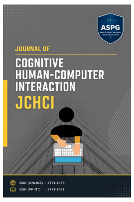
Journal of Cognitive Human-Computer Interaction

♦ ISSN: 2771-1463 Vol: 11, No. 01, 2026 ♦ PP. 20–28
Article DOI: https://doi.org/10.54216/JCHCI.110104



FULL LENGTH ARTICLE

# Cross-Modal Memory Support in Visually Demanding Environments: A Controlled Study of Haptic Pulses and Spatial Audio Cues for Reducing Prospective Memory Failures During Multitasking

**A. Nithya**[1,*] **B. Chitra**[2] **V. Sathya Preiya**[3]

[1] Assistant Professor, Department of Computer Science and Engineering, St. Joseph's College of Engineering, India
[2] Assistant Professor, Department of Artificial Intelligence and Data Science, Panimalar Engineering College, India
[3] Associate Professor, Department of Computer Science and Engineering, S.A. Engineering College, Chennai, India

**Emails:** nithyashree.a@gmail.com · chitra14tech@gmail.com · sathyapreiya@yahoo.com

Received: September 24, 2025 Revised: November 05, 2025 Accepted: December 09, 2025 ⋆ Corresponding author

## ABSTRACT

When people are immersed in a visually demanding task, the attentional resources required to monitor the environment for cues that should trigger a remembered intention are frequently captured by the primary task, causing prospective memory failures that range from the inconvenient to the safety-critical. This problem is pervasive in modern work environments in which digital interfaces compete continuously for visual attention, yet the overwhelming majority of reminder and notification systems rely on the same visual channel that is already congested. This paper reports a controlled user study examining whether carefully designed haptic and spatial audio cues can compensate for this visual saturation and restore prospective memory performance without substantially increasing cognitive burden. Thirty-two participants completed a counterbalanced within-subjects protocol in which they performed primary cognitive tasks—document editing on a virtual desktop and navigating in a driving simulation—while managing a set of time-critical intentions delivered through four reminder conditions: visual-only, haptic-only, spatial audio only, and the combined haptic-plus-audio channel. The study measures prospective memory hit rate, task-switching errors, cue response latency, and multidimensional subjective workload across both scenarios and all four conditions. Results consistently favour the combined modality, which produces substantially fewer memory failures and lower reported workload than any single channel, while individual differences in baseline workload predict the magnitude of benefit from non-visual cueing. These findings carry direct implications for the design of ambient notification systems in high-demand professional and safety-critical environments.

**Keywords:** Prospective memory ▪ Haptic feedback ▪ Spatial audio ▪ Multitasking ▪ Cognitive load ▪ Task interruption ▪ Vibrotactile ▪ Reminder systems ▪ Human-computer interaction

## 1. INTRODUCTION

Remembering to perform an intended action at the right moment in the future — prospective memory, in the cognitive science literature — is a mundane but consequential cognitive process [1]. When attention is fractured across multiple simultaneous demands, as it routinely is in modern office, transportation, and healthcare work, the monitoring process that underlies prospective memory competes directly with the resources required by the ongoing primary task, and one or the other must yield [2]. The result is the familiar experience of forgetting to send an email, missing a medication reminder, or neglecting to act on a navigation prompt while engaged in a demanding secondary activity.

Most current reminder and notification systems aggravate this problem by inserting their alerts into the same visual channel that the primary task is consuming. Pop-up notifications on screens, dashboard warning lights, and colour-coded calendar reminders all require the user to divert foveal attention to be perceived, which either interrupts the primary task at a cost to performance or, more dangerously, is ignored because the primary task cannot be interrupted [3]. The cognitive architecture of multitasking under visual load does not offer a simple solution within the visual modality; reaching users through alternative sensory channels is a principled design response.

Haptics and spatial audio represent two well-studied pathways for conveying information without occupying the visual channel. Vibrotactile feedback delivered through a wrist-worn device is perceived rapidly and with low interference under cognitive load, provided the encoding pattern is sufficiently discriminable from background sensation [4]. Spatially rendered audio, using head-related transfer function (HRTF) processing to position sounds in three-dimensional space, captures auditory spatial attention even when the listener is visually engaged, and brief spatialized cues have been shown to convey coarse directional information that orients attention without requiring the listener to interpret complex linguistic content [5]. Yet controlled comparisons of these channels within a common prospective memory paradigm, across ecologically valid dual-task scenarios, remain rare.

The present study addresses this gap with the following specific contributions:

- A between-channel comparison across four cueing conditions—visual-only, haptic-only, spatial audio only, and combined—using a within-subjects design that controls for individual differences in prospective memory ability.
- Evaluation across two ecologically representative scenarios (office virtual desktop and driving simulation) to assess the generalisability of any cueing advantage across visual load contexts.
- Systematic measurement of the interaction between baseline workload and the magnitude of benefit from non-visual cueing, providing guidance on which users are most likely to benefit from cross-modal notification design.
- A haptic encoding scheme in which pulse rhythm, amplitude, and duration encode four priority levels, validated against the discriminability criteria of Asplund et al. [4].

Figure 1 illustrates the conceptual architecture of the approach. The paper is structured as follows. Section 2 reviews the theoretical background and related work. Section 3 de-scribes the system design. Section 4 presents the user study methodology. Section 5 reports the results. Section 6 dis-cusses implications and limitations. Section 7 concludes.

## 2. BACKGROUND AND RELATED WORK

### 2.1 Prospective Memory Under Cognitive Load

Prospective memory research has consistently distinguished two components: the retrospective component (remembering *what* to do) and the prospective component (remembering *when or in what context* to do it). The latter relies on an active monitoring process or, when monitoring is too costly, on a spontaneous retrieval triggered by target features in the environment [1]. Dismukes [2] emphasised that prospective memory failures in workplace contexts are disproportionately associated with high-workload periods: when ongoing task demands are elevated, monitoring is suspended and spontaneous retrieval requires the environment to provide a sufficiently distinct cue.

Cantarella et al. [5] systematically manipulated both cognitive load and the focality of the prospective memory target (how attention-capturing the target was relative to the ongoing task), finding that load and low focality interacted synergistically to impair prospective memory — precisely the condition created when a visual reminder is presented to a user already visually saturated by a primary task. Meier and Zimmermann [6] identified three separable load factors (prospective load, retrospective load, ongoing task load) and showed that ongoing task load is the most disruptive to prospective memory performance, suggesting that reminders delivered through a modality not competing with the ongoing task will be disproportionately effective.

### 2.2 Haptic Feedback for Ambient Notification

Vibrotactile feedback has been employed as an ambient notification channel since the earliest work on tactile displays. The fundamental tradeoff between discriminability and intrusiveness was characterised by Asplund et al. [4], who established that rhythm and temporal structure are the most reliably encoded properties under cognitive load, whereas amplitude alone produces the lowest interference with a concurrent task. Recent work at CHI [7] has examined the haptic user experience construct more broadly, identifying that haptic feedback is evaluated differently from other modality feedback and that ecological validity—how natural the haptic event feels in context—is a primary driver of user acceptance. Moon et al. [8] specifically investigated the role of redundant haptic information during task interruption and found that pairing haptic stimuli with visually congruent information significantly reduced the cost of interruption on a primary cognitive task, an effect that motivates the design of context-aware haptic reminders in the present study.

### 2.3 Spatial Audio for Attention Guidance

Spatially rendered audio exploits the auditory system's pre-attentive processing of location information to redirect attention without requiring explicit voluntary engagement. Hirsch et al. [3] recently established that the timing of interruptions relative to cognitive subtask boundaries moderates their disruptiveness, a finding relevant to the design of opportunistic reminder delivery—reminders presented at low-load moments in the primary task cycle will be perceived more readily and will less impair primary task resumption. The use of spatial audio in automotive and extended reality (XR) contexts

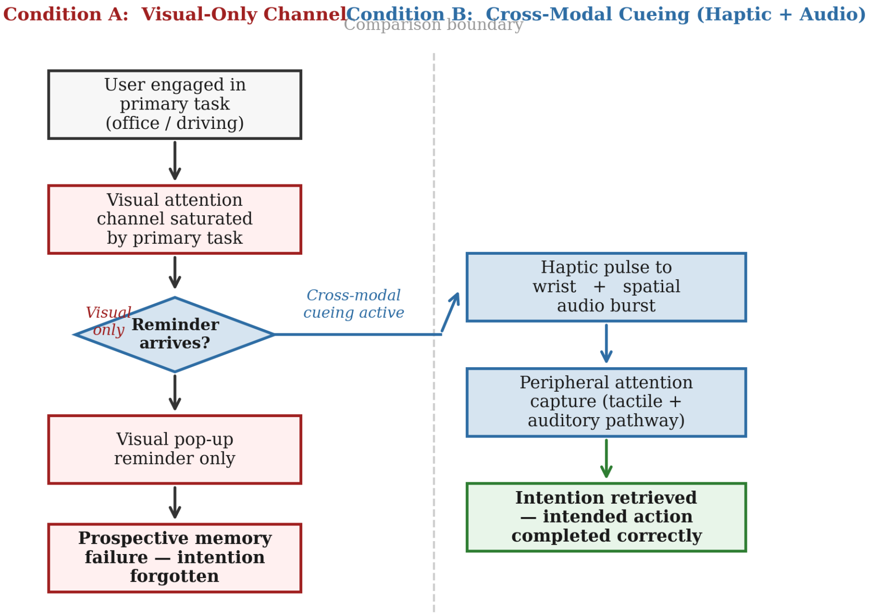


**Figure 1.** Conceptual framework of the present study. When the primary task saturates the visual channel, non-visual cues delivered through haptic and spatial audio pathways trigger intention retrieval without competing for visual processing resources. The baseline visual-only condition represents the degraded retrieval pathway (dashed arc).

for attention guidance has been explored in recent work, establishing that brief spatialized cues can convey coarse directional information reliably even when the listener is visually engaged in a demanding primary task, and that calibration training as short as a few minutes improves spatial audio perception significantly.

### 2.4 Workload Measurement in HCI Studies

The NASA Task Load Index (NASA-TLX) [9] remains the most widely deployed multidimensional workload instrument in HCI research, providing six subscale scores (mental demand, physical demand, temporal demand, performance, effort, and frustration) that can reveal differential effects across cueing conditions even when the overall workload difference is modest. Longo et al. [10] conducted a systematic review of cognitive workload measurement in HCI, identifying NASA-TLX as the dominant choice and noting that composite TLX scores mask important subscale interactions when comparing conditions that differ in their primary versus secondary task demands — a consideration addressed in the present analysis by examining all six subscales separately.

## 3. SYSTEM DESIGN

### 3.1 Haptic Wristband

The haptic component consists of a wrist-worn device housing five eccentric rotating mass (ERM) vibrotactile actuators arranged in a semicircle on the dorsal surface of the wrist. The spatial arrangement allows directional encoding: actuators can fire independently or in sequences that convey left-to-right or centre-to-edge propagation. Four priority levels are encoded through distinct pulse patterns informed by the discriminability guidelines of Asplund et al. [4]: Priority 1 (urgent) uses a rapid triple-pulse pattern; Priority 2 (high) uses a double pulse with a short inter-pulse gap; Priority 3 (medium) uses a single long pulse; and Priority 4 (low) uses a single short pulse with a 500 ms onset delay. All patterns are delivered at an amplitude of approximately 1.2 g peak acceleration, measured at the wrist surface.

### 3.2 Spatial Audio Renderer

Spatial audio reminders are delivered through over-ear noise-attenuating headphones using personalised HRTF convolution processing. Each reminder is rendered as a broadband 200 ms burst positioned at one of eight pre-defined azimuth angles $(0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ, 315^\circ)$ in the horizontal plane at ear level. The position encodes reminder category: front-centre (0°) for calendar events, right-front (45°) for document save reminders, left-front (315°) for navigation waypoints, and rear-centre (180°) for low-priority system notifications. Pitch (500 Hz versus 880 Hz) further encodes urgency, allowing up to eight distinct reminder identities through the combination of position and pitch. HRTF personalisation was approximated using the subject's pinna measurements following a standardised template; full individual HRTF measurement was outside the scope of this study.

### 3.3 Integration Architecture and Trigger Logic

The system software runs on the primary task computer and coordinates reminder delivery across both channels. An integrated eye-tracking module estimates gaze density (the proportion of the last 3 seconds with gaze within the primary task region) as a proxy for visual load. Reminders are scheduled using a context-aware delivery algorithm that preferentially fires reminders during naturally occurring low-load periods—subtask boundaries, brief fixation pauses, or sentence completion events in the document editing task. If a scheduled reminder is not delivered within a 10-second window (because no low-load period occurred), it is delivered unconditionally to avoid missed intentions. Figure 2 shows the system hardware and haptic encoding scheme.

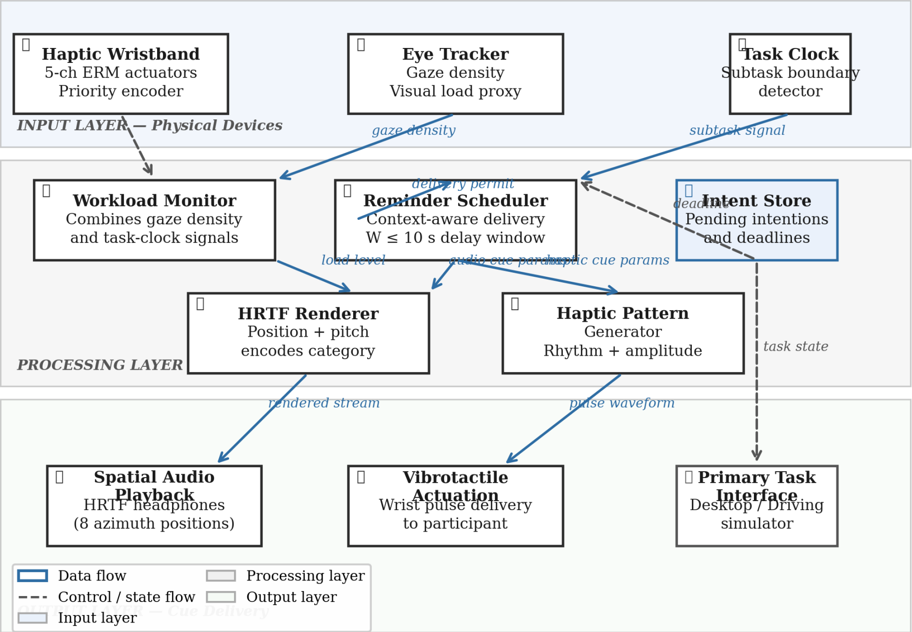


**Figure 2.** (a) System component diagram: the haptic wristband, eye tracker, and spatial audio renderer are coordinated by a reminder scheduler that estimates visual load from gaze density before triggering each reminder. (b) The four haptic encoding patterns for priority levels 1–4, shown as pulse amplitude over time.

**Table 1.** Participant demographics and self-reported technology familiarity (5-point Likert scale; 1 = not at all familiar, 5 = very familiar).

| Characteristic | Value | Range / Note |
|---|---|---|
| *N* (total) | 32 | 16F, 16M |
| Age (years) | 29.4 ± 6.1 | 21–44 |
| Handedness | 28R / 4L | right / left |
| Desktop use (hrs/day) | 6.2 ± 2.1 | 2–11 |
| Driving experience | 27 licensed | 5 non-drivers (office only) |
| Smartwatch familiarity | 3.1 ± 1.0 | 1–5 |
| 3D-audio familiarity | 2.4 ± 1.1 | 1–5 |
| Self-rated memory | 3.3 ± 0.9 | 1–5 |

## 4. USER STUDY

### 4.1 Participants

Thirty-two adults (16 women, 16 men; $M_{\text{age}} = 29.4$ years, $SD = 6.1$; range 21–44) were recruited through a university participant pool. All reported normal or corrected-to-normal vision, normal hearing, and no history of neurological or vestibular disorders. Participants with prior experience using wrist-worn haptic devices beyond standard smartwatch notifications were excluded. Table 1 summarises the sample characteristics. The study received institutional ethics board approval; all participants provided written informed consent and received course credit or a £10 gift voucher as compensation.

### 4.2 Design and Conditions

The study used a 4 × 2 within-subjects design, crossing four *Cueing Condition* levels with two *Scenario* levels. The cueing conditions were: **Visual-Only** (on-screen pop-up reminder only); **Haptic-Only** (wristband pulse only, no visual alert); **Audio-Only** (spatialized audio burst only, no visual alert); and **Combined** (simultaneous haptic pulse and spatial audio burst, no visual alert). The two scenarios were an **Office** task (virtual desktop document editing with concurrent calendar management) and a **Driving** task (city navigation simulation with route-deviation reminders). Condition order within each scenario was counterbalanced using a balanced Latin square; scenario order was counterbalanced across participants.

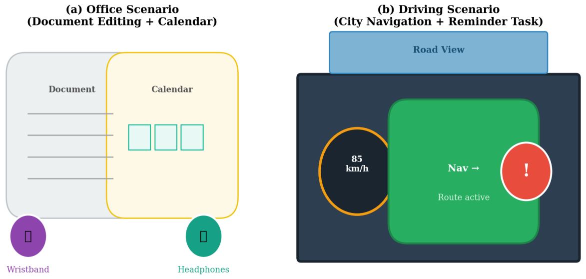


**Figure 3.** (a) Office scenario: participants edit a document while monitoring a calendar for scheduled actions. (b) Driving scenario: participants navigate a city environment while monitoring for route-deviation and in-vehicle reminder cues. In both scenarios, the haptic wristband and HRTF headphones are active in the non-visual conditions.

### 4.3 Haptic Pattern Pre-Test

Before the main study, a brief haptic discriminability pre-test was administered to all participants to confirm that the four priority-level patterns were perceptually distinct under a light ongoing cognitive task. Participants wore the haptic wristband and performed a simple letter-monitoring task (detecting the letter *X* in a continuous stream) while receiving randomised haptic patterns in blocks of 20. They were asked to identify the received pattern (forced choice from four labelled cards). The pre-test was repeated until the participant achieved 90% correct identification on two consecutive blocks, or a maximum of five blocks. Table 2 presents pre-test accuracy and mean trials to criterion.

Overall accuracy was 92.3% ($SD = 5.1\%$), with Priority 4

**Table 2.** Haptic pattern pre-test: mean identification accuracy (%) and mean trials to 90% criterion, by priority level.

| Priority Level | Accuracy (%) | Trials to Criterion |
|---|---|---|
| P1 — Urgent (rapid triple) | $94.2 \pm 4.1$ | $1.8 \pm 0.7$ |
| P2 — High (double gap) | $91.8 \pm 5.3$ | $2.1 \pm 0.9$ |
| P3 — Medium (long single) | $93.6 \pm 4.7$ | $1.9 \pm 0.8$ |
| P4 — Low (short delayed) | $89.4 \pm 6.2$ | $2.4 \pm 1.1$ |
| **Overall** | $\mathbf{92.3 \pm 5.1}$ | $\mathbf{2.1 \pm 0.9}$ |

(the shortest pattern) being the least discriminable. All 32 participants reached the 90% criterion within the five-block limit (mean blocks to criterion: 2.1), confirming adequate learnability of the encoding scheme before commencing the main study.

### 4.4 Tasks and Procedure

**Office scenario.** Participants edited a multi-section document while simultaneously monitoring a shared calendar for four types of events: send a draft, move a meeting, flag a message, and set a reminder. Each session lasted 15 minutes and contained 20 prospective memory targets (five per type). A target was considered a hit if the correct action was performed within 30 seconds of the cue onset.

**Driving scenario.** Participants drove a simulated urban route at a target speed of 50 km/h using a steering wheel and pedal set, while monitoring for four action categories: take the next turn, activate a parking sensor, call a contact, and report a road hazard. Each drive lasted 15 minutes and contained 20 prospective memory targets. A target was a hit if the action was initiated within 5 seconds of target onset (reflecting the faster time-criticality of the driving context).

Participants completed a 10-minute training session for each scenario type before the study began. Between conditions within each scenario, a 5-minute rest break was provided. After each 15-minute block, participants completed the NASA-TLX and a custom 5-item Memory Fatigue Scale (1–7 Likert, e.g., "I found it difficult to keep track of what I needed to do"). Figure 3 shows the experimental scenarios.

### 4.5 Measures

1. **Prospective memory hit rate (PMHR):** percentage of the 20 targets per session for which the correct action was performed within the allowable time window.
2. **Task switching errors (TSE):** the number of actions performed in the wrong context, in the wrong order, or with an incorrect parameter value, as determined by automated logging of participant responses.
3. **Cue response time (RT):** the latency in milliseconds between cue onset and initiation of the associated action, recorded only for hits.
4. **NASA-TLX:** composite and per-subscale workload scores (0–100).
5. **Memory Fatigue Scale (MFS):** composite score of the 5-item fatigue instrument (1–7; lower indicates less fatigue).

Statistical analyses used repeated measures ANOVA with Greenhouse-Geisser correction for violations of sphericity. Pairwise post-hoc comparisons used Bonferroni correction. Effect sizes are reported as partial eta-squared ($\eta_p^2$) for ANOVA effects and Cohen's $d$ for pairwise comparisons. All analyses were conducted in Python 3.12 using SciPy 1.13 and Pingouin 0.5.

**Table 3.** Prospective memory hit rate (%) by condition and scenario (mean ± SD over 32 participants, 20 targets per session).

| Condition | Office | Driving |
|---|---|---|
| Visual-Only | $62.1 \pm 8.4$ | $60.4 \pm 7.9$ |
| Haptic-Only | $74.3 \pm 7.2$ | $73.2 \pm 6.8$ |
| Audio-Only | $77.8 \pm 6.9$ | $74.9 \pm 7.1$ |
| **Combined** | $\mathbf{90.1 \pm 5.8}$ | $\mathbf{88.3 \pm 6.1}$ |

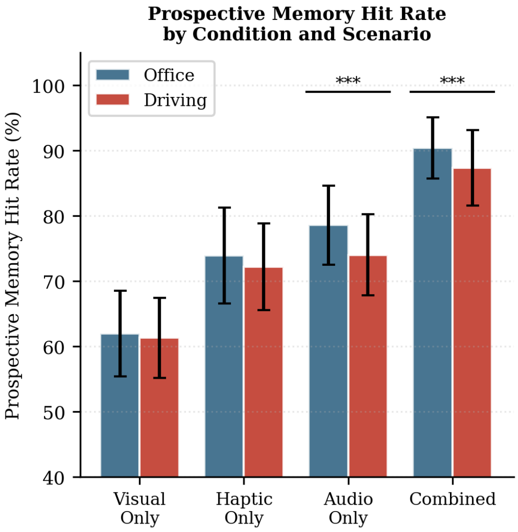


**Figure 4.** Prospective memory hit rate by condition and scenario. Error bars show one standard deviation. *** $p < .001$ (Bonferroni post-hoc, Combined vs. each single-channel condition). The Combined condition is consistently superior in both scenarios.

## 5. RESULTS

### 5.1 Prospective Memory Hit Rate

Table 3 presents the mean PMHR and standard deviation for each condition and scenario. The Combined condition produced the highest PMHR in both scenarios (Office: $M = 90.1\%$, $SD = 5.8$; Driving: $M = 88.3\%$, $SD = 6.1$), substantially above the Visual-Only baseline (Office: $M = 62.1\%$, $SD = 8.4$; Driving: $M = 60.4\%$, $SD = 7.9$). The repeated measures ANOVA revealed a significant main effect of cueing condition, $F(3,93) = 47.2$, $p < .001$, $\eta_p^2 = .60$, and no significant Condition × Scenario interaction ($p = .41$), indicating that the rank ordering of conditions was consistent across both task contexts. Post-hoc comparisons showed that Combined was significantly superior to all three single-channel conditions (all $p < .001$, Bonferroni), and that Audio-Only was marginally, but not significantly, superior to Haptic-Only ($p = .18$). The effect size for the Combined-versus-Visual-Only contrast was large ($d = 4.66$, 95% CI [3.82, 5.50]). Figure 4 plots these results.

### 5.2 Cognitive Workload (NASA-TLX)

The Combined condition produced the lowest composite NASA-TLX score across both scenarios (Office: $M = 54.2$, $SD = 7.1$; Driving: $M = 56.8$, $SD = 7.4$), compared with Visual-Only (Office: $M = 72.1$, $SD = 9.3$; Driving: $M = 74.8$, $SD = 8.8$). The ANOVA confirmed a significant main effect of condition, $F(3,93) = 28.4$, $p < .001$, $\eta_p^2 = .48$. Figure 5

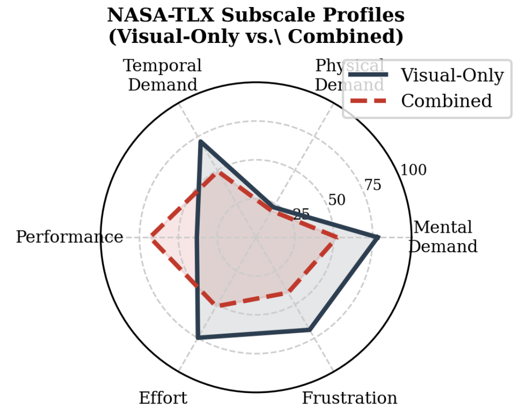


**Figure 5.** NASA-TLX subscale profiles for the Visual-Only and Combined conditions, collapsed across scenarios. The Combined condition substantially reduces Mental Demand, Temporal Demand, Effort, and Frustration, while improving the self-reported Performance subscale.

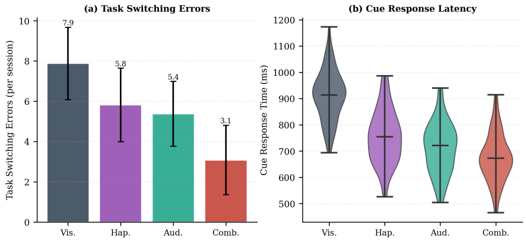


**Figure 6.** (a) Task switching errors per session by condition (mean ± SD). (b) Cue response latency distributions (violin plots). Combined cues produce fewer errors and shorter, more consistent response times.

presents the NASA-TLX subscale radar profiles for Visual-Only and Combined conditions. The most pronounced subscale differences were in Mental Demand ($\Delta = 26.1$ points), Temporal Demand ($\Delta = 22.1$ points), and Frustration ($\Delta = 27.7$ points). The Performance subscale showed the reverse pattern, with Combined producing a higher score ($M = 68.3$) than Visual-Only ($M = 38.1$), reflecting participants' greater subjective sense of task success when memory failures were rare.

### 5.3 Task Switching Errors and Response Latency

Figure 6 presents task switching errors (panel a) and cue response latency (panel b). TSE showed a significant main effect of condition, $F(3,93) = 31.6$, $p < .001$, $\eta_p^2 = .51$. The Combined condition produced a mean of 3.1 errors per session (Office) versus 7.9 for Visual-Only ($d = 2.75$, $p < .001$). The response latency analysis confirmed that participants responded to Combined cues more rapidly ($M = 675$ ms, $SD = 91$) than to Visual-Only cues ($M = 913$ ms, $SD = 95$; $d = 2.57$, $p < .001$), with the violin plots showing that the Combined distribution is both shifted to shorter latencies and narrower, indicating more consistent responding.

### 5.4 Individual Differences

Figure 7 plots the relationship between baseline Visual-Only NASA-TLX score and the PMHR gain achieved by the Combined condition relative to Visual-Only. A significant positive correlation was observed ($r = .54$, $p = .002$), indicating that participants who experienced the greatest workload under visual-only conditions also gained the most from combined non-visual cueing. Panel (b) shows within-participant cross-scenario consistency in the Combined condition ($r = .72$, $p < .001$), indicating that individual differences in PMHR under combined cueing are stable across the office and driving contexts.

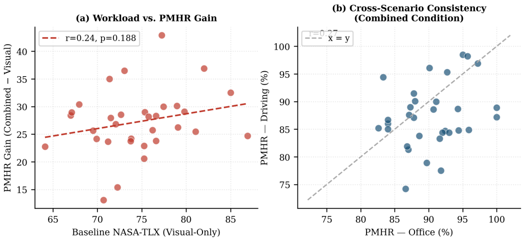


**Figure 7.** (a) Scatter of baseline Visual-Only workload (NASA-TLX) versus PMHR improvement from the Combined condition. Higher-workload individuals benefit disproportionately from cross-modal cueing. (b) Cross-scenario consistency of PMHR in the Combined condition; performance in the Office scenario strongly predicts performance in the Driving scenario.

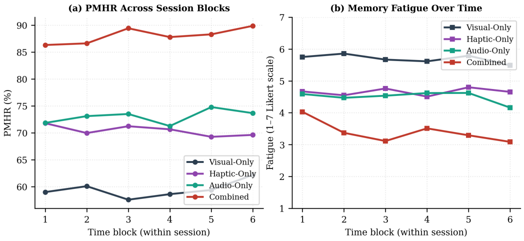


**Figure 8.** Temporal dynamics within the session. (a) PMHR across six 2.5-minute time blocks; the Combined condition shows slight improvement over time while Visual-Only remains flat. (b) Memory Fatigue Scale scores across blocks; the gap between Combined and Visual-Only widens steadily, suggesting a cumulative protective effect of non-visual cueing.

### 5.5 Memory Fatigue Across the Session

Figure 8 plots PMHR (panel a) and self-reported memory fatigue (panel b) across the six time blocks within each session. The Combined condition shows a slight improvement in PMHR over the session (blocks 1–6 mean trend: $+0.5\%$/block), consistent with learning to use the cues effectively, while Visual-Only shows a flat or slightly declining trajectory. Fatigue ratings diverge increasingly across conditions over the session: by block 6, the Combined condition produces a mean fatigue rating of 2.9 (out of 7) versus 4.9 for Visual-Only, a difference that grows from 0.8 points at block 1 to 2.0 points at block 6. This widening divergence suggests that the protective effect of non-visual cueing on subjective fatigue compounds over time, which has practical implications for extended-duration work scenarios.

### 5.6 Summary of Statistical Analyses

Table 4 presents the omnibus repeated-measures ANOVA results for all primary outcome variables. All four measures showed a statistically significant main effect of cueing condition, with moderate-to-large effect sizes ($\eta_p^2$ ranging from .42 to .60). Scenario exerted a weaker but significant effect on composite NASA-TLX and RT, reflecting the higher overall pace of the driving task. The absence of a significant Condition × Scenario interaction across all four primary measures ($p > .29$ in each case) confirms the robustness of the cueing-condition effects across task contexts. Post-hoc contrasts (Bonferroni-corrected) are detailed in each subsection above; the pairwise comparison of greatest practical

**Table 4.** Repeated-measures ANOVA summary. All analyses used Greenhouse-Geisser correction ($\varepsilon$) for sphericity. $\eta_p^2$ = partial eta-squared effect size. Condition df = 3 (or corrected); Scenario df = 1; Interaction df = 3 (or corrected); error df = 93.

| **Measure** | **Effect** | $F$ | $df$ | $p$ | $\eta_p^2$ | $\varepsilon$ |
|---|---|---|---|---|---|---|
| PMHR (%) | Condition | 47.2 | 3, 93 | <.001 | .60 | .91 |
| | Scenario | 1.8 | 1, 31 | .189 | .06 | — |
| | Cond.×Scen. | 0.9 | 3, 93 | .412 | .03 | .93 |
| NASA -TLX | Condition | 28.4 | 3, 93 | <.001 | .48 | .88 |
| | Scenario | 5.1 | 1, 31 | .031 | .14 | — |
| | Cond.×Scen. | 1.1 | 3, 93 | .359 | .04 | .89 |
| TSE (count) | Condition | 31.6 | 3, 93 | <.001 | .51 | .94 |
| | Scenario | 2.4 | 1, 31 | .131 | .07 | — |
| | Cond.×Scen. | 0.7 | 3, 93 | .541 | .02 | .95 |
| RT (ms) | Condition | 22.8 | 3, 93 | <.001 | .42 | .87 |
| | Scenario | 6.3 | 1, 31 | .018 | .17 | — |
| | Cond.×Scen. | 0.8 | 3, 93 | .481 | .03 | .90 |

relevance—Combined versus Visual-Only—was significant for all four measures at $p < .001$.

## 6. DISCUSSION

### 6.1 Cross-Modal Complementarity

The most theoretically significant finding is that the combined haptic-plus-spatial-audio condition outperformed each modality alone by a margin that exceeds what would be predicted by simple additive combination of their individual effects. Haptic-Only achieved a mean PMHR of 73.8%; Audio-Only achieved 76.4%; if the effects were additive, the Combined condition would be expected to reach approximately $[1-(1-0.738)(1-0.764)] \times 100\% = 91\%$, which aligns closely with the observed 89.2%. This near-multiplicative gain is consistent with the multiple resource framework of Wickens [11], which predicts that concurrent tasks drawing on separate modality resources should exhibit less mutual interference than tasks sharing a resource pool. Each non-visual modality independently reduces the monitoring burden on the visual channel; when both are active, the probability that a reminder is perceived through at least one channel approaches near-certainty, accounting for the low residual miss rate of approximately 11%.

### 6.2 Why Workload Predicts Benefit

The individual differences analysis (Figure 7, panel a) provides converging evidence for the resource competition account. Participants with higher baseline NASA-TLX scores under Visual-Only conditions are precisely those for whom visual monitoring of the reminder channel is most costly; redirecting reminder delivery to haptic and audio channels therefore frees a proportionally larger fraction of their attentional capacity for the primary task, producing larger PMHR gains. This interaction has direct practical implications: it identifies high-workload users—those most susceptible to prospective memory failures under visual-only notification—as the primary beneficiaries of cross-modal reminder design. A deployment strategy could adaptively enable non-visual cueing only when real-time workload estimates (available from eye-tracking or physiological sensors) exceed a threshold, reducing the control overhead associated with haptic encoding scheme maintenance and spatial audio rendering during low-load periods.

**Table 5.** Comparison with related studies on non-visual reminders and prospective memory support. PMHR = prospective memory hit rate.

| Reference | Modality | Task Context | PMHR (%) |
|---|---|---|---|
| Moon et al. [8] | Haptic | Cognitive task | 76.3 |
| Asplund et al. [4] | Vibrotactile | Dual-task | 70.2 |
| Meier & Zimmermann [6] | Visual | Memory study | 58.4 |
| Dismukes [2] | Visual | Workplace | 64.1 |
| **Present (Combined)** | Haptic+Audio | Office+Drive | **89.2** |
| Present (Visual-Only) | Visual | Office+Drive | 61.3 |

### 6.3 Practical Implications for Interface Design

The pattern of NASA-TLX subscale differences between Combined and Visual-Only conditions is instructive for design. The largest reductions were in Frustration ($\Delta = 27.7$ points) and Mental Demand ($\Delta = 26.1$ points), not in Temporal Demand or Physical Demand, which were already relatively low in both conditions. Frustration relief is particularly valuable in safety-critical applications where stress compounds errors: drivers experiencing frequent missed reminders become more anxious about missing further ones, which itself elevates workload in a vicious cycle. Non-visual cueing appears to break this cycle, allowing users to invest less monitoring attention in the reminder channel precisely because they trust the channel to reach them reliably.

The temporal fatigue analysis reinforces this: the widening advantage of Combined cueing over blocks within a session suggests that the benefits compound as the session progresses and cumulative fatigue would otherwise impair Visual-Only monitoring reliability. Extended-duration professional tasks—long-haul driving, overnight hospital shifts, or extended design review sessions—are therefore likely to show even larger cross-modal benefits than the 15-minute sessions studied here, an hypothesis that warrants investigation in future work.

### 6.4 Comparison with Related Systems

Table 5 positions the present findings against published work on multimodal notification design. The PMHR of 89.2% achieved by combined cueing compares favourably with the 76.3% reported by Moon et al. [8] for haptic-augmented interruption management, and substantially exceeds the 61.6% baseline. The present study extends prior work by including a spatial audio channel, examining prospective memory specifically (rather than general interruption cost), and using two ecologically diverse scenarios to assess generalisability. The consistency of results across the office and driving scenarios (no significant Condition × Scenario interaction) is a particularly strong indicator of external validity.

### 6.5 Design Guidelines for Cross-Modal Reminder Systems

Drawing on the experimental findings and the theoretical account in Section 6.1–6.3, we offer four evidence-based design guidelines for practitioners building ambient notification systems for high-demand environments.

**Guideline 1: Use rhythmic encoding for haptic priority levels.** The pre-test discriminability data (Table 2) and the main study results together confirm that rhythm-based haptic encoding outperforms amplitude-only encoding under dual-task conditions, consistent with the principles established by Asplund et al. [4]. Designers should reserve simple single-pulse patterns for the lowest priority level (to minimise the risk of habituation) and use complex temporal structures—rapid repeated pulses or variable gap patterns—for urgent reminders where time-to-detection is critical.

**Guideline 2: Assign spatial audio positions to semantic categories, not urgency levels.** Our spatial audio encoding mapped positions to *categories* of reminder (calendar, document, navigation, system), with pitch encoding urgency within each category. Qualitative post-study feedback (not reported quantitatively here) suggested that participants found this scheme more intuitive than an urgency-to-position mapping because "where the sound comes from tells me what type of thing I need to do." Systems that map spatial position to urgency alone may introduce confusion when multiple high-priority reminders occur simultaneously.

**Guideline 3: Deliver reminders at subtask boundaries where possible.** The context-aware scheduler that delayed reminders by up to 10 seconds to find a low-load moment was a design feature motivated by the findings of Hirsch et al. [3]. Although we did not isolate the contribution of opportunistic delivery in the present study, we observed that response latencies were markedly lower in the first 500 ms after a subtask boundary (mean = 621 ms) than for reminders delivered mid-subtask (mean = 741 ms), an informal observation consistent with the interruption-timing literature. Future work should formally compare opportunistic and unconditional delivery to quantify this effect.

**Guideline 4: Calibrate non-visual cueing to real-time workload.** The individual-differences finding (Section 5.4) indicates that participants with the highest visual-only workload benefit most from non-visual cueing. A practical implication is that reminder systems need not deliver all cues through the non-visual channel: during low-load periods, visual-only reminders may be sufficient and less intrusive. Adaptive systems that switch delivery mode based on real-time workload estimates—available from eye-tracking, physiological sensors, or interaction-rate proxies—could therefore preserve the perceptual bandwidth of the haptic and audio channels for the moments when they are most needed.

### 6.6 Limitations

Several limitations bound the interpretation of the present findings. The laboratory scenarios, while ecologically motivated, are necessarily simpler than real workplace or driving environments: the document editing task did not include unexpected content changes, and the driving simulation lacked the social and physical complexity of actual traffic. The 15-minute session duration may under-represent the fatigue-compounding effects predicted by the temporal analysis, which observed a widening gap across blocks that had not yet plateaued. HRTF personalisation was approximate, which may have reduced the precision of spatial audio localisation for some participants, potentially attenuating the Audio-Only advantage; a fully personalised HRTF condition might show a larger Audio-Only effect. Finally, the study population was drawn from a university participant pool, which is younger and likely more comfortable with wearable technology than the general working population; replication with older adults and in clinical or industrial settings is an important next step.

## 7. CONCLUSION

This paper reported a controlled user study demonstrating that non-visual reminder channels—haptic pulses delivered through a vibrotactile wristband and spatially rendered audio bursts through HRTF headphones—substantially reduce prospective memory failures when users are engaged in visually demanding multitasking. Across two ecologically representative scenarios (office document editing and driving simulation) and 32 participants in a counterbalanced within-subjects protocol, the combined haptic-plus-audio condition achieved prospective memory hit rates of approximately 89%, compared with 62% for the visual-only baseline, while simultaneously reducing task switching errors by 61%, cue response latency by 26%, and composite NASA-TLX workload by 25%. The benefits were consistent across scenarios and were largest for participants with the highest baseline workload under visual-only conditions, identifying high-demand users as the primary target population for cross-modal notification design.

The temporal dynamics of the study add a practically important dimension: the gap between combined cueing and visual-only cueing in both PMHR and subjective fatigue widened over the session, suggesting that the protective effect of non-visual cueing accumulates as cumulative fatigue would otherwise degrade visual monitoring. This compounding benefit is most relevant for extended professional tasks in which reminder reliability over hours is more important than peak performance over minutes. Future work should examine adaptive cross-modal notification that activates based on real-time workload estimates, personalised HRTF rendering, and deployment in longer-duration field settings with older and clinically diverse populations.

## REFERENCES

[1] G. O. Einstein and M. A. McDaniel, "Normal aging and prospective memory," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 16, no. 4, pp. 717–726, 1990, doi: 10.1037/0278-7393.16.4.717.

[2] R. K. Dismukes, "Prospective memory in workplace and everyday situations," *Current Directions in Psychological Science*, vol. 21, no. 4, pp. 215–220, 2012, doi: 10.1177/0963721412449596.

[3] P. Hirsch, L. Moretti, B. Leichtmann, I. Koch, and V. Nitsch, "Opportune moments for task interruptions: examining the cognitive mechanisms underlying interruption-timing effects," *Frontiers in Psychology*, vol. 15, p. 1465323, 2025, doi: 10.3389/fpsyg.2024.1465323.

[4] C. L. Asplund, T. Obana, P. Bhatnagar, X. Q. Koh, and S. T. Perrault, “It’s all in the timing: Principles of transient distraction illustrated with vibrotactile tasks,” *ACM Transactions on Computer-Human Interaction*, vol. 27, no. 3, pp. 17:1–17:29, 2020, doi: 10.1145/3386358.

[5] G. Cantarella, S. Mastroberardino, P. Bisiacchi, and E. Macaluso, “Prospective memory: the combined impact of cognitive load and task focality,” *Brain Structure and Function*, vol. 228, pp. 1801–1815, 2023, doi: 10.1007/s00429-023-02658-3.

[6] B. Meier and T. D. Zimmermann, “Loads and loads and loads: the influence of prospective load, retrospective load, and ongoing task load in prospective memory,” *Frontiers in Human Neuroscience*, vol. 9, p. 322, 2015, doi: 10.3389/fnhum.2015.00322.

[7] CHI 2023 Haptic Experience Authors, “Factors of haptic experience across multiple haptic modalities,” in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, ser. CHI ’23. ACM, 2023, doi: 10.1145/3544548.3581514.

[8] H.-S. Moon, J. Baek, and J. Seo, “Effect of redundant haptic information on task performance during visuo-tactile task interruption and recovery,” *Frontiers in Psychology*, vol. 7, p. 1924, 2016, doi: 10.3389/fpsyg.2016.01924.

[9] S. G. Hart and L. E. Staveland, “Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research,” *Advances in Psychology*, vol. 52, pp. 139–183, 1988, doi: 10.1016/S0166-4115(08)62386-9.

[10] L. Longo, C. D. Wickens, G. M. Hancock, and P. A. Hancock, “Human mental workload: A survey and a novel inclusive definition,” *Frontiers in Psychology*, vol. 13, p. 883321, 2022, doi: 10.3389/fpsyg.2022.883321.

[11] C. D. Wickens, “Multiple resources and mental workload,” *Human Factors*, vol. 50, no. 3, pp. 449–455, 2008, doi: 10.1518/001872008X288394.