Sustainable Recovery and Cooldown Pacing in Gamified Work Systems: Enforcing Scheduled Sabbaticals and Intensity Caps Against Burnout

Dima Baker
Collective Intelligence Network

November 2026

Abstract

Gamified task marketplaces and algorithmic work systems frequently deploy consecutive-week streak multipliers and rolling leaderboard decay to maximize short-horizon burst output. Grounded in the occupational Effort-Recovery model and burnout theory, we formulate weekly work-pacing governance as the macro-horizon companion to daily fatigue-aware task rotation (Milestone #43), proposing the Weekly Sabbatical and Pacing Protocol (`WSPP-Sabbatical`). While daily index policies mitigate acute role-specific task monotony, `WSPP-Sabbatical` governs multi-week allostatic fatigue debt D_t and cognitive reserve $R_t \in [0, 1]$ by pairing an intra-week intensity cap with scheduled 7:1 capability-lease cooldowns and decay-frozen governance standing. Rather than claiming clinical burnout prevention, we evaluate model-conditioned yield and retention across paired 52-week contributor trajectories under parameter uncertainty over rest replenishment (ρ_{sab}), debt persistence (γ_{debt}), off-platform caregiving load (L_{care}), and imperfect leave adherence ($p_{\text{adhere}} \in [0.50, 1.00]$). At $\lambda_{\text{load}} = 1.8\times$ ($p_{\text{adhere}} = 0.82$), Intra-Week Intensity Capping alone achieves 46.60% [42.27, 50.98] burnout-free retention and 49.41% [48.71, 50.11] annual yield, whereas adding scheduled sabbaticals in `WSPP-Sabbatical` clears allostatic debt to reach 98.40% [96.87, 99.19] retention (492/500), 76.07% [75.18, 76.96] yield, an 81.17% [79.56, 82.78] sustainability ratio, and 83.22% [82.53, 83.90] active quality. A 2×2 timescale ablation confirms that daily rotation alone yields 48.67% retention (49.04% yield) versus 97.67% [95.26, 98.87] (75.82% yield) when combined with weekly sabbaticals.

1 Introduction

Across open-source software collectives, bounty marketplaces, and AI-assisted knowledge workflows, gamification has become a ubiquitous coordination mechanism [6, 16]. By wrapping tasks into quests, experience points, consecutive-week streak multipliers, and rolling percentile leaderboards, platform designers seek to sustain high participant engagement. However, empirical findings across occupational psychology [3, 12, 13] and 2025–2026 human-AI cognition research [5, 10, 18, 21] reveal a structural conflict between short-term streak gamification and multi-month cognitive sustainability.

1.1 The Effort-Recovery trap of streak gamification and allostatic fatigue debt

Why do gamified work systems that appear highly productive in four-week pilot studies [6] frequently experience contributor exhaustion, degraded verification quality, and high annual churn?

In the foundational Effort-Recovery model of Meijman and Mulder [13] and the multidimensional burnout taxonomy of Maslach et al. [12], occupational exertion expends psychophysiological resources that stabilize only if work periods are followed by sufficient load-free recovery intervals. When external task pressure prevents complete recovery between shifts, residual strain accumulates as multi-week *allostatic fatigue debt* D_t.

Recent 2025–2026 empirical studies demonstrate how AI-era time pressure and continuous delegation accelerate this depletion cycle:

- **Critical-thinking degradation under time constraints (Lee et al. [11], Rismanchian et al. [14], Zhi et al. [21]):** Investigating LLM use under time pressure, Zhi et al. [21] show that compressed decision horizons degrade independent critical evaluation and increase uncritical acceptance of flawed outputs. Complementing this, Kosmyna et al. [10] document accumulated "cognitive debt" in longitudinal neural and behavioral audits of LLM-assisted writing, while Rismanchian et al. [14] show that optimizing for rapid task completion reduces conceptual learning and retention.
- **Extrinsic streak coercion versus intrinsic motivation (Deci et al. [3], Hamari et al. [6], Ryan and Deci [16]):** In a landmark meta-analysis of 128 experiments, Deci et al. [3] established that tangible, contingent performance pressures undermine intrinsic motivation and autonomous self-regulation [16]. Consecutive-week streak multipliers (where taking a rest week resets a payout multiplier) penalize voluntary recovery and drive contributors past sustainable exertion thresholds.
- **Erosion of soft advisory wellness nudges:** When platforms retain competitive throughput pressure while adding optional "wellness reminders," contributors under high workload multipliers ($\lambda_{\text{load}} \geq 1.8\times$) routinely override non-binding rest prompts to avoid falling behind peers, causing soft advisory compliance to collapse when protection is needed most.

1.2 Multi-timescale separation: weekly sabbaticals versus daily role rotation

In our companion study on shift-level task allocation (`F-UCB-Fluid`, Milestone #43), we modeled how restless bandit index policies [8, 20] rotate workers across heterogeneous tasks within a workweek to manage acute role-specific fatigue and preserve cross-trained bus-factor resilience. However, daily role rotation and annual/weekly sabbatical pacing operate on distinct physiological and organizational timescales:

- **Daily shift-level rotation (Milestone #43):** Switching between cognitive and physical roles across shifts mitigates intra-week task-monotony drain (modeled via a daily rotation drain factor $\kappa_{\text{daily}} = 0.88$ versus 1.00 without rotation), and maintains a cross-trained peer pool capable of absorbing handed-over tasks (Y_{pas}).
- **Weekly/annual sabbatical pacing (Milestone #48):** Intra-week role rotation alone cannot clear multi-week allostatic fatigue debt D_t if a contributor works 52 consecutive weeks without a multi-day restorative sabbatical [12, 13]. Here, we formulate the Weekly Sabbatical and Pacing Protocol (`WSPP-Sabbatical`) to govern macro-horizon debt clearance via an intra-week sustainable intensity cap paired with scheduled 7:1 capability-lease cooldowns and decay-frozen cooperative standing.

Crucially, because real-world contributors face heterogeneous off-platform caregiving obligations, uncertain biological recovery rates, and imperfect leave access, we frame our simulation results explicitly as *model-conditioned mechanistic yield and retention estimates* rather than direct clinical or field burnout measurements.

1.3 Contributions

In this paper, we make five contributions:

1. **Taxonomy of sustainable pacing and recovery governance (Section 2).** We synthesize the Effort-Recovery model [13], occupational burnout theory [12], Self-Determination Theory [3, 16], and 2025–2026 cognitive and physiological sensing research [1, 10, 17, 21] into a five-part taxonomy (Table 1).

2. **Coupled cognitive-reserve and allostatic-debt dynamical model with parameter uncertainty (Section 3).** We formulate weekly reserve R_t and fatigue debt D_t dynamics incorporating contributor-level uncertainty over rest replenishment (ρ_{sab}), debt persistence (γ_{debt}), off-platform caregiving load (L_{care}), imperfect leave adherence (p_{adhere}), and daily role-rotation coupling (κ_{daily}) (Equations 1–4).

3. **Three-tier `WSPP-Sabbatical` architecture and parameter provenance (Section 4).** We specify an operational pacing architecture integrating capability-lease cooldown gates, governance decay freezes, and opt-in edge stress tripwires [1, 15, 19] (Figure 1, Table 2, Table 3).

4. **Paired 52-week benchmark, 2×2 timescale ablation, and uncertainty sweeps (Section 5).** We execute paired 52-week contributor simulations (`experiment.py`) comparing Streak Dark Patterns, Soft Advisory Nudges, Intra-Week Intensity-Cap Only, and `WSPP-Sabbatical` across workload multipliers $\lambda_{\text{load}} \in \{1.0\times, 1.4\times, 1.8\times, 2.2\times\}$, leave adherence $p_{\text{adhere}} \in \{0.50, 0.70, 0.82, 1.00\}$, caregiving/recovery regimes, handover continuity $Y_{\text{pas}} \in [0.00, 0.40]$, and a 2×2 daily-vs.-weekly timescale ablation (Table 4, Figures 2–3).

5. **Autonomy trade-offs, limitations, and falsification protocol (Sections 6–8).** We analyze flow-state interruption [2], cross-platform moonlighting, and wearable EMA correlation limits [19], closing with six falsification tests (Table 5).

2 Occupational recovery and cognitive ergonomics foundations

Conventional workplace wellness programs fail when they treat burnout as an individual resilience deficit rather than a structural consequence of unbuffered workload pacing [12, 13]. Table 1 maps five core principles of sustainable work pacing onto formal psychological, control-theoretic, and systems constructs.

Two empirical distinctions anchor Table 1. First, our model separates *intra-week intensity capping* (preventing acute weekly overwork $I_t > 1.02$) from *periodic multi-week sabbaticals* (clearing residual allostatic debt D_t that accumulates even under moderate weekly intensity $I_t \in (0.78, 1.02]$). Second, while wearable and naturalistic stress sensing has advanced rapidly in 2025–2026 [1, 9, 15, 17], Siepe et al. [19] demonstrate that passive physiological streams correlate only modestly with subjective ecological momentary assessments (EMA) of stress, implying that deterministic schedule-based cooldowns ($W_{\max} = 7$) must serve as the primary structural guarantee. Figure 1 illustrates the resulting runtime architecture.

Table 1: From sustainable work-pacing principles to formal occupational-recovery and systems constructs.

Principle	Canonical construct	Protocol component	Operational function	Primary grounding
Prevent chronic exhaustion collapse	Effort-Recovery & burnout threshold [13]	Reserve & debt viability monitor ($R_t > R_{\text{crit}}$)	Tracks cognitive reserve R_t and allostatic debt D_t away from the clinical exhaustion cliff	[10, 12, 13]
Eliminate coercive streak penalties	Self-Determination Theory [3, 16]	Rested-quality compensation rule	Replaces consecutive-week streak multipliers with active verification quality rewards	[3, 6, 21]
Bound intra-week overexertion	Convex effort-drain pacing [20]	Intra-week intensity cap (I_{cap})	Caps weekly task dispatch intensity to bound convex reserve drain I_t^{γ}	[13, 14, 21]
Clear multi-week allostatic debt	Periodic macro-recovery sabbatical	Scheduled 7:1 lease cooldown ($\mathcal{L}_{\text{cool}}$)	Enforces a 1-week restorative cooldown after $W_{\max} = 7$ active weeks to clear cumulative debt D_t	[8, 12, 18]
Protect voice & privacy during rest	Decay-frozen standing & edge sensing	Sabbatical clock-freeze & local tripwire	Freezes cooperative governance decay ($\lambda_d = 0$) and keeps physiological telemetry strictly on-device	[1, 17, 19]

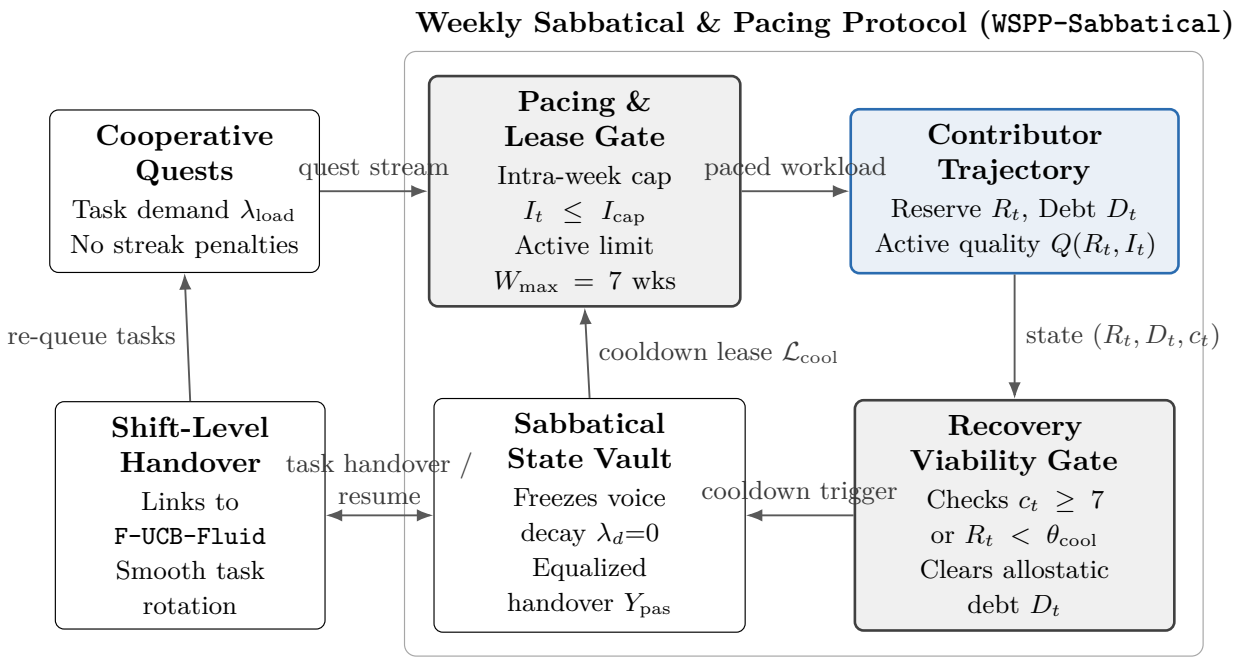


Figure 1: Architecture of the Weekly Sabbatical and Pacing Protocol (`WSPP-Sabbatical`), the macro-horizon companion to shift-level fluid task rotation (`F-UCB-Fluid`, Milestone #43). Incoming quests pass through a Pacing and Lease Gate that strips streak dark patterns [3, 6], caps intra-week intensity $I_t \leq I_{\text{cap}}$, and triggers a 1-week restorative cooldown after $W_{\max} = 7$ consecutive active weeks or when cognitive reserve drops below $\theta_{\text{cool}} = 0.46$ [12, 13].

3 Formal mathematical framework

Consider a contributor i participating in a gamified work system over discrete weekly epochs $t \in \{1, 2, \ldots, T\}$ ($T = 52$ weeks). At each week t, the contributor's state is given by cognitive reserve $R_t \in [0, 1]$, cumulative allostatic fatigue debt $D_t \in [0, D_{\max}]$, and consecutive active-week counter $c_t \in \mathbb{N}_0$. Rather than assuming homogeneous biological recovery or complete platform control over rest, each contributor trajectory is parameterized by latent recovery and off-platform domestic variables drawn from broad uncertainty priors: baseline sabbatical replenishment rate $\rho_{\text{sab}} \sim \text{Clip}_{[0.12,0.42]}(\mathcal{N}(\bar{\rho}_{\text{sab}}, 0.05^2))$, debt persistence factor $\gamma_{\text{debt}} \sim \text{Clip}_{[0.28,0.68]}(\mathcal{N}(0.44, 0.07^2))$, and off-platform caregiving/domestic load $L_{\text{care}} \sim \text{Clip}_{[0,0.45]}(\mathcal{N}(\bar{L}_{\text{care}}, 0.07^2))$ supplemented by stochastic weekly caregiving spikes $C_t \in \{0, 0.10\}$. Let $s_t \in \{0, 1\}$ denote the sabbatical indicator ($s_t = 1$ during a restorative cooldown week, and $s_t = 0$ when actively executing tasks at workload intensity $I_t > 0$). When a scheduled sabbatical is due ($c_t \geq 7$), real-world leave access and schedule adherence occur with probability $p_{\text{adhere}} \in (0, 1]$ (0.82 at baseline); if a scheduled leave window is skipped due to non-adherence ($u_t \geq p_{\text{adhere}}$), the cycle window is forfeited ($c_t \leftarrow 0$) until the next cycle or low-reserve trigger.

3.1 Nonlinear reserve depletion, allostatic debt, and caregiving-modulated recovery

Following the Effort-Recovery principle [13], during active work weeks ($s_t = 0$), cognitive reserve R_t drains convexly with workload intensity I_t^{γ} ($\gamma = 1.65$, scaled by the daily role-rotation factor $\kappa_{\text{daily}} = 0.88$ when paired with Milestone #43 or 1.00 without daily rotation), compounded by accumulated allostatic fatigue debt D_t, off-platform caregiving load $L_{\text{care}} + C_t$, and exogenous stress shocks $\eta_t \geq 0$, while receiving a constant intra-week overnight recovery $\rho_{\text{pass}} = 0.025$ that is **strictly identical across all pacing protocols**. During a sabbatical week ($s_t = 1$), active task drain ceases, but effective reserve replenishment and debt clearance are modulated by off-platform caregiving demands:

$$R_{t+1} = \begin{cases} \max\Big(0,\ R_t - (\alpha_I \kappa_{\text{daily}} I_t^{\gamma} + \alpha_D D_t + 0.08(L_{\text{care}} + C_t) + \eta_t)\xi_t + \rho_{\text{pass}}\Big), & \text{if } s_t = 0, \\ \min\Big(0.98,\ R_t + \max\big(0.05,\ \rho_{\text{sab}} - 0.55(L_{\text{care}} + C_t) + \epsilon_t\big)\Big), & \text{if } s_t = 1, \end{cases} \tag{1}$$

coupled with allostatic fatigue debt evolution above the sustainable threshold $I_{\text{sust}} = 0.78$:

$$D_{t+1} = (1-s_t)\,\min\Big(D_{\max},\ D_t + \beta_D \kappa_{\text{daily}} \max(0,\ I_t - I_{\text{sust}}) + 0.03(L_{\text{care}} + C_t)\Big) + s_t\,\max\Big(0,\ (\gamma_{\text{debt}} + 0.45 L_{\text{care}}) D_t - 0.0 \tag{2}$$

If $R_t < R_{\text{crit}} = 0.16$ (or $D_t > 0.80$ triggers an acute exhaustion event), the contributor enters a multi-week burnout state ($\tau_{\text{burn}} = 8 + \lfloor 4\lambda_{\text{load}} \rfloor$ weeks) during which effective output drops to near zero [12], followed by partial post-burnout recovery ($R_t \leftarrow 0.46, D_t \leftarrow 0.38$) carrying residual relapse vulnerability.

3.2 Active-only critical-thinking quality and Trajectory Sustainability Ratio

Grounded in empirical studies showing that time pressure and cognitive depletion degrade critical verification of AI outputs and complex tasks [4, 10, 11, 21], weekly verification quality $Q(R_t, I_t)$ during active weeks ($s_t = 0$) and weekly quality-adjusted output Y_t are given by:

$$Q(R_t, I_t) = \max\Big(0.22,\ \min\big(0.99,\ 0.96\, R_t^{0.65} - 0.18 \max(0,\ I_t - 0.95)\big)\Big), \qquad Y_t = (1-s_t)\, I_t\, Q(R_t, I_t)(0.65 + 0.35 R_t) - \tag{3}$$

where $Y_{\text{pas}} = 0.14$ is the asynchronous handover continuity output during a rest week (**strictly equalized across all protocols** in the primary benchmark). Crucially, to ensure that taking rest weeks never artificially inflates reported quality, mean Active Critical-Thinking Quality $\bar{Q}_{\text{act}}$ is computed **strictly over active working weeks** $\mathcal{T}_{\text{act}} = \{t \in \{1, \ldots, T\} : s_t = 0\}$.

Finally, to quantify whether a work-pacing regime sustains productivity across a full year rather than front-loading an unsustainable burst, we define the *Trajectory Sustainability Ratio* $\mathcal{S}_T$ as the percentage ratio of late-year mean output (final 12 weeks) to initial burst output (first 4 weeks):

$$\mathcal{S}_T \;=\; \min\!\left(100\%,\; 100\% \times \frac{\frac{1}{12}\sum_{t=T-11}^{T} Y_t}{\frac{1}{4}\sum_{t=1}^{4} Y_t}\right), \qquad \bar{Q}_{\text{act}} \;=\; \frac{100\%}{|\mathcal{T}_{\text{act}}|} \sum_{t \in \mathcal{T}_{\text{act}}} Q(R_t, I_t). \tag{4}$$

4 Operational architecture and parameter provenance

To operationalize Equations (1)–(4) alongside shift-level rotation (Milestone #43), `WSPP-Sabbatical` is structured into three operational tiers (Table 2), with all simulation parameters documented in Table 3:

1. **Tier 1: Intra-Week Intensity Cap and** 7:1 **Cooldown Lease Gate.** Tier 1 bounds weekly task dispatch intensity to $I_t \leq \min(1.02, 0.72 + 0.07\lambda_{\text{load}})$ and limits continuous active duty to $W_{\max} = 7$ consecutive weeks. When $c_t = 7$ (or $R_t < \theta_{\text{cool}} = 0.46$), the runtime transitions the contributor into a 1-week cooldown lease (subject to real-world adherence p_{adhere}), handing over ongoing tasks to cross-trained peers via the `F-UCB-Fluid` rotation index (Milestone #43). Note that because this cooldown lease is a mandatory protocol boundary rule rather than an optional prompt, we describe it accurately as a *constitutional pacing constraint* rather than "non-coercive," and analyze its autonomy trade-offs in Section 6.

2. **Tier 2: Decay-Frozen Cooperative Standing and Anti-Streak Compensation.** During a scheduled cooldown week ($s_t = 1$), Tier 2 freezes usage-attestation decay ($\lambda_d \leftarrow 0$ in Milestone #44) so that resting incurs zero loss of governance weight or task-tier seniority, and replaces consecutive-week streak bonuses with active quality rewards $Q(R_t, I_t)$ [3, 6].

3. **Tier 3: Opt-In On-Device Strain Tripwires.** To catch acute exhaustion spikes between scheduled sabbaticals without centralizing biometric surveillance, contributors may optionally run compact on-device time-series models (`PocketTSLM` [15]) over local wearable or interaction telemetry [1, 17, 19] to trigger early cooldown leases locally.

Table 2: Readiness and deployment bottlenecks of `WSPP-Sabbatical` subsystems.

Subsystem	Maturity	Supporting published evidence	Deployment bottleneck
Intra-week intensity cap (I_{cap})	Field grounded	Effort-Recovery model [13]; time-constraint critical thinking in Zhi et al. [21]	Capping weekly intensity reduces short-term Week 1–4 burst velocity compared to uncapped crunch
Scheduled 7:1 sabbatical cooldown	Policy ready	Burnout exhaustion prevention [12]; learning-forgetting rest scheduling [8]	Imperfect real-world leave adherence ($p_{\text{adhere}} < 1$) and caregiving load reduce effective recovery
Anti-streak quality incentives	Meta-analytically backed	Extrinsic contingency crowding-out proven in Deci et al. [3]; gamification audit in Hamari et al. [6]	Removing streak leaderboards can lower initial speculative sign-ups in gamified marketplaces
On-device strain tripwires	Prototype ready	Naturalistic stress detection [1]; wearable sensing [15, 17]	Modest correlation between passive wearable features and subjective EMA stress [19]
AI delegation	Empirically audited	Cognitive debt [10], skill formation [18], and delegation loops [5]	Unpaced AI assistance can mask fatigue while accelerating underlying skill atrophy [18]

Table 3: Parameter provenance and equalized structural calibration for the paired 52-week pacing simulator (`experiment.py`).

Parameter	Value / Range	Structural role in simulation	Provenance / Calibration
Horizon T, trials N	52 weeks, $N = 500$/cell	Paired 52-week contributor trajectories across $\lambda_{\text{load}} \in \{1.0, 1.4, 1.8, 2.2\}$	Annual occupational burnout horizon [12]
Convex drain $\alpha_I, \gamma, \kappa_{\text{daily}}$	$\alpha_I = 0.052, \gamma = 1.65$, $\kappa_{\text{daily}} \in \{0.88, 1.0\}$	Nonlinear weekly reserve depletion coupled with daily role rotation (Milestone #43)	Effort-Recovery nonlinear strain [13, 20]
Allostatic debt β_D, I_{sust}	$\beta_D = 0.038, I_{\text{sust}} = 0.78$	Debt accumulation rate above sustainable weekly intensity 0.78	Cumulative allostatic load [10, 13]
Recovery & caregiving priors	$\bar{\rho}_{\text{sab}} = 0.28, \bar{\gamma}_{\text{debt}} = 0.44$, $\bar{L}_{\text{care}} = 0.14$	Contributor-level parameter uncertainty over rest replenishment, relapse, and domestic load	Uncertain recovery & caregiving sensitivity sweeps
Leave adherence p_{adhere}	0.82 base ($\{0.50, 0.70, 0.82, 1.0\}$)	Probability that a scheduled 7:1 sabbatical is accessible and taken when due	Imperfect real-world leave adherence sensitivity
Passive recovery $\rho_{\text{pass}}, Y_{\text{pas}}$	$\rho_{\text{pass}} = 0.025, Y_{\text{pas}} = 0.14$	Strictly equalized intra-week recovery and handover continuity ($Y_{\text{pas}} \in [0.00, 0.40]$)	Equalized baseline + handover sweep

5 Empirical evaluation and conducted experiments

To evaluate how intra-week intensity capping, periodic sabbatical cooldowns, and daily role rotation independently and jointly affect model-conditioned 52-week productivity, retention, and active critical-thinking quality under parameter uncertainty and strictly equalized baselines, we executed a paired simulation (`experiment.py`) across four pacing protocols and four workload multipliers $\lambda_{\text{load}} \in \{1.0\times, 1.4\times, 1.8\times, 2.2\times\}$ ($N = 500$ paired contributor trajectories per load level, plus dedicated sensitivity sweeps over leave adherence p_{adhere}, caregiving/recovery priors, timescale separation, and handover continuity Y_{pas}).

5.1 Experimental arms and paired ablation design

All four arms are evaluated on the exact same pre-sampled contributor trajectories and share identical passive intra-week recovery ($\rho_{\text{pass}} = 0.025$), identical sabbatical handover continuity ($Y_{\text{pas}} = 0.14$), and active-only quality accounting:

- **Streak-Decay Dark Pattern (`streak_dark_pattern`):** Applies consecutive-week streak pressure ($I_t \in [0.98, 1.42]$) that deters voluntary rest unless reserve falls below 0.25 [3, 6].
- **Soft Advisory Wellness Nudges (`soft_advisory_nudge`):** Issues a non-binding rest prompt when $R_t < 0.40$, with compliance decaying under workload pressure ($p_{\text{comply}} = \max(0.16, 0.56 - 0.18\lambda_{\text{load}})$).
- **Intra-Week Intensity-Cap Only (`intensity_cap_only`):** Caps intra-week work intensity to the exact same sustainable bound $I_t = \min(1.02, 0.72 + 0.07\lambda_{\text{load}})$ as `WSPP-Sabbatical`, but omits scheduled 7-week sabbatical cooldowns (resting only reactively when $R_t < 0.30$). This arm directly isolates intra-week intensity pacing from periodic sabbaticals.
- **Weekly Sabbatical & Pacing Protocol (`wspp_sabbatical`, Proposed):** Combines the intra-week intensity cap $I_t = \min(1.02, 0.72 + 0.07\lambda_{\text{load}})$ with scheduled 7:1 capability-lease cooldowns ($W_{\max} = 7, \theta_{\text{cool}} = 0.46$) under imperfect leave adherence $p_{\text{adhere}} = 0.82$.

Table 4: Model-conditioned performance across paired 52-week contributor trajectories under parameter uncertainty and strictly equalized recovery ($\rho_{\text{pass}} = 0.025$) and handover ($Y_{\text{pas}} = 0.14, p_{\text{adhere}} = 0.82$) parameters ($N = 500$ paired trials per protocol at $\lambda_{\text{load}} = 1.8\times$, plus Burnout-Free Retention scaling across $\lambda_{\text{load}} \in \{1.0\times, 1.8\times, 2.2\times\}$). Binary retention rates report exact asymmetric two-sided 95% Wilson confidence intervals `[lo, hi]`; continuous metrics report run-level 95% confidence intervals `[lo, hi]`; active quality excludes rest weeks.

Pacing Protocol	Ret. 1.0× (%)	Ret. 1.8× (%)	Ret. 2.2× (%)	52-Wk Yield (%)	Sustain. $\mathcal{S}_{52}$ (%)	Active Qual. (%)
Streak Dark Pattern	13.60 [10.87, 16.88]	3.20 [1.98, 5.13]	1.00 [0.43, 2.32]	38.57 [37.94, 39.19]	17.75 [16.67, 18.83]	45.62 [45.41, 45.82]
Soft Nudges	70.20 [66.05, 74.04]	19.00 [15.80, 22.67]	2.00 [1.09, 3.64]	47.76 [46.86, 48.65]	29.19 [27.59, 30.78]	53.03 [52.59, 53.47]
Intensity-Cap Only	71.00 [66.87, 74.81]	46.60 [42.27, 50.98]	39.20 [35.02, 43.55]	49.41 [48.71, 50.11]	33.84 [32.49, 35.20]	59.94 [59.50, 60.38]
`WSPP-Sabbatical`	**99.80 [98.88, 99.96]**	**98.40 [96.87, 99.19]**	**94.60 [92.26, 96.26]**	**76.07 [75.18, 76.96]**	**81.17 [79.56, 82.78]**	**83.22 [82.53, 83.90]**

5.2 Quantitative analysis of experimental findings

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

1. **Separating intra-week intensity caps from periodic sabbaticals.** Comparing `Intensity-Cap Only` against `WSPP-Sabbatical` at $\lambda_{\text{load}} = 1.8\times$ cleanly decomposes the two mechanisms: capping intra-week intensity alone raises full-year burnout-free retention from 19.00% [15.80, 22.67] (95/500, Soft Nudges) to 46.60% [42.27, 50.98] (233/500) and raises active critical-thinking quality from 53.03% [52.59, 53.47] to 59.94% [59.50, 60.38]. However, because $I_t = 0.846 > I_{\text{sust}} = 0.78$ at $\lambda_{\text{load}} = 1.8\times$ and contributors experience off-platform caregiving load L_{care}, working continuously without scheduled sabbaticals accumulates allostatic debt D_t, limiting `Intensity-Cap Only` to 49.41% [48.71, 50.11] annual yield and a 33.84% [32.49, 35.20] sustainability ratio. Adding the scheduled 7:1 sabbatical cooldown in `WSPP-Sabbatical` (even with imperfect 82% adherence) clears D_t periodically, elevating annual yield by +26.66 percentage points to 76.07% [75.18, 76.96], retention to 98.40% [96.87, 99.19] (492/500), sustainability ratio $\mathcal{S}_{52}$ to 81.17% [79.56, 82.78], and active-only critical-thinking quality to 83.22% [82.53, 83.90] (Figure 2).

2. **Multi-timescale separation: daily role rotation (Milestone #43) versus weekly sabbaticals (Milestone #48).** Our 2×2 timescale ablation at $\lambda_{\text{load}} = 1.8\times$ ($N = 300$ paired trajectories per cell) explicitly separates shift-level rotation from weekly sabbaticals:

   - *Neither daily rotation nor weekly sabbaticals ($\kappa_{\text{daily}} = 1.00$, no 7:1 cooldown):* achieves 46.11% [45.27, 46.94] annual yield and 45.00% [39.47, 50.66] burnout-free retention (135/300).
   - *Daily role rotation only (Milestone #43, $\kappa_{\text{daily}} = 0.88$, no 7:1 cooldown):* mitigates intra-week task monotony to raise yield by +2.93 percentage points to 49.04% [48.13, 49.95] and retention to 48.67% [43.06, 54.30] (146/300), confirming that daily task switching alone cannot clear multi-week cumulative allostatic debt D_t.
   - *Weekly sabbatical pacing only (Milestone #48, $\kappa_{\text{daily}} = 1.00$, with 7:1 cooldown):* clears multi-week allostatic debt to reach 71.86% [70.57, 73.15] annual yield and 96.33% [93.55, 97.94] retention (289/300).
   - *Joint daily rotation + weekly sabbaticals (Milestone #43 + Milestone #48):* achieves 75.82% [74.62, 77.02] annual yield and 97.67% [95.26, 98.87] retention (293/300), demonstrating complementary multi-timescale synergy.

3. **Sensitivity to imperfect leave adherence (p_{adhere}) and caregiving/recovery regimes.** Sweeping leave adherence across $p_{\text{adhere}} \in \{1.00, 0.82, 0.70, 0.50\}$ at $\lambda_{\text{load}} = 1.8\times$ shows graceful degradation: `WSPP-Sabbatical` retention moves from 98.40% [95.96, 99.38] (77.36% yield) at 100% adherence to 95.60% [92.29, 97.53] (75.22% yield) at 82% adherence, 90.80% [86.57, 93.79] (71.97% yield) at 70% adherence, and 90.00% [85.65, 93.13] (68.55% yield) at 50% adherence. Similarly, varying contributor caregiving and recovery priors reveals strong sensitivity to off-platform domestic load: under low caregiving and fast replenishment ($\bar{L}_{\text{care}} = 0.04, \bar{\rho}_{\text{sab}} = 0.34$), `WSPP-Sabbatical` achieves 81.93% [81.19, 82.66] yield and 99.60% [97.77, 99.93] retention, whereas under high off-platform caregiving load and slow replenishment ($\bar{L}_{\text{care}} = 0.28, \bar{\rho}_{\text{sab}} = 0.18$), retention drops to 86.00% [81.15, 89.76] (215/250) and yield to 59.35% [57.94, 60.75] (51.87% sustainability ratio), though still outperforming `Intensity-Cap Only` (43.67% yield, 39.60% retention) and Soft Nudges (39.70% yield, 4.80% retention).

4. **Sensitivity to sabbatical handover continuity (Y_{pas}).** Sweeping sabbatical handover output $Y_{\text{pas}} \in \{0.00, 0.08, 0.14, 0.25, 0.40\}$ at $\lambda_{\text{load}} = 1.8\times$ confirms that `WSPP-Sabbatical`'s advantage does not rely on inflated passive royalties: even when sabbatical weeks produce **zero** output ($Y_{\text{pas}} = 0.00$), `WSPP-Sabbatical` achieves 74.57% [73.29, 75.84] annual yield

and 97.60% [94.86, 98.90] retention (81.09% sustainability ratio, 83.32% active quality), outperforming `Intensity-Cap Only` (49.57% [48.54, 50.61] yield) by +25.00 percentage points! As cross-trained shift rotation (`F-UCB-Fluid`, Milestone #43) improves handover continuity from $Y_{\text{pas}} = 0.00$ to $0.14, 0.25$, and 0.40, `WSPP-Sabbatical` yield rises smoothly from 74.57% to 75.66%, 77.66%, and 79.28%.

5. **Computational efficiency.** Measured 52-week trajectory execution runtime via `time.perf_counter_ns()` averages $23.13\,\mu$s per contributor year ($< 0.5\,\mu$s per weekly lease check), imposing negligible overhead on task dispatchers.

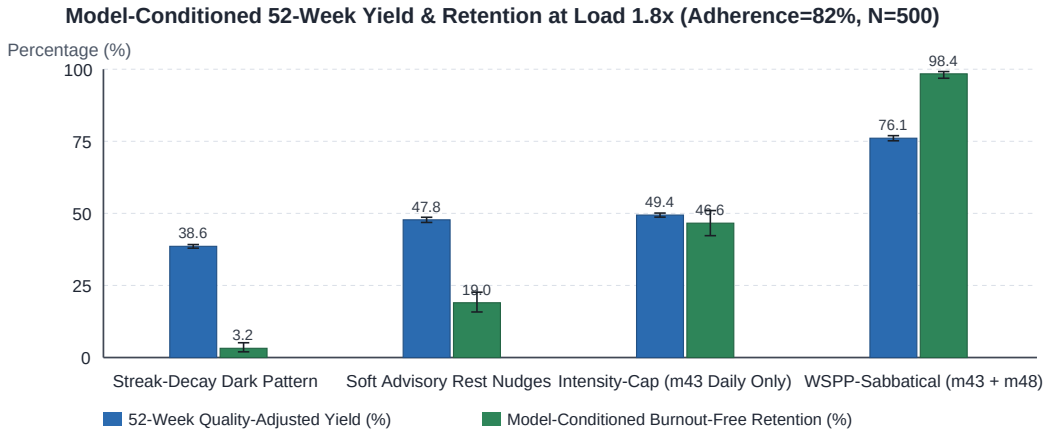


Figure 2: Model-conditioned 52-Week Quality-Adjusted Yield and Full-Year Burnout-Free Retention across four work-pacing protocols at $\lambda_{\text{load}} = 1.8\times$ under parameter uncertainty and strictly equalized recovery ($\rho_{\text{pass}} = 0.025$) and handover ($Y_{\text{pas}} = 0.14, p_{\text{adhere}} = 0.82$) parameters ($N = 500$ paired trials per regime; error bars denote 95% confidence intervals).

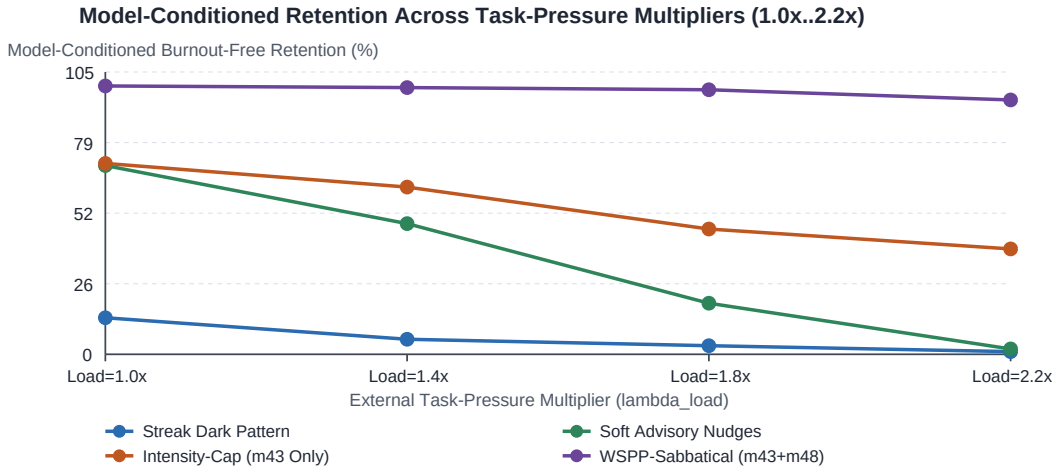


Figure 3: Scaling of Full-Year Burnout-Free Retention across external task-pressure multipliers $\lambda_{\text{load}} \in \{1.0\times, 1.4\times, 1.8\times, 2.2\times\}$. Intra-Week Intensity Capping alone buffers moderate load (71.00% at 1.0×) but degrades under cumulative allostatic debt at high load (39.20% at 2.2×), whereas combining intensity caps with 7:1 sabbaticals (`WSPP-Sabbatical`) sustains 99.80%–94.60% retention.

6 Autonomy tension, moonlighting, and biometric limits

A rigorous assessment of mandatory cooldown schedules must examine four counter-arguments and structural failure modes established in psychological and systems research:

1. **Tension between mandatory lease lockouts and personal autonomy (Csikszentmihalyi [2], Deci et al. [3], Ryan and Deci [16]).** Self-Determination Theory [3, 16] and flow theory [2] emphasize that perceived autonomy and uninterrupted deep absorption are primary drivers of intrinsic satisfaction. A rigid capability-lease gate that locks a contributor out of their workspace at the end of Week 7 regardless of personal rhythm is an explicit paternalistic constraint, not a "non-coercive" nudge. To mitigate psychological reactance without re-introducing streak pressure, `WSPP-Sabbatical` should permit a bounded ±1-week flexible window around $W_{\max} = 7$ so stewards can complete an active creative sprint before initiating their cooldown week.

2. **Unfulfilled excess demand and cross-platform moonlighting.** Because `WSPP-Sabbatical` caps intra-week intensity ($I_t \leq 1.02$) and schedules 12.5% sabbatical time (1 in 8 weeks), a single contributor never absorbs the raw demanded load when $\lambda_{\text{load}} = 1.8\times$ or $2.2\times$; meeting aggregate organizational demand requires expanding the cross-trained contributor pool via `F-UCB-Fluid` (Milestone #43). Furthermore, if liquidity-constrained workers are locked out of Platform A during Week 8 without baseline income smoothing (Milestone #45), they may moonlight on unpaced commercial gig platforms, defeating recovery (as reflected in our $p_{\text{adhere}} \in \{0.50, 0.70\}$ non-adherence sweeps).

3. **Weak alignment between wearable sensors and subjective stress (Sathyanarayana et al. [17], Siepe et al. [19]).** Although naturalistic stress classification from wearables has improved [1, 15, 17], Siepe et al. [19] demonstrate that passive physiological indicators explain only a small fraction of variance in self-reported ecological momentary assessment (EMA) stress. Relying on biometric tripwires as a mandatory gate creates both false-positive

interruptions and reward-hacking incentives [7]; hence physiological sensing must remain an opt-in local aid rather than a mandatory organizational oracle.

4. **Delegation feedback loops during rushed handovers (Eliav [5], Shen and Tamkin [18]).** If contributors rush to offload unfinished tasks to autonomous AI agents immediately prior to a scheduled sabbatical, unverified delegation can accumulate technical and cognitive debt [5, 10, 18].

7 Scope constraints and empirical calibration boundaries

- **Model-conditioned mechanistic simulation vs. clinical burnout trials.** While Equations (1)–(3) incorporate parameter uncertainty over replenishment ρ_{sab}, debt persistence γ_{debt}, and off-platform caregiving load L_{care}, all reported yield and retention figures are model-conditioned mechanistic projections rather than clinical or field burnout measurements. Empirical calibration to longitudinal field data—such as multi-month organizational 4-day workweek trials or experience-sampling cohorts [19]—is required to estimate occupation-specific drain exponents γ and relapse hazards.
- **Off-platform caregiving and domestic inequality.** As demonstrated in our caregiving sensitivity sweep ($\bar{L}_{\text{care}} = 0.28$), heavy unpaid caregiving obligations significantly erode sabbatical replenishment (59.35% yield and 86.00% retention vs. 81.93% yield and 99.60% retention under low caregiving load), showing that workplace pacing rules cannot substitute for broader childcare and social support infrastructure.
- **Seasonal and project-milestone cadence mismatch.** Fixed 7:1 weekly cycles may conflict with academic terms, hardware tape-out deadlines, or agricultural seasons that require longer low-intensity recovery intervals following seasonal peaks.

8 Pre-registerable falsification protocol

Table 5 specifies six pre-registerable empirical tests designed to falsify the core claims of `WSPP-Sabbatical` in organizational trials.

Table 5: Six pre-registerable tests that could falsify the `WSPP-Sabbatical` work-pacing framework.

Protocol	Hypothesis	Design and primary metrics	Falsified if
P1. 52-week cumulative yield parity ($Y_{\text{pas}} = 0$)	Even with zero sabbatical output ($Y_{\text{pas}} = 0$), 7:1 cooldowns outperform unpaused pacing over 52 weeks	12-month randomized trial across 160 contributors comparing `Intensity-Cap Only`, `Soft Nudges`, and `WSPP-Sabbatical`	`WSPP-Sabbatical` cumulative verified annual output trails `Intensity-Cap Only` or `Soft Nudges`
P2. Trajectory Sustainability Ratio ($\mathcal{S}_{52} \geq 0.75$)	Periodic sabbaticals prevent Quarter-4 productivity and Maslach burnout degradation [12]	Measure ratio of weeks 41–52 output and MBI exhaustion scores to weeks 1–4 baseline	Mean Sustainability Ratio $\mathcal{S}_{52}$ falls below 75.0% or annual burnout attrition exceeds 10.0%
P3. Active-only critical verification quality	Rested contributors detect significantly more subtle AI errors during active weeks [11, 21]	Seed 200 subtle factual/code errors into active review queues during Week 6 of continuous work vs. Week 1 post-sabbatical	Post-sabbatical active error detection fails to exceed unpaused Week 6+ detection by ≥ 10 percentage points
P4. Daily rotation vs. weekly sabbatical decomposition	Weekly sabbaticals provide > 20 percentage points higher retention than daily role rotation (Milestone #43) alone at high load	2×2 factorial trial crossing daily task rotation (`F-UCB-Fluid`) with 7:1 weekly sabbaticals under high demand ($\lambda_{\text{load}} \geq 1.8\times$)	Retention or active quality gain of adding weekly sabbaticals on top of daily rotation is < 15 percentage points
P5. Decay-frozen cooperative standing	Sabbatical clock-freeze ($\lambda_d = 0$) prevents loss of patronage governance weight (Milestone #44)	Audit `UACS-Ostrom` voting weights of 200 stewards before and after scheduled 1-week cooldowns	Any steward loses $> 0.0\%$ relative voting weight or task-tier access as a result of taking a scheduled cooldown
P6. Flexible $\pm$1-week window autonomy audit	Allowing a $\pm$1-week flexible cooldown window preserves intrinsic motivation [2, 16]	Randomized comparison of rigid day-49 lockout vs. $\pm$7-day flexible cooldown window on Intrinsic Motivation Inventory (IMI)	$> 10\%$ of contributors report flow-interruption reactance under the flexible cooldown window

9 Conclusion

Gamified work systems that rely on consecutive-week streak multipliers, leaderboard decay, or optional wellness nudges induce cumulative allostatic fatigue debt that degrades both contributor retention and active critical-thinking quality [12, 13, 21]. Formulated as the macro-horizon companion to daily shift-level fluid task rotation (`F-UCB-Fluid`, Milestone #43), `WSPP-Sabbatical` combines an intra-week sustainable intensity cap with scheduled 7:1 capability-lease cooldowns and decay-frozen cooperative standing. Across paired 52-week contributor trajectories incorporating parameter uncertainty over rest replenishment, debt relapse, off-platform caregiving load, and imperfect leave adherence ($p_{\text{adhere}} = 0.82$), our model-conditioned simulation demonstrates that while intra-week intensity capping alone raises retention from 19.00% $[15.80, 22.67]$ to 46.60% $[42.27, 50.98]$ (49.41% yield), adding periodic sabbaticals clears cumulative fatigue debt to achieve 98.40% $[96.87, 99.19]$ retention (492/500), 76.07% $[75.18, 76.96]$ quality-adjusted annual yield (74.57% even at $Y_{\text{pas}} = 0$), an 81.17% $[79.56, 82.78]$ Trajectory Sustainability Ratio, and 83.22% $[82.53, 83.90]$ active critical-thinking quality.

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] Mihaly Csikszentmihalyi. Flow: The psychology of optimal experience (autotelic experience). Harper & Row, 1990. URL ISBN:978-0060920432.

[3] Edward L. Deci, Richard Koestner, and Richard M. Ryan. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. Psychological Bulletin, 125(6):627–668, 1999. URL https://doi.org/10.1037/0033-2909.125.6.627.

[4] 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.

[5] 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.

[6] Juho Hamari, Jonna Koivisto, and Harri Sarsa. Does gamification work? — a literature review of empirical studies on gamification. Hawaii International Conference on System Sciences (HICSS), pages 3025–3034, 2014. URL https://doi.org/10.1109/HICSS.2014.377.

[7] Yue Huang, Zhangchen Xu, Yuchen Ma, and Wenjie Wang. Reward hacking challenges oversight of autonomous research agents. arXiv preprint arXiv:2609.28614, 2026. URL https://arxiv.org/abs/2609.28614.

[8] 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.

[9] 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.

[10] 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.

[11] Hao-Ping Lee, Advait Sarkar, Lev Tankelevitch, Ian Drosos, Sean Rintel, Richard Banks, and Nicholas Wilson. The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 2025. doi: 10.1145/3706598.3713778.

[12] Christina Maslach, Wilmar B. Schaufeli, and Michael P. Leiter. Job burnout. Annual Review of Psychology, 52(1):397–422, 2001. URL `https://doi.org/10.1146/annurev.psych.52.1.397`.

[13] 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`.

[14] Sina Rismanchian, Hasan Uzun, Jeffrey Matayoshi, Eric Cosyn, and Eyad Kurd-Misto. Faster completion, less learning: Generative AI reduced study time on math problems and the knowledge they build, 2026. URL `https://arxiv.org/abs/2605.21629`.

[15] Max Rosenblattl, Nikhil Krishnaswamy, Leon Nissen, Manuel Milling, Nicolas Zumarraga, Ning Wang, Thomas Kaar, Robert Jakob, Björn W. Schuller, Elgar Fleisch, Patrick Langer, and Paul Schmiedmayer. PocketTSLM: On-device time-series language models for physiological monitoring. arXiv preprint arXiv:2610.21308, 2026. URL `https://arxiv.org/abs/2610.21308`.

[16] 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.

[17] 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.

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

[19] Björn S. Siepe, Rayyan Tutunji, Carlotta L. Rieble, Ricarda K. K. Proppert, and Eiko I. Fried. Associations between ecological momentary assessment and passive sensor data in a large student sample. *Journal of Psychopathology and Clinical Science*, 134(8):912–925, 2025. doi: 10.1037/abn0001013.

[20] 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`.

[21] 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.