AI at your Fingertips: Wearable Ring as a Low-Friction Interface for Agentic AI

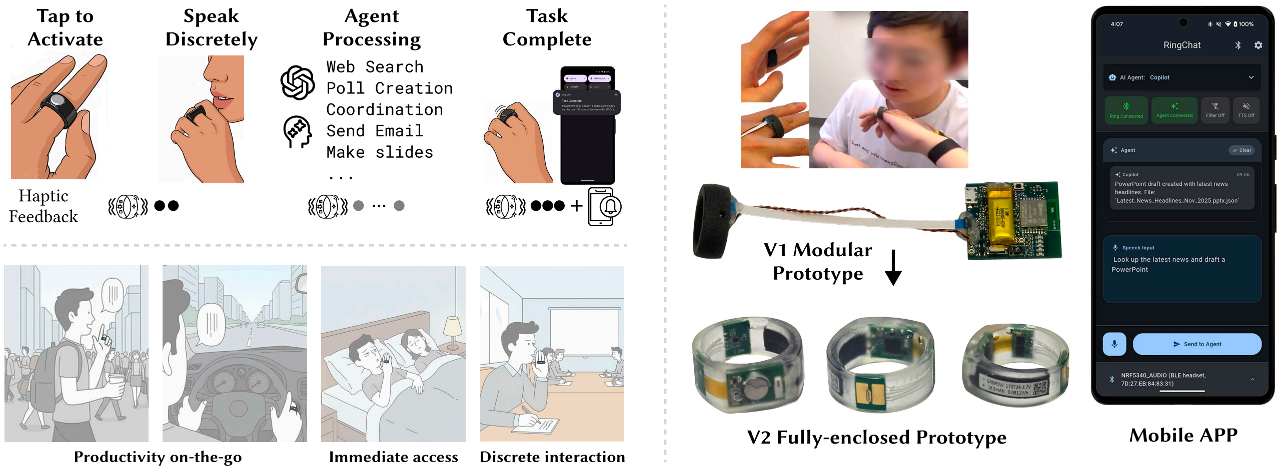
Minghui Zhao*
Electrical Engineering
Columbia University
New York, New York, USA
mz2866@columbia.edu

Judith Amores
Microsoft Research
Cambridge, Massachusetts, USA
judithamores@microsoft.com

Vaishnavi Nattar Ranganathan
Microsoft Research
Microsoft
Redmond, Washington, USA
vnattar@microsoft.com

Xiaofan Jiang
Department of Electrical Engineering
Columbia University
New York, New York, USA
jiang@ee.columbia.edu

Bodhi Priyantha
Microsoft Research
Microsoft
Redmond, Washington, USA
bodhip@microsoft.com



PowerPoint draft created with latest news headlines. File: `Latest_News_Headlines_Nov_2025.pptx.json`

Look up the latest news and draft a PowerPoint

NRF5340_AUDIO (BLE headset, 7D:27:EB:84:83:31)

Figure 1: AI at your Fingertips: A wearable ring designed for discrete, screenless task delegation to LLM agents. (Left) The low-friction interaction loop and use cases: a user (a) taps to activate, (b) whispers a command (e.g., "draft slides"), and (c) receives haptic feedback once the agent completes the workflow. (Right) The hardware design and companion mobile app.

Abstract

While "Agentic AI" can now execute complex, multi-step workflows, human interaction with these agents remains tethered to high-friction screens. We present a technology probe designed to explore the experience of screenless, "fire-and-forget" task delegation. Our system consists of a custom-built wearable ring with touch input and haptic feedback, paired with an agentic pipeline that autonomously recovers from failures to ensure robust execution. Through an exploratory user study ($N = 11$) involving real-world scenarios, we identify design tensions in screenless interaction. Our findings reveal a conflict between delegation and verification: participants valued the efficiency of screenless interaction for simple tasks but lacked the confidence to delegate complex workflows without audio/visual feedback. We further highlight the social tension of public voice input, where users prefer whispering to maintain privacy and social acceptability. This work contributes a functional prototype and initial design implications for the future of always-available AI agents that minimize visual attention.

CCS Concepts

• **Human-centered computing → Ubiquitous and mobile computing**; **Interaction devices**; • **Computer systems organization** → *Embedded and cyber-physical systems.*

Keywords

Agentic AI, Wearable Devices, Technology Probes, Interaction Design, Large Language Models, Screenless Interfaces

---

[*] This research was conducted during an internship at Microsoft Research.

CHI EA '26, Barcelona, Spain

ACM ISBN 979-8-4007-2281-3/26/04
https://doi.org/10.1145/3772363.3798736

ACM Reference Format:
Minghui Zhao, Judith Amores, Vaishnavi Nattar Ranganathan, Xiaofan Jiang, and Bodhi Priyantha. 2026. AI at your Fingertips: Wearable Ring as a Low-Friction Interface for Agentic AI. In *Extended Abstracts of the 2026 CHI Conference on Human Factors in Computing Systems (CHI EA '26), April 13–17, 2026, Barcelona, Spain.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3772363.3798736

1 Introduction

Recent advancements in Large Language Models (LLMs) have expanded the capabilities of AI beyond conversational chat. We are entering the era of "Agentic AI", as these systems are now capable of planning, tool use, and executing multi-step workflows [13, 14, 18]. However, while the backend capabilities of these agents have grown exponentially, the interfaces we use to engage them have not kept pace. Users remain tethered to high-friction, screen-based interactions on smartphones, where delegating a complex task (e.g., "Find a trail for Saturday and email the itinerary to my partner") still requires unlocking a device, navigating apps, and visually verifying outputs. This friction undermines the promise of an "always-available" agent.

To address this, researchers and industries have looked to wearables, which offer persistent, hands-free access that aligns with users' natural mobility. Recognizing this potential, a recent wave of commercial AI wearables, including lapel pins, pendants, and smart glasses, has emerged to untether users from their screens. However, current form factors often struggle to support complex, eyes-free delegation while maintaining social acceptability. Smartwatches replicate the smartphone's visual paradigm on a constrained display, while speaking aloud to smart glasses or pins can be socially awkward in public. Furthermore, voice interfaces are historically designed for simple, synchronous commands (e.g., "turn on the lights"), leaving a gap in our understanding of how to design for asynchronous, complex task delegation without a visual feedback loop.

These limitations point to a broader gap in our understanding of screenless agency and how users initiate, monitor, and trust complex agentic behavior when visual inspection is minimized or unavailable. In particular, it remains unclear what interaction structures, feedback strategies, and on-body affordances are needed to support such delegation in everyday and public contexts. To investigate these questions, we introduce a wearable technology probe and study its use in task delegation scenarios, using a ring form factor as a concrete instantiation of screenless, always-available interaction. This paper makes the following contributions:

1. **A system artifact**: Design and implementation of a wearable ring as a technology probe, with an agentic backend featuring an LLM Supervisor capable of orchestrating tools (web search, email, and calendar) to support robust execution of asynchronous, multi-step tasks and handle AI agent failures.
2. **Empirical insights**: Findings from an exploratory lab study ($N = 11$) simulating realistic scenarios (travel planning, content creation, and finance) that reveal key trade-offs in always-available, non-visual AI interaction, including a confidence gap arising from the absence of visual feedback.
3. **Design implications for future wearable AI agents**: A set of design implications derived from system use and study findings that highlight how feedback, interaction structure, and input modality shape users' comfort, trust, and discretion when delegating tasks to screenless AI in everyday and public contexts.

2 Related Work

Smart Rings and Wearable Input: Prior work on wearability emphasizes that interaction design for on-body interfaces is fundamentally shaped by body location, which mediates functional capabilities, technical constraints, and social interpretation [20]. From this perspective, rings occupy a distinctive position among wearable form factors. Their proximity to the hands affords frequent, low-effort access and fine-grained input, while their small size and continuous wear support persistent availability with minimal disruption to ongoing activity. Compared to wrist-worn or head-worn devices, rings also enable subtle, embodied gestures (such as bringing the hand toward the mouth) that can support discreet voice input without requiring a sustained shift of visual attention. In this work, we adopt the ring not as an optimized interface for task delegation, but as a design probe that foregrounds how on-body affordances influence initiation, feedback, and user confidence when interacting with agentic AI in the absence of a screen.

Smart rings offer a discreet, always-available form factor for ubiquitous computing [17]. Current commercial products primarily fall into two categories: passive sensing rings (e.g., Oura) and voice-enabled rings (e.g., Amazon Echo Loop, Wizpr, Pebble). These voice-enabled rings are currently designed to tackle simple, synchronous commands or note-taking, typically relying on primitive haptic buzzes merely to confirm a button press or indicate that the device is listening, rather than conveying complex system states. In parallel, HCI research has explored rings as active input for other devices through mid-air gestures [5, 10, 21], or directly manipulating the ring itself through touch [3, 8] and buttons [6]. Furthermore, researchers have demonstrated the viability of haptics for notifications [12], navigation cues [7], and rich communication [19]. While the constrained form factors of rings have historically been viewed as a limitation for complex tasks, we believe these low-bandwidth modalities hold significant potential in the AI era. Building directly upon this foundational prior work in wearable input and haptics, our system artifact investigates how carefully structured feedback strategies can be leveraged to provide the transparency and trust required to delegate complex, asynchronous AI workflows without a screen.

Agentic AI and LLM-Based Systems: Recent advances in large language models have enabled agentic AI systems capable of reasoning, planning, and tool use [13, 18]. Despite backend advancements, user interaction remains largely tethered to screens or limited to synchronous, single-turn voice commands [1, 4]. Luger et al. [11] documented a significant "gulf between user expectation and experience" when users interact with conversational agents, finding that users struggle to understand system state and verify action success without visual feedback. Our work extends this literature by examining how this gap manifests when screenless interfaces are paired with complex, multi-step agentic AI rather than simple commands.

3 Design: The Ring Probe

We developed our wearable ring as a technology probe to investigate the boundaries of what screenless, voice-first interaction can support for agentic AI. We chose the ring form factor to support always-available interaction in a location that balances comfort, social acceptability, low attentional demand, and potential for "fire-and-forget" input, consistent with prior wearability research [20]. Our probe constrains feedback modalities to haptic and optional audio and employs three core principles: (1) *minimal friction* via single-handed gestures; (2) *social discretion* for public use; and (3) *persistent availability*.

3.1 Hardware Prototype

We developed two prototypes: a V1 ring paired with a wrist-worn control unit for user studies, and a fully integrated, standalone V2 (Fig. 2). The finger-worn unit integrates a **capacitive touch sensor** for "push-to-talk" activation, a **MEMS PDM Microphone** for voice capture, and an **Eccentric Rotating Mass (ERM) haptic motor** for feedback.

The ERM was generally sufficient for conveying basic binary states, such as success or failure. However, its relatively low fidelity resulted in mixed user experiences when interpreting more complex haptic patterns. Some participants found it challenging to distinguish the distinct vibrations. This difficulty was particularly pronounced in attention-constrained scenarios, such as when participants were occupied with the simulated driving task and not actively focusing on the ring's physical sensations. In V2, we used a dedicated haptic driver (DRV2625) to make the haptic effects sharper, stronger, and more consistent.

The finger-worn unit connects to a wrist-worn unit housing an nRF5340 SoC. Using **Bluetooth LE Audio**, the system achieves 9.2 hours of continuous audio streaming with a 0.3 Wh (80 mAh) battery. A secondary ultra-low-power microcontroller (STM32L071) monitors the touch sensor and manages nRF5340 sleep states to maintain responsiveness while maximizing battery life.

We validated the audio pipeline by measuring Word Error Rate (WER) across distances (Fig. 2(d)). The system maintains a WER $< 5\%$ within 125cm. Since the intended interaction is hand-to-mouth (approx. 10-15cm), the device captures whispered speech robustly, with minor transcription errors correctable by the LLM backend.

3.2 Interaction Loop

To enable intuitive, eyes-free operation with a minimal learning curve, we designed the interaction loop around a baseline haptic vocabulary, grounded in mechanical button metaphors and smartphone conventions, while following prior guidelines [16]. The interaction follows a fire-and-forget model: **(1) Touch**: the user presses the ring thumb pad (*haptic: short buzz*); **(2) Speak**: the user issues a command while holding the pad; **(3) Release**: the audio is transmitted for processing (*haptic: double click*); **(4) Feedback**: the agent executes the task and signals the outcome (*haptic: three short clicks for success, four long buzzes for failure*).

3.3 System Architecture & LLM Supervisor

To enable complex delegation, we constructed a pipeline connecting the ring to an agentic backend (Fig. 3(a)). Voice input is transcribed locally on the paired smartphone using OpenAI's Whisper to ensure privacy. The resulting text command is then transmitted over a WebSocket connection to a remote backend. There, an LLM agent (Visual Studio Code's Copilot in agent mode, powered by GPT-4.1) coordinates task execution via the Model Context Protocol (MCP) [2], which exposes unified interfaces for tools such as browser automation, email, and file management.

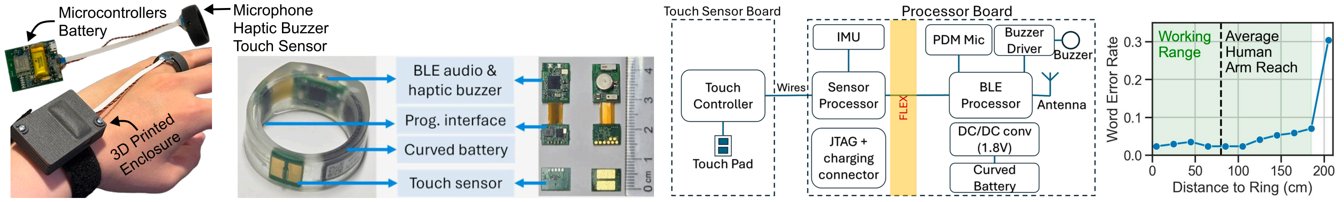
During our pilot study, we found that these LLM agents are often brittle, prone to transcription errors or "premature termination" (stating a plan but failing to execute it fully). To support a true "fire-and-forget" experience where users cannot visually debug errors, we developed a supervisor layer using LangGraph [9] to enforce robustness through three stages, as shown in Fig. 3(b): **(1) Speech Correction**: Before passing the command to the LLM agent, we use an LLM to correct semantic transcription errors (e.g., fixing "media stock" to "NVIDIA stock" based on context). **(2) Execution Monitoring**: The supervisor LLM monitors the agent's tool usage. If the agent stops without completing the plan, the supervisor injects intervention prompts (e.g., *"You stated you would send the email but did not call the tool. Continue execution."*) to recover the workflow transparently. **(3) Outcome Synthesis**: Finally, an LLM translates the complex execution log into a success/failure signal for the ring's haptic feedback, and generates a summary.

While routing the text command from the smartphone to the backend contributes negligibly to system delay, the multi-step supervisor architecture introduces a noticeable latency overhead of 4 to 10 seconds. However, this trade-off enables a highly reliable assistant capable of handling multi-minute workflows (such as researching and running programs) without requiring user intervention.

4 User Study

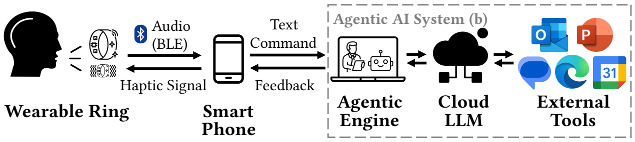
We conducted an exploratory lab study to investigate how users experience confidence, trust, and social dynamics when delegating tasks to a screenless AI agent with varying levels of feedback. The study focused on understanding when and why users require additional feedback to feel comfortable delegating increasingly complex tasks, particularly in mobile or socially situated contexts. The study protocol was reviewed and approved by the institutional review board (IRB), and all participants provided informed consent prior to participation.

We recruited 11 participants (6 female, 5 male; ages 18-54). As the first phase of the study, participants completed a short questionnaire about their current practices for capturing ideas on the go, prior use of voice assistants, familiarity with wearable devices, and attitudes toward always-available voice interaction (see Appendix A). Responses from this questionnaire were used to contextualize participants' expectations and to inform follow-up questions during the subsequent interview and task discussion.

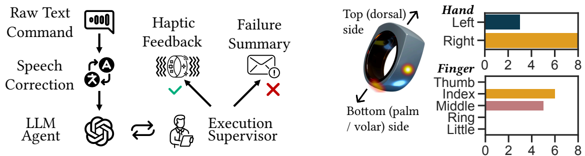


(a) Prototype ring used in user study (V1) **(b) Fully integrated ring (V2)** **(c) Hardware architecture** **(d) Speech working range**

Figure 2: (a-c) Hardware System Overview. (d) Word error rate (WER) versus distance from ring to mouth. The working range (0-185cm, shaded green) maintains low WER and can be corrected by LLM, well beyond natural arm reach (dashed line at ~80 cm).



(a) System pipeline



(b) Supervisor layer **(c) User interaction preference**

Figure 3: (a) The end-to-end flow from ring input to task execution. (b) The supervisor layer logic that intercepts commands to validate intent, correct STT errors, and inject continuation prompts if the agent terminates prematurely. (c) User interaction preference derived from user-marked mockup rings. Left: a heatmap of user preferred ring activation location; right: preferred wearing hand and finger.

4.1 Study Design

We conducted a developmental/progressive exposure design. The system supported three feedback modalities that were progressively revealed during each task as opposed to a within-subjects comparison. Haptic-only feedback provided vibration-based acknowledgments indicating that the system had received and was processing a command. Audio feedback added spoken confirmations and progress cues. Finally, visual feedback revealed the completed task output on a screen, serving as a baseline comparison to conventional interaction.

This staged interaction was intentionally ordered rather than randomized. Rather than treating feedback modalities as independent experimental conditions, our goal was to examine how users' confidence, expectations, and perceived sense of control evolved as progressively richer feedback became available. Randomizing modality order would have disrupted the cumulative nature of the delegation experience and obscured how trust is gradually established in screenless interaction.

4.2 Procedure and Tasks

We conducted 60-minute sessions that began with a contextual interview and system training. Participants were introduced to the prototype and trained on its press-and-hold interaction and vibration cues. The main study consisted of three scenario-based tasks designed to vary in complexity and perceived risk, while simulating attention-constrained contexts such as walking or driving. The session concluded with a semi-structured interview focusing on perceived value, moments of delight or frustration, social comfort, and preferences related to feedback and form factor.

The tasks in the main study included: **(1) Capture (*simple task*)**: summarizing meeting notes and emailing them to themselves; **(2) Coordination (*moderate*)**: researching hiking trails, creating a poll, and emailing it to friends; **(3) Creation (*complex*)**: retrieving stock data, analyzing trends using code, and generating a PowerPoint presentation.

For each task, participants first interacted using haptic-only feedback and rated their confidence in the system's performance on a 7-point Likert scale. The same interaction was then replayed with audio feedback enabled, followed by a final reveal of the visual output. This staged interaction enabled us to identify when additional feedback altered users' confidence, expectations, and perceived sense of control.

5 Results

Our thematic analysis revealed three critical tensions for the user experience of screenless agentic AI: the gap between haptic ambiguity and user confidence, the paradox of one-shot delegation, and the social dynamics of public use.

5.1 The Confidence Gap: Haptics vs. Audio

The most prominent finding was a fundamental *confidence gap* rooted in insufficient feedback. While participants appreciated the discreet nature of haptics, they struggled to distinguish patterns during complex workflows. As P5 noted, *"I can tell it heard me, but I can't tell the difference between the finishing buzz and the intermediate buzz."* This ambiguity created anxiety; users knew the agent was *active*, but not if it was *succeeding*.

Fig. 4 quantifies this pattern. With haptic feedback alone, participants were fairly confident their command was submitted ($M = 6.5, 6.4, 5.7$ for simple, moderate, and complex tasks) but far less confident it was completed correctly ($M = 4.5, 3.7, 3.7$). The drop was

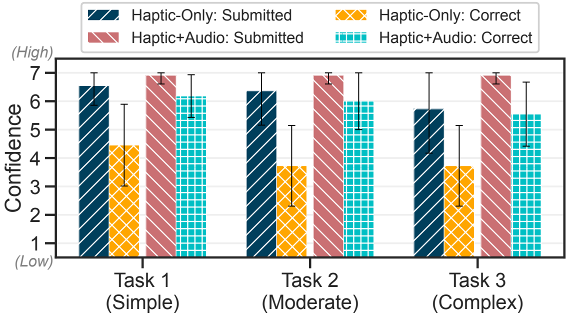


Figure 4: Self-reported confidence across feedback modalities and task complexity (N = 11, 7-point scale). "Submitted" = confidence the command was received; "Correct" = confidence it was completed correctly. Error bars show ±1 SD, clipped to scale bounds.

most pronounced for tasks with external consequences. P2 noted, *"Sending something to someone else is risky. If it's not my friends but my manager, I would be very worried."* Haptic distinguishability also degraded under divided attention; P4 found that while simulating driving, *"I cannot distinguish between [the vibration patterns]."*

Audio as the Confidence Bridge. Adding audio raised correctness confidence by roughly 2 points on average (Fig. 4), though complex tasks still lagged (M=5.5 vs. 6.2 for simple). Audio allowed users to verify that the agent had correctly interpreted their intent (e.g., looking up "AllTrails"). P10 remarked, *"Hearing it say 'task complete' and specify **which** tasks were completed makes me feel pretty confident."* Even when the agent hit a tool error, hearing it explain recovery was reassuring. However, this introduced a new tension: verbosity. P7 noted, *"I don't want all those updates. If I'm driving, I don't want to hear all the tool use commands."* This points to a need for adaptive verbosity (concise for success, detailed for failure).

User-Designed Haptics. When asked to design ideal feedback, responses fell into two distinct models: *symbolic counts* (1 buzz=listening, 3=success) and *temporal dynamics*. P5 suggested a "heartbeat" that speeds up as task completion nears, mirroring visual progress bars. Some example participant sketches and quotes are included in Appendix B.

5.2 The One-Shot Interaction Paradox

We observed a sharp divergence in usability based on task type. **Success in "Capture and Delegate."** The ring excelled at short, asynchronous tasks. P3 praised it for *"quick, directive tasks"* like adding reminders or capturing thoughts. The "fire-and-forget" model shone here; P3 described waking up to capture a thought: *"I can make a quick note and go back to sleep; I don't wear my watch to bed, but the ring I can deal with."*

Friction in Complex Coordination. Conversely, multi-step tasks (e.g., planning a hike) exposed the brittleness of one-shot commands. Participants found it cognitively taxing to formulate a perfect prompt in a single breath. P7 argued, *"For something like this, I would want to have a conversation; it has to be iterative."* Users did not want to program the agent; they wanted to collaborate with it. The agent did complete all tasks, sometimes with the supervisor recovering from mid-task failures (e.g., a hallucinated URL corrected via fallback search). But complex tasks still received lower satisfaction even after visual verification (M=5.5 vs. 6.3 for simple), suggesting the problem is one-shot specification, not execution alone.

5.3 Social Acceptability & Form Factor

The Whisper Paradigm. Participants preferred the ring over smart glasses or watches for public use because the "hand-to-mouth" gesture allowed for whispering. P3 noted, *"It's a sweet spot... Social norms are not violated."* However, audio *output* in public remained a barrier. P9 emphasized, *"I don't necessarily want people around me hearing it,"* suggesting a need for private audio modalities.

Wearability. While participants valued the ring's persistent availability, its sizing and thickness were barriers. Fig. 3(c) shows a varied preference on activation region, with a split between thumb side (N = 6) and palm side (N = 5).

6 Discussion: Design Implications

Our probe demonstrates that removing the screen does not simply remove friction; it shifts the burden of verification to other modalities. Based on our findings, we propose four design implications for future wearable agents:

(1) Adaptive Feedback Granularity: Feedback must scale with task complexity. Our confidence data (Fig. 4) shows that haptic-only feedback left correctness confidence near the scale midpoint even for simple tasks, and adding audio closed much of the gap. A simple reminder may need only a binary "success" vibration, but a complex workflow (e.g., stock analysis) requires a "heartbeat" to signal ongoing processing (as suggested by P5) and a summarized audio report upon completion.

(2) Conversational Scaffolding: The "fire-and-forget" model fails for ambiguous tasks. Screenless agents should support a hybrid mode: accepting a one-shot trigger, but proactively initiating a clarification turn (e.g., *"I found three trails; do you want the hardest or easiest?"*) rather than guessing.

(3) The "Ring-to-Ear" Gesture: To solve the social tension of public audio feedback, we propose a gesture-gated audio mode. The ring should remain silent by default (haptic only) but switch to low-volume audio output when the user raises their hand to their ear, mimicking the privacy of a phone call without the device.

(4) Transparent Failure Recovery: Trust in screenless agents is fragile. In our study, the supervisor recovered from mid-task failures in several sessions (e.g., correcting a hallucinated URL via fallback search), and participants who heard these recoveries via audio found them reassuring. The system should optionally communicate this resilience (e.g., *"The web search failed, so I checked the cached data instead"*) rather than hiding it.

7 Conclusion

We presented a technology probe investigating the boundaries of screenless interaction for Agentic AI. By combining a custom wearable ring with a robust "LLM Supervisor" backend, we enabled users to delegate complex, multi-step tasks without a screen. Our exploratory study (N = 11) revealed a critical "confidence gap":

while users value the efficiency of "fire-and-forget" delegation, they lack the trust to leave high-stakes tasks to an invisible agent without richer feedback loops. We contribute this system, empirical insights on screenless delegation, and a set of design implications to move wearable AI beyond simple commands toward genuine, trustworthy task delegation.

References

[1] Tawfiq Ammari, Jofish Kaye, Janice Y Tsai, and Frank Bentley. 2019. Music, search, and IoT: How people (really) use voice assistants. *ACM Transactions on Computer-Human Interaction (TOCHI)* 26, 3 (2019), 1–28.

[2] Anthropic. [n. d.]. *Model Context Protocol.* https://github.com/modelcontextprotocol

[3] Daniel Ashbrook, Patrick Baudisch, and Sean White. 2011. Nenya: subtle and eyes-free mobile input with a magnetically-tracked finger ring. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* 2043–2046.

[4] Frank Bentley, Chris Luvogt, Max Silverman, Rushani Wirasinghe, Brooke White, and Danielle Lottridge. 2018. Understanding the long-term use of smart speaker assistants. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (2018), 1–24.

[5] Taizhou Chen, Tianpei Li, Xingyu Yang, and Kening Zhu. 2023. Efring: Enabling thumb-to-index-finger microgesture interaction through electric field sensing using single smart ring. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 4 (2023), 1–31.

[6] Satabdi Das, Arshad Nasser, and Khalad Hasan. 2023. Fingerbutton: Enabling controller-free transitions between real and virtual environments. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR).* IEEE, 533–542.

[7] Jan BF Van Erp, Hendrik AHC Van Veen, Chris Jansen, and Thomas Dobbins. 2005. Waypoint navigation with a vibrotactile waist belt. *ACM Transactions on Applied Perception (TAP)* 2, 2 (2005), 106–117.

[8] Yizheng Gu, Chun Yu, Zhipeng Li, Zhaoheng Li, Xiaoying Wei, and Yuanchun Shi. 2020. Qwertyring: Text entry on physical surfaces using a ring. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4 (2020), 1–29.

[9] LangChain Inc. 2024. *LangGraph: Build Resilient Language Agents as Graphs.* https://github.com/langchain-ai/langgraph

[10] Chen Liang, Chun Yu, Yue Qin, Yuntao Wang, and Yuanchun Shi. 2021. DualRing: Enabling subtle and expressive hand interaction with dual IMU rings. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 1–27.

[11] Ewa Luger and Abigail Sellen. 2016. " Like Having a Really Bad PA" The Gulf between User Expectation and Experience of Conversational Agents. In *Proceedings of the 2016 CHI conference on human factors in computing systems.* 5286–5297.

[12] Karon E MacLean. 2008. Haptic interaction design for everyday interfaces. *Reviews of Human Factors and Ergonomics* 4, 1 (2008), 149–194.

[13] Timo Schick, Jane Dwivedi-Yu, Roberto Dessì, Roberta Raileanu, Maria Lomeli, Eric Hambro, Luke Zettlemoyer, Nicola Cancedda, and Thomas Scialom. 2023. Toolformer: Language models can teach themselves to use tools. *Advances in Neural Information Processing Systems* 36 (2023), 68539–68551.

[14] Significant Gravitas. [n. d.]. *AutoGPT.* https://github.com/Significant-Gravitas/AutoGPT

[15] Gareth Terry, Nikki Hayfield, Victoria Clarke, Virginia Braun, et al. 2017. Thematic analysis. *The SAGE handbook of qualitative research in psychology* 2, 17-37 (2017), 25.

[16] Yulin Wang, Barbara Millet, and James L Smith. 2016. Designing wearable vibrotactile notifications for information communication. *International Journal of Human-Computer Studies* 89 (2016), 24–34.

[17] Zeyu Wang, Ruotong Yu, Xiangyang Wang, Jiexin Ding, Jiankai Tang, Jun Fang, Zhe He, Zhuojun Li, Tobias Röddiger, Weiye Xu, et al. 2025. Computing with Smart Rings: A Systematic Literature Review. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 3 (2025), 1–54.

[18] Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik R Narasimhan, and Yuan Cao. 2022. React: Synergizing reasoning and acting in language models. In *The eleventh international conference on learning representations.*

[19] Gareth W Young, Néill O'Dwyer, Mauricio Flores Vargas, Rachel Mc Donnell, and Aljosa Smolic. 2023. Feel the music!—audience experiences of audio–tactile feedback in a novel virtual reality volumetric music video. In *Arts*, Vol. 12. MDPI, 156.

[20] Clint Zeagler. 2017. Where to wear it: functional, technical, and social considerations in on-body location for wearable technology 20 years of designing for wearability. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers.* 150–157.

[21] Cheng Zhang, Xiaoxuan Wang, Anandghan Waghmare, Sumeet Jain, Thomas Ploetz, Omer T Inan, Thad E Starner, and Gregory D Abowd. 2017. FingOrbits: interaction with wearables using synchronized thumb movements. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers.* 62–65.

A Initial Questionnaire

Prior to the lab session, participants completed an initial questionnaire to establish baseline context about their current practices and attitudes. The questionnaire captured: (1) methods for capturing ideas and reminders when away from computers (e.g., phone notes apps, voice dictation, paper, memory); (2) frequency of voice assistant usage and satisfaction ratings (1-7 scale); (3) primary devices for voice interaction; (4) current wearable device ownership and usage patterns; (5) comfort level with always-listening wake-word devices (1-7 scale); and (6) open-ended hopes and concerns about always-available voice command devices. These baseline measures provided context for interpreting participants' reactions to our ring prototype and helped personalize the initial contextual interview.

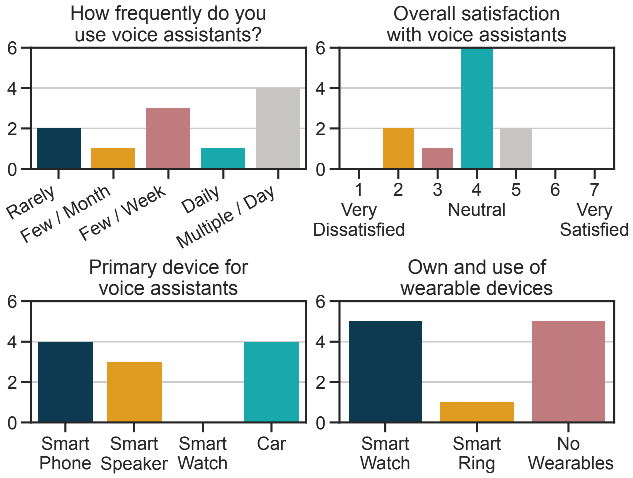


Figure 5: Results from the initial questionnaire.

The initial questionnaire revealed diverse prior experience with voice assistants and wearable devices (Fig. 5). Most participants used voice assistants regularly (8 of 11 at least weekly), with diverse satisfaction levels. Smartphones ($N = 4$) and cars ($N = 4$) were the primary voice interaction contexts, followed by smart speakers ($N = 3$). Half owned smartwatches ($N = 5$), while only one participant owned a smart ring.

B Participant-Designed Vibration Languages

To understand user expectations for haptic communication, we asked participants to sketch their ideal vibration patterns for different system states. The resulting designs revealed both the diversity of individual preferences and common principles for effective haptic feedback.

Participants' designs fell into three categories: **minimal patterns** using simple counts (e.g., 1 buzz = listening, 2 buzzes = working, 3 buzzes = complete); **temporal patterns** using rhythm and duration variations; and **sophisticated patterns** incorporating tone, frequency, and intensity changes to create rich vocabularies.

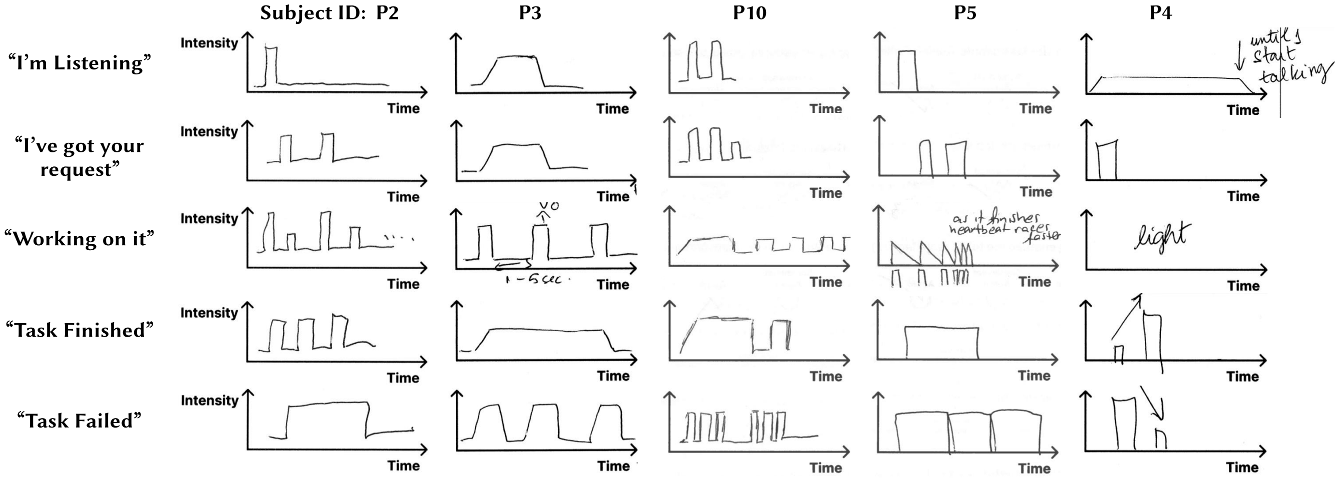


Figure 6: Sample vibration patterns sketched by participants.

A subset of vibration pattern sketches by participants is shown in Fig. 6.

P5 proposed dynamic feedback: *“I have a new idea: you should start speeding up the heartbeat as it gets closer to finishing, so that I have an idea of how far it’s got, even without hearing it. Or if it errors out, it could be a different heartbeat.”* P4 emphasized duration-based adaptation: *“For simple tasks, start and finish should be good. But if it’s a task that’s gonna take five minutes, I might need an intermittent... or it buzzing every two seconds.”* P1 suggested feedback should scale with task duration: *“If I’m going to send it off on some 30-minute mission, then... I may not even want it to buzz at all. I may just want to go look at the end.”* P6 designed detailed specifications with different patterns for each state: listening (short buzz gradually increasing intensity), working (different tone/frequency), success (quick sharp upward buzz), and failure (longer low-tone buzz).

However, the sophistication of designed patterns often contrasted with participants’ actual ability to distinguish subtle haptic differences in practice, highlighting the gap between desired and perceivable haptic complexity. This tension underscores the challenge of designing haptic vocabularies that are both expressive enough to convey rich information and simple enough to be reliably perceived across diverse users and contexts.

C Use of Generative AI

This research involved the use of generative AI tools in the following specific capacities:

User Study Transcript Summarization: Claude 4.5 Sonnet was used to generate preliminary summaries of user study interview transcripts to support the initial stages of thematic analysis. The AI was prompted to summarize key themes, concerns, and observations from each participant interview. These AI-generated summaries served only as a starting point to aid researchers in identifying potential patterns; all final thematic codes, theme definitions, and interpretations were determined independently through manual qualitative analysis by the research team following established thematic analysis procedures [15].

Illustrative Figure Creation: Some illustrations in the teaser figure (Figure 1) were created using GPT-image-1 (OpenAI) and gemini-2.5-flash-image (Google). The images were generated from text prompts describing the concept of the ring in various use cases. The generated images were then composed and annotated by the authors to create the final teaser figure.

Writing Assistance: Claude 4.5 Sonnet was used for language editing and refinement of author-written text in portions of the manuscript. The AI was prompted to improve clarity, grammar, and conciseness of existing sentences and paragraphs. All substantive content, arguments, technical descriptions, and interpretations were originally authored by the research team; the AI was used solely for linguistic polishing and did not generate original content or ideas.

All other aspects of this work, including study design, data collection, quantitative analysis, system implementation, and interpretation of findings, were conducted by the authors without AI assistance.