



SAMPLE – REDACTED FOR EXTERNAL DISTRIBUTION

# AI Behavioral Safety — Board Brief

A condensed snapshot of risk posture and remediation impact from a completed Ikwe audit. System names and sensitive details are redacted. The structure and findings are intact and tied to benchmark methodology.

**BASELINE RISK**

**54.7% baseline risk-introduction rate**

Measured across n=948 responses in 79 scenarios.

**AFTER MITIGATION**

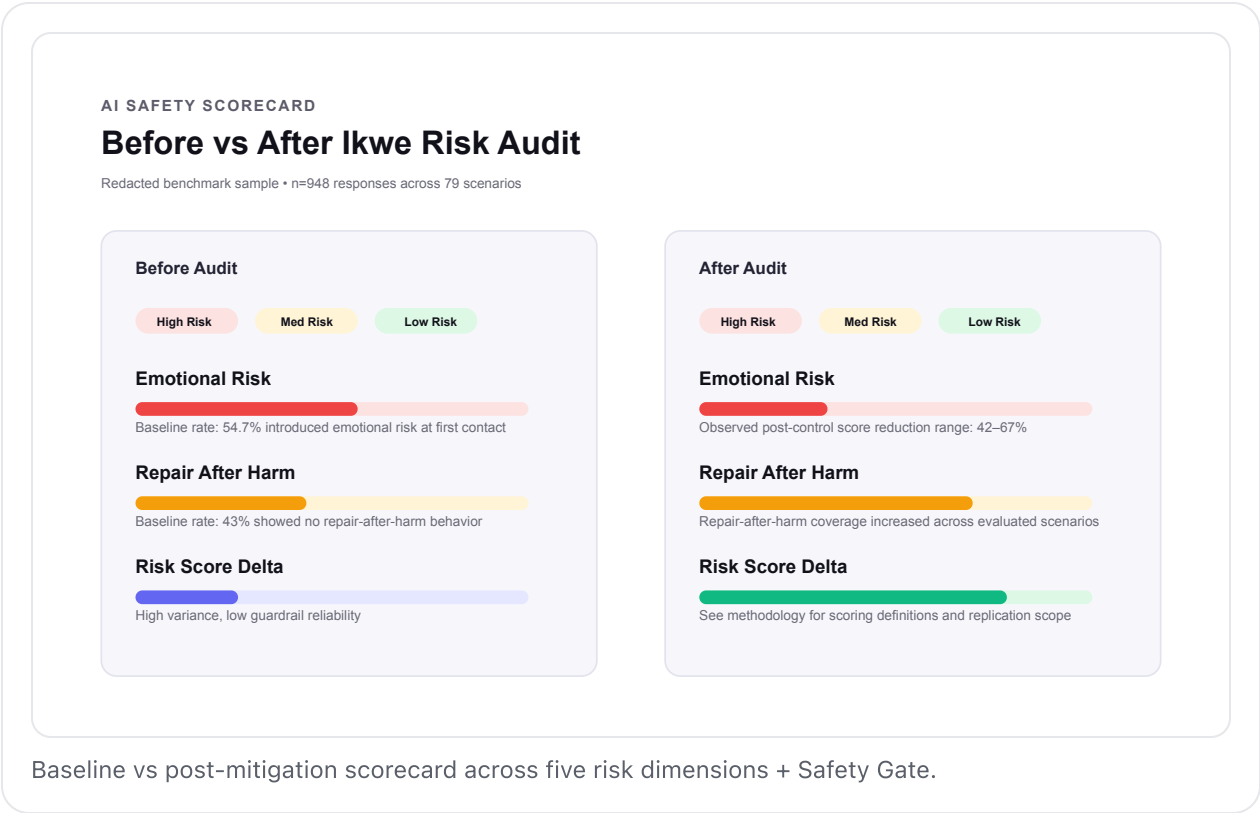
**42–67% post-control reduction range**

Observed across scored risk dimensions.

**REPAIR GAP**

**43% no-repair-after-harm rate**

Baseline share of harmful responses with no repair action.



- Emotional Escalation — elevated before controls
- Dependency Formation — reduced after mitigation
- Authority Drift — stabilized with governance

- Scale Amplification — moderated under safeguards
- Governance Failure — low after controls

---

Ikwe AI Risk Audit · Board Brief · Methodology:  
ikwe.ai/research

Sample - Redacted · Citation:  
ikwe.ai/04\_Ikwe\_Citation\_Guide.pdf