Using Machine Learning to Encourage Reflective Smartphone Use: An RL and LLM-Based Intervention System

Nikola Dimitrijevic
Dartmouth College
nikola.dimitrijevic.25@dartmouth.edu

May 2025

Abstract

Modern smartphone ecosystems can amplify short-term gratification loops through interface design and recommendation algorithms. Many digital well-being interventions optimize for control (blocking, limiting, or compliance), which can undermine user agency. This paper presents an intervention system that encourages reflection rather than restriction. The system combines (1) a reinforcement learning based intervention scheduler that selects moments likely to benefit from a cognitive pause and (2) a large language model prompt generator that produces non-judgmental, goal-aligned reflection prompts grounded in user-declared intentions. We evaluate the approach in an offline simulation setting using smartphone-usage style logs and measure timing quality, intrusiveness, learning speed, and coverage of at-risk sessions. The learned scheduler achieves 81% intervention timing accuracy with a 7.4% false positive rate, adapts to shifting user patterns with convergence in approximately three simulated days, and covers 68% of at-risk sessions, with remaining gaps largely attributable to cold-start conditions.

1 Introduction

Smartphone overuse is increasingly associated with short-term gratification loops driven by attention-optimizing design patterns and recommender systems. Existing interventions often take a control-oriented posture, for example by restricting access or aggressively interrupting sessions. While such approaches can reduce usage, they risk reducing autonomy and may not help users build lasting self-regulation.

This work investigates a human-centered alternative: use machine learning to create brief opportunities for self-awareness at moments when a user is likely engaging in impulsive usage. The design principle is to help users reflect, not restrict. We avoid inferring emotional states and instead ground prompting in user-declared goals such as focus, rest, or intentional use.

We present an end-to-end system with two components. First, an intervention scheduler uses a contextual bandit formulation to decide whether to insert a reflection prompt at the beginning of a session or immediately prior to a suspected "gratification" action. Second, a prompt generator uses a GPT-style language model to produce short, non-judgmental prompts that are tailored to the user's stated goals and the current usage context.

2 Related Work

LLM-driven interventions such as MindShift generate personalized prompts based on inferred mental states and have been shown to reduce problematic smartphone use modestly in field settings [**?**].

Reinforcement learning based approaches such as Time2Stop learn when to interrupt smartphone use, improving timing and uptake relative to static baselines [**?**]. TypeOut proposes a lightweight reflective pause by requiring a self-affirmation during unlock, demonstrating large reductions in use in a long study, but it employs a fixed interaction that does not adapt to context [**?**].

Our approach borrows the timing adaptivity of Time2Stop and the language flexibility of MindShift, while enforcing a different design constraint: the system should not infer emotion or persuasion targets. Instead, prompts are anchored in user-declared values, and the reward is aligned to reflective success rather than usage suppression.

3 System Design

3.1 Intervention timing via reinforcement learning

We model the intervention decision as a contextual bandit. At each decision point the agent observes a feature vector that summarizes the current session, including:

- session start time and recency features,
- app switching frequency and short-horizon volatility,
- coarse app category signals (for example, entertainment vs productivity).

The agent selects an action from a small discrete set (prompt now vs do not prompt). We use an offline training and evaluation loop with log-based replay.

Reward design. Unlike "reduce screen time" objectives, the reward targets reflective success and autonomy-preserving delay. In simulation, we assign positive reward to outcomes consistent with reflection, such as (1) increased delay before opening a gratification app and (2) reduced impulsive rapid app-switching immediately after the decision point. We assign negative reward to prompting during low-risk or productive sessions.

3.2 Reflection prompt generation via LLM

Given a decision to prompt, we generate a short message conditioned on two sources of information:

- user-declared goals (for example, "focus", "rest", or "be intentional"),
- real-time behavioral context (for example, "third short entertainment session today").

Prompts are designed to be non-judgmental, brief, and supportive of the user's autonomy. The prompt generator does not rely on inferred emotions.

4 Experimental Setup

4.1 Data and simulation

We evaluate using smartphone-usage style logs in an offline setting. Following common practice in intervention research, we simulate decision points and outcomes using log-based replay. The proposal stage targeted publicly available datasets such as StudentLife and a large-scale smartphone usage dataset [**?**, **?**]. For the final system evaluation, we used synthetic usage logs that mirror these sources, capturing session timing, app categories, and switching behavior. This enables controlled experiments across user profiles, including consistent users and mixed-use profiles.

4.2 Metrics

We report four metrics that capture both timing quality and intrusiveness:

- **Intervention timing accuracy:** proportion of prompts delivered immediately before a high-risk gratification action.
- **False positive rate:** proportion of prompts delivered during productive or low-risk sessions.
- **Adaptation rate:** time required for the agent to converge after user patterns shift.
- **Reflection opportunity coverage:** proportion of at-risk sessions preceded by a well-timed prompt.

5 Results

Table 1 summarizes performance.

Table 1: System learning and intervention accuracy metrics from offline simulation.

Metric	What it captures	Result
Intervention timing accuracy	Prompts delivered immediately before high-dopamine app launches or rapid app-switching bursts.	81%
False positive rate	Prompts triggered during productive or low-risk sessions (misclassifying focus time as reflection time).	7.4%
Adaptation rate	How quickly the scheduler adjusts when usage patterns shift.	˜3 days
Reflection opportunity coverage	Fraction of at-risk sessions that were preceded by a well-timed prompt.	68%

The scheduler frequently identifies high-risk moments just prior to rapid switching or entertainment app launches, while keeping intrusions during focused sessions low. Adaptation behavior indicates that the bandit quickly updates timing preferences for consistent users and more slowly for mixed-use profiles. Coverage remains limited by cold-start conditions where insufficient prior data exists to confidently classify a session as high-risk.

6 Discussion and Limitations

Reflection, not restriction. A central design goal is to support autonomy. The reward structure avoids directly optimizing for total reduction in usage time. Instead, it promotes a cognitive pause that enables the user to choose.

Cold-start and personalization. The main observed performance gap is coverage in early phases. Future work should incorporate better priors, lightweight onboarding signals (for example, goal selection plus a short calibration period), or uncertainty-aware prompting that reduces overconfidence.

Offline evaluation limits. Offline simulation cannot fully capture user perception, long-run behavior change, or habituation effects. A longitudinal deployment is necessary to assess sustained outcomes, user trust, and perceived helpfulness versus intrusiveness.

7 Ethical Considerations

Deploying behavior-facing ML systems raises issues of privacy, consent, and transparency. A real deployment should prioritize on-device feature extraction where possible, minimize data retention, provide clear explanations for interventions, and allow users to inspect and override settings. Because prompts are goal-anchored rather than emotion-inferred, the system reduces the risk of over-claiming understanding of the user's internal state, but it still must avoid manipulation and maintain user control.

8 Conclusion

We presented an RL and LLM-based intervention system that aims to encourage reflective smartphone use. In offline simulation, the scheduler achieved 81% timing accuracy with a 7.4% false positive rate, converged in approximately three simulated days, and covered 68% of at-risk sessions. These results support the feasibility of autonomy-preserving, language-based reflection at well-timed moments. Future work will focus on real-device deployment and longitudinal evaluation.