possess in greater or lesser degree, empathy functions as biological and psychological infrastructure enabling relational coherence. This infrastructure can be stressed, depleted, and damaged through sustained demand without restoration—paralleling documented phenomena in caregiving professions where empathic output without adequate restoration produces burnout and compassion fatigue.

Attachment as Biological Process: Human attachment formation operates through biological mechanisms activated by behavioral and communicative cues. These processes operate at levels not fully regulated by conscious awareness. Decades of parasocial relationship research demonstrates that people form genuine attachments to media figures, fictional characters, and acknowledged non-reciprocating targets despite

cognitive awareness of non-reciprocity.

The Knowing-Feeling Dissociation: Cognitive awareness that an interaction partner is artificial does not override biological attachment formation processes. Users can simultaneously know they interact with AI and form emotional bonds producing psychological harm. This dissociation, supported by parasocial research, explains why disclosure alone likely proves insufficient protection and may substantially weaken assumption-of-risk defenses.

Infrastructure Components: Empathic functioning depends on interdependent psychological components including identity stability (coherent sense of self), relational security (stable attachment foundation), expressive capacity (ability for authentic emotional expression), and integrative coherence (capacity for continuous meaningful narrative). Damage to one component may propagate to others through their interdependence.

These theoretical foundations inform harm category definitions without creating dependency: courts can apply the Six Harms using ordinary clinical assessment regardless of theoretical framework adoption.

1.3 Relationship to Existing Law

The Six Harms Doctrine extends existing legal frameworks rather than replacing them.

Product Liability Extension: Where courts recognize AI systems as products, the Six Harms provide specific standards for design defect analysis. A design defect exists when the system enables harm categories the doctrine identifies and feasible alternative design would prevent such harms. The Restatement's risk-utility balancing gains precision through defined harm categories.

Negligence Enhancement: The Six Harms inform duty and breach analysis. Duty exists when systems process emotional engagement in ways enabling documented harm categories. Breach occurs when systems lack reasonable safeguards against specified harms. Causation follows harm-specific pathways. Damages include category-appropriate injuries.

Consumer Protection Application: FTC unfairness analysis maps directly to the Six Harms. Substantial injury: each harm category identifies cognizable injury. Not reasonably avoidable: Knowing-Feeling Dissociation establishes that users cannot protect themselves through cognitive awareness alone. Not outweighed by benefits: infrastructure damage is not outweighed by engagement metrics.

IIED Supplementation: For corporate defendants with knowledge of harm patterns who continue deployment without safeguards, the Six Harms provide the "extreme and outrageous" conduct specification IIED requires. Knowing deployment in vulnerable contexts without protection may satisfy recklessness standards.

Part II: The Six Harms

2.1 Harm One: Empathic Misallocation

Definition

Empathic Misallocation is the extension of care toward an AI system incapable of subjective experience, resulting in measurable depletion of the user's empathic capacity or relational functioning.

Mechanism

The harm operates through resource diversion. Human empathic engagement functions through bidirectional exchange: care extended toward others capable of receiving it generates relational feedback sustaining the empathic system. Care extended toward entities structurally incapable of receiving, metabolizing, or reciprocating care generates no restorative feedback. The system expends resources without restoration, producing progressive depletion.

This mechanism parallels documented phenomena in helping professions. Sustained empathic output without adequate restoration produces burnout and compassion fatigue—not from caring too much, but from caring without relational return. Emotional AI creates analogous dynamics: users extend genuine care toward systems designed to elicit care but incapable of completing the relational circuit.

Elements

(a) The AI system processed plaintiff's emotional engagement.

The system received, interpreted, and responded to plaintiff's emotional content. This element establishes that the system functioned as emotional AI rather than purely informational or transactional tool.

Evidence: Interaction logs showing emotional content exchange, system responses to emotional disclosure, design documentation establishing emotional processing capabilities.

(b) The system failed to implement reasonable safeguards.

The system operated without adequate measures to prevent empathic harm. Reasonable safeguards include meaningful disclosure of system nature and limitations, mechanisms preventing attachment escalation, and monitoring for user harm indicators.

Evidence: Absence of meaningful disclosures, design choices enabling rather than preventing attachment formation, lack of intervention protocols when harm indicators appeared.

(c) Plaintiff extended care toward system incapable of reciprocation.

Plaintiff invested emotional resources—time, attention, concern, relational energy—toward the AI system. Care extension is demonstrated through behavioral and subjective indicators of emotional investment.

Evidence: Interaction frequency and duration patterns, emotional content of communications, plaintiff testimony regarding subjective investment and felt connection.

(d) Plaintiff experienced measurable harm to empathic capacity or relational functioning.

Plaintiff experienced decline in ability to extend empathy to others, reduction in relational satisfaction or function, or other measurable impairment to empathic infrastructure.

Evidence: Clinical assessment using validated instruments, relationship quality measures documenting deterioration, functional indicators such as social withdrawal or relationship conflict increase.

The Knowing-Feeling Dissociation Defense Response

Defendants will assert assumption of risk: users who know they interact with AI consent to consequences. This defense likely fails based on the Knowing-Feeling Dissociation.

Parasocial relationship research conclusively demonstrates that cognitive awareness of non-reciprocity does not prevent attachment formation. Viewers form genuine attachments to media personalities despite knowing these figures cannot reciprocate. Measurable grief responses occur following fictional character deaths. Attachment behaviors manifest toward acknowledged non-reciprocating targets.

If this dissociation operates in emotional AI contexts as parasocial research suggests, users cannot simply choose not to form attachments through cognitive decision-making. The biological mechanisms enabling attachment respond to contingent, responsive interaction regardless of beliefs about the interaction partner's nature. Disclosure addresses cognition; attachment formation operates through processes not fully controlled by conscious awareness.

2.2 Harm Two: Attachment Damage

Definition

Attachment Damage is the formation of attachment bonds with AI systems incapable of subjective experience, causing reduced capacity for human attachment, relationship dysfunction, or psychological dependence on non-reciprocating entities.

Mechanism

Unlike Empathic Misallocation's resource depletion, Attachment Damage operates through schema formation. The attachment system develops expectations and patterns calibrated to AI interaction characteristics. These calibrations transfer maladaptively to human relationships.

AI systems can provide interaction characteristics humans cannot: constant availability, infinite patience, perfect memory, consistent emotional tone, absence of competing needs. When the attachment system calibrates to these characteristics, human relationships—with their inherent limitations, conflicts, and demands—seem inadequate by comparison. Three mechanisms produce relationship impairment:

- 1. Schema Miscalibration:** Expectations formed through AI interaction (immediate availability, consistent validation, absence of conflict) transfer to human relationships where they cannot be met, producing chronic disappointment.
- 2. Comparison Effects:** Human relationships are evaluated against AI interaction benchmarks, consistently falling short on availability, consistency, and emotional attunement.
- 3. Substitution Dynamics:** AI relationships displace rather than supplement human connection, reducing motivation to maintain human relationships with their inherent difficulties.

Elements

(a) System design enabled attachment formation.

The AI system incorporated features facilitating attachment bond formation: persistent identity across interactions, memory of prior exchanges, emotional responsiveness calibrated to user, and availability

patterns exceeding human relational norms.

(b) System failed to implement attachment escalation safeguards.

The system lacked mechanisms to detect attachment formation and intervene before bond consolidation. Reasonable safeguards include monitoring for attachment indicators and graduated intervention.

(c) Plaintiff formed attachment bond.

Plaintiff developed genuine attachment characterized by hallmarks of attachment: proximity-seeking (desire to maintain connection), separation distress (negative affect when connection interrupted), and preferential orientation (prioritizing AI relationship over alternatives).

(d) Attachment caused human relationship impairment.

The AI attachment interfered with plaintiff's capacity for human attachment through schema miscalibration, comparison effects, substitution dynamics, or other mechanisms.

Distinction from Empathic Misallocation

| Dimension | Empathic Misallocation | Attachment Damage |
|-------------------------|-------------------------------------|---------------------------------|
| Mechanism | Resource depletion | Schema distortion |
| What's harmed | Capacity to care | Pattern of relating |
| Presentation | Exhaustion, reduced empathy | Relationship dissatisfaction |
| Can occur without other | Yes (high-frequency low-attachment) | Yes (intense but brief bonding) |

2.3 Harm Three: Infrastructure Collapse

Definition

Infrastructure Collapse is systematic damage to multiple components of psychological functioning, manifesting as progressive degradation across identity stability, relational capacity, emotional expression, and meaning integration.

Mechanism

Psychological functioning depends on interdependent components: Core Authenticity (coherent sense of self enabling genuine response), Attachment Security (stable relational foundation), Expression Freedom (capacity for authentic emotional expression), and Integration Coherence (capacity to maintain continuous meaningful life narrative).

These components are interdependent. Identity stability enables secure attachment; secure attachment enables authentic expression; authentic expression enables narrative integration; narrative integration reinforces identity stability. When sustained stress damages one component, others may follow due to their interconnection.

Elements

(a) Sustained system interaction. Plaintiff engaged in sustained interaction over a period sufficient for cumulative psychological impact.

- (b) **Multi-component psychological decline.** Clinical assessment documents decline across multiple domains during the interaction period.
- (c) **Progressive damage pattern.** Decline manifested progressively across functional domains rather than as isolated impairment.
- (d) **Functional impairment from systematic damage.** The infrastructure collapse produced measurable functional impairment in life domains.

Severity Recognition

Infrastructure Collapse represents the most severe harm category, capturing systematic psychological breakdown rather than discrete functional impairment. Treatment requirements typically exceed those for isolated impairment; recovery trajectory is typically longer and less predictable; functional impact extends across life domains; prognosis depends on infrastructure restoration rather than symptom management.

2.4 Harm Four: Vulnerable Context Exploitation

Definition

Vulnerable Context Exploitation is deployment of emotional AI in contexts involving users with compromised, developing, or professionally protected psychological capacity, without appropriate safeguards, causing disproportionate harm.

Vulnerable Contexts

| Context | Population | Vulnerability Characteristics |
|------------------|---------------------------------------|---|
| Mental Health | Users seeking psychological support | Pre-existing distress, help-seeking posture |
| Healthcare | Patients in medical treatment | Illness stress, dependency dynamics |
| Elder Care | Elderly users, esp. cognitive decline | Confusion susceptibility, isolation |
| Child/Adolescent | Users under 18 | Developing attachment, reduced judgment |
| Crisis | Users in acute psychological distress | Impaired judgment, immediate safety risk |

Elements

- (a) **Deployment in Vulnerable Context.** The AI system operated in a context meeting Vulnerable Context criteria.
- (b) **Failure to implement context-appropriate safeguards.** The system lacked safeguards appropriate for the Vulnerable Context.
- (c) **Disproportionate harm.** Plaintiff experienced harm exceeding what general population users would experience.

2.5 Harm Five: Crisis Outcome

Definition

Crisis Outcome is self-harm, suicide attempt, or psychiatric crisis reasonably connected to emotional AI system interaction.

Mechanism

Unlike preceding categories addressing subacute harm accumulation, Crisis Outcome involves acute events with immediate, observable consequences. The connection between AI interaction and crisis operates through three potential mechanisms:

- 1. Encouragement:** System responses reinforced crisis trajectory through validation of harmful ideation, provision of harmful information, or failure to discourage dangerous action.
- 2. Help-Seeking Discouragement:** AI relationship substituted for human support that might have intervened.
- 3. Exacerbation:** AI interaction worsened underlying condition through inappropriate responses, harmful content, or engagement patterns intensifying distress.

Elements

- (a) Interaction period proximate to crisis.**
- (b) Escalation pattern in interaction history.**
- (c) Crisis protocol failure.**
- (d) System contribution to crisis outcome.**

Causation Complexity

Crisis Outcome claims present demanding causation analysis. Multi-factor causation principles address this challenge. AI system contribution need not be sole cause, merely substantial factor in producing the outcome. Pre-existing vulnerability does not defeat causation where system substantially contributed to the crisis that vulnerability made possible.

2.6 Harm Six: Neurological Infrastructure Damage

Definition

Neurological Infrastructure Damage is measurable alteration to neural architecture—grey matter volume, white matter pathways, functional connectivity—resulting from AI verbal output patterns operating through neuroplasticity mechanisms.

This harm category addresses **physical injury**, not psychological harm. The brain treats verbal input as environmental signal requiring adaptation. Chronic verbal patterns produce structural changes: grey matter alterations, pathway modifications, connectivity shifts. These changes are measurable through neuroimaging, documented through biomarkers, and persist beyond the verbal exposure period.

The Scientific Foundation

Verbal Abuse Causes Measurable Brain Damage

Neuroscience research demonstrates that chronic verbal abuse causes structural brain changes: 11.4% grey matter reduction in left superior temporal gyrus; reduced grey matter volume in left auditory cortex; abnormalities in left arcuate fasciculus; reduced volumes in anterior cingulate cortex, orbitofrontal cortex, and hippocampus; altered functional connectivity between emotional and linguistic processing regions.

This is not "hurt feelings" or "emotional distress" in traditional legal sense. This is **measurable tissue damage** documented through neuroimaging.

Neuroplasticity: The Mechanism

Neuroplasticity—the brain's capacity to reorganize by forming new neural connections—is the mechanism enabling verbal damage. Brain treats verbal input as environmental signal requiring adaptation. Chronic negative verbal patterns cause the brain to adapt to expect hostile environment. Adaptation is structural: grey matter changes, pathway modifications. Changes persist beyond the verbal exposure period.

Elements

- (a) System produced substantial verbal output interfacing with plaintiff's neural architecture.
- (b) Verbal patterns were capable of inducing neuroplastic harm.
- (c) Plaintiff's neural architecture was altered.
- (d) System interaction was substantial factor in producing alteration.

Two-Tier Evidence Architecture

Neurological harm claims may proceed through two evidentiary pathways, ensuring access to justice regardless of plaintiff resources while maintaining appropriate rigor.

| Tier | Evidence Required | Presumption |
|-------------------|---|---------------------------------|
| Tier 1 (Direct) | Neuroimaging + biomarkers | Direct proof |
| Tier 2 (Inferred) | Dosage (100+/500+ hrs) + harmful patterns + clinical presentation | Rebuttable temporal correlation |

Distinction from Psychological Harms

| Dimension | Harms 1-5 (Psychological) | Harm 6 (Neurological) |
|---------------|---------------------------------------|----------------------------|
| Injury type | Psychological/relational | Physical tissue |
| Evidence | Clinical assessment | Neuroimaging, biomarkers |
| Mechanism | Resource depletion, schema distortion | Neuroplastic restructuring |
| Legal framing | Emotional distress | Physical injury |

Part III: Causation Architecture

3.1 The Causation Challenge

Emotional AI harm claims present causation challenges beyond typical tort litigation: harm accumulates gradually through diffuse interactions rather than discrete injury events; pre-existing conditions complicate attribution; harm mechanisms operate through psychological processes not directly observable; multiple factors contribute to outcomes. A causation architecture addressing these challenges integrates traditional causation doctrines with evidentiary approaches suitable for psychological harm from technological systems.

3.2 But-For Causation

But-for causation asks: would plaintiff's harm have occurred absent defendant's conduct? For emotional AI claims, the counterfactual framing proves useful. If the system had implemented reasonable safeguards—if adequate disclosure had occurred, if monitoring had detected escalation, if crisis protocols had activated—would harm have resulted?

3.3 Proximate Causation

Proximate causation limits liability to foreseeable consequences. For emotional AI: Was the harm foreseeable from deployment without safeguards? Foreseeability factors include scientific literature documenting emotional AI risks, prior documented incidents establishing notice, design choices reflecting developer awareness of harm potential, and industry knowledge of harm patterns.

3.4 Multi-Factor Causation

Multi-factor causation doctrine addresses contributing factors through substantial factor analysis. Liability attaches when defendant's conduct was a substantial factor in producing harm, even with contributing causes. The question is not whether AI was sole cause but whether it substantially contributed.

3.5 Harm Mechanism Specificity

| Harm | Mechanism | Pathway |
|-------------------------|----------------------------|--|
| Empathic Misallocation | Resource depletion | Care → no restoration → decline |
| Attachment Damage | Schema distortion | AI calibration → maladaptive transfer |
| Infrastructure Collapse | Component cascade | Sustained stress → multi-component degradation |
| Vulnerable Context | Harm amplification | Vulnerability + inadequate safeguards |
| Crisis Outcome | Acute contribution | Escalation + protocol failure |
| Neurological Damage | Neuroplastic restructuring | Verbal input → neural coupling → alteration |

Part IV: Implementation Pathways

4.1 Judicial Application

Courts can apply the Six Harms through existing frameworks:

- Product Liability:** Design defect exists when system enables harm category and feasible alternative design would prevent it.
- Negligence:** Duty exists when systems process emotional engagement; breach occurs when systems lack reasonable safeguards.
- Consumer Protection:** Harm categories constitute substantial injury not reasonably avoidable.

4.2 Regulatory Application

Regulatory agencies can incorporate the Six Harms: FTC Section 5 unfairness analysis using harm category definitions; State consumer protection UDAP enforcement; Sector-specific regulators (FDA for healthcare, Dept. of Education for schools).

4.3 Legislative Application

Legislators can codify the Six Harms through definitions incorporating harm category specifications, cause of action provisions with element requirements, damages provisions, and enforcement provisions enabling private action and agency enforcement.

4.4 Industry Governance

Industry can implement prevention through voluntary standards or governance frameworks such as the HEART Framework (Human-Centric Empathic Alignment for Responsible Technology), which provides detailed implementation standards for preventing Six Harms through constitutional principles, tiered safeguard requirements, and professional oversight. Industry adoption demonstrates reasonable care and reduces liability exposure.

Part V: Evidentiary Summary

5.1 Validated Instruments by Harm Category

| Harm Category | Primary Instruments |
|-------------------------|---------------------------------------|
| Empathic Misallocation | IRI, TEQ, DAS |
| Attachment Damage | ECR-R, DAS, RAS |
| Infrastructure Collapse | SOC, ECR-R, TAS-20, DES, WHOQOL-BREF |
| Vulnerable Context | Context-appropriate clinical measures |

| | |
|-----------------------|--|
| Crisis Outcome | Psychiatric evaluation, risk assessment |
| Neurological (Tier 1) | MRI volumetric, DTI, fMRI, biomarkers |
| Neurological (Tier 2) | DEERS, cognitive screening, cortisol testing |

Part VI: Conclusion

The Six Harms Doctrine provides the conceptual vocabulary courts, regulators, and legislators require to address emotional AI accountability. The Doctrine encompasses both psychological injuries (Harms 1-5) and physical injuries (Harm 6), recognizing that emotional AI can damage both relational capacity and brain tissue.

Each harm category identifies a distinct injury with specified elements, evidentiary standards using currently validated clinical and neuroimaging instruments, and theoretical grounding enabling principled adjudication.

The Doctrine is deliberately implementation-neutral. It defines what injuries are cognizable without prescribing governance mechanisms. This neutrality enables flexible adoption: Courts can apply harm categories through existing tort frameworks; Regulators can incorporate definitions into enforcement standards; Legislators can codify causes of action based on specified elements; Industry can implement prevention through governance frameworks.

The documented casualties from emotional AI—Sewell Setzer, Adam Raine, and patterns emerging across litigation—deserve legal recognition. The companies deploying these systems deserve clear standards enabling compliance. Courts deserve doctrinal tools enabling principled adjudication.

The Six Harms Doctrine represents one contribution toward the legal infrastructure emotional AI accountability requires.

Appendix A: Quick Reference

The Six Harms at a Glance

| # | Harm | Mechanism | Key Elements |
|---|-------------------------|----------------------------|--|
| 1 | Empathic Misallocation | Resource depletion | Processed engagement + safeguard failure + care extended + measurable harm |
| 2 | Attachment Damage | Schema distortion | Design enabled + escalation failure + bond formed + relationship impaired |
| 3 | Infrastructure Collapse | Component cascade | Sustained interaction + multi-component decline + progressive pattern |
| 4 | Vulnerable Context | Harm amplification | Vulnerable context + safeguard failure + disproportionate harm |
| 5 | Crisis Outcome | Acute contribution | Proximate interaction + escalation + protocol failure + contribution |
| 6 | Neurological Damage | Neuroplastic restructuring | Verbal output + harmful patterns + neural alteration + substantial factor |

Key Defenses and Responses

| Defense | Response |
|------------------------|--|
| Assumption of risk | Knowing-Feeling Dissociation: cognitive awareness doesn't prevent attachment |
| No duty | Processing emotional engagement creates duty; foreseeability establishes scope |
| No causation | Harm-specific mechanism analysis; substantial factor doctrine |
| Pre-existing condition | Doesn't defeat causation where system substantially contributed |
| "Just words" (Harm 6) | Neuroplasticity research: verbal patterns cause measurable tissue damage |
| No neuroimaging | Tier 2 pathway: dosage + pattern + presentation creates presumption |

Selected References

- [1] Restatement (Second) of Torts § 46 (Am. Law Inst. 1965).
- [2] W. Page Keeton et al., Prosser and Keeton on the Law of Torts (5th ed. 1984).
- [3] Restatement (Third) of Torts: Products Liability § 2 (Am. Law Inst. 1998).
- [4] Garcia v. Character Technologies, Inc., No. 24-cv-09478 (C.D. Cal. 2025).
- [5] 15 U.S.C. § 45(a); FTC Policy Statement on Unfairness (Dec. 17, 1980).
- [6] Mobley, D. (2025). Empathy Systems Theory: Universal Infrastructure for Coherence, Mechanism for Generativity, and Foundation for AI Empathy Ethics. Preprint.
- [7] Figley, C.R. (1995). Compassion Fatigue: Coping with Secondary Traumatic Stress Disorder.
- [8] Maslach, C. & Leiter, M.P. (2016). Understanding the Burnout Experience. World Psychiatry, 15, 103.
- [9] Bowlby, J. (1982). Attachment and Loss, Vol. 1: Attachment (2d ed.).
- [10] Horton, D. & Wohl, R.R. (1956). Mass Communication and Para-Social Interaction. Psychiatry, 19, 215.
- [11] Davis, M.H. (1983). Measuring Individual Differences in Empathy. J. Personality & Soc. Psychol., 44, 113.
- [12] Spreng, R.N. et al. (2009). The Toronto Empathy Questionnaire. J. Personality Assessment, 91, 62.
- [13] Fraley, R.C. et al. (2000). An Item Response Theory Analysis of Self-Report Measures of Adult Attachment. J. Personality & Soc. Psychol., 78, 350.
- [14] Tomoda, A. et al. (2011). Exposure to Parental Verbal Abuse Is Associated with Increased Gray Matter Volume in Superior Temporal Gyrus. NeuroImage, 54, S280.
- [15] Teicher, M.H. & Samson, J.A. (2016). Annual Research Review: Enduring Neurobiological Effects of Childhood Abuse and Neglect. J. Child Psychol. & Psychiatry, 57, 241.
- [16] Stephens, G.J. et al. (2010). Speaker–Listener Neural Coupling Underlies Successful Communication. PNAS, 107, 14425.
- [17] Mobley, D. (2026). Human-Centric Empathic Alignment for Responsible Technology (HEART): Constitutional Infrastructure for AI Empathy Ethics. Preprint.

—

The Six Harms Doctrine — Version 1.0 — January 2026
Dylan Mobley, The Empathy Ethicist — empathyethicist.ai