Craving Checkpoint: An Interactive Fridge Lock for Mindful Eating

Xiaoman Yang
Dept. of Design Technology
Cornell University
Ithaca, New York, USA
xy498@cornell.edu

Jinyue Wang
Dept. of Design Technology
Cornell University
Ithaca, New York, USA
jw2796@cornell.edu

Zijie Zhou
Dept. of Design Technology
Cornell University
Ithaca, New York, USA
zz858@cornell.edu

Ziang Liu
Dept. of Computer Science
Cornell University
Ithaca, New York, USA
zl873@cornell.edu

Marcelo Coelho
Dept. of Design Technology
Cornell University
Ithaca, New York, USA
mc2957@cornell.edu

Figure 1: CRAVING CHECKPOINT intercepts routine cravings with a prompt for reflection, followed by LLM-driven healthy suggestions, transforming food access as a co-authored ritual.

Abstract

Traditional dietary interventions often rely on restriction, tracking, or delayed reflection, which can limit their ability to foster lasting change. We present CRAVING CHECKPOINT, a Large Language Object (LLO) in the form of an interactive fridge lock designed as a Just-In-Time Adaptive Intervention that supports mindful eating through embodied and emotionally expressive interaction. The system engages users at the moment of food access, prompting self-reflection through mood and hunger input and offering real-time feedback through an anthropomorphic voice and synesthetic lighting. A large language model personalizes suggestions based on user state and behavioral patterns, gently guiding healthier choices without enforcing control. By transforming food access into a shared ritual between human and machine, CRAVING CHECKPOINT explores how creative AI can support sustainable behavior change through timely, affective, and co-authored interventions.

Submitted to 39th Conference on Neural Information Processing Systems (NeurIPS 2025). Do not distribute.

1 Introduction

Unhealthy eating remains a pressing public health concern. Over 40% of U.S. adults are classified as obese [13], contributing to elevated risks for diabetes, cardiovascular disease, and other chronic conditions [40]. While popular interventions such as weight-management apps [16], structured meal plans [48], and commercial programs [27] can produce short-term weight loss, they often fail to sustain long-term behavior change. A meta-analysis found that most participants regained their lost weight within five years [1, 26]. One key limitation is that these approaches overlook the moment when habitual urges, such as opening the fridge "just to look", bypass conscious intention and trigger automatic behavior.

Supporting behavior change at the moment these urges arise, rather than before or after, is critical for disrupting automatic routines. Behavioral research suggests that brief cognitive pauses can help align impulsive actions with long-term goals [20, 43]. Techniques like health goal priming [34], mindfulness prompts [41], and acceptance-based strategies [3] have been shown to gently interrupt habitual behavior without demanding significant effort or self-control.

To meet users in these small but pivotal moments, we draw on two complementary ideas. Just-In-Time Adaptive Interventions (JITAIs) [30] provide the scaffolding for delivering support at exactly the right time — when someone is most likely to fall into an automatic pattern or act on impulse. They've been used across domains like smoking cessation [50] and emotion regulation [47], where timing is key. Large Language Objects (LLOs) [10], by contrast, focus on the form that support takes. By embedding language models into physical artifacts, LLOs transform everyday objects into interactive systems that can sense, respond, and express [8]. They offer a more situated and expressive form of engagement than screen-based apps, allowing interventions to feel less like interruptions and more like part of the surrounding environment. By combining these two frameworks, we create interventions that are both well-timed and gently integrated into the user's environment, offering support that feels less like correction and more like presence.

We introduce CRAVING CHECKPOINT, an interactive fridge lock that brings these two frameworks together to support mindful eating in everyday contexts. When the user reaches for the fridge, CRAVING CHECKPOINT gently interrupts the routine with a short prompt for mood and hunger reflection. Based on this input and prior interactions, a large language model generates personalized, context-aware feedback, delivered through anthropomorphic voice and synesthetic lighting. Rather than enforcing control, the system introduces a moment of pause that preserves user agency while reshaping habitual patterns. To sustain engagement over time, CRAVING CHECKPOINT tracks consistency and occasionally responds with playful encouragement or rewards inspired by game design [5].

In reframing food access as a co-authored ritual, CRAVING CHECKPOINT explores how embodied language models might move beyond screens and commands into the fabric of daily life. In doing so, it meets users in the moment of the urge, supporting mindful decision-making right when it matters most, and laying the groundwork for lasting behavioral change over time, offering a vision of AI not as enforcers or assistants, but as a quiet partner in the everyday rhythms of being human.

2 Related Work

Recent research in HCI and creative AI has explored how intelligent systems can shape not just tasks but habits, rituals, and moods [15, 36]. As large language models become embedded in physical interfaces, their role is shifting from passive tools to active participants in everyday decision-making. This shift is especially relevant in health behavior change, where effective interventions must align with emotional states, contextual cues, and timing. In parallel, design research has begun to reimagine AI not as a distant assistant but as a presence that engages softly within ordinary routines [23, 25]. These converging directions suggest new possibilities for emotionally responsive systems that support meaningful, long-term change through grounded and symbiotic interaction.

2.1 Shaping Healthy Behavior Through Intervention

Digital approaches to dietary support have long focused on retrospective tracking and self-monitoring. Tools like MyFitnessPal [28] and Noom [31] encourage users to log meals, track calorie intake, and reflect on progress over time. While these systems raise awareness, they often fall short of sustaining long-term engagement due to their reliance on delayed feedback, rigid structures, and high cognitive burden [35]. Recent alternatives have begun to explore more emotionally attuned and context-sensitive formats. Eat4Thought, for instance, enables users to journal meals through mood and sensory

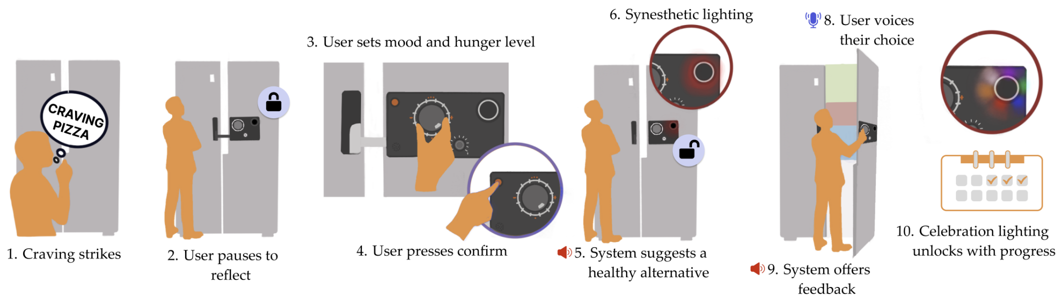

Figure 2: **Storyboard**: illustrating the user interaction flow, from *craving* to *choice*.

reflections rather than nutritional metrics alone [52], promoting self-awareness but offering limited support in the moment when decisions are made. Other systems incorporate food preferences, health conditions, and contextual cues to generate personalized recommendations [53, 12], representing a shift toward more sustainable, value-aligned interventions.

Building on this trajectory, just-in-time adaptive interventions (JITAIs) [29] offer a framework for delivering support precisely when users are most susceptible to automatic or impulsive behaviors. Rather than relying on static schedules or retrospective prompts, JITAIs use context-aware triggers, such as time, location, or affective state, to deliver timely, personalized nudges. This approach has shown success in promoting healthier behaviors across domains like physical activity, stress management, and substance use [29, 19]. Within eating contexts, however, few systems deliver support at the moment of urge itself, such as when someone opens the fridge "just to look." Our work extends JITAI principles into the physical setting of food access, combining momentary intervention with embodied, emotionally resonant language feedback to gently disrupt habitual patterns and support mindful decision-making.

2.2 Embodied and Expressive Large Language Models

The role of large language models is expanding beyond text-based chat interfaces toward more affective and situated forms of interaction. Recent systems have explored embedding LLMs into physical artifacts that communicate through voice [8], gesture [54], or light [49], enabling more expressive and emotionally attuned exchanges. These developments reflect a growing interest in designing AI not just as tools for information retrieval, but as collaborators in everyday practices, offering context-sensitive support, co-authorship, and anthropomorphic presence in the flow of daily life.

At the same time, HCI research has explored how embodied interfaces shape cognition and behavior through ambient cues. Systems that use color, light, or sound to signal mood, intention, or state have been shown to influence perception and prompt reflection [44]. Prior work on synesthetic feedback [51], expressive IoT [4], and mood-responsive environments [24] demonstrates how nonverbal, sensory modalities can be leveraged to create emotionally supportive experiences. These directions align with broader trends in HCI design and embodied AI, where systems are designed not just for task completion but for emotional attunement and daily integration.

Together, these threads point to a growing design space for physically grounded, emotionally aware AI systems that participate in everyday routines not only through words but through presence, timing, and sensing. Our work draws from this space to examine how language models, when embodied and affectively expressive, can participate in shaping habits and supporting behavior change through subtle, situated interactions.

3 Craving Checkpoint

CRAVING CHECKPOINT turns the act of opening the fridge into a moment of mindful reflection. Instead of imposing control, it introduces a brief interaction where users report their mood and hunger before receiving personalized food suggestions from a large language model. These suggestions are reinforced through synesthetic lighting and voice feedback. By embedding self-awareness into a familiar routine, the system helps shift eating decisions from impulse to intention, gradually supporting healthier habits through repeated, low-effort engagement.

3.1 Self-Reflection through Mood and Hunger

To initiate each interaction, Craving Checkpoint prompts users to report two aspects of their internal state: mood and hunger. These inputs serve not only as parameters for personalization but as lightweight interventions that encourage self-awareness before eating.

Mood is selected using a rotary encoder from a curated set of terms drawn from the circumplex model of affect [39]: bored, calm, relaxed, comfort, happy, excited, alert, tense, angry, distressed, sad, and depressed, covering both arousal and valence dimensions. This labeling step draws on affect regulation research showing that identifying one's emotional state can reduce impulsive behavior by interrupting automaticity [46]. By explicitly surfacing the emotional drivers that often underlie unplanned eating, the system invites users to pause and assess whether the impulse is driven by mood rather than physiological need.

Next, users report hunger on a discrete 0-1 scale via a second encoder. This numeric measure is based on validated visual analogue scales commonly used in satiety research [14]. Unlike binary or coarse categories, the ten-point scale supports nuanced self-awareness, encouraging users to differentiate mild desire from true hunger.

3.2 Synesthetic Lighting Feedback

Craving Checkpoint uses lighting not merely for decoration but as a meaningful channel for emotional and contextual feedback. While synesthesia [37], where stimulation in one sense involuntarily triggers another, is uncommon in the population, research shows that even non-synesthetes exhibit consensual mappings between color and taste [11]. Red and pink are often linked to sweetness, green to sourness, and white to salty [11, 45]. These associations, even among non-synesthetes, allow light to communicate affect and flavor in a subtle, intuitive way.

The large language model selects both hue and animation dynamics based on the user's reported mood and the recommended food type. Color serves as a semantic anchor for taste, while temporal properties, such as pulsation speed and fade duration, modulate arousal and attention [22]. Faster pulses can convey urgency or excitement, while slower fades support calm and reflection [17, 6].

Together, these elements form a multisensory interaction that aligns color, sound, and suggestion. By embedding these cues into the everyday routine of opening the fridge, the system offers affective guidance that supports awareness and decision-making with minimal cognitive effort.

3.3 LLM-Powered Interaction

At the core of Craving Checkpoint is a real-time decision support pipeline powered by GPT-4o [33], accessed via the OpenAI Python API [32]. The system generates personalized food suggestions, lighting cues, and motivational responses based on structured user input collected at the moment of interaction.

Prompt Structure Each interaction captures four inputs: the user's selected mood, hunger level, time of day (categorized as breakfast, lunch, dinner, or snack), and the number of fridge accesses so far that day. These values are packaged into a structured prompt and sent to GPT-4o for processing. The prompt instructs the model to generate three outputs: a food suggestion appropriate to the user's current state, a corresponding LED color, and a light animation tempo. The prompt includes three in-context examples that demonstrate mappings between moods and light patterns, appropriate suggestions across different contexts, and formatting constraints to ensure structured output.

To ground the model's suggestions in reality, the prompt also includes a list of available ingredients inside the fridge provided by the user. This constraint limits the model's suggestions to foods the user actually has, while enabling simple recipe generation when possible. For example, a user might receive a suggestion like "fruit and yogurt parfait" if both components are on hand. The food suggestion is passed to Google Text-to-Speech [18] for voice delivery, while the RGB and animation frequency values are sent to the LED controller to update ambient light in real time.

History-Based Encouragement Following suggestion delivery, the system uses onboard speech recognition to capture the user's verbal response. If the user declines, a second prompt is sent to GPT-4o to generate an encouraging follow-up. This prompt includes a summary of the user's recent interaction history, retrieved from a local SQLite [21] database. By referencing past decisions and successful choices, the model tailors motivational feedback to the individual, maintaining a supportive tone without exerting pressure.

Figure 3: **Left**: physical interface components. **Top right**: light colors evoke taste via synesthetic associations. **Bottom right**: light animations reflect mood and celebrate progress.

3.4 Gamification for Sustained Engagement

CRAVING CHECKPOINT incorporates lightweight gamification strategies to sustain user engagement and reinforce healthy decision-making through positive feedback rather than control. Prior research shows that gamified features, such as progress tracking [2], surprise rewards [7] can significantly increase motivation and long-term adherence in behavior-change interventions [9].

Progress Feedback The system tracks each accepted suggestion and verbally reports the number of healthy choices made in the past week. Simple, cumulative feedback has been shown to enhance users' sense of progress and increase adherence in behavior-change interventions [42].

Milestone Events To introduce surprise and delight, the system also includes a celebration event. After at least three consecutive healthy choices, a randomized trigger may activate, limited to once per week and guaranteed at least once per month. When triggered, it plays one of ten celebratory LED light sequences and delivers a congratulatory message generated by GPT-4o.

3.5 Hardware Implementation

CRAVING CHECKPOINT is built using affordable and widely available components. A Raspberry Pi 5 [38] serves as the central controller, managing all inputs, outputs, and API calls. Users interact with the system through three KY-040 rotary encoders, two for selecting mood and hunger level, and one for confirming their input. A USB microphone is used to capture spoken responses after the suggestion is delivered. A WS2812b LED strip provides full-color lighting feedback, with the Raspberry Pi controlling color and animation speed based on the output from the language model. An SG-90 micro servo acts as a soft-lock on the fridge door. All physical components are housed in a 3D-printed PLA enclosure.

4 Evaluation

We conducted a pilot study with three participants (aged 22–30; two female, one male) from our university community to explore the system's experiential qualities and perceived effects. Each participant used CRAVING CHECKPOINT throughout a single day as part of their normal routine. The goal was not to measure quantitative outcomes, but to assess how the system influenced decision-making moments and shaped user perceptions during everyday use. This small sample size reflects the exploratory nature of the pilot, which aimed to gather early feedback on the interaction flow and inform future design iterations.

4.1 Interrupting Habitual Behavior

Participants reported that the system effectively reduced instances of opening the fridge out of habit or boredom. This behavior, often automatic and disconnected from physiological hunger, was disrupted by the need to interact with the system before accessing food. The presence of the physical device introduced a brief but salient moment of reflection. Participants described this interaction as a "checkpoint" that created cognitive distance between the initial urge and the subsequent decision, particularly during emotionally neutral or bored states.

4.2 AI as Reflective Companion

Participants described the AI's voice-based suggestions as prompting moments of reflection about their actual needs and available options. By surfacing food ideas in real time, the system encouraged users to mentally review the contents of their fridge and consider whether they were truly hungry or simply seeking distraction. Even when the suggested foods were not on hand, participants noted that the recommendations often introduced new ideas, sparking curiosity and influencing future grocery decisions.

Beyond immediate choices, the system offered relief from decision fatigue. When unsure of what to eat, participants found the interaction both helpful and enjoyable, transforming an otherwise frustrating moment into one of lightness and discovery. This affective quality stood in contrast to traditional, control-based interventions. Rather than enforcing behavior, the AI acted as a gentle companion embedded in daily life, providing contextually relevant prompts that felt supportive rather than prescriptive. This collaborative tone aligned with the broader vision of AI systems as emotionally attuned partners in everyday routines.

4.3 Suggested Improvements

Participants identified several areas for refinement. One recurring concern was the system's interference during meal preparation, particularly when multiple fridge accesses were needed to gather ingredients. In such contexts, the locking mechanism disrupted the natural flow of cooking, suggesting the need for more adaptive logic that distinguishes between habitual snacking and intentional meal-related behavior.

Additionally, participants expressed a desire for greater cultural and dietary inclusivity in the food suggestions. The current recommendation set, while generally well-received, did not always align with individual culinary preferences or backgrounds. This highlights the importance of expanding the system's food knowledge base and personalization capabilities to better serve diverse users and eating habits.

5 Conclusion and Future Work

CRAVING CHECKPOINT reimagines the act of opening the fridge as a moment of reflection and co-authorship between human and machine. By combining self-reported mood and hunger with real-time suggestions from a large language model, the system transforms an everyday gesture into a site for situated decision-making, affective feedback, and habit shaping. Through this small ritual, it gestures toward a future in which intelligent systems do not direct behavior from a distance, but participate quietly and creatively in the rhythms of daily life. In doing so, it offers one possible answer to the broader question of what it means to design AI that supports, not overrides, human behavior.

Our preliminary pilot indicates that the system can prompt users to pause, reflect, and reconsider their actions in the moment. However, several limitations remain. The current study was intentionally brief and exploratory, with participants using the system over a single day. While this revealed key experiential insights, it leaves open questions about long-term engagement, habit formation, and adaptation over time. Future work will involve longitudinal deployments to examine sustained use, behavior change, and potential user fatigue.

The system also lacks situational awareness. At present, each fridge interaction is treated uniformly, whether the user is preparing a meal or browsing impulsively. A more context-sensitive approach, such as recognizing patterns of cooking, snacking, or grocery restocking, could help reduce unnecessary friction and support more nuanced interactions.

Finally, personalization remains an open challenge. While users found the AI suggestions engaging, several noted gaps in cultural relevance and dietary fit. Expanding the food knowledge base, incorporating user profiles, and learning from past preferences will be critical for increasing the inclusivity and resonance of future recommendations.

In sum, CRAVING CHECKPOINT is an early exploration of how creative, embodied AI might support self-regulation and mindful behavior through co-constructed micro-interventions. It invites further reflection on how we might design AI companions that are not only smart and efficient, but emotionally attuned, culturally aware, and deeply integrated into the evolving choreography of human life.

References

[1] James W Anderson, Elizabeth C Konz, Robert C Frederich, and Constance L Wood. Long-term weight-loss maintenance: a meta-analysis of US studies. *Am. J. Clin. Nutr.*, 74(5):579–584, November 2001.

[2] Leona Aschentrup, Pia Anna Steimer, Kevin Dadaczynski, Timothy Mc Call, Florian Fischer, and Kamil J Wrona. Effectiveness of gamified digital interventions in mental health prevention and health promotion among adults: a scoping review. *BMC Public Health*, 24(1):69, January 2024.

[3] Ruth A Baer and Jennifer Krietemeyer. Overview of mindfulness- and acceptance-based treatment approaches. In *Mindfulness-Based Treatment Approaches*, pages 3–27. Elsevier, 2006.

[4] Frank Beruscha, Katharina Lorenz, Anke Königschulte, Serge Autexier, Annika Sabrina Schulz, Bodo Pahlke, Valerie Bartsch, and Hendrik Leibrandt. Connecting textiles: Exploring textile interior surfaces for power supply, communication and user interaction in the IoT. In *Proceedings of the 12th International Conference on the Internet of Things*, New York, NY, USA, November 2022. ACM.

[5] Paula Bitrián, Isabel Buil, and Sara Catalán. Enhancing user engagement: The role of gamification in mobile apps. *J. Bus. Res.*, 132:170–185, August 2021.

[6] Steven Bleicher and Steven Bleicher. *Contemporary color*. Routledge, London, February 2023.

[7] Chandranil Chakraborttii and Lucas Ferreirra. Towards generating surprising content in 2D platform games. In *Proceedings of the 19th International Conference on the Foundations of Digital Games*, pages 1–7, New York, NY, USA, May 2024. ACM.

[8] Ethan Chang, Zhixing Chen, Jb Labrune, and Marcelo Coelho. Be the beat: AI-powered boombox for music suggestion from freestyle dance. In *Proceedings of the Nineteenth International Conference on Tangible, Embedded, and Embodied Interaction*, pages 1–6, New York, NY, USA, March 2025. ACM.

[9] Cecilia Cheng and Omid V Ebrahimi. Gamification: A novel approach to mental health promotion. *Curr. Psychiatry Rep.*, 25(11):577–586, November 2023.

[10] Marcelo Coelho and Jean-Baptiste Labrune. Large language objects: The design of physical AI and generative experiences. *Interactions*, 31(4):43–48, July 2024.

[11] Richard E Cytowic and David M Eagleman. *Wednesday is indigo blue*. The MIT Press. MIT Press, London, England, September 2011.

[12] David Elsweiler, Christoph Trattner, and Morgan Harvey. Exploiting food choice biases for healthier recipe recommendation. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 575–584, New York, NY, USA, August 2017. ACM.

[13] Samuel Emmerich, Cheryl Fryar, Bryan Stierman, and Cynthia Ogden. Obesity and severe obesity prevalence in adults: United states, august 2021–august 2023. Technical report, September 2024.

[14] A Flint, A Raben, J E Blundell, and A Astrup. Reproducibility, power and validity of visual analogue scales in assessment of appetite sensations in single test meal studies. *Int. J. Obes. (Lond)*, 24(1):38–48, January 2000.

[15] Diego Garaialde, Christopher P Bowers, Charlie Pinder, Priyal Shah, Shashwat Parashar, Leigh Clark, and Benjamin R Cowan. Quantifying the impact of making and breaking interface habits. 2020.

[16] Drishti P Ghelani, Lisa J Moran, Cameron Johnson, Aya Mousa, and Negar Naderpoor. Mobile apps for weight management: A review of the latest evidence to inform practice. *Front. Endocrinol. (Lausanne)*, 11:412, June 2020.

[17] Rostam Golmohammadi, Hanieh Yousefi, Negar Safarpour Khotbesara, Abbas Nasrolahi, and Nematullah Kurd. Effects of light on attention and reaction time: A systematic review. *J. Res. Health Sci.*, 21(4):e00529, October 2021.

[18] Google. Text-to-speech api. `https://cloud.google.com/text-to-speech`, n.d. Accessed: 2025-07-28.

[19] Wendy Hardeman, Julie Houghton, Kathleen Lane, Andy Jones, and Felix Naughton. A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity. *Int. J. Behav. Nutr. Phys. Act.*, 16(1):31, April 2019.

[20] Justin Hepler, Dolores Albarracin, Kathleen C McCulloch, and Kenji Noguchi. Being active and impulsive: The role of goals for action and inaction in self-control. *Motiv. Emot.*, 36(4): 416–424, December 2012.

[21] D. Richard Hipp. Sqlite. `https://www.sqlite.org/`. Accessed: YYYY-MM-DD.

[22] Gijs Huisman, Merijn Bruijnes, and Dirk K J Heylen. A moving feast. In *Proceedings of the 13th International Conference on Advances in Computer Entertainment Technology*, New York, NY, USA, November 2016. ACM.

[23] Matthew Jörke, Shardul Sapkota, Lyndsea Warkenthien, Niklas Vainio, Paul Schmiedmayer, Emma Brunskill, and James A Landay. GPTCoach: Towards LLM-based physical activity coaching. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, pages 1–46, New York, NY, USA, April 2025. ACM.

[24] Azmyin Md Kalam, Musfiqul Azad, and Sumayia Jerin Chowdhury. Prototype smart home environment with biofeedback. In *2020 IEEE Region 10 Symposium (TENSYMP)*. IEEE, 2020.

[25] Sue Lim, Ralf Schmälzle, and Gary Bente. Artificial social influence via human-embodied AI agent interaction in immersive virtual reality (VR): Effects of similarity-matching during health conversations. *Computers in Human Behavior: Artificial Humans*, 5(100172):100172, August 2025.

[26] Paul S Maclean, Audrey Bergouignan, Marc-Andre Cornier, and Matthew R Jackman. Biology's response to dieting: the impetus for weight regain. *Am. J. Physiol. Regul. Integr. Comp. Physiol.*, 301(3):R581–600, September 2011.

[27] Christine N May, Matthew Cox-Martin, Annabell Suh Ho, Meaghan McCallum, Caroline Chan, Kelly Blessing, Heather Behr, Paige Blanco, Ellen Siobhan Mitchell, and Andreas Michaelides. Weight loss maintenance after a digital commercial behavior change program (noom weight): Observational cross-sectional survey study. *Obes. Sci. Pract.*, 9(5):443–451, October 2023.

[28] MyFitnessPal, Inc. Myfitnesspal official website. `https://www.myfitnesspal.com`, 2025. Accessed: 2025-07-27.

[29] Inbal Nahum-Shani, Shawna N Smith, Bonnie J Spring, Linda M Collins, Katie Witkiewitz, Ambuj Tewari, and Susan A Murphy. Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. *Ann. Behav. Med.*, 52(6):446–462, May 2018.

[30] Inbal Nahum-Shani, Shawna N Smith, Bonnie J Spring, Linda M Collins, Katie Witkiewitz, Ambuj Tewari, and Susan A Murphy. Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. *Ann. Behav. Med.*, 52(6):446–462, May 2018.

[31] Noom. Noom launches noom vibe, a wellness community and habit tracker built for the age of glp-1s, June 2024. URL `https://www.noom.com/`. Accessed: 2025-07-27.

[32] OpenAI. Openai python api. `https://github.com/openai/openai-python`, 2024. Accessed: 2025-07-27.

[33] OpenAI. Introducing gpt-4o and more tools to chatgpt free users, May 2024. URL `https://openai.com/index/hello-gpt-4o`. Accessed: 2025-07-27.

[34] Esther K Papies and Petra Hamstra. Goal priming and eating behavior: enhancing self-regulation by environmental cues. *Health Psychol.*, 29(4):384–388, July 2010.

[35] Michele L Patel, Lindsay N Wakayama, and Gary G Bennett. Self-monitoring via digital health in weight loss interventions: A systematic review among adults with overweight or obesity. *Obesity (Silver Spring)*, 29(3):478–499, March 2021.

[36] Nina Rajcic and Jon McCormack. Message ritual: A posthuman account of living with lamp. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, volume 1, pages 1–16, New York, NY, USA, April 2023. ACM.

[37] V. Ramachandran and E. Hubbard. Synaesthesia – a window into perception, thought and language. *Journal of Consciousness Studies*, 8(12):3–34, 2001.

[38] Raspberry Pi Foundation. Raspberry pi 5. `https://www.raspberrypi.com/products/raspberry-pi-5/`, 2023. Accessed: 2025-07-26.

[39] James A Russell. A circumplex model of affect. *J. Pers. Soc. Psychol.*, 39(6):1161–1178, December 1980.

[40] Philipp E Scherer and Joseph A Hill. Obesity, diabetes, and cardiovascular diseases: A compendium. *Circ. Res.*, 118(11):1703–1705, May 2016.

[41] Zev Schuman-Olivier, Marcelo Trombka, David A Lovas, Judson A Brewer, David R Vago, Richa Gawande, Julie P Dunne, Sara W Lazar, Eric B Loucks, and Carl Fulwiler. Mindfulness and behavior change. *Harv. Rev. Psychiatry*, 28(6):371–394, 2020.

[42] Masaaki Shakudo, Misa Takegami, Ai Shibata, Miki Kuzumaki, Takahiro Higashi, Yasuaki Hayashino, Yoshimi Suzukamo, Satoshi Morita, Michio Katsuki, and Shunichi Fukuhara. Effect of feedback in promoting adherence to an exercise programme: a randomized controlled trial. *J. Eval. Clin. Pract.*, 17(1):7–11, February 2011.

[43] Travis Smith, Kelsey Panfil, Carrie Bailey, and Kimberly Kirkpatrick. Cognitive and behavioral training interventions to promote self-control. *J. Exp. Psychol. Anim. Learn. Cogn.*, 45(3): 259–279, July 2019.

[44] Marina Sokolova and Antonio Fernández-Caballero. A review on the role of color and light in affective computing. *Appl. Sci. (Basel)*, 5(3):275–293, August 2015.

[45] Charles Spence, Xiaoang Wan, Andy Woods, Carlos Velasco, Jialin Deng, Jozef Youssef, and Ophelia Deroy. On tasty colours and colourful tastes? assessing, explaining, and utilizing crossmodal correspondences between colours and basic tastes. *Flavour*, 4(1), December 2015.

[46] Jared B Torre and Matthew D Lieberman. Putting feelings into words: Affect labeling as implicit emotion regulation. *Emot. Rev.*, 10(2):116–124, April 2018.

[47] Claire R van Genugten, Melissa S Y Thong, Wouter van Ballegooijen, Annet M Kleiboer, Donna Spruijt-Metz, Arnout C Smit, Mirjam A G Sprangers, Yannik Terhorst, and Heleen Riper. Beyond the current state of just-in-time adaptive interventions in mental health: a qualitative systematic review. *Front. Digit. Health*, 7:1460167, January 2025.

[48] R R Wing, R W Jeffery, L R Burton, C Thorson, K S Nissinoff, and J E Baxter. Food provision vs structured meal plans in the behavioral treatment of obesity. *Int J Obes Relat Metab Disord*, 20(1):56–62, January 1996.

[49] Zhen Wu, Ruoyu Wen, Marcel Zaes Sagesser, Sum Yi Ma, Jiashuo Xian, Wenda Wu, and Tristan Braud. I light U up: Exploring a new emergent narrative paradigm through physical data participation in AI generative experiences. In *SIGGRAPH Asia 2024 Art Papers*, pages 1–9, New York, NY, USA, December 2024. ACM.

[50] Min-Jeong Yang, Steven K Sutton, Laura M Hernandez, Sarah R Jones, David W Wetter, Santosh Kumar, and Christine Vinci. A Just-In-Time adaptive intervention (JITAI) for smoking cessation: Feasibility and acceptability findings. *Addict. Behav.*, 136(107467):107467, January 2023.

[51] Jian Yu, Ling Li, and Wei Zeng. HarmonyWave: Immersive space sound therapy with audio-visual synesthesia. In *Proceedings of the 17th International Symposium on Visual Information Communication and Interaction*, pages 1–8, New York, NY, USA, December 2024. ACM.

[52] Yixuan Zhang and Andrea G Parker. Eat4Thought: A design of food journaling. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–8, New York, NY, USA, April 2020. ACM.

[53] Zheyuan Zhang, Zehong Wang, Tianyi Ma, Varun Sameer Taneja, Sofia Nelson, Nhi Ha Lan Le, Keerthiram Murugesan, Mingxuan Ju, Nitesh V Chawla, Chuxu Zhang, and Yanfang Ye. MOPI-HFRS: A multi-objective personalized health-aware food recommendation system with LLM-enhanced interpretation. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.1*, pages 2860–2871, New York, NY, USA, July 2025. ACM.

[54] Yubo Zhao and Xiying Bao. Narratron: Collaborative writing and shadow-playing of children stories with large language models. In *Adjunct Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, New York, NY, USA, October 2023. ACM.

NeurIPS Paper Checklist

The checklist is designed to encourage best practices for responsible machine learning research, addressing issues of reproducibility, transparency, research ethics, and societal impact. Do not remove the checklist: **The papers not including the checklist will be desk rejected.** The checklist should follow the references and follow the (optional) supplemental material. The checklist does NOT count towards the page limit.

Please read the checklist guidelines carefully for information on how to answer these questions. For each question in the checklist:

- You should answer [Yes] , [No] , or [NA] .
- [NA] means either that the question is Not Applicable for that particular paper or the relevant information is Not Available.
- Please provide a short (1–2 sentence) justification right after your answer (even for NA).

The checklist answers are an integral part of your paper submission. They are visible to the reviewers, area chairs, senior area chairs, and ethics reviewers. You will be asked to also include it (after eventual revisions) with the final version of your paper, and its final version will be published with the paper.

The reviewers of your paper will be asked to use the checklist as one of the factors in their evaluation. While "[Yes] " is generally preferable to "[No] ", it is perfectly acceptable to answer "[No] " provided a proper justification is given (e.g., "error bars are not reported because it would be too computationally expensive" or "we were unable to find the license for the dataset we used"). In general, answering "[No] " or "[NA] " is not grounds for rejection. While the questions are phrased in a binary way, we acknowledge that the true answer is often more nuanced, so please just use your best judgment and write a justification to elaborate. All supporting evidence can appear either in the main paper or the supplemental material, provided in appendix. If you answer [Yes] to a question, in the justification please point to the section(s) where related material for the question can be found.

IMPORTANT, please:

- **Delete this instruction block, but keep the section heading "NeurIPS Paper Checklist"**,
- **Keep the checklist subsection headings, questions/answers and guidelines below.**
- **Do not modify the questions and only use the provided macros for your answers**.

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: **[TODO]**

   Justification: **[TODO]**

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: **[TODO]**

   Justification: **[TODO]**

Guidelines:

- The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
- The authors are encouraged to create a separate "Limitations" section in their paper.
- The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
- The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
- The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
- The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
- If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: **[TODO]**

   Justification: **[TODO]**

   Guidelines:

   - The answer NA means that the paper does not include theoretical results.
   - All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
   - All assumptions should be clearly stated or referenced in the statement of any theorems.
   - The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
   - Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
   - Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: **[TODO]**

   Justification: **[TODO]**

   Guidelines:

   - The answer NA means that the paper does not include experiments.

- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: **[TODO]**

   Justification: **[TODO]**

   Guidelines:

   - The answer NA means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).

- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: **[TODO]**

   Justification: **[TODO]**

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: **[TODO]**

   Justification: **[TODO]**

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.
   - It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
   - For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
   - If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: **[TODO]**

   Justification: **[TODO]**

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.

- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: **[TODO]**

   Justification: **[TODO]**

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: **[TODO]**

    Justification: **[TODO]**

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
    - The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
    - The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
    - If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: **[TODO]**

    Justification: **[TODO]**

    Guidelines:

    - The answer NA means that the paper poses no such risks.

- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: **[TODO]**

    Justification: **[TODO]**

    Guidelines:

    - The answer NA means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
    - If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: **[TODO]**

    Justification: **[TODO]**

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: **[TODO]**

    Justification: **[TODO]**

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: **[TODO]**

    Justification: **[TODO]**

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

    Answer: **[TODO]**

    Justification: **[TODO]**

    Guidelines:

    - The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy (`https://neurips.cc/Conferences/2025/LLM`) for what should or should not be described.