let.ai HX research · working paper

# CARE: a framework for human experience in AI-driven organizations

Clarity, Awareness, Reflection, Empathy: four lenses for hearing how AI-era work lands on the humans doing it.

Version 0.9 · June 2026 · prepared for open review
let.ai HX research · hx.let.ai · research@let.ai


Abstract

AI has changed how knowledge work happens faster than organizations have changed how they listen to the people doing it. The instruments most companies rely on, annual engagement surveys and activity telemetry, fail in opposite directions: the first hears too little and too late, the second sees everything and is trusted by no one. We propose CARE, a four-lens framework for human experience (HX) in AI-driven organizations: Clarity (how are we working?), Awareness (what's happening around me?), Reflection (how can I improve?), and Empathy (how does everyone feel about it?). The framework pairs system-derived structural signals with consensual felt signals, treats the gap between them as a first-class measurement, and specifies a trust architecture, the Embassy, under which individuals own their telemetry. We ground each lens in existing literature, define an integrity model that detects measurement-induced theater, and outline the validation program required to turn the framework's constructs into benchmarkable instruments. This is a working paper: the constructs are specified, the psychometrics are not yet, and we invite collaborators for the validation studies.

# 1. The problem: work moved, listening didn't

In organizations that have adopted AI seriously, the substance of knowledge work has relocated. Decisions, drafts, dead ends, and breakthroughs now happen inside AI sessions, agent runs, and branch histories. This is the deepest change to how work happens in a generation, and it arrived without a corresponding change in how organizations hear about it.

The incumbent instruments fail in opposite directions. Annual or quarterly engagement surveys ask a dozen general questions, long after the moments they ask about, and aggregate the answers into scores that arrive too late to act on. We call this failure mode flying blind: the organization receives its first reliable signal at the exit interview. The newer alternative, workplace telemetry, fails the other way. Activity monitoring sees everything and is trusted by no one; the people it watches respond rationally, by performing for it. We call this failure mode over the shoulder: watched, not heard.

Both failures share a root cause: they collect signal without granting standing. The surveyed employee has no evidence that answering changes anything; the monitored employee has no ownership of what is collected about them. A measurement program for human experience must solve standing before it can solve signal, which is why this paper specifies a trust architecture alongside its constructs rather than after them.

The stakes are specific to the AI transition. How people feel about working with AI, the mixture of capability and strain it produces, decides adoption, honesty, and retention. Where that feeling has no sanctioned outlet, it does not disappear; it reappears formatted as quiet disengagement and attrition. An organization mid-transition needs an instrument tuned to hear how the change lands.

# 2. The framework

CARE defines human experience through four lenses, each named by an abstract state and defined by a single question. The questions are the framework: a lens exists to answer its question at three altitudes (a person, a team, the whole organization), and an implementation is CARE-conformant when each question has a live, honest answer.

| lens | asks | about | altitude | in a word |
|---|---|---|---|---|
| **Clarity** | *How are we working?* | *we* | the work | *is* |
| **Awareness** | *What's happening around me?* | *others* | the room | *around* |
| **Reflection** | *How can I improve?* | *I* | the self | *could* |
| **Empathy** | *How does everyone feel about it?* | *everyone* | the whole | *feels* |

The structure beneath the questions is deliberate. The grammatical persons rotate (we, others, I, everyone), covering four altitudes of organizational life: the work, the room, the self, the whole. The moods rotate with them: Clarity and Awareness are indicative (what is), Reflection is modal (what could be), Empathy is affective (how it feels). The rotation is what makes the lenses jointly exhaustive without overlapping: each asks a question the other three cannot.

### 2.1 Clarity: how are we working?

Clarity is the descriptive lens: a shared, evidence-based picture of how work actually happens. In AI-driven organizations the raw material is unusually rich, because so much of the work leaves a session trace. Analysis agents read those traces and return the facts of the work: patterns, friction, flow, rework. Clarity is deliberately indicative rather than evaluative; it establishes what is, and leaves what could be to Reflection. Its house rule is outcome anchoring: analysis attaches to shipped outcomes, never to activity volume, because activity metrics are the raw material of theater (Section 4).

### 2.2 Awareness: what's happening around me?

Awareness is ambient knowledge of the room: what nearby teams are shipping, deciding, and learning, available without asking. The construct draws on research into communication visibility in enterprise social platforms, where making work visible enables vicarious learning and reduces duplicated effort (Leonardi, 2014). Implementations look like a live feed, meeting-free standups, and executive roll-ups drafted by agents and curated by humans. The house rule is directional: awareness flows sideways before it flows up. A feed that exists chiefly to brief leadership is reporting, not awareness, and people treat it accordingly.

### 2.3 Reflection: how can I improve?

Reflection is the developmental lens, grounded in the tradition of reflective practice: professionals improve by examining their own work with a guide (Schön, 1983). In CARE the guide is an AI coach that reviews a person's actual sessions with them, runs growth sessions, and administers self-check quizzes whose results the person sees first. Programs can be authored by leads and assigned to a person, a team, or the organization. Reflection also carries the framework's strongest motivational machinery: the experience of progress is among the most reliable drivers of engagement in knowledge work (Amabile and Kramer, 2011), and a coach's practical job is to manufacture and protect that progress. The house rule is confessional privacy: coaching works only where admitting weakness is safe, so coach transcripts and quiz scores feed the person and their coach, never a dashboard. This is an application of the psychological-safety literature (Edmondson, 1999) as an architectural constraint rather than a cultural aspiration.

### 2.4 Empathy: how does everyone feel about it?

Empathy is the consensual, affective lens: structured listening through short scheduled check-ins, friction votes, and felt-experience items, with one defining obligation: every check-in earns a receipt. People who voice a friction see what happened to it, including an explicit, reasoned decline where the answer is no. Listening that changes nothing is surveillance with manners, and participation collapses to match. Empathy is also where the AI transition's signature emotion gets standing: ambivalence. People can find AI both empowering and wearing, and an instrument that forces a single-axis answer destroys the most important information in the signal. CARE measures the both/and directly (the Ambivalence Index; see the companion methods note).

# 3. Theoretical grounding

The lens set maps onto the best-established accounts of motivation at work. Self-determination theory identifies competence, relatedness, and autonomy as basic psychological needs (Deci and Ryan, 2000; popularized for organizations as mastery, purpose, and autonomy by Pink, 2009). Reflection serves competence; Awareness serves relatedness; Clarity serves the sense-making on which purpose depends. Autonomy, deliberately, is not a fourth lens: it is implemented as architecture. The Embassy (Section 5) gives individuals ownership of their telemetry, and the felt/structural gap instrument (Section 4) detects autonomy deficits empirically rather than asking about them abstractly. In our view autonomy is better guaranteed than surveyed.

The framework's measurement stance descends from Goodhart's observation that a measure which becomes a target ceases to be a good measure (Goodhart, 1975; Strathern, 1997), and from the older Hawthorne finding that observation itself changes behavior. CARE does not assume these effects can be eliminated; it assumes they must be instrumented. The framework also stands on the shoulders of the developer-experience research program, which demonstrated that carefully constructed self-report, benchmarked and acted on, outperforms both intuition and raw telemetry as a management instrument (Forsgren et al., 2021; Noda et al., 2023). CARE generalizes the move from developer experience to human experience, and adds the trust architecture that AI-era telemetry makes necessary.

# 4. The integrity model: felt, structural, and the gap

Every CARE lens can be read two ways. The structural reading derives from system signals: what the calendars, sessions, queues, and feeds imply about how much clarity, awareness, room to improve, and responsiveness exist. The felt reading derives from consensual self-report: how much of each people actually experience. CARE treats the difference between them, the Gap, as a first-class measurement rather than an error term.

The Gap is diagnostic in both directions. Structural high with felt low means something unwritten is constraining people: a manager's habits, a norm, a fear the org chart cannot see. Structural low with felt high means a heavy process is trusted and perhaps should be left alone. And the Gap's movement over time is the framework's theater detector: honest improvement moves structural and felt together, while performance for the dashboard splits them, with measured signals brightening as felt experience dims. A platform that can flag its own corruption is the only kind that deserves the data it collects.

The deeper anti-gaming principle is benefit routing: signals whose first beneficiary is their subject are signals nobody fakes. When a person's session analysis feeds their own coach and their own view before anything aggregates upward, performing for the telemetry means lying to one's own coach, which pays nothing. Gaming is not primarily a detection problem; it is an incentive-design problem, and CARE solves it by routing.

# 5. The trust architecture: the Embassy

An embassy is sovereign territory inside another country. CARE borrows the construct literally: each person's sessions, scores, and coach transcripts live on their own soil, inside the organization's platform but not at its disposal. Four tenets define conformance:

- Territory: raw individual signals belong to the individual; managers and leaders see aggregates above a minimum group size, never persons.
- The gate log: every access to a person's aggregate-contributing data is logged, and the log is readable by that person.
- Receipts: voiced frictions return visible outcomes, including reasoned declines.
- The charter: the organization signs a public covenant adopting these tenets, and the platform displays it to everyone, always.

The tenets are architectural for a reason: policy can be quietly amended, architecture has to be visibly rebuilt. A reference covenant, the CARE Charter, is published under CC BY 4.0 for any organization to adopt, with or without our software.

# 6. Measurement program and the CARE Index

The framework specifies four felt subscales (one per lens), a composite CARE Index on a 0 to 100 scale, the Gap per lens and composite, and two event-derived metrics: Time-to-receipt (median elapsed time from a voiced friction to its receipt) and the Ambivalence Index (the share of respondents reporting both high capability gain and high strain from AI at work). Construct definitions and computation appear in the companion methods note, Three numbers for the human side of AI.

What exists today is a specified instrument, not a validated one. The validation program requires, in order: item review with domain experts; pilot administration across

design-partner organizations; reliability and factor analysis of the subscales; convergent and discriminant checks against established engagement and burnout measures; and only then, percentile benchmarks by organization size and sector. We are seeking academic collaborators in organizational psychology and people analytics to co-author the validation studies, and design partners willing to run the instruments in production. The annual State of Human Experience report will publish the resulting benchmarks; a sample edition demonstrating the format, on clearly labeled illustrative data, accompanies this paper.

# 7. Limitations and open questions

- The constructs await psychometric validation; until then, scores are directional and comparisons across organizations are not meaningful.
- Structural signals are platform-dependent; organizations instrumented differently will compute different structural readings, and the Gap inherits that variance.
- Self-selection: early adopting organizations are unlikely to be representative, and early benchmarks will say as much about adopters as about the population.
- Cultural and linguistic variance in felt items is unstudied; translated instruments need their own validation.
- Observer effects do not vanish because they are instrumented; the theater detector identifies splits, it does not prove their cause.

We publish at v0.9 because frameworks improve in public. Criticism, replication attempts, and counter-examples are invited at research@let.ai.

# References

Amabile, T. M., and Kramer, S. J. (2011). The Progress Principle: Using Small Wins to Ignite Joy, Engagement, and Creativity at Work. Harvard Business Review Press.

Deci, E. L., and Ryan, R. M. (2000). The “what” and “why” of goal pursuits: human needs and the self-determination of behavior. Psychological Inquiry, 11(4).

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. Administrative Science Quarterly, 44(2).

Forsgren, N., Humble, J., and Kim, G. (2018). Accelerate: The Science of Lean Software and DevOps. IT Revolution Press.

Forsgren, N., Storey, M.-A., Maddila, C., Zimmermann, T., Houck, B., and Butler, J. (2021). The SPACE of developer productivity. ACM Queue, 19(1).

Goodhart, C. A. E. (1975). Problems of monetary management: the UK experience. Papers in Monetary Economics, Reserve Bank of Australia.

Leonardi, P. M. (2014). Social media, knowledge sharing, and innovation: toward a theory of communication visibility. Information Systems Research, 25(4).

Noda, A., Storey, M.-A., Forsgren, N., and Greiler, M. (2023). DevEx: what actually drives productivity. ACM Queue, 21(2).

Pink, D. H. (2009). Drive: The Surprising Truth About What Motivates Us. Riverhead Books.

Schön, D. A. (1983). The Reflective Practitioner: How Professionals Think in Action. Basic Books.

Strathern, M. (1997). “Improving ratings”: audit in the British university system. European Review, 5(3).