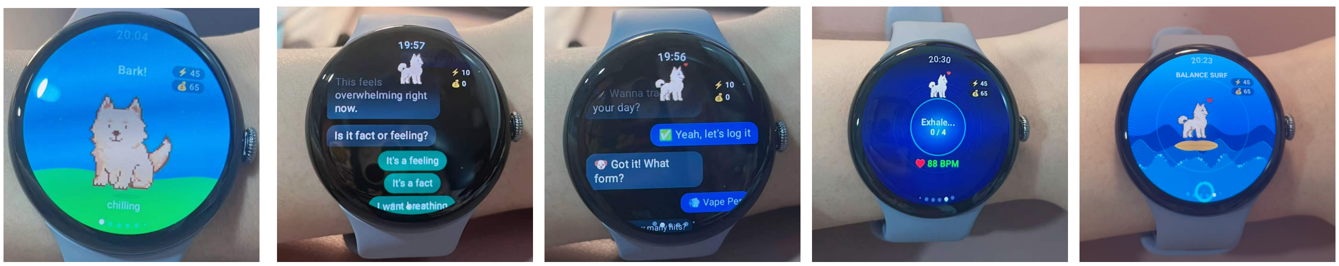
PAWS: Empowering Everyday Cannabis Use Disorder Support through a Personalized AI Digital Pet on Smartwatches

Zhihan Jiang, Mengyuan "Millie" Wu, Ruishi Zou, Shiyu Xu, Emma Macmanus, Steven Liao, Ping Zhang, Dakuo Wang, James L. David, Nabila El-Bassel, Lena Mamykina, Frances R. Levin, Ryan Sultan, Xuhai Xu
Columbia University, New York, New York, USA
{zj2445, xx2489}@cumc.columbia.edu



(a) Home Playground (b) MI/CBT Dialogue (c) Daily Log (d) Biofeedback Breathing (e) Urge Surfing

Figure 1: *PAWS* on Wear OS. (a) Home playground with a reactive digital pet whose mood reflects the user's physiological state; (b) Multi-agent LLM dialogue grounded in Motivational Interviewing (MI) and Cognitive Behavioral Therapy (CBT); (c) Guided substance use logging with CBT reflections and harm-reduction goals; (d) Just-in-time biofeedback breathing exercise triggered by elevated heart rate; (e) Urge Surfing, a motion-controlled mindfulness game for craving management.

Abstract

Cannabis Use Disorder (CUD) significantly affects youth, yet traditional interventions suffer from high stigma and attrition. We demonstrate *PAWS*, a multimodal AI system on Wear OS that transforms evidence-based CUD treatment into the everyday ritual of caring for a personalized digital pet. Grounded in the *Tamagotchi Effect*, *PAWS* reframes clinical exercises as "care actions," fostering intrinsic motivation and a non-judgmental therapeutic bond. The system integrates (1) a multi-agent LLM architecture with intent-based routing grounded in Motivational Interviewing (MI) and Cognitive Behavioral Therapy (CBT) principles, (2) real-time biometric sensing and embodied interactions for just-in-time adaptive interventions (JITAIs), and (3) a layered safety architecture ensuring clinical alignment. In a within-subjects comparison ($N = 52$ current cannabis users; *PAWS* vs. a baseline chatbot), youth aged 18–30 ($n = 38$) rated *PAWS* significantly higher on both engagement (Wilcoxon signed-rank, $p < .001$, $r > 0.8$) and acceptability/usability ($p < .003$, $r > 0.8$).

1 Introduction

Cannabis remains the most prevalent illicit substance among youth [4], and its use has been shown to impair attention, planning, and impulse control [1]. While over 13% of 16-25 US youth meet the criteria for CUD [14], 83.5% of young adults with CUD do not receive any substance-use treatment [5]. Furthermore, 30-60% of those in care drop out due to stigma associated with seeking treatment and limited youth-tailored daily support for managing substance use triggers [2, 13]. While LLM-powered conversational AI offers a scalable framework for evidence-based care, generic models remain clinically unsuitable for CUD due to safety risks, social-evaluative anxiety, and inadequate contextual adaptation [7].

We developed *PAWS* (Pawsitive Companion), a multimodal AI system that embeds evidence-based CUD treatment within a digital pet metaphor on smartwatches (Figure 1). Grounded in the *Tamagotchi Effect*, i.e., the tendency for virtual pet care to foster genuine affective bonds [8], *PAWS* creates a reciprocal care loop: the user feels responsible for the pet's well-being, and this responsibility becomes an intrinsic motivator for sustained treatment adherence. Therapeutic activities, such as practicing coping skills or logging cravings, are reframed as daily care actions that directly influence the pet's state, transforming clinical exercises into a non-judgmental, supportive relationship that avoids the social-evaluative paradox of human-like chatbots. The wrist-worn form factor further isolates the intervention from smartphone-associated social triggers while enabling passive biometric sensing for personalized, context-aware support [9, 11, 12]. We describe the *PAWS* architecture and report pilot findings demonstrating that this relational metaphor significantly increases user engagement and acceptability of the system among youth compared to standard chatbots.

2 System Description

PAWS is a Wear OS smartwatch application backed by a GPT-4o-mini inference server. It comprises four components.

(1) Multi-Agent Dialogue. A multi-agent LangGraph architecture [15] routes user intent to four specialized agents: (a) a *Chat* agent that delivers empathetic [10], CUD-specific support grounded in MI principles and CBT techniques [6], (b) a *Craving Intervention* agent that detects urges and directs the user to a motion-based Urge Surfing game, (c) a *Stress Intervention* agent that detects stress cues to guide users through biofeedback-driven breathing exercises, and (d) a *Summary* agent that provides a personalized welcome-back message with progress context. To ensure glanceability on smartwatch interfaces, agent responses are strictly constrained to ≤7 words. These are supplemented by contextual first-person quick-replies, updates on pet state (selected from 10 emotional profiles), and voice and on-screen text input support for wrist-based interaction. The system employs contingency management-inspired reinforcement: users earn in-app currency and "energy" by logging substance use (form, amount, timing)—which triggers personalized CBT reflections and optional harm-reduction goal setting—and by completing therapeutic exercises. Conversation history is persisted across sessions to maintain a relational bond.

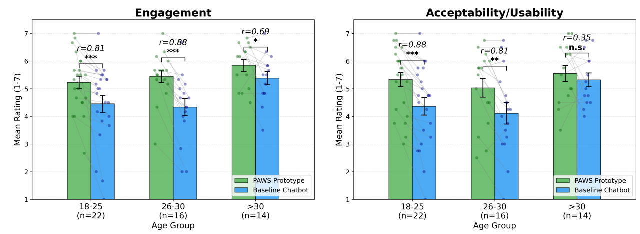
(2) Tiered Sensing & Just-in-Time Adaptive Intervention. To optimize therapeutic efficacy against wearable battery constraints, *PAWS* employs a two-tier sensing strategy via the Wear OS Health Services API. The *passive tier* continuously monitors heart rate, steps, calories, and activity state in a low-power background mode. When passive heart rate exceeds a configurable threshold (i.e., 120 BPM) during a non-exercise activity state, the system triggers a JITAI: the *active tier* automatically navigates the user to a biofeedback screen for high-fidelity heart rate monitoring and guided breathing exercise with real-time visual feedback. To mitigate intervention fatigue, the system incorporates a five-minute cooldown and hysteresis-based thresholds to prevent repeated triggering. The digital pet serves as a visual proxy for this sensing–intervention–feedback loop. Additionally, passive activity recognition (e.g., sleep, active, passive) synchronizes the pet's behavioral state with the user's circadian rhythm.

(3) Embodied Interaction: Urge Surfing. *PAWS* leverages the smartwatch's inertial sensors to transform the mindfulness-based urge-surfing technique [3] into a tangible, motion-based interaction. Using real-time accelerometer data, users physically tilt their wrist to balance the digital pet on a virtual surfboard amidst "craving waves." Haptic feedback activates when the pet deviates beyond a threshold, using physical sensation to ground the user and redirect focus from substance-use triggers. A 3-minute scoring window with performance-based rewards reinforces in-the-moment engagement with craving management.

(4) Safety & Platform Integration. Clinical safety is enforced through two complementary mechanisms. The *alignment layer* embeds non-negotiable behavioral rules directly into each agent's system prompt, restricting the pet persona to clinically appropriate language and bounded affective expressions. Reward policies reinforce only genuine therapeutic engagement, preventing gaming of the care-action loop. The *structural layer* enforces type-safe outputs via structured generation schemas, ensuring every LLM response conforms to the expected clinical format before delivery to the user.

3 Preliminary Evaluation

We conducted a within-subjects study ($N = 60$ self-reported current cannabis users, 78% daily or near-daily use, recruited via CloudResearch; 8 excluded for failing attention checks; $N = 52$ analyzed) comparing the *PAWS* prototype against a baseline GPT-4o-mini model. This study was approved by the Columbia University Institutional Review Board. After providing informed consent,



(a) Engagement (b) Acceptability/Usability

Figure 2: Within-subjects comparison of *PAWS* vs. Baseline across age groups ($N = 52$). Error bars represent standard error; gray lines connect paired observations. *** $p < .001$, ** $p < .01$, * $p < .05$, n.s. = not significant.

participants experienced both systems in counterbalanced order. Each interaction lasted 10.2 ± 3.6 minutes. Participants then rated each on 10 items (7-point Likert scale) spanning two constructs: *Engagement* (worthwhileness, goal alignment, task importance, trust, reliability) and *Acceptability/Usability* (ease of use, capability fit, social comfort, reuse intention, attractiveness). Data were analyzed by age group using Wilcoxon signed-rank tests.

As shown in Figure 2, among youth aged 18–25 ($n = 22$), *PAWS* was rated significantly higher than the baseline chatbot on both engagement (*PAWS*: $M = 5.23$, $SD = 1.06$; Baseline: $M = 4.45$, $SD = 1.44$; $p < .001$, $r = 0.81$) and acceptability/usability (*PAWS*: $M = 5.33$, $SD = 1.21$; Baseline: $M = 4.36$, $SD = 1.47$; $p < .001$, $r = 0.88$), both with large effect sizes. Among participants aged 26–30 ($n = 16$), the pattern was consistent: engagement (*PAWS*: $M = 5.45$, $SD = 0.89$; Baseline: $M = 4.33$, $SD = 1.24$; $p < .001$, $r = 0.88$) and acceptability/usability (*PAWS*: $M = 5.03$, $SD = 1.35$; Baseline: $M = 4.11$, $SD = 1.55$; $p < .003$, $r = 0.81$) both showed significant advantages with large effect sizes. Among participants over 30 ($n = 14$), the engagement advantage persisted ($p = .031$, $r = 0.69$); the acceptability difference was in the same direction but did not reach significance ($p = .326$).

4 Conclusion and Future Work

We presented *PAWS*, a multimodal AI system integrating evidence-based CUD treatment into digital pet care on a smartwatch. Our evaluation demonstrates that this relational metaphor significantly increases engagement and acceptability among youth compared to standard chatbots—addressing a primary barrier to CUD intervention adoption. We are conducting co-design sessions with youth with lived experience of CUD and clinicians to incorporate community input into the pet's personality, therapeutic content, and safety boundaries, followed by an 8-week feasibility trial measuring usability, safety, engagement, and preliminary efficacy. At the workshop, we will demonstrate the live *PAWS* prototype on Google Pixel Watch 4, allowing attendees to experience the AI-driven pet conversations, biofeedback breathing exercises, and urge-surfing game firsthand, and to discuss design implications for wearable health interventions targeting vulnerable populations.

References

[1] Matthew D Albaugh, Jonatan Ottino-Gonzalez, Amanda Sidwell, Claude Lepage, Anthony Juliano, Max M Owens, Bader Chaarani, Philip Spechler, Nicholas Fontaine, Pierre Rioux, et al. 2021. Association of cannabis use during adolescence with neurodevelopment. *JAMA psychiatry* 78, 9 (2021), 1031–1040.

[2] Robert D Ashford, Austin M Brown, and Brenda Curtis. 2018. Systemic barriers in substance use disorder treatment: A prospective qualitative study of professionals in the field. *Drug and alcohol dependence* 189 (2018), 62–69.

[3] Sarah Bowen and Alan Marlatt. 2009. Surfing the urge: brief mindfulness-based intervention for college student smokers. *Psychology of Addictive Behaviors* 23, 4 (2009), 666.

[4] Jonathan P Caulkins. 2024. Changes in self-reported cannabis use in the United States from 1979 to 2022. *Addiction* 119, 9 (2024), 1648–1652.

[5] Namkee G Choi, John Moore, and Bryan Y Choi. 2024. Cannabis use disorder and substance use treatment among US adults. *Journal of substance use and addiction treatment* 167 (2024), 209486.

[6] Peter J Gates, Pamela Sabioni, Jan Copeland, Bernard Le Foll, and Linda Gowing. 2016. Psychosocial interventions for cannabis use disorder. *Cochrane database of systematic reviews* 5 (2016).

[7] Zhijun Guo, Alvina Lai, Johan H. Thygesen, Joseph Farrington, Thomas Keen, and Kezhi Li. 2024. Large Language Models for Mental Health Applications: Systematic Review. *JMIR Mental Health* 11 (2024), e57400.

[8] Andreas Holzinger, Arnold Pichler, Wolfgang Almer, and Hermann Maurer. 2001. TRIANGLE: A Multi-Media test-bed for examining incidental learning, motivation and the Tamagotchi-Effect within a Game-Show like Computer Based Learning Module. In *EdMedia+ Innovate Learning*. Association for the Advancement of Computing in Education (AACE), 766–771.

[9] Zhihan Jiang, Adrienne YL Chan, Dawn Lum, Kirstie HTW Wong, Janice CN Leung, Patrick Ip, David Coghill, Rosa S Wong, Edith CH Ngai, and Ian CK Wong. 2024. Wearable Signals for Diagnosing Attention-Deficit/Hyperactivity Disorder in Adolescents: A Feasibility Study. *JAACAP Open* (2024).

[10] Zhihan Jiang, Qianhui Chen, Chu Zhang, Yanheng Li, and RAY LC. 2026. Hear You in Silence: Designing for Active Listening in Human Interaction with Conversational Agents Using Context-Aware Pacing. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*. ACM. doi:10.1145/3772318.3791238

[11] Zhihan Jiang, Lin Lin, Xinchen Zhang, Jianduo Luan, Running Zhao, Longbiao Chen, James Lam, Ka-Man Yip, Hung-Kwan So, Wilfred HS Wong, et al. 2023. A data-driven context-aware health inference system for children during school closures. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 1 (2023), 1–26.

[12] Zhihan Jiang, Vera Van Zoest, Weipeng Deng, Edith CH Ngai, and Jiangchuan Liu. 2023. Leveraging machine learning for disease diagnoses based on wearable devices: A survey. *IEEE Internet of Things Journal* 10, 24 (2023), 21959–21981.

[13] Sara N Lappan, Andrew W Brown, and Peter S Hendricks. 2020. Dropout rates of in-person psychosocial substance use disorder treatments: a systematic review and meta-analysis. *Addiction* 115, 2 (2020), 201–217.

[14] Substance Abuse and Mental Health Services Administration. 2025. *Key Substance Use and Mental Health Indicators in the United States: Results from the 2024 National Survey on Drug Use and Health*. Technical Report. Center for Behavioral Health Statistics and Quality, SAMHSA.

[15] Mengyuan Wu, Zhihan Jiang, Yuang Fan, Richard Feng, Sahiti Dharmavaram, Mathew Polowitz, Shawn Fallon, Bashima Islam, Lizbeth Benson, Irene Tung, David Creswell, and Xuhai Xu. 2026. MindfulAgents: Personalizing Mindfulness Meditation via an Expert-Aligned Multi-Agent System. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*.