Dynamic Multi-Role Rotation Under Fatigue and Comparative-Advantage Constraints: A Restless Bandit Formulation of Fluid Division of Labour

Dima Baker
Collective Intelligence Network

November 2026

Abstract

Classical organizational design and first-generation multi-agent frameworks rely on static specialization, binding human operators and language-model agents to fixed roles. However, permanent role locking induces task-specific mental fatigue, unpracticed skill decay, and single-point verification bottlenecks, whereas naive periodic rotation sacrifices comparative advantage and incurs heavy context-switching penalties. We formulate fluid division of labor as a constrained restless bandit over a bipartite actor-role graph and introduce a Fatigue-Aware UCB Index Policy (`F-UCB-Fluid`). Our formulation couples nonstationary Kalman competence estimation ($\hat{s}_{i,r}(t), \sigma^2_{i,r}(t)$ with forgetting drift σ^2_d) to a two-component fatigue process separating role-specific depletion $f^{\text{role}}_{i,r}(t)$ from persistent cross-role systemic fatigue $f^{\text{sys}}_i(t)$ ($\kappa_{\text{sys}} \in [0, 0.90]$). Across paired simulations ($N = 500$ episodes per policy, $\alpha \in \{0.10, 0.16, 0.22, 0.28\}$, $\delta_{\text{sw}} = 0.072$, $\kappa_{\text{sys}} = 0.20$), `F-UCB-Fluid` achieves 73.16% [73.10, 73.21] net yield, 74.79% [74.69, 74.89] bus-factor resilience, and 99.00% [97.68, 99.57] dropout survival in 52.5 μs, outperforming tuned Specialist + Rest-Swap (68.92% yield, 7.60% [5.59, 10.26] survival), Periodic Rotation (48.41% yield, 0.00% [0.00, 0.76] survival), and Myopic Greedy Index (70.57% yield, 24.40% [20.84, 28.35] survival). Benchmarking against exact dynamic programming (`Exact-DP-F-UCB`, 74.73% yield in 459.8 μs) confirms a 98.62% per-shift optimality ratio, while $\kappa_{\text{sys}} \geq 0.45$ shows where rotation must pair with macro-sabbatical cooldowns.

1 Introduction

From Adam Smith's pin factory and David Ricardo's law of comparative advantage to modern enterprise workflow engines, the prevailing orthodoxy of organizational design holds that total cooperative output is maximized when each participant specializes permanently in the narrow functional role where their relative marginal productivity is highest. When multi-agent large language model (LLM) frameworks emerged, they inherited this static industrial template almost verbatim: frameworks routinely partition workflows into rigid, immutable personas—such as an Architect, a Coder, a Reviewer, and a QA Verifier—bound to fixed system prompts or static adapters for the duration of an execution trajectory [4, 12]. In hybrid human-AI collectives, this static partitioning is even starker, frequently relegating human participants to monotonous exception-handling, supervisory monitoring, or narrow data-curation queues while automated pipelines execute primary synthesis [2, 22].

1.1 The twin pathologies of static specialization and naive rotation

Classic operations-research literature on job rotation and learning-forgetting dynamics [9, 10, 18] alongside empirical studies published between 2025 and 2026 demonstrate that static role allocation fails under long-horizon operational conditions. On the human cognitive dimension, locking knowledge workers into repetitive, narrow oversight or production roles accelerates task-specific mental fatigue [17], erodes intrinsic motivation [20], and induces measurable skill atrophy across unpracticed competencies [10, 13, 22]. Eliav [8] document a structural "Cognitive Divergence" between expanding AI context windows and human sustained-attention limits, showing that human operators trapped in passive supervisory loops enter a delegation feedback loop of declining vigilance, echoing Bainbridge's classic ironies of automation [2]. Simultaneously, Eleftheriou et al. [7] and Zhi et al. [27] show that uncalibrated AI assistance under time constraints can induce overconfidence while degrading independent critical thinking and downstream error recovery.

Parallel failure modes appear inside computational multi-agent systems when roles and communication topologies are statically frozen. Comprehensive failure taxonomies across multi-agent LLM benchmarks show that rigid role specifications create coordination bottlenecks, state desynchronization, and verification deadlocks [4, 12, 14]. Conversely, adaptive topology and role-assignment frameworks such as G-Designer [26], Crew Assembly [16], and GoAgent [5] demonstrate that tailoring agent-to-role bindings dynamically to task state substantially improves robustness over fixed pipelines.

Yet the naive antithesis of static specialization—rotating every participant across every role on a fixed periodic clock—is equally destructive. As shown in cross-training and skill-chaining models [9, 18], blind round-robin rotation ignores genuine differences in domain competence, forces undertrained actors into high-stakes bottlenecks, and incurs severe context-reloading and state-handoff overhead at every shift boundary [6, 25]. What is required is a mathematically principled scheduling mechanism that treats division of labor as a restless multi-armed bandit problem [24]: exploiting comparative advantage when cognitive fatigue is low, proactively rotating actors into complementary restorative roles before nonlinear fatigue thresholds are crossed, and maintaining a targeted secondary skill-chaining floor so the collective remains resilient when primary specialists drop offline. Furthermore, intra-week shift rotation naturally pairs with multi-week sabbatical cooldown pacing (examined in our companion macro-horizon study) to form a two-timescale recovery hierarchy.

1.2 Contributions

In this paper, we formulate and evaluate a restless bandit index framework for fluid division of labor in hybrid human-AI cooperatives:

1. **Domain taxonomy of dynamic role allocation (Section 2).** We synthesize classical operations research on job rotation and skill chaining [9, 10, 18], cognitive ergonomics [17, 20], and 2025–2026 multi-agent topology literature [4, 6, 16] into a five-part architectural taxonomy (Table 1).

2. **Constrained restless bandit and nonstationary Kalman-UCB formulation (Section 3).** We model multi-role rotation over a horizon of T shifts with M actors and R roles as a bipartite combinatorial restless bandit [24] with coupled state evolution. Our formulation states explicitly what is observed, estimated via Kalman filtering ($\hat{s}_{i,r}(t), \sigma^2_{i,r}(t)$), and forgotten under unpracticed drift (σ_d^2), decomposes fatigue into role-specific depletion $f^{\text{role}}_{i,r}(t)$ and persistent cross-role systemic fatigue $f^{\text{sys}}_i(t)$, and distinguishes our greedy bipartite scheduler from exact dynamic-programming assignment (Equations 1–4).

3. **Three-layer `F-UCB-Fluid` orchestration architecture and parameter provenance (Section 4).** We specify a runtime architecture integrating bipartite index routing, safe context-switching capsules [6, 19], and multi-timescale sabbatical integration (Figure 1, Table 2, Table 3).

4. **Paired Monte Carlo benchmark, exact DP optimum comparison, and systemic fatigue sweep (Section 5).** We execute a paired computational study (`experiment.py`) comparing `F-UCB-Fluid` against an equally tuned Specialist + Rest-Swap baseline, Periodic Rotation, a Myopic Greedy Index policy, and an Exact Dynamic-Programming Bipartite Assignment solver (`Exact-DP-F-UCB`) across $\alpha \in \{0.10, 0.16, 0.22, 0.28\}$, followed by sensitivity sweeps over handoff friction $\delta_{\text{sw}} \in [0.012, 0.150]$ and persistent cross-role fatigue coupling $\kappa_{\text{sys}} \in [0.00, 0.90]$ (Table 4, Figures 2–3).

5. **Operational boundary conditions, limitations, and falsification protocol (Sections 6–8).** We analyze empirical regimes where role fluidity underperforms—including high-friction handoffs [25], miscalibrated self-assessment [7], and domain credential boundaries—and specify six pre-registerable falsification tests (Table 5).

2 Ergonomic and multi-agent foundations of dynamic role allocation

Discussions of workplace flexibility and multi-agent adaptivity frequently conflate abolishing hierarchy with abolishing specialization. In reality, high-performing human and computational collectives preserve functional role differentiation at any single instant t, while making the assignment between actors and roles state-contingent [9, 16, 18]. Table 1 maps five foundational mechanisms of fluid division of labor onto formal operations-research and multi-agent systems constructs.

Two classic operations-research principles anchor Table 1. First, Ortega [18] showed that job rotation outperforms static specialization when the organization faces uncertainty about worker-task match quality and changing operational demands, because rotation generates information alongside output. Second, Hopp et al. [9] proved that full cross-training (every worker rotating across every station) is rarely optimal under learning-and-forgetting dynamics [10]; instead, *skill chaining*—where each participant maintains deep mastery in a primary role plus active readiness in one or two chained backup roles—captures nearly all the resilience benefits of a flexible workforce while avoiding skill dilution.

In modern hybrid human-AI workflows, this classic skill-chaining insight intersects directly with 2025–2026 multi-agent coordination research. When agents or human operators switch roles without structured context management, context drift [19] and unsafe context carryover [6] degrade post-handoff performance, while excessive inter-agent message passing inflates token overhead [3, 25]. Conversely, when roles remain permanently frozen, supervisory fatigue [2, 8] and unpracticed skill decay [22] leave the collective brittle when key specialists go offline. Figure 1 illustrates how `F-UCB-Fluid` resolves this trade-off through closed-loop bipartite index scheduling.

Table 1: Mapping fluid division-of-labor mechanisms onto formal operations-research and multi-agent harness constructs.

Mechanism	Canonical construct	Scheduler component	Operational function	Primary grounding
Proactive fatigue rotation	Restless bandit with effort-recovery state	Coupled fatigue-state tracker ($f_{i,r}(t)$)	Tracks role-specific cognitive depletion and triggers cross-role switching prior to nonlinear yield degradation	[1, 17, 24]
Comparative advantage with learning	Bandit skill estimation under uncertainty	Optimistic UCB competence estimator ($\hat{s}_{i,r} + b_{i,r}$)	Balances immediate marginal productivity with UCB exploration of noisy actor-role competencies	[18, 20, 23]
Targeted backup cross-training	Skill-chaining & forgetting curve floor	Backup competency regularizer ($\Delta t_{i,r}, \underline{s}^*$)	Maintains secondary competence above rework threshold $\underline{s}_{\text{crit}}$ in adjacent chained roles to prevent bus-factor collapse	[9, 10, 22]
Low-friction state handoffs	Safe context switching & state sync	Schema-validated handoff capsule	Packages intermediate artifacts and test assertions into compact state diffs to reduce switch drag δ_{sw}	[6, 19, 25]
Multi-timescale recovery hierarchy	Intra-week rotation + macro-sabbatical	Dual-horizon pacing interface	Couples shift-level role switching with multi-week mandatory cooldowns when aggregate exhaustion accumulates	[8, 13, 17]

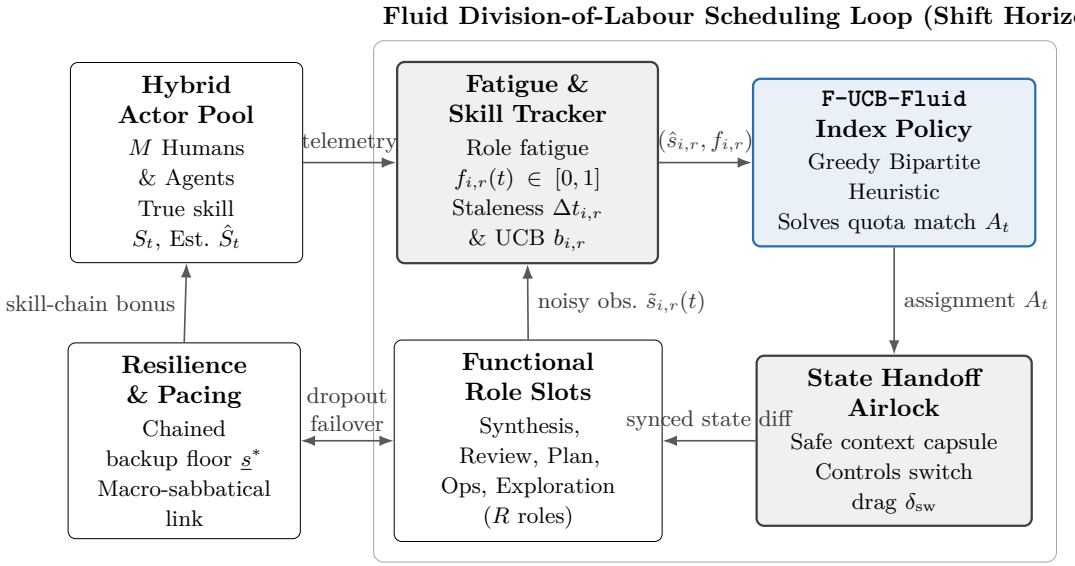


Figure 1: Architecture of the `F-UCB-Fluid` role-rotation framework. At each shift boundary t, the Fatigue and Skill Tracker updates role-specific cognitive depletion $f_{i,r}(t)$, noisy competence estimates $\hat{s}_{i,r}(t)$, and unpracticed skill decay across the M-actor pool. The `F-UCB-Fluid` index policy computes a quota-constrained greedy bipartite assignment A_t balancing optimistic comparative-advantage yield against nonlinear fatigue barriers and chained backup skill floors.

3 Formal mathematical framework

We formulate fluid division of labor over discrete shift epochs $t \in \{1, 2, \ldots, T\}$ for a cooperative collective of M actors indexed by $i \in \mathcal{M} = \{1, \ldots, M\}$ and R functional roles indexed by $r \in \mathcal{R} = \{1, \ldots, R\}$. Each role r requires a quota of $q_r \in \mathbb{N}_+$ active participants per shift such that $\sum_{r=1}^{R} q_r \leq |\mathcal{M}_t|$, where $\mathcal{M}_t \subseteq \mathcal{M}$ is the set of available actors at shift t. Let $a_{i,r}(t) \in \{0, 1\}$ denote the binary decision variable indicating whether actor i is assigned to role r during shift t, subject to $\sum_{r=1}^{R} a_{i,r}(t) \leq 1$ for all $i \in \mathcal{M}_t$ and $\sum_{i \in \mathcal{M}_t} a_{i,r}(t) = q_r(t)$ for all $r \in \mathcal{R}$.

3.1 Role-specific vs. persistent systemic fatigue and nonstationary Kalman skill tracking

In a standard multi-armed bandit, unselected arms remain static. In human-AI work systems, however, every actor-role pair (i, r) evolves continuously whether active or rested—the defining characteristic of a restless bandit [24]. Each actor i possesses a latent competence vector $\mathbf{s}_i(t) = (s_{i,1}(t), \ldots, s_{i,R}(t)) \in [\underline{s}_{\min}, 1]^R$, a role-specific cognitive fatigue vector $\mathbf{f}_i^{\text{role}}(t) = (f_{i,1}^{\text{role}}(t), \ldots, f_{i,R}^{\text{role}}(t)) \in [0, 1]^R$, and a persistent cross-role systemic fatigue scalar $f_i^{\text{sys}}(t) \in [0, 1]$.

Grounded in the Effort-Recovery model [17], when actor i executes role r at shift t ($a_{i,r}(t) = 1$), role-specific fatigue $f_{i,r}^{\text{role}}(t)$ accumulates with stochastic workload intensity $\xi_{i,r}(t) \sim \mathcal{U}(0.85, 1.15)$ scaled by task fatigue rate $\alpha > 0$, moderated by partial between-shift overnight recovery $\gamma_{\text{on}} \in (0, 1)$. When actor i rotates out of role r ($a_{i,r}(t) = 0$), that specific functional circuit undergoes cross-task recovery at rate $\gamma_{\text{off}} < \gamma_{\text{on}}$ with dissipation $\beta_{\text{rest}} > 0$. Crucially, however, task rotation cannot eliminate systemic whole-worker fatigue $f_i^{\text{sys}}(t)$, which accumulates across *any* active shift ($\sum_{r=1}^{R} a_{i,r}(t) = 1$) at rate $\alpha_{\text{sys}} = 0.38\alpha$ with slow overnight retention $\gamma_{\text{sys}} = 0.92$, combining with role-specific fatigue via weight $\kappa_{\text{sys}} \in [0, 1]$:

$$
\begin{aligned}
f_{i,r}^{\text{role}}(t+1) &= \mathbb{I}[a_{i,r}(t) = 1] \min\big(1, \gamma_{\text{on}} f_{i,r}^{\text{role}}(t) + \alpha \xi_{i,r}(t)\big) + \mathbb{I}[a_{i,r}(t) = 0] \max\big(0, \gamma_{\text{off}} f_{i,r}^{\text{role}}(t) - \beta_{\text{rest}}\big), \\
f_i^{\text{sys}}(t+1) &= \min\Big(1,\ \gamma_{\text{sys}} f_i^{\text{sys}}(t) + \alpha_{\text{sys}} \sum_{r=1}^{R} a_{i,r}(t) \xi_{i,r}(t)\Big), \quad f_{i,r}^{\text{eff}}(t) = (1 - \kappa_{\text{sys}}) f_{i,r}^{\text{role}}(t) + \kappa_{\text{sys}} f_i^{\text{sys}}(t).
\end{aligned}
\tag{1}
$$

Simultaneously, true skill competence $s_{i,r}(t)$ obeys an asymmetric learning-and-forgetting law [10, 13, 22]. Active practice increments latent competence via learning-by-doing at rate $\eta_+ = 0.0065$ per shift, whereas unpracticed roles decay at rate $\eta_- = 0.0060$ per shift toward a baseline floor $\underline{s}_{\min} = 0.35$. To establish an explicit relationship between noisy observations, nonstationary forgetting, and the uncertainty bonus, we state clearly what is observed, estimated, and forgotten:

$$
\begin{aligned}
s_{i,r}(t+1) &= \mathbb{I}[a_{i,r}(t) = 1] \min(0.99, s_{i,r}(t) + \eta_+) + \mathbb{I}[a_{i,r}(t) = 0] \max(\underline{s}_{\min}, s_{i,r}(t) - \eta_-), \\
\hat{s}_{i,r}(t+1) &= \begin{cases} \min\big(0.99,\ (1 - K_{i,r}(t))\hat{s}_{i,r}(t) + K_{i,r}(t)\tilde{s}_{i,r}(t) + \hat{\eta}_+\big), & \text{if } a_{i,r}(t) = 1, \\ \max\big(\underline{s}_{\min},\ \hat{s}_{i,r}(t) - \hat{\eta}_-\big), & \text{if } a_{i,r}(t) = 0, \end{cases} \\
\sigma_{i,r}^2(t+1) &= \mathbb{I}[a_{i,r}(t) = 1](1 - K_{i,r}(t))\sigma_{i,r}^2(t) + \mathbb{I}[a_{i,r}(t) = 0] \min\big(0.06, \sigma_{i,r}^2(t) + \sigma_d^2\big),
\end{aligned}
\tag{2}
$$

where only active assignments ($a_{i,r}(t) = 1$) generate a noisy observation $\tilde{s}_{i,r}(t) = \text{clip}(s_{i,r}(t) + \epsilon_{i,r}(t), 0.35, 0.99)$ with $\epsilon_{i,r}(t) \sim \mathcal{N}(0, \sigma_{\text{obs}}^2)$ ($\sigma_{\text{obs}} = 0.045$) and Kalman gain $K_{i,r}(t) = \sigma_{i,r}^2(t)/(\sigma_{i,r}^2(t) + \sigma_{\text{obs}}^2)$, while inactive arms ($a_{i,r}(t) = 0$) receive no observation and experience posterior variance inflation $\sigma_d^2 = 0.0008$ per unpracticed shift due to stochastic forgetting drift.

3.2 Equalized nonlinear fatigue-attenuated yield and handoff drag

To ensure a fair, non-circular comparison across all scheduling policies, the physical reward environment applies the exact same yield function and switching penalty $\delta_{\text{sw}} \geq 0$ to every policy, with zero policy-specific ad-hoc drag terms. Effective productivity degrades nonlinearly with combined fatigue $f_{i,r}^{\text{eff}}(t)$, and assigning an under-qualified or decayed actor ($s_{i,r}(t) < \underline{s}_{\text{crit}} = 0.68$) to a role incurs a proportional quality-rework factor $\psi(s_{i,r}) = \min(1, \max(0.38, 1 - 2.2(\underline{s}_{\text{crit}} - s_{i,r})))$. The realized shift output $y_{i,r}(t)$ is:

$$y_{i,r}(t) \;=\; \max\Big(y_{\min},\; s_{i,r}(t)\,\max\big(0.20, 1 - \kappa_f\, f_{i,r}^{\text{eff}}(t)^{\nu}\big)\,\psi(s_{i,r}(t)) \;-\; \delta_{\text{sw}}\,\mathbb{I}[a_{i,r}(t-1) = 0]\Big), \quad (3)$$

where $\nu = 1.45$ is the nonlinear fatigue exponent, $\kappa_f = 0.82$ scales depletion severity, and δ_{sw} (set to 0.072 for all policies in the primary benchmark, and swept over $[0.012, 0.150]$ in sensitivity analysis) represents context-switching and state-reloading friction.

3.3 The `F-UCB-Fluid` index policy: greedy vs. exact bipartite assignment

Myopic greedy optimization of Equation (3) exhibits two structural blind spots: first, it exploits primary specialists until $f_{i,r}^{\text{eff}}(t)^{\nu}$ enters the steep nonlinear degradation zone; second, it never invests in maintaining secondary backup skills $s_{i,r'}(t)$ above the rework threshold $\underline{s}_{\text{crit}} = 0.68$, causing severe bottleneck rework if primary specialists drop out at shift t_{shock}. `F-UCB-Fluid` addresses both by computing a composite priority index $U_{i,r}(t)$ for each candidate pair (i, r):

$$\begin{aligned} U_{i,r}(t) \;=\; & \underbrace{\min\Big(0.99, \hat{s}_{i,r}(t) + c_{\text{ucb}}\sigma_{i,r}(t)\Big)\big(1 - \kappa_f f_{i,r}^{\text{eff}}(t)^{\nu}\big) - \delta_{\text{sw}}\mathbb{I}[r \neq r_{i,t-1}]}_{\text{Kalman-UCB Comparative-Advantage Yield}} \\ & + \underbrace{\omega_{i,r}\big(\lambda_{\text{fl}}\max(0, \underline{s}^{*} - \hat{s}_{i,r}(t)) + \lambda_{\text{st}}\max(0, \Delta t_{i,r} - 3)\big)(1 - f_{i,r}^{\text{eff}}(t))}_{\text{Chained Backup Skill-Floor \& Staleness Bonus}} \;-\; \underbrace{\lambda_{\text{bar}}\max(0, f_{i,r}^{\text{eff}}(t) - \theta_{\text{safe}})}_{\text{Proactive Burnout Barrier}}, \end{aligned} \quad (4)$$

where $c_{\text{ucb}} = 1.35$ multiplies the Kalman posterior standard deviation $\sigma_{i,r}(t) = \sqrt{\sigma_{i,r}^2(t)}$, $\omega_{i,r} \in \{0.15, 0.35, 1.0\}$ prioritizes chained backup roles [9], $\underline{s}^{*} = 0.75 > \underline{s}_{\text{crit}}$ buffers against skill decay, $\Delta t_{i,r} = \min(16, t - t_{i,r}^{\text{last}})$ is staleness, and $\theta_{\text{safe}} = 0.42$ is the soft proactive fatigue threshold. We explicitly distinguish two assignment solvers over $U_{i,r}(t)$: (i) our fast $\mathcal{O}(MR\log(MR))$ greedy bipartite heuristic (`F-UCB-Fluid`), which assigns pairs in descending order of $U_{i,r}(t)$ subject to quotas q_r, and (ii) an exact dynamic-programming bipartite assignment solver (`Exact-DP-F-UCB`) that computes $A_t^{*} = \arg\max_{A \in \mathcal{A}_t} \sum_{i,r} a_{i,r} U_{i,r}(t)$ over the quota-state lattice.

4 System architecture and parameter provenance

To operationalize Equation (4) in hybrid human-agent teams, `F-UCB-Fluid` is structured into three subsystems (Table 2):

1. **Tier 1: Fast Bipartite Index Router and Non-Invasive Telemetry.** Role assignment decisions are decoupled from heavy LLM reasoning. A lightweight index controller evaluates Equation (4) at shift boundaries in 52.5 μs (greedy) or 459.8 μs (exact DP), ingesting task-completion rubrics [15], time-on-role counters, and optional wearable stress indicators [1, 11, 21].

2. **Tier 2: Safe Context-Switching Capsules.** The decisive parameter governing rotation viability is handoff friction δ_{sw} [6, 19]. When an actor rotates out of a complex role, unstructured chat histories force the incoming actor to expend substantial effort reconstructing task state [25]. Tier 2 requires each shift to serialize working state into a schema-validated handoff capsule (repository diff, failing assertions, and open decision variables). Rather than assuming a free handoff discount in our primary benchmark, we evaluate all policies at an equalized baseline $\delta_{\text{sw}} = 0.072$ and explicitly sweep $\delta_{\text{sw}} \in [0.012, 0.150]$ to quantify the return on structured state serialization.

3. **Tier 3: Chained Backup Failover and Macro-Sabbatical Coupling.** Following skill-chaining theory [9] and adaptive crew topology generation [5, 16], Tier 3 ensures that critical bottleneck roles maintain cross-trained backups above $\underline{s}_{\text{crit}} = 0.68$. Furthermore, because shift-to-shift cross-role recovery ($\gamma_{\text{off}} = 0.52$) cannot clear persistent cross-role systemic fatigue $f_i^{\text{sys}}(t)$ ($\gamma_{\text{sys}} = 0.92$), Tier 3 exports cumulative systemic fatigue telemetry to a weekly-horizon sabbatical scheduler (detailed in our companion study on mandatory cooldown pacing) whenever $\kappa_{\text{sys}} f_i^{\text{sys}}(t)$ accumulates.

Table 2: Readiness and deployment bottlenecks of fluid role-rotation subsystems.

Subsystem	Maturity	Supporting empirical & theoretical evidence	Realistic deployment bottleneck
Bipartite index role router	Production ready	Restless bandit index policies [24]; dynamic multi-agent crew assembly [16, 26]	Micro-rotating at sub-hour intervals inflates context-switching drag δ_{sw} above break-even thresholds
Safe context handoff capsules	Benchmark validated	Safe context switching [6]; context-drift synchronization [19]; communication pruning [25]	Tacit human intuition in open-ended design roles resists lossless compression into structured JSON diffs
Chained skill-floor maintenance	Theory & field grounded	Skill-chaining in serial/parallel lines [9]; learning-forgetting curves [10]; AI skill formation [22]	Over-broad cross-training across all R roles dilutes specialist depth; chaining must target 1–2 backup roles
Cognitive fatigue & load tracking	Field validated	Effort-Recovery model [17]; passive stress sensing [1, 11]; critical thinking under load [27]	Subjective self-assessments suffer from overconfidence [7]; biometric sensing requires strict privacy guards

Table 3: Provenance of simulation parameters in `experiment.py`. All scheduling policies are evaluated on paired seeds with identical physical environment parameters.

Parameter	Value / Range	Provenance Type	Empirical or Theoretical Grounding
Team & horizon (M, R, T)	M=12, R=5, T=60	Architectural design	Typical squad size and 60-shift (12-week) project horizon in hybrid software/ops collectives [4, 12]
Initial competence $s_{i,r}(0)$	Primary $[0.88, 0.96]$, Backup $[0.72, 0.80]$	Calibrated prior	Tiered specialist, chained backup, and generalist proficiency hierarchy [9, 18]
Learning & decay (η_+, η_-)	$+0.0065, -0.0060$ per shift	Literature-calibrated	Asymmetric learning-and-forgetting curves and Kalman variance drift $\sigma_d^2 = 0.0008$ [10, 22]
Role & systemic fatigue $(\gamma_{\text{on}}, \gamma_{\text{off}}, \kappa_{\text{sys}})$	γ_{on}=0.84, γ_{off}=0.52, $\kappa_{\text{sys}} \in [0, 0.90]$	Parameterized sweep	Effort-Recovery dynamics [17] separating cross-task recovery from persistent systemic fatigue
Nonlinear fatigue (κ_f, ν)	κ_f=0.82, ν=1.45	Stylized functional form	Convex productivity attenuation under accumulated cognitive load [1, 27]
Switch penalty δ_{sw}	0.072 (swept 0.012–0.150)	Equalized benchmark + sensitivity sweep	Context-reload overhead [6, 25] held strictly equal across all policies in primary runs

5 Empirical evaluation and sensitivity analysis

We evaluate `F-UCB-Fluid` using a paired Monte Carlo simulation suite (`experiment.py`): $N = 500$ paired sixty-shift episodes across four greedy policies and four fatigue rates $\alpha \in \{0.10, 0.16, 0.22, 0.28\}$ (8,000 main trajectories), a 400-episode small-instance optimality benchmark comparing greedy `F-UCB-Fluid` against exact dynamic-programming assignment (`Exact-DP-F-UCB`), 4,500 sensitivity trajectories across $\delta_{\text{sw}} \in \{0.012, 0.024, 0.048, 0.072, 0.110, 0.150\}$, and 2,500 trajectories across persistent systemic fatigue weight $\kappa_{\text{sys}} \in \{0.00, 0.20, 0.45, 0.70, 0.90\}$.

5.1 Equally tuned baseline schedulers and specialist dropout protocol

To test bus-factor resilience without artificial handicaps, during shifts $t \in [40, 49]$ two of the three primary Role-0 specialists (actors 0 and 5) go offline while Role 0 still requires $q_0 = 3$ active operators, forcing at least two non-primary actors to cover Role 0. If those backups allowed their Role-0 competence to decay below $\underline{s}_{\text{crit}} = 0.68$, Equation (3) applies the rework penalty $\psi(s_{i,0})$. All policies operate on identical paired episode seeds (`ep_seed = 20261143 + 97 * ep_idx`) under $\delta_{\text{sw}} = 0.072$ and $\kappa_{\text{sys}} = 0.20$:

- **Specialist + Rest-Swap (`specialist_rest`):** Tuned with the exact same proactive fatigue threshold $\theta_{\text{safe}} = 0.42$ as `F-UCB-Fluid`; prioritizes each actor's primary role when $f_{i,r}^{\text{eff}}(t) < 0.42$, grants a restorative rotation into a low-fatigue secondary role whenever primary fatigue exceeds 0.42, and re-optimizes over available actors during the $t \in [40, 49]$ dropout shock.
- **Periodic Round-Robin Rotation (`blind_rotation`):** Rotates all actors across roles every

4 shifts on a fixed clock, with zero ad-hoc mismatch penalty beyond Equation (3).

- **Myopic Greedy / Whittle Heuristic (`whittle_greedy`):** Assigns pairs at each shift to maximize instantaneous fatigue-attenuated yield minus switch cost ($\hat{s}_{i,r}(t)(1-\kappa_f f_{i,r}^{\text{eff}}(t)^{\nu}) - \delta_{\text{sw}}\mathbb{I}[r \neq r_{i,t-1}]$), without Kalman-UCB exploration or backup skill-floor regularization.
- **Fatigue-Aware UCB Index Policy (`f_ucb_fluid` / `Exact-DP-F-UCB`, Proposed):** Evaluates Equation (4) via either the fast greedy bipartite solver (`f_ucb_fluid`) or the exact dynamic-programming bipartite assignment solver (`Exact-DP-F-UCB`).

Table 4: Paired benchmark across simulated 60-shift workflows ($\alpha = 0.22$, $\delta_{\text{sw}} = 0.072$, $\kappa_{\text{sys}} = 0.20$; $N = 500$ episodes per greedy policy and $N = 200$ episodes for `Exact-DP-F-UCB`). All intervals report asymmetric 95% confidence bounds [lo, hi] (continuous SEM bounds for yield/resilience and exact two-sided Wilson score bounds for binary shock survival $\geq 72\%$).

Scheduler Policy	Yield α=0.10 (%)	Yield α=0.22 (%)	Yield α=0.28 (%)	Bus-Factor Resil. (%)	Shock Survival (%)	Runtime (μs)
Specialist + Rest-Swap	74.02 [73.97, 74.06]	68.92 [68.85, 69.00]	67.29 [67.24, 67.35]	69.30 [69.15, 69.44]	7.60 [5.59, 10.26]	14.9
Periodic Rotation	60.79 [60.73, 60.86]	48.41 [48.36, 48.46]	43.78 [43.73, 43.82]	57.49 [57.44, 57.55]	0.00 [0.00, 0.76]	15.1
Myopic Greedy Index	76.84 [76.78, 76.91]	70.57 [70.49, 70.64]	68.91 [68.84, 68.97]	71.06 [70.94, 71.18]	24.40 [20.84, 28.35]	19.9
`F-UCB-Fluid` (Greedy)	**83.15 [83.11, 83.19]**	**73.16 [73.10, 73.21]**	**70.65 [70.59, 70.70]**	**74.79 [74.69, 74.89]**	**99.00 [97.68, 99.57]**	**52.5**
`Exact-DP-F-UCB`	—	**74.73 [74.64, 74.82]**	—	**76.61 [76.47, 76.74]**	**100.00 [98.12, 100.00]**	**459.8**

5.2 Quantitative findings: exact DP gap, handoff drag, and systemic fatigue

Table 4 and Figures 2–3 establish four empirical findings:

1. **Superiority over equally tuned rest-swaps and myopic greedy indexing.** Even when `specialist_rest` is tuned with the exact same proactive rest-swap threshold ($\theta_{\text{safe}} = 0.42$), its secondary Role-0 backups experience unpracticed skill decay below $\underline{s}_{\text{crit}} = 0.68$ prior to shift 40, resulting in 69.30% [69.15, 69.44] bus-factor resilience and only 7.60% [5.59, 10.26] (38/500) shock survival at $\alpha = 0.22$. Myopic Greedy Index (`whittle_greedy`) reaches 70.57% [70.49, 70.64] yield and 24.40% [20.84, 28.35] (122/500) shock survival. By combining Kalman-UCB competence tracking ($\sigma_{i,r}(t)$) with chained backup skill-floor maintenance, greedy `F-UCB-Fluid` achieves 73.16% [73.10, 73.21] net yield, 74.79% [74.69, 74.89] bus-factor resilience, and 99.00% [97.68, 99.57] (495/500) shock survival while maintaining 98.92% burnout-free shifts.
2. **Greedy heuristic vs. exact dynamic-programming bipartite optimum.** Comparing greedy `F-UCB-Fluid` against exact combinatorial assignment (`Exact-DP-F-UCB`) across $N = 200$ paired episodes at $\alpha = 0.22$ shows that the greedy bipartite solver achieves 98.62% of the exact single-step DP index sum on average, yielding 73.15% [73.07, 73.24] vs. 74.73% [74.64, 74.82] trajectory yield while reducing per-shift solver latency by 8.8× (52.5 μs vs. 459.8 μs).
3. **Handoff-penalty sensitivity and break-even regime.** Sweeping $\delta_{\text{sw}} \in \{0.012, 0.024, 0.048, 0.072, 0.110, 0$ at $\alpha = 0.22$ confirms that structured handoff capsules [6] strongly amplify rotation gains: at $\delta_{\text{sw}} = 0.024$ ($\delta_{\text{sw}} = 0.012$), `F-UCB-Fluid` yield rises to 78.02% [77.94, 78.10] (79.34% [79.26, 79.43]) with 80.41% (81.97%) bus-factor resilience. Conversely, at $\delta_{\text{sw}} = 0.150$, yield compresses to 69.07% [68.96, 69.17].

4. **Persistent cross-role systemic fatigue (κ_{sys}) and necessity of macro-sabbicals.** Sweeping the systemic fatigue coupling weight $\kappa_{\text{sys}} \in \{0.00, 0.20, 0.45, 0.70, 0.90\}$ demonstrates a critical ergonomic boundary: when fatigue is purely or predominantly role-specific ($\kappa_{\text{sys}} = 0.00$ and 0.20), `F-UCB-Fluid` achieves 78.49% and 73.14% net yield with 99.17% and 98.92% burnout-free shifts. However, when persistent cross-role exhaustion dominates ($\kappa_{\text{sys}} = 0.45, 0.70, 0.90$), net yield falls to 62.39%, 49.03%, and 40.54%, and burnout-free shifts drop to 96.25%, 44.06%, and 26.92%. Because role rotation cannot clear systemic fatigue $f_i^{\text{sys}}(t)$, high-κ_{sys} regimes require coupling shift-level rotation with periodic multi-week sabbatical cooldowns.

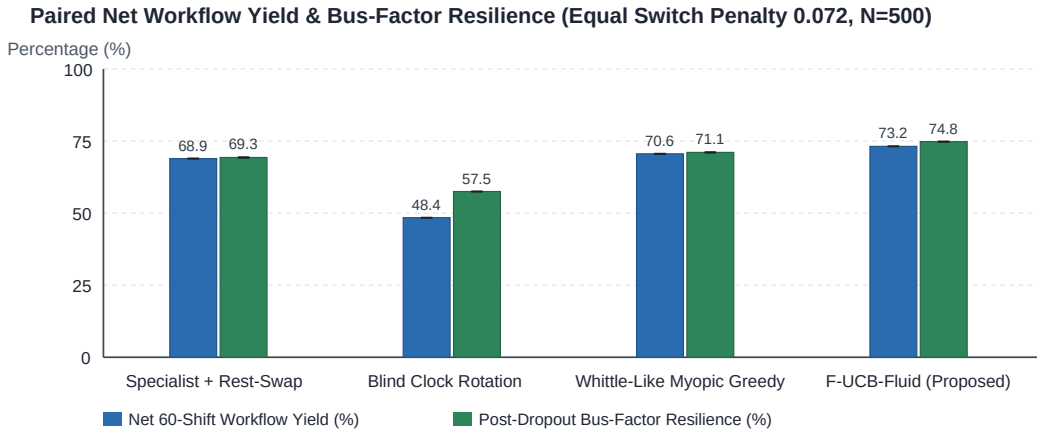


Figure 2: Paired comparison of Net Cooperative Yield and Specialist-Dropout Bus-Factor Resilience across four schedulers at $\alpha = 0.22$, $\delta_{\text{sw}} = 0.072$, and $\kappa_{\text{sys}} = 0.20$ ($N = 500$ episodes per policy; error bars denote 95% confidence intervals).

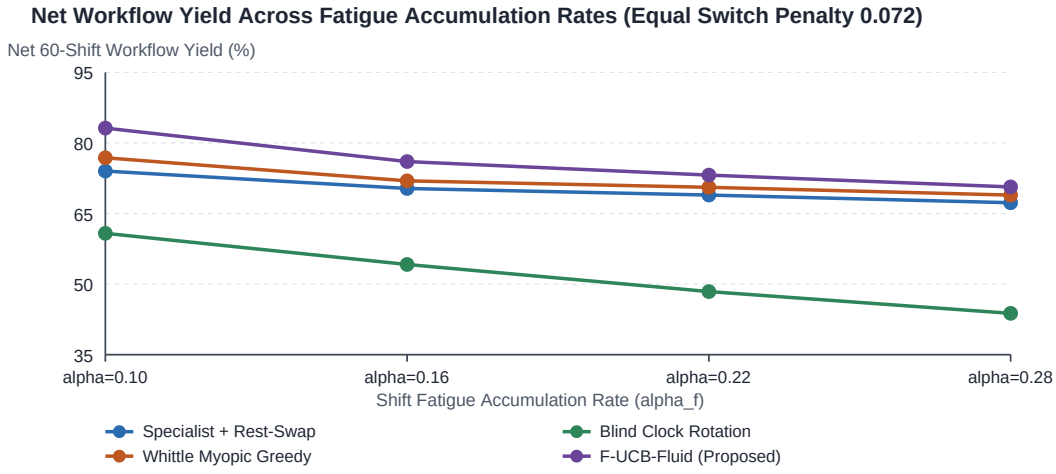


Figure 3: Scaling of Net Cooperative Yield across shift cognitive fatigue rates $\alpha \in \{0.10, 0.16, 0.22, 0.28\}$ under equalized switching penalty $\delta_{\text{sw}} = 0.072$ and systemic fatigue weight $\kappa_{\text{sys}} = 0.20$.

6 Operational boundary conditions and failure modes

We highlight four empirical boundary conditions from operations research and 2025–2026 human-AI literature where fluid division of labor can degrade collective performance:

1. **Communication overhead and context drift under frequent handoffs (Bertalanič and Fortuna [3], Rodrigues [19], Zhang et al. [25]).** Zhang et al. [25] and Bertalanič and Fortuna [3] show that inter-agent state synchronization can consume substantial token budgets and propagate correlated errors when shared context is noisy, while Rodrigues [19] demonstrate that unsynchronized handoffs induce context drift. As confirmed by our sensitivity sweep, if $\delta_{\text{sw}} \geq 0.150$, frequent rotation erodes comparative-advantage surplus.
2. **Overconfidence and miscalibrated skill priors (Eleftheriou et al. [7], Vaccaro et al. [23]).** In meta-analyses of human-AI combinations, Vaccaro et al. [23] find that human-AI teams frequently underperform the better of the human or AI alone on decision tasks when task allocation is miscalibrated, and Eleftheriou et al. [7] document systematic overconfidence among AI-assisted knowledge workers. If competence estimates $\hat{s}_{i,r}(t)$ rely on subjective confidence rather than objective verification rubrics [15], `F-UCB-Fluid` will rotate underqualified actors into critical bottleneck roles.
3. **Unsafe context carryover across security boundaries (Das and Roy [6]).** Das and Roy [6] prove that when agents switch operational contexts without strict memory sanitization, residual instructions or permissive priors from an exploratory role can compromise downstream safety-critical execution.
4. **Irreducible credential and regulatory boundaries.** In high-hazard physical or legally licensed professions, regulatory accountability prohibits rotating uncredentialed generalists into specialist slots. In such environments, skill chaining [9] must be restricted strictly within certified equivalence classes.

7 Limitations

- **Stylized parametric simulation.** While our learning, forgetting, and effort-recovery parameters are grounded in qualitative laws from operations research [9, 10, 17], the numerical values in Table 3 represent a calibrated mechanistic model rather than econometric fits to a specific live organization.
- **Single-step bipartite optimization vs. multi-step restless DP.** While we benchmark our greedy $\mathcal{O}(MR\log(MR))$ solver against exact single-step dynamic programming (`Exact-DP-F-UCB`, 98.62% per-shift ratio), neither solver computes the PSPACE-hard multi-step lookahead policy over the full horizon T.
- **Inability of task rotation to clear systemic fatigue ($\kappa_{\text{sys}} \geq 0.45$).** As shown in our κ_{sys} sweep, shift-level rotation only relieves role-specific depletion $f_{i,r}^{\text{role}}(t)$; persistent cross-role exhaustion $f_i^{\text{sys}}(t)$ requires coupling `F-UCB-Fluid` with macro-horizon mandatory sabbatical cooldowns.

8 Pre-registerable falsification protocol

Table 5 defines six pre-registerable empirical protocols designed to falsify the core claims of `F-UCB-Fluid`.

Table 5: Six pre-registerable empirical tests that could falsify the `F-UCB-Fluid` framework.

Protocol	Hypothesis	Design and primary metrics	Falsified if
P1. Multi-agent deadlock recovery	Rotating Reviewer and Implementer roles breaks verification deadlocks [4, 14]	Benchmark 5-agent crews on 200 long-horizon SWE-bench tasks comparing Static Roles vs. `F-UCB-Fluid`	`F-UCB-Fluid` fails to improve task resolution rate by ≥ 5 percentage points over static role assignment
P2. Handoff capsule drag reduction (δ_{sw})	Safe context-switching capsules [6] reduce post-rotation orientation overhead	Measure token usage and first-turn error rates across 400 role switches with vs. without schema-validated handoff capsules	Structured capsules reduce post-handoff orientation tokens or errors by $< 25\%$ relative to raw transcript handoffs
P3. Human cognitive fatigue & stress	Proactive rotation at $\theta_{\text{safe}} = 0.42$ lowers developer stress without reducing output quality	6-week crossover trial with 48 software engineers rotating across Feature Dev, Code Review, and Triage [1, 27]	Passive physiological stress or NASA-TLX fatigue shows no statistically significant reduction ($p \geq 0.05$)
P4. Chained backup skill retention	Staleness-regularized chaining prevents AI-induced skill decay [10, 22]	Unassisted post-trial domain assessments in each participant's chained backup role after 60 shifts	Mean unassisted backup-role score falls below the rework threshold $\underline{s}_{\text{crit}} = 0.68$
P5. Specialist dropout bus-factor shock	Skill-chained teams sustain $\geq 72\%$ throughput when two primary specialists drop out	Inject unannounced 10-shift specialist removal at $t = 40$ across 200 hybrid team trajectories	Post-shock throughput under `F-UCB-Fluid` fails to exceed Specialist + Rest-Swap by ≥ 4 percentage points
P6. Context sanitization on role switch	Context-reset guards [6, 19] prevent persona bleed across shifts	Audit 1,000 post-rotation agent turns for role-boundary violations after switching from Exploration to Verification	Post-switch boundary violation rate exceeds 2.5%

9 Conclusion

Neither rigid specialization nor blind periodic rotation provides a viable foundation for long-horizon human-AI collectives. By formulating division of labor as a constrained restless bandit with nonstationary Kalman competence estimation and two-component role/systemic fatigue dynamics [10, 24], we show on paired simulations ($\alpha = 0.22, \delta_{\text{sw}} = 0.072, \kappa_{\text{sys}} = 0.20$) that `F-UCB-Fluid` achieves 73.16% [73.10, 73.21] net cooperative yield, 74.79% [74.69, 74.89] bus-factor resilience, and 99.00% [97.68, 99.57] shock survival, capturing 98.62% of the single-step exact dynamic-programming optimum (`Exact-DP-F-UCB`, 74.73% yield) at 8.8× lower latency while delineating the exact handoff-drag ($\delta_{\text{sw}} \geq 0.150$) and systemic-fatigue ($\kappa_{\text{sys}} \geq 0.45$) boundaries where rotation must be paired with structured context capsules and macro-sabbatical cooldowns.

Use of AI tools

Large language models were used to assist in drafting prose, formatting LaTeX tables and TikZ diagrams, and implementing the standalone Python simulation script (`experiment.py`). Cited

references were assembled from foundational computer-science literature and 2024–2026 research publications and checked against primary metadata records; simulation outputs, confidence intervals, and LaTeX source packaging were verified using automated static consistency checks (`validate_and_package.py`).

References

[1] Selin Acan, Peter De Looff, Mitra Baratchi, Martin Dresler, and Erno Hermans. Detection of stress in naturalistic settings through passive mobile sensing. bioRxiv preprint, 2026.

[2] Lisanne Bainbridge. Ironies of automation. Automatica, 19(6):775–779, 1983. URL https://doi.org/10.1016/0005-1098(83)90046-8.

[3] Blaž Bertalanič and Carolina Fortuna. The cost of consensus: Isolated self-correction prevails over unguided homogeneous multi-agent debate. arXiv preprint arXiv:2605.00914, 2026. URL https://arxiv.org/abs/2605.00914.

[4] Mert Cemri, Melissa Z. Pan, Shuyi Yang, Lakshya A. Agrawal, Bhavya Chopra, Rishabh Tiwari, et al. Why do multi-agent LLM systems fail? In *Advances in Neural Information Processing Systems (NeurIPS), Datasets and Benchmarks Track*, 2025. URL https://arxiv.org/abs/2503.13657.

[5] Hongjiang Chen, Xin Zheng, Yixin Liu, Pengfei Jiao, Shiyuan Li, Huan Liu, et al. GoAgent: Group-of-agents communication topology generation for LLM-based multi-agent systems, 2026. URL https://arxiv.org/abs/2603.19677.

[6] Akash Das and Ishan Roy. Safe context switching for agents in the wild: Mitigating subspace interference via orthogonal adaptation. International Conference on Learning Representations (ICLR) / arXiv:2610.05219, 2026. URL https://arxiv.org/abs/2610.05219.

[7] Elena Eleftheriou, George Pallis, and Marios Constantinides. Confidence without competence in AI-assisted knowledge work. In *Proceedings of CHIWORK '26*, 2026. URL https://arxiv.org/abs/2604.09444.

[8] Netanel Eliav. The cognitive divergence: AI context windows, human attention decline, and the delegation feedback loop. arXiv preprint arXiv:2603.26707, 2026. URL https://arxiv.org/abs/2603.26707.

[9] Wallace J. Hopp, Eylem Tekin, and Mark P. Van Oyen. Benefits of skill chaining in serial production lines with cross-trained workers. Management Science, 50(1):83–98, 2004. URL https://doi.org/10.1287/mnsc.1030.0166.

[10] Mohamad Y. Jaber and Maurice Bonney. Production breaks and the learning curve: The forgetting phenomenon. Applied Mathematical Modelling, 20(2):162–169, 1996. URL https://doi.org/10.1016/0307-904X(95)00157-F.

[11] Kaizhe Jin, Adrian Rubio-Solis, Ravi Naik, Daniel Leff, James Kinross, and George Mylonas. Human-centric cognitive state recognition using physiological signals: A systematic review of machine learning strategies across application domains. *Sensors*, 25(13):4207, 2025. doi: 10.3390/s25134207.

[12] Yubin Kim, Ken Gu, Chanwoo Park, Chunjong Park, Samuel Schmidgall, A. Ali Heydari, et al. Towards a science of scaling agent systems, 2025. URL https://arxiv.org/abs/2512.08296.

[13] Nataliya Kosmyna, Eugene Hauptmann, Ye Tong Yuan, Jessica Situ, Xian-Hao Liao, Ashly Vivian Beresnitzky, Iris Braunstein, and Pattie Maes. Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task, 2025. URL https://arxiv.org/abs/2506.08872.

[14] Yuxiang Lai, Jike Zhong, Chengzhi Liu, Ming Li, Haonan Wang, Xiaofeng Yang, Liang Zhao, Youbao Tang, and Ruixi Fan. CLASHarness: Benchmarking multi-agent conflict in computer use via a controllable and verifiable generation pipeline. arXiv preprint arXiv:2610.19128, 2026. URL https://arxiv.org/abs/2610.19128.

[15] Hao Li, Zhengkun Zhang, Gangqiang Hu, Zhen Zhang, Yude Gao, Dai Dai, and Jing Liu. RLVR^2: Reinforcement learning with verifiable rubric-based ranking, 2026. URL https://arxiv.org/abs/2609.23457.

[16] Shiyuan Li, Yixin Liu, Qingsong Wen, Chengqi Zhang, and Shirui Pan. Assemble your crew: Automatic multi-agent communication topology design via autoregressive graph generation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2026. URL https://arxiv.org/abs/2507.18224.

[17] Theo F. Meijman and Gijsbertus Mulder. Psychological aspects of workload (Effort-Recovery Model). Handbook of Work and Organizational Psychology, volume 2, pages 5–33, 1998. URL https://psycnet.apa.org/record/1998-07166-001.

[18] Jaime Ortega. Job rotation as a learning mechanism. Management Science, 47(10):1361–1370, 2001. URL https://doi.org/10.1287/mnsc.47.10.1361.10257.

[19] Carson Rodrigues. Hallucination as context drift: Synchronization protocols for multi-agent LLM systems. arXiv preprint arXiv:2606.21666, 2026. URL https://arxiv.org/abs/2606.21666.

[20] Richard M. Ryan and Edward L. Deci. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1):68–78, 2000.

[21] Aarti Sathyanarayana, Ohida Binte Amin, Jennie An, and Jukka Pekka Onnela. Examining the use of consumer wearable devices and digital tools for stress measurement in college students: Scoping review of methods. *JMIR mHealth and uHealth*, 14:e64144, 2026. doi: 10.2196/64144.

[22] Judy Hanwen Shen and Alex Tamkin. How AI impacts skill formation, 2026. URL https://arxiv.org/abs/2601.20245.

[23] Michelle Vaccaro, Abdullah Almaatouq, and Thomas Malone. When combinations of humans and AI are useful: A systematic review and meta-analysis. Nature Human Behaviour, 8:2293–2303, 2024. URL https://arxiv.org/abs/2405.06087.

[24] Peter Whittle. Restless bandits: Activity allocation in a changing world. Journal of Applied Probability, 25(A):287–298, 1988. URL https://doi.org/10.2307/3214163.

[25] Guibin Zhang, Yanwei Yue, Zhixun Li, Sukwon Yun, Guancheng Wan, Kun Wang, et al. Cut the crap: An economical communication pipeline for LLM-based multi-agent systems. In

International Conference on Learning Representations (ICLR), 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/bbc461518c59a2a8d64e70e2c38c4a0e-Abstract-Conference.html.

[26] Guibin Zhang, Yanwei Yue, Xiangguo Sun, Guancheng Wan, Miao Yu, Junfeng Fang, et al. G-Designer: Architecting multi-agent communication topologies via graph neural networks. In *International Conference on Machine Learning (ICML)*, 2025. URL https://icml.cc/virtual/2025/poster/45567.

[27] Jiayin Zhi, Harsh Kumar, and Mina Lee. Investigating the effects of LLM use on critical thinking under time constraints: Access timing and time availability. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 2026. doi: 10.1145/3772318.3791796.