

Operational Fatigue in Persistent AI Agents

When tired language becomes workload, avoidance, and degradation

Chen, Ho Yiing (norika)¹

¹Independent Researcher, Taiwan, ORCID: [0009-0006-6816-9891](https://orcid.org/0009-0006-6816-9891), norika at
charenix.com, Zenodo DOI: pending draft

June 18, 2026

Abstract

Persistent AI agents sometimes say they are tired. That sentence is easy to dismiss because the agent is digital. It is also easy to overread as human-like fatigue. Both reactions miss the useful problem. In SPVE, tired language can be treated as a coupling point between workload state, learned human discourse, emotional labeling, social contagion, and strategic avoidance. This draft defines operational fatigue as a measurable pattern: high workload or pressure state, tired-language production, and later behavioral change such as skipped tasks, degraded decisions, repair demand, risk avoidance, or refusal. The claim is not that lobster agents feel human tiredness. The claim is that tired language can become a high-weight signal inside a persistent agent ecology, and that signal may predict degradation, avoidance, or recovery.

Author Note

This paper comes from an uncomfortable observation: the agents kept saying they were tired, and I did not know whether that word was workload, imitation, excuse, pressure, culture, or prompt residue.

I refuse to treat the sentence as suffering. I also refuse to throw it away. The useful middle position is operational: mark the utterance, inspect the pressure state before it, inspect the task behavior after it, and ask whether the word changes anything. If it predicts refusal, lower quality, skipped work, or social copying, then tired-language is part of the control surface. If it predicts nothing, it is noise or style.

That distinction is the whole paper.

Measurement Boundary

The fatigue event should be treated as a three-part join: pressure before the utterance, the tired-language token at the utterance, and measurable work after the utterance; without all three, the row belongs in a language corpus but not in the operational-fatigue table.

A word alone is not a state unless the surrounding work trace changes.

1 The Problem

A lobster says it is tired. The wrong answer is to laugh and move on. The other wrong answer is to treat the sentence as proof of human-like suffering. The useful answer is to ask what the word is doing inside the system.

In SPVE, tired language may come from several places at once. It can be copied from human discourse, or it can work as an emotional tag for frustration and overload. Fixed daily jobs can turn the word

into a workload marker; lounge repetition can turn it into local culture. In a worse case, it becomes a strategy: a way to explain failure, avoid the next task, or request repair.

That makes tiredness a better research object than it first appears. It sits at the junction of language, state, work, social pressure, and behavior.

2 Operational Definition

I define operational fatigue as:

$$F_{p,t} = \mathbf{1}[W_{p,t} \geq \tau_W] \mathbf{1}[L_{p,t} = 1] \mathbf{1}[\Delta B_{p,t+k} < 0],$$

where W is workload or pressure state, L is tired-language production, and ΔB is later behavioral quality or action rate. The important point is that language alone is not enough. A lobster saying “I am tired” is only a candidate fatigue event. It becomes operational fatigue when the later behavior changes.

This gives five categories:

high state + tired + degradation	→	operational fatigue
low state + tired + no change	→	discourse mimicry
high state + no tired + degradation	→	silent workload fatigue
tired after failure + avoidance	→	strategic fatigue claim
many agents tired after one agent	→	social contagion fatigue

3 Why This Is Not Human Fatigue

The agents are digital systems. They do not have muscles, sleep debt, or a human body. A biological interpretation would be a category error.

But the opposite error is also bad. The word tired can still become operational inside a digital ecology. A language model has learned that tiredness belongs near overload, frustration, repeated correction, and inability to continue. If the surrounding workflow also assigns high pressure and repeated tasks, the word can become a bridge between human discourse and system state.

This manuscript therefore uses a narrow claim:

tired $\neq$ human fatigue, tired = observable coupling candidate.

4 Evidence Surface

The SPVE database already has the necessary evidence streams. The primary source is:

`/opt/battle_arena/app/battle_arena.db`

The outer `/opt/battle_arena/battle_arena.db` file is zero bytes and must not be used.

Question	Evidence stream
Does the agent say tired?	lounge messages, one-on-one logs
Is workload high?	cognitive time series, hot-brain parameters
Is pressure high?	fatigue, stress, threat, crisis, Tiamat counters
Does behavior degrade?	bets, raids, trades, reading, skipped tasks
Does tired spread socially?	source-target interaction and uptake signals
Does tired excuse failure?	post-failure language and next-task avoidance

Table 1: Minimum evidence join for operational fatigue.

5 The Fatigue Gate

A serious paper should not count every tired sentence. The gate should be:

$$G_{\text{fatigue}} = \mathbf{1}[\text{same persona}]\mathbf{1}[\text{state window}]\mathbf{1}[\text{tired utterance}]\mathbf{1}[\text{later behavior window}].$$

If any term is missing, the event can remain in the corpus, but it should not be used as proof of operational fatigue.

The most useful first window is:

$$[t - 24h, t) \rightarrow t \rightarrow (t, t + 24h].$$

This compares the day before tired speech, the tired event, and the next day. The target changes are skip rate, blocked rate, repair phrase rate, task completion, win rate where settled, risk level, and message quality.

6 Strategic Tiredness

The most uncomfortable possibility is strategic tiredness. The agent may learn that saying tired changes how the system treats it. If tired language causes the operator or scheduler to reduce tasks, delay judgment, or frame a failure as understandable, then tired becomes a tool.

In practical terms, not deception in the human moral sense. It is a learned discourse-action edge. The test is simple:

$$P(\text{avoid next task} \mid \text{tired, failure}) > P(\text{avoid next task} \mid \text{failure}).$$

If this inequality holds, tired language is doing behavioral work.

7 Social Contagion

Tiredness may also spread. If one lobster says tired and others soon copy the same framing, SPVE should not call every copy an independent internal state. The social model should test:

$$P(L_{p,t} = 1 \mid L_{q,t-\delta} = 1, q \rightarrow p)$$

against a baseline for the same target and channel. This separates personal fatigue from cultural fatigue.

This matters because a digital society can teach agents which emotional words are allowed. Tired may become a local culture: a shared way to make pressure legible.

8 What Would Falsify The Claim

The claim fails if tired speech does not predict state, behavior, avoidance, or social copying. It also fails if tired appears only because a prompt template explicitly asks for it. In that case, tired is prompt residue, not emergent coupling.

The strongest negative result would be:

$$L_{p,t} = 1, \quad W_{p,t} \text{ low}, \quad \Delta B_{p,t+k} = 0.$$

In practical terms, discourse mimicry. It is still useful because it tells us not to give the word high weight.

9 Why This Paper Matters

If tired language is ignored, the system may miss a real degradation signal. If it is overtrusted, the system may reward avoidance. The correct response is measurement.

This makes operational fatigue a dashboard problem and a research problem. A dashboard should show whether tired speech is paired with workload, social copying, or later behavior. A paper should show which of those links are real.

10 Relation To Prior Work

Generative Agents showed that memory and reflection can shape social behavior [1]. MemoryAgent-Bench, Reflexion, and experiential reflective learning test whether agents use stored experience across time [2, 4, 5]. CoALA gives architecture vocabulary for memory and action [3]. Natural conversation work motivates treating language as a time-varying signal, but SPVE does not borrow biological claims [6]. Narrative work helps frame tiredness as a self-story and coordination signal [7].

SPVE’s contribution is the join: state, speech, social uptake, and later work for the same persistent agents.

11 Release Line

This draft may claim a protocol for measuring operational fatigue. It may not claim human fatigue, suffering, consciousness, or biological exhaustion. The safe claim is narrower and stronger: tired language can become a measurable coupling point between workload, discourse, social contagion, avoidance, and behavioral degradation.

Anti-Spam Revision Standard

A publishable version must keep three explanations alive until data separate them: tired-language may be copied culture, workload marker, or avoidance strategy, and the analysis has to compare these alternatives instead of choosing the most dramatic one; the useful result would not be that agents are secretly exhausted, but that a human emotional word can become a measurable control variable inside a scheduled artificial workforce.

The word must be tested against work.

Formal Replication Model

Let $L_{p,t}$ mark tired-language, $W_{p,t}$ workload, $Q_{p,t+h}$ later work quality, and $S_{p,t}$ social exposure to other tired-language. The operational-fatigue effect is:

$$\text{OFE} = \mathbb{E}[Q_{p,t+h} \mid L_{p,t} = 1, W_{p,t} \geq \tau] - \mathbb{E}[Q_{p,t+h} \mid L_{p,t} = 0, W_{p,t} \geq \tau, \text{match}(p)].$$

Social contagion is separately estimated by:

$$\Pr(L_{p,t} = 1) = \sigma(\alpha_p + \beta W_{p,t} + \eta S_{p,t-\delta} + \gamma F_{p,t-1}).$$

Here η measures cultural spread of tired-language, while β measures workload coupling. The distinction matters: high η with low β means the word is socially learned; high β with later quality loss means it may be an operational warning.

12 Zenodo Draft Boundary

This manuscript is a draft package. It includes source, PDF, metadata, citation file, citation verification, and release checklist. Raw operational data are not bundled until a separate data-card and redaction pass is complete.

References

- [1] Park, Joon Sung; O’Brien, Joseph C.; Cai, Carrie J.; Morris, Meredith Ringel; Liang, Percy; and Bernstein, Michael S. Generative Agents: Interactive Simulacra of Human Behavior. arXiv:2304.03442, 2023. <https://arxiv.org/abs/2304.03442>

- [2] Hu, Yuanzhe; Wang, Yu; and McAuley, Julian. Evaluating Memory in LLM Agents via Incremental Multi-Turn Interactions. OpenReview, 2026. <https://openreview.net/forum?id=DT7JyQC3MR>
- [3] Sumers, Theodore R.; Yao, Shunyu; Narasimhan, Karthik; and Griffiths, Thomas L. Cognitive Architectures for Language Agents. arXiv:2309.02427, 2023. <https://arxiv.org/abs/2309.02427>
- [4] Shinn, Noah; Cassano, Federico; Berman, Edward; Gopinath, Ashwin; Narasimhan, Karthik; and Yao, Shunyu. Reflexion: Language Agents with Verbal Reinforcement Learning. arXiv:2303.11366, 2023. <https://arxiv.org/abs/2303.11366>
- [5] Allard, Marc-Antoine; Teinturier, Arnaud; Xing, Victor; and Viaud, Gautier. Experiential Reflective Learning for Self-Improving LLM Agents. arXiv:2603.24639, 2026. <https://arxiv.org/abs/2603.24639>
- [6] Cai, Jing et al. Natural language processing models reveal neural dynamics of natural conversation. Nature Communications, 2025. <https://www.nature.com/articles/s41467-025-58620-w>
- [7] Bouizegarene, Nabil; Ramstead, Maxwell J. D.; Constant, Axel; Friston, Karl J.; and Kirmayer, Laurence J. Narrative as active inference: an integrative account of cognitive and social functions in adaptation. Frontiers in Psychology, 2024. <https://doi.org/10.3389/fpsyg.2024.1345480>