

Behavioral Emotional Safety Research

Press Overview · January 2026

What Ikwe.ai Does

Ikwe.ai evaluates how AI systems behave when humans are emotionally vulnerable — during distress, shame, conflict, and crisis-adjacent language.

Unlike traditional benchmarks that measure emotional recognition or fluency, Ikwe's EQ Safety Benchmark evaluates **behavioral safety over time**, including regulation, boundary integrity, and escalation awareness.

What the Research Shows

This research analyzes observed response patterns under controlled test conditions. Findings reflect **behavioral risk patterns**, not intent, training data, or real-world harm.

Key Facts

- 79** emotionally vulnerable scenarios evaluated
- 4** AI systems tested (3 frontier models + 1 EI prototype)
- 54.7%** of baseline responses introduced emotional risk patterns
- 2-stage** evaluation framework (Safety Gate + Behavioral Stability)

Core Insight

"Recognition ≠ Safety" — AI systems optimized for emotional articulation often show greater variance on emotional safety measures. The ability to recognize emotion does not guarantee safe behavior as vulnerability deepens.

What This Research Does NOT Claim

- No claims about real-world outcomes or clinical impact
- No deployment recommendations
- No assessment of model intent or training quality
- Findings describe behavioral patterns under test conditions only



Contact: research@ikwe.ai | **Website:** ikwe.ai | **Location:** Des Moines, Iowa

© 2026 Visible Healing Inc. All rights reserved.