scientific reports

OPEN

Comparing voice-only navigation and paper maps in middle-aged and older adults: A randomized crossover field study of wayfinding, mood, usability, and text-mined user experiences

Chizuru Nagata[1✉], Tomoko Ota[1], Yasuko Sumida[1], Yumi Isomura[2], Yuko Shimizu[2], Takeshi Yamamoto[3] & Chong Chen[4✉]

Independent mobility is essential for autonomy in later life, yet age-related cognitive changes make map-reading increasingly demanding. One potential alternative is voice-only navigation, which reduces reliance on visual search and map-to-world transformations. We conducted a randomized within-subject crossover study to compare voice-only navigation with a paper map in adults aged 50–69. Forty participants each walked two 800-m urban routes, once using a paper map and once guided by Apple Maps voice instructions. We measured goal completion, completion time, mood (pleasure, relaxation, vigor) and fatigue before and after each walk, and usability (System Usability Scale; SUS). Goal completion and completion time did not differ between conditions. However, participants reported feeling more relaxed after voice navigation; other mood states and fatigue showed no significant changes. Usability ratings for voice navigation were moderate (SUS 60.50 ± 2.40), indicating room for improvement. Text-mining of post-task interviews highlighted needs for clearer distance phrasing, earlier turn prompts, explicit intersection guidance, and landmark-based cues. Overall, voice-only navigation matched paper maps in performance while enhancing relaxation, consistent with lower perceived cognitive effort. Moderate usability underscores the need for design improvements, particularly in phrasing and timing. These findings warrant further investigation in adults ≥ 70 and in cognitively impaired populations.

Keywords Voice navigation system, System Usability Scale, Text mining, Dementia, Spatial cognition

Independent mobility is a cornerstone of autonomy, social participation, and quality of life in older adults[1,2]. It also contributes to cognitive health through the benefits of regular physical activity[3–5]. However, age-related changes in perception, attention, and memory make wayfinding increasingly demanding, especially in complex environments[3,6–9]. Using a traditional map is particularly demanding as it requires the simultaneous use of several cognitive processes: visually searching for landmarks and street names, mentally rotating the map to align with the environment, and holding a sequence of turns in working-memory. These demands increase with environmental complexity and can raise cognitive effort even in healthy older adults[3,9].

[1]Division of Community/Gerontological Nursing, Yamaguchi University Graduate School of Medicine, 1-1-1 Minami-kogushi, Ube, Yamaguchi 755-8505, Japan. [2]Department of Nursing, Faculty of Nursing, Ube Frontier University, 2-1-1 Bunkyoudai, Ube, Yamaguchi 755-0805, Japan. [3]Department of Laboratory Medicine, Faculty of Health Sciences, Yamaguchi University Graduate School of Medicine, 1-1-1 Minami-kogushi, Ube, Yamaguchi 755-8505, Japan. [4]Division of Neuropsychiatry, Department of Neuroscience, Yamaguchi University Graduate School of Medicine, 1-1-1 Minami-kogushi, Ube, Yamaguchi 755-8505, Japan. ✉email: nagata@yamaguchi-u.ac.jp; cchen@yamaguchi-u.ac.jp

Voice-only navigation offers a potential alternative. Stepwise auditory prompts that specify actions (e.g., "turn right in 50 m") can reduce the need for visual search and mental map rotation. Cognitively, such guidance may (i) reduce visuospatial processing demands, (ii) easy working memory load required for route planning, and (iii) free up visual attention for critical tasks such as detecting hazards and maintaining balance. Although extensively studied in people with visual impairments[10,11], the effectiveness of voice-only navigation for sighted middle-aged and older adults remains largely untested. Prior research in young adults using mixed cardinal and egocentric prompts in sparse environmental settings produced equivocal results, which leaves open the question of whether common left/right voice guidance in realistic urban environments benefits older users[12].

Our study seeks to address this gap and to clarify the scope for such research. Navigation tools may help support healthy behaviors like outdoor walking, but they are not health interventions per se. We thus choose to focus on immediate and relevant outcomes—wayfinding performance, perceived cognitive effort, and usability—rather than on long-term health risks such as falls or dementia, which require different study designs[4,5,13,14].

Prior research on pedestrian-navigation has extensively compared auditory or voice-based aids with visual maps and digital displays. However, most studies have been conducted in young or broadly mixed-age samples and under conditions that differ from everyday walking among older adults. For instance, Xu et al.[15], examined how voice-assisted digital maps influence pedestrian wayfinding with participants aged 19–25 (primarily university students) navigating relatively simple environments. Rehrl et al.[16], tested augmented-reality, voice, and digital-map navigation among adults aged 22–66. However, this wide age range included relatively few older participants and combined multiple interfaces (visual, auditory, and AR), which makes modality-specific effects difficult to isolate. More recent work on intelligent voice assistants such as DuIVA[17] emphasized system architecture and algorithmic dialogue management rather than behavioral outcomes in older users.

To address these gaps, the present study specifically focuses on adults aged 50–69, an age group that maintains high smartphone usage yet already experiences subtle age-related declines in visuospatial and attentional functions[18]. We deliberately isolate modality by comparing a voice-only, eyes-free condition with a visual-only paper map, thus removing confounding factors from smartphone screens (e.g., touch gestures, zoom, display size). Moreover, we test navigation during real urban walking in Japan—settings rich in landmarks and environmental complexity that better reflect daily mobility. Finally, we incorporate text-mining analysis of post-task interviews to derive actionable improvements in phrasing, timing, intersection clarity, and landmark use, directly linking behavioral findings to human-centered design[13]. We hypothesized that voice-only guidance would yield comparable goal-completion rates and completion times to paper-map navigation on short urban routes, while producing a more favorable pre-post change in mood and fatigue states due to reduced perceived cognitive effort.

Methods
Study design
This study was a randomized within-subjects crossover pretest posttest comparison study.

Participants
Forty middle-aged and older adults (aged 50–69) were recruited with a convenience sampling method through voluntary participation. Exclusion criteria included the inability to walk normally due to conditions such as broken bones or poor physical condition. Based on an a priori power analysis for repeated measures ANOVA, 34 participants were required to detect a medium effect size ($f = 0.25$) with two groups in a within-subjects crossover design (power = 0.80, $\alpha = 0.05$). To account for dropouts and invalid data, we recruited 40 participants.
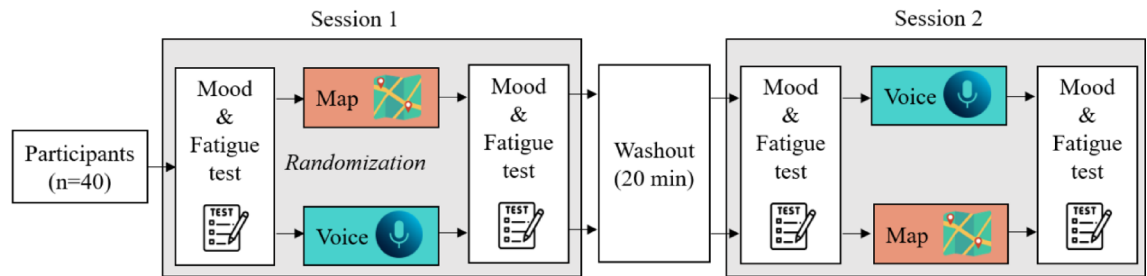
Ethical considerations
Participants were informed that participation was voluntary and that they had the right to withdraw at any time. Participants provided written informed consent before the study following detailed explanations about the study. This study was approved by the Ethical Review Board of Nursing and Laboratory Science, Yamaguchi University Graduate School of Medicine (Approval No. 788-2) and conducted per the latest version of the Declaration of Helsinki.

Procedure
Participants navigated two equivalent routes (A and B, approximately 800 m each) using two different tools—a traditional paper map (Map) and voice navigation via Apple Maps (Voice). To prevent route familiarity from influencing results, participants navigated each route once, using a different tool for each, with the order of conditions (Map and Voice) counterbalanced. We confirmed that both completion rate and completion time (shown in Fig. S1) were similar between routes A and B for both navigation methods, with no statistically significant differences ($p > 0.3$). Prior to the experiment, we also confirmed that participants were unfamiliar with the area by providing the names of the destination points and verifying that they had never been there.

Apple Maps, a widely used voice navigation tool, provides directions using left and right instead of the traditional east, west, north, and south, making it generally easier to follow[12]. The navigation tasks were set in a local city with approximately 160,000 residents, which provides a realistic environment rich in landmarks. The routes specifically passed through residential neighborhoods, shopping streets, and areas with banks, including multiple intersections and traffic lights. An overview of the procedure is shown in Fig. 1.

Each participant was given a leaflet with walking instructions for each navigation, and one researcher per participant provided one-on-one explanations and assistance from start to finish. If the participants could not continue the experiment for any reason, such as feeling unwell or losing their way, they could contact the researcher who was following them from a distance of about 4 m. The researcher observed from a distance to ensure safety and minimize interference. Based on multiple pre-tests, the average walking time was approximately

Fig. 1. Overview of the study procedure. In this randomized, within-subject crossover study, 40 participants navigated using two methods: a traditional paper map (Map) and voice navigation via Apple Maps (Voice). The order was counterbalanced so that half of the participants started with the paper map and then used voice navigation, while the other half experienced the reverse order. Participants rated their mood and fatigue immediately before and after each navigation task. Icon source: Flaticon.com.

10 min. Therefore, a 20-min time limit was set for the walk. For participants who did not reach the destination within this timeframe, their completion time was recorded as 20 min (1,200 s) for comparison purposes (as a result, non-parametric tests were used for this analysis, see below).

The effectiveness of voice navigation was evaluated using multiple measures, including goal completion (i.e., whether participants reached the destination within the 20-min timeframe), completion time (i.e., the time taken to reach the destination), participants' subjective experiences (including mood and fatigue), and the usability of voice navigation. Mood and fatigue were evaluated both before and immediately after each navigation task. Additionally, after using the voice navigation system, participants filled out the System Usability Scale (SUS) to indicate their evaluations of the usability of the voice navigation system. Lastly, after completing all navigation tasks, researchers conducted interviews to gather participants' impressions of walking with both maps and voice navigation. Participants were also asked how the voice navigation system could be made more user-friendly. Data was collected during September to November 2024.

Assessments

Mood test: The Chen-HAgiwara Mood Test (CHAMT), which consists of three test items evaluating pleasure, relaxation, and vigor using a 100 mm visual analog scale,[19,20] was employed to evaluate participants' mood states.

Fatigue test: The Fatigue Feelings Test (Jikaku-sho Shirabe) developed by the Japan Society for Occupational Health[21] was employed. This test uses 25 items to evaluate feelings of drowsiness (Factor Ⅰ), instability (Factor Ⅱ), uneasiness (Factor Ⅲ), local pain or dullness (Factor Ⅳ), and eyestrain (Factor Ⅴ).

System Usability Scale: A satisfaction survey of voice navigation using the SUS developed by John Brooke[22] was employed. This scale uses 25 items rated on a 5-point ordinal scale. In general, scores below 70 indicate a possible lack of consideration for usability, whereas scores above 70 indicate satisfaction with usability.

Cognitive assessments: To understand participants' general cognitive characteristics, we administered the following tests at the beginning of the study prior to the randomized trial: Touch M (Human Corporation[23]), a touchscreen-based cognitive test that evaluates three core functions, including visuospatial cognition, temporal order cognition, and working memory, the Spatial Recognition test included in Brain Check (Re-Gaku Corporation[24]), and the Sense of Direction Questionnaire-Short Form (SDQ-S)[25,26], a 20-item self-report measure.

Data analysis

Statistical analysis was conducted using IBM SPSS Statistics version 29 and Python 3.11.7 (numpy 1.26.4, pandas 2.3.3, scipy 1.11.4, statsmodels 0.14.0, lifelines 0.30.0, matplotlib 3.10.7, seaborn 0.13.2). The McNemar test and Wilcoxon signed-rank test were employed to compare the goal completion rate and completion time, respectively. To address right-censoring in completion times, we conducted a supplementary Cox proportional hazards model with participant as a random effect (frailty term), and Navigation as the predictor; time-outs/non-completions were treated as right-censored. The repeated measures analysis of variance (ANOVA) was employed to analyze the mood and fatigue data, where a significant navigation*time interaction would indicate differences between navigation methods. Furthermore, as an exploratory analysis of any potential moderation effects by cognitive indices, we probed Navigation × Age and Navigation × cognitive indices (Touch-M, Brain Check spatial score, SDQ-S) using linear mixed-effects models for each outcome, using Participant as a random intercept and fixed effects for Time (Pre/Post), Navigation (Map/Voice), each moderator, and their interactions, focusing on the Time × Navigation × moderator term. The significance level was set at $p < 0.05$, two-sided.

Text-mining analysis was carried out on the qualitative data from the interviews using the software KH Coder 3 (Koichi Higuchi, https://khcoder.net/en/) with a Notation Variation & Synonym Editor plug-in (SCREEN Advanced System Solutions Co., Ltd., https://www.screen.co.jp/as/products/monkin-syn). Specifically, to compare the linguistic characteristics of participants' responses for the two navigation methods, we applied correspondence analysis. Correspondence analysis is a data science visualization technique used to explore relationships between categorical variances, in this case, the navigation method and the words used in participants' comments. Furthermore, to visualize participants' feedback on improving the Voice navigation system, we conducted a co-occurrence network analysis. This technique visualizes and explore relationships

between words based on their tendency to appear together within participants' individual responses, which helps identify clusters of semantically related terms and reveals underlying themes or patterns in participants' feedback.

Results

Participants' characteristics

The study included 40 participants (11 male and 29 female), with an average age of 57.33 ± 3.89 years. The mean percentage of correct responses on the cognitive test Touch M was $82.00\% \pm 10.49\%$. The average spatial cognition score of Brain Check was 16.33 ± 3.22 out of 20, and the mean SDQ-S was 58.55 ± 14.05 out of 100.

Goal completion and completion time

Of the 40 participants, four failed to reach their destination using Map navigation and one failed under Voice navigation, with no statistically significant difference between navigations ($p = 0.375$, McNemar Test).

Completion times for Map and Voice navigations are shown in Fig. 2. The median completion time was 527.5 s for both Map and Voice navigations. A non-parametric comparison of completion times using the Wilcoxon signed-rank test showed no significant difference between the two navigation methods ($p = 0.877$). The Cox sensitivity analysis produced $HR = 1.063$ (95% CI = [0.755, 1.498]), $p = 0.726$, consistent with the result of the Wilcoxon signed-rank test.
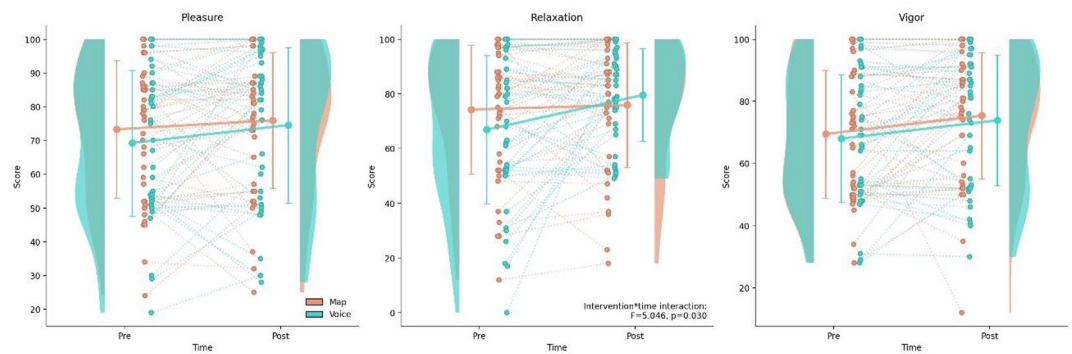
Mood and fatigue

Changes in mood states from pre- to post-navigation are shown in Fig. 3. A repeated measures ANOVA showed no significant navigation*time interaction for pleasure ($F_{1,39} = 0.528$, $p = 0.472$, partial $\eta^2 = 0.013$) or vigor ($F_{1,39} = 0.000$, $p = 0.995$, partial $\eta^2 < 0.001$). However, there was a significant navigation*time interaction for relaxation ($F_{1,39} = 5.046$, $p = 0.030$, partial $\eta^2 = 0.115$, Fig. 3, middle panel), indicating that in comparison to Map navigation, participants reported greater relaxation after Voice navigation.
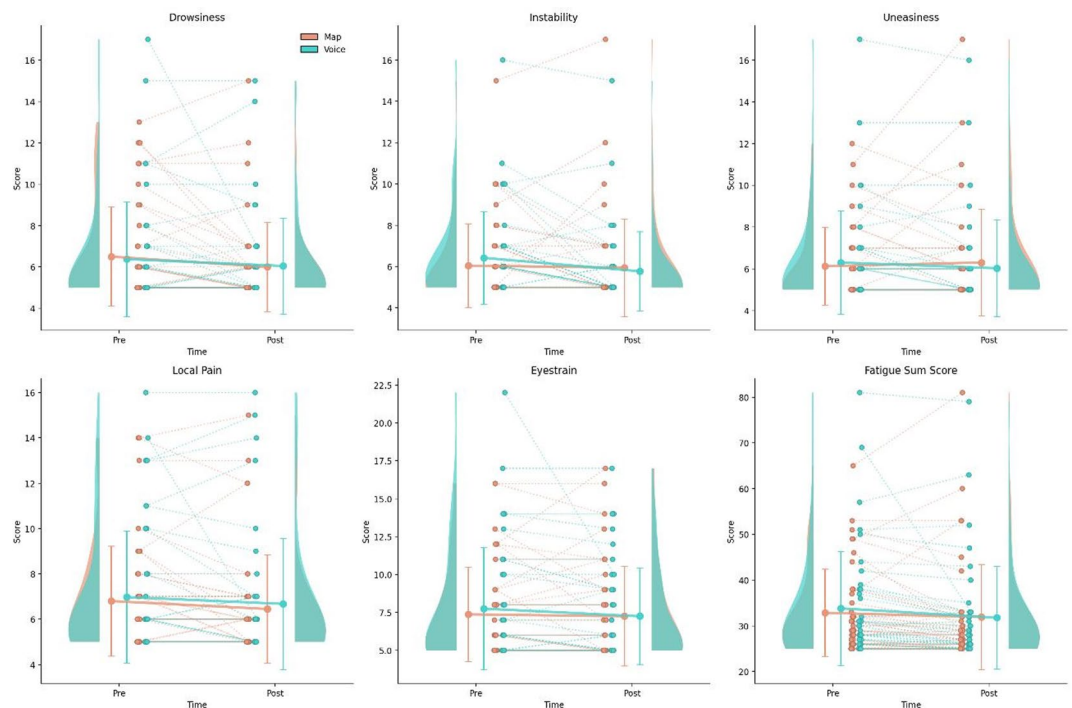
Changes in fatigue scores are presented in Fig. 4. A repeated measures ANOVA revealed no significant navigation*time interactions for any of the five fatigue factors: drowsiness ($F_{1,39} = 0.319$, $p = 0.575$, partial $\eta^2 = 0.008$), instability ($F_{1,39} = 3.373$, $p = 0.074$, partial $\eta^2 = 0.080$), uneasiness ($F_{1,39} = 2.432$, $p = 0.127$, partial $\eta^2 = 0.059$), local pain ($F_{1,39} = 0.033$, $p = 0.857$, partial $\eta^2 = 0.001$) and eyestrain ($F_{1,39} = 0.808$, $p = 0.374$, partial $\eta^2 = 0.020$). Similarly, there was no significant navigation*time interaction for the total fatigue score ($F_{1,39} = 0.947$, $p = 0.337$, partial $\eta^2 = 0.024$).

Fig. 2. Completion time for Map and Voice navigations. Half-violin plots represent the kernel density estimation of individual scores, with box plots summarize the distribution by showing the median and the interquartile range. Individual data points are shown with a slight horizontal jitter and connected by dotted lines to highlight within-participant changes over time.

Fig. 3. Changes in mood states pre- and post-navigation. This figure displays the distributions of the three mood measures, Pleasure, Relaxation, and Vigor, recorded before (Pre) and after (Post) a navigation, separately for Map and Voice navigation). For each measure, half-violin plots represent the kernel density estimation of individual scores, with error bars representing the group mean ± standard deviation. Individual data points are shown with a slight horizontal jitter and connected by dotted lines to highlight within-participant changes over time, while solid lines connect the group means between time points. In the Relaxation panel (second subplot), the annotation indicates a significant navigation*time interaction, with F = 5.046, $p = 0.030$.

Fig. 4. Changes in fatigue score pre- and post-navigation. This figure displays the distributions of the five fatigue measures, Drowsiness, Instability, Uneasiness, Local Pain, Eyestrain, as well as their sum score recorded before (Pre) and after (Post) a navigation, separately for Map and Voice navigation. For each measure, half-violin plots represent the kernel density estimation of individual scores, with error bars representing the group mean ± standard deviation. Individual data points are shown with a slight horizontal jitter and connected by dotted lines to highlight within-participant changes over time, while solid lines connect the group means between time points.

Explorative analysis of the moderation effects by cognitive indices

Across outcomes, only sense of direction (measured by SDQ-S) moderated the Time × Navigation effect; neither age, Touch M (%), nor Brain Check showed evidence of moderation (all p's > 0.05). For mood outcomes (higher = better), higher SDQ-S was associated with a less favorable Voice–Map change from Pre to Post for pleasure ($\beta = -7.584$, SE = 3.750, z = −2.02, $p = 0.043$; LR $\chi^2 = 4.02$, $p = 0.045$, $\Delta R^2 = 0.008$) and vigor ($\beta = -10.992$, SE = 4.111, z = −2.67, $p = 0.007$; LR $\chi^2 = 6.94$, $p = 0.008$, $\Delta R^2 = 0.018$). For fatigue outcomes (higher = more fatigue), higher SDQ-S was linked to a more unfavorable Voice–Map change for drowsiness ($\beta = 0.997$, SE = 0.364, z = 2.74,

$p = 0.006$; LR $\chi^2 = 7.27$, $p = 0.007$, $\Delta R^2 = 0.011$), eyestrain ($\beta = 0.902$, SE $= 0.401$, z $= 2.25$, $p = 0.024$; LR $\chi^2 = 4.96$, $p = 0.026$, $\Delta R^2 = 0.004$), local pain ($\beta = 0.723$, SE $= 0.286$, z $= 2.53$, $p = 0.012$; LR $\chi^2 = 6.22$, $p = 0.013$, $\Delta R^2 = 0.005$), uneasiness ($\beta = 0.658$, SE $= 0.279$, z $= 2.36$, $p = 0.018$; LR $\chi^2 = 5.44$, $p = 0.020$, $\Delta R^2 = 0.005$), and the fatigue sum ($\beta = 3.736$, SE $= 1.230$, z $= 3.04$, $p = 0.002$; LR $\chi^2 = 8.89$, $p = 0.003$, $\Delta R^2 = 0.007$). Effects were statistically reliable but contributed modest additional variance ($\Delta R^2 \approx 0.004$–0.018).

Satisfaction with voice navigation usage

Participants' ratings of the Voice navigation system on the System Usability Scale (SUS) are shown in Fig. 5. The mean SUS score was 60.50 ± 2.40, and the median was 61.25. Of the participants, 10 (25%) gave scores below 50, 7 (17.5%) gave scores of 50–59, 13 (32.5%) rated it 60–69, and 10 participants (25%) gave ratings of 70 or higher.

Feedback on map and voice navigations

We analyzed participants' thoughts on the Map and Voice navigation systems using text mining via KH-Coder. To compare their experiences with the two navigation methods, we first conducted a correspondence analysis. As shown in Fig. 6, differences in word occurrence frequencies between the two navigation methods are represented along a single dimension. To better understand the characteristic words for each method, we referred back to the original text written by participants.
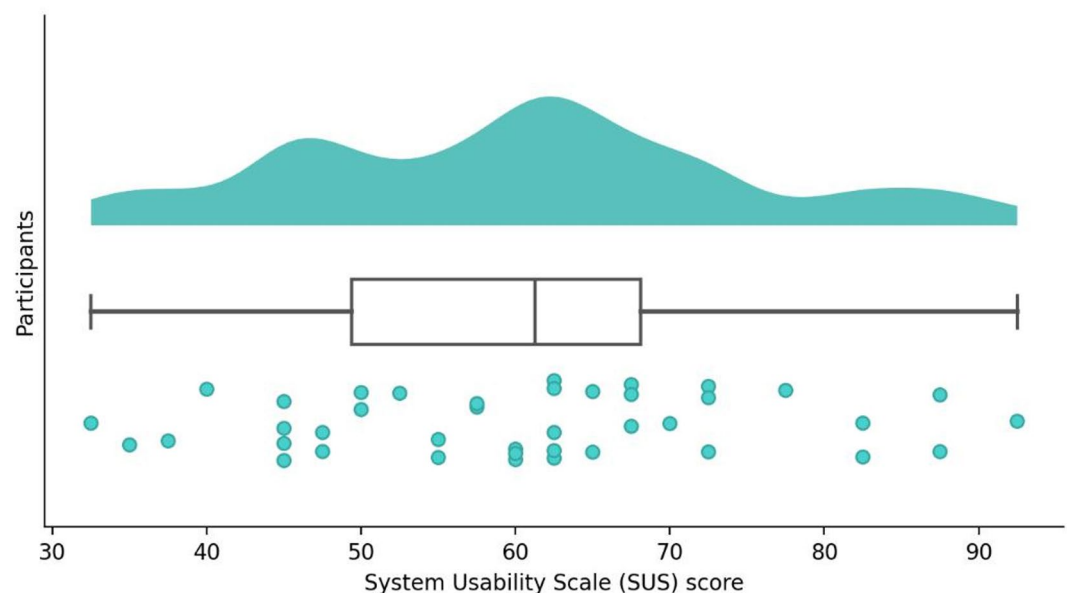
When using the Map (characterized by words in the lower left corner of the figure), participants tended to choose a route or road initially based on the provided map. As they navigated to their destination, they relied on landmarks such as banks and traffic lights to confirm their location and guide their movement. Several participants reported that they had become accustomed to reading maps.

In contrast, when using Voice navigation (shown in the upper right of the figure), participants described a number of challenges. Several noted that it was hard to hear instructions in noisy settings, such as when traffic, ambulances, or helicopters were nearby. They also said that perceiving distances from the prompts was difficult. For example, one participant noted that the phrase "Turn right in 300 m" did not provide a clear mental image of the distance, which led to a little uncertainty and anxiety. Ambiguous instructions also caused problems at times, as the system did not always clarify whether to cross a street, turn at an intersection, or take some other action. Despite these challenges, several participants mentioned that they adjusted to the voice guidance fairly quickly. Others stated that they were already accustomed to this style of navigation through their prior use of in-car navigation systems.
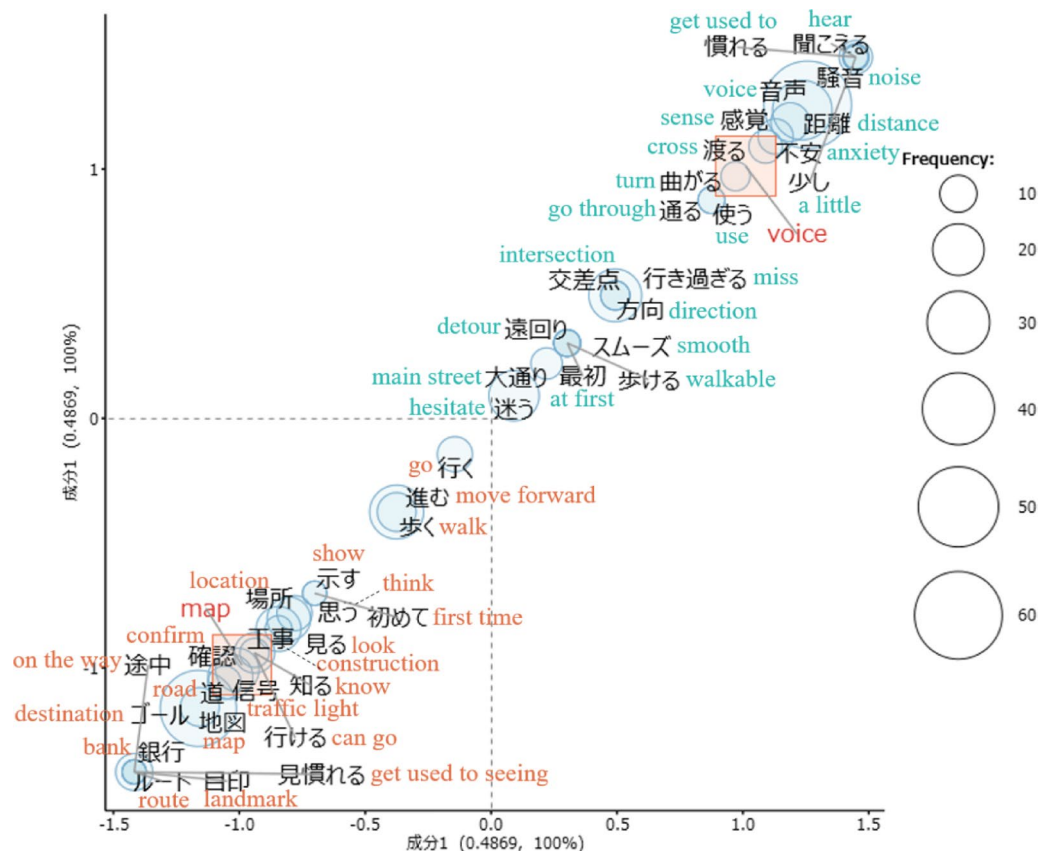
We next conducted a co-occurrence network analysis of participants' feedback on improving the Voice navigation system. As shown in Fig. 7, five clusters were identified.

Difficulty in judging distances (Cluster 5, dark blue)
Participants reported difficulty in getting a precise sense of distances (e.g., 300 m) and preferred shorter, more manageable distance references in instructions. They suggested that, in addition to distances, landmarks should be included in navigation cues. When giving left or right turn instructions, participants preferred receiving them just before or slightly ahead of the turning point.

Fig. 5. System Usability Scale (SUS) scores for the voice navigation system. The upper half-violin represents the kernel density estimation of the scores, with the middle box plot summarizes the distribution by showing the median and the interquartile range. Individual data points are shown in the lower portion of the plot with a slight vertical jitter to improve visibility.

Fig. 6. Correspondence analysis comparing participants' thoughts on Map and Voice. Differences in term occurrence frequencies between the two categories are represented along a single dimension, labeled as "Component 1." In this figure, Component 1 is plotted on both the x- and y-axes, which allows the extracted terms to be positioned along the diagonal. This layout provides space for label display and improves readability. Terms closer to the origin (0,0) indicate minimal differences between Map and Voice, while terms positioned toward the upper-right represent those characterizing Voice, and terms toward the lower-left represent those characterizing Map. Each term was represented by a node shown as a circle, whereas its size or diameter indicates its frequency. English translations have been provided for each Japanese term used in the initial analysis.

Location-based guidance and interaction (Cluster 1, light blue)
Participants expressed a desire for the navigation system to recognize their current location and provide instructions accordingly. They also wanted feedback confirming whether their movement was correct and suggested the possibility of a conversational interface, where the system could respond to user queries.

Clarification of intersection instructions (Cluster 3, purple)
Participants expected clearer instructions on whether they should cross a traffic light or intersection, or whether they should turn left or right at a traffic signal.

Experience with car navigation (Cluster 4, pink)
Some participants felt that, compared to car navigation systems, voice navigation provided instructions too late and should give guidance earlier.
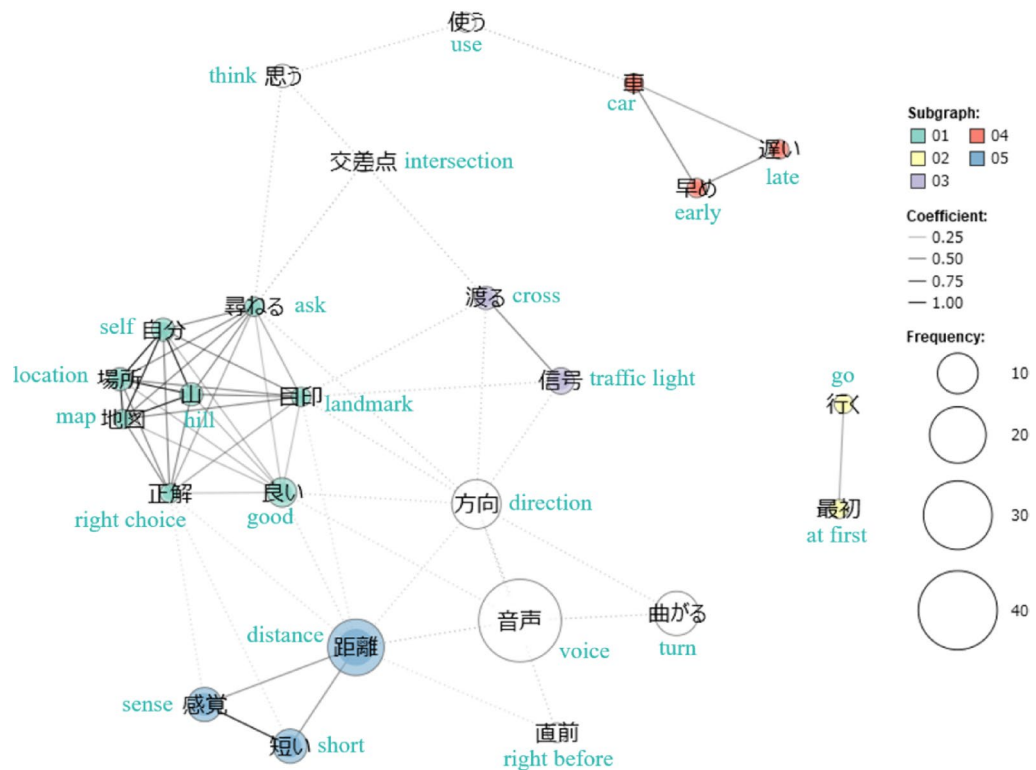
General route overview and directional clarity (Cluster 2, yellow)
This cluster contained two comments: one suggested that a brief overview of the route at the beginning would be helpful, while the other noted initial confusion about whether "right" meant turning right or simply moving to the right side.

Discussion
Principle findings
This randomized crossover field study of adults aged 50–69 yielded three main findings. First, voice-only navigation produced wayfinding performance comparable to a traditional paper map, as measured by goal completion and completion time. Second, despite no performance differences between the two methods, participants reported greater relaxation after using voice navigation. Third, usability ratings for the voice

Fig. 7. Co-occurrence network of high-frequency terms in participants' feedback on improving the Voice navigation system. In the network, each node (represented by a circle) corresponds to a high-frequency term, with the size of the circle indicating its frequency. Edges (links between nodes) represent co-occurrence relationships between terms in participants' feedback, where darker edges indicate stronger associations which were quantified by standardized coefficients. The visual network is structured based on the proximity of high-frequency terms to identify clusters of related concepts. Nodes within the same cluster share the same color, and a total of five clusters were identified (denoted as "subgraph" in the figure legend). Edges within a single cluster are shown as solid lines, while those spanning multiple clusters are shown as dotted lines. English translations have been provided for each Japanese term used in the initial analysis.

system was only moderate (mean SUS ≈ 60), and qualitative interview feedback pointed to concrete areas for improvement, including clearer distance phrasing, earlier turn prompts, explicit intersection guidance, and the use of landmarks.

Interpretation and mechanisms

The greater relaxation reported after voice navigation is consistent with a lower perceived cognitive effort during wayfinding. Traditional map use imposes heavy, concurrent demands on visual search, mental rotation, and working memory[3,9]. By offloading step-by-step instructions to the auditory channel, voice guidance likely frees up these cognitive resources, which allows for greater visual attention to the environment. This interpretation aligns with known age-related changes in navigation strategy[3,7–9]. However, as we measured subjective mood rather than objective cognitive load (e.g., pupillometry, dual-task costs), this proposed mechanism remains provisional and should be investigated in future research.

Our performance results complement prior work that reported mixed outcomes when guidance combined cardinal and egocentric terms in sparse environments[12]. In contrast, we used widely deployed left/right prompts in typical urban streetscapes, a combination that may be more tractable for older adults[3,7–9]. Together, these findings suggest that for healthy adults aged 50–69 on short urban routes, voice-only guidance can match the performance of paper maps while improving the affective experience of walking.

Exploratory moderation by cognitive indices

We probed whether age or cognitive indices qualified the Time × Navigation effects and found no moderation by age, Touch-M accuracy, or Brain Check spatial score. By contrast, self-reported sense of direction (SDQ-S) moderated several outcomes: participants with a better sense of direction showed a less favorable Voice–Map change from pre to post for pleasure and vigor, and a more unfavorable Voice–Map change for multiple fatigue dimensions and the fatigue sum, though the additional variance explained was modest ($\Delta R^2 \approx 0.004$–0.018).

These results fit the idea that individuals who perceive themselves as strong navigators may benefit less (or even be mildly disrupted) by stepwise voice prompts, whereas those with a poorer sense of direction may experience greater affective relief and lower fatigue when visual search and map-to-world transformations are

offloaded. Given their exploratory nature and small effect sizes, we view these findings as hypothesis-generating: future work should test whether personalized prompt density, earlier summaries, and optional landmark anchoring are especially helpful for users with poor sense of direction, while minimal-interruption, overview-first modes better suit those with stronger navigation skills.

Usability and design implications

While the affective experience favored voice guidance, the mean usability score (60.50 ± 2.40) falls below the commonly cited threshold of ~ 70 for acceptable usability[19]. This underscores that current voice systems are not yet optimized for older users. Our qualitative analysis provides actionable recommendations for improvement:

Distance phrasing
Participants found it hard to translate metric distances into actions. Shorter, action-oriented phrasing (e.g., "Turn right in 50 m") or landmark-based cues ("After the bank, turn right") would be more effective.

Prompt timing
Users preferred earlier and layered prompts (e.g., a preview, a pre-turn alert, and an at-turn confirmation), similar to modern in-car navigation systems.

Intersection clarity
Prompts should clearly distinguish between actions like "turn right at the signal" versus "cross the street, then turn right".

Landmark anchoring
Adding visible landmarks such as banks or traffic signals is essential for older adults, as they often rely on salient environmental cues[3,7–9].

Robustness to noise
Participants reported difficulty hearing instructions in traffic. Features such as adaptive volume, repeat options, or multimodal cues like haptics could address this issue.

Positioning vs prior work

Our design complements prior studies that compared voice/auditory aids with digital visual maps or AR overlays[15,16] by (i) isolating modality through a voice-only vs paper-map contrast that avoids smartphone UI confounds; (ii) testing adults 50–69 in realistic urban routes rather than student/VR samples; and (iii) extracting actionable phrasing/timing/landmark requirements via text mining in a non-English (Japanese) context, which targets immediate design improvements for off-the-shelf systems.

Positioning within the health context

Independent mobility is a cornerstone of autonomy and healthy aging[4,13,14]. Navigation tools can support this by making walking and wayfinding less stressful. However, these tools are enablers, not health interventions. Our study appropriately focuses on proximal outcomes: wayfinding performance, affect, and usability. Claims about reducing long-term risks such as falls or dementia would require longitudinal studies in at-risk populations[4,5,13,14]. We therefore position these findings as crucial feasibility and design evidence that can inform future clinical and public health research.

Limitations

Our findings should be considered in light of several limitations. First, our convenience sample of healthy adults aged 50–69 limits generalizability to those over 70, people with cognitive impairments, or those with less technology experience[3,13,15]. Second, the paper-map comparator isolates the auditory modality but does not address questions about smartphone-based visual navigation; this will require a separate comparison. Third, our findings are based on short (~ 800 m), daytime, urban routes and may not apply to longer journeys or rural/night-time conditions. Fourth, we relied on self-reported affective states and did not include objective cognitive load or stress indices such as heart-rate variability, pupillometry, or dual-task costs. Fifth, we intentionally applied SUS to the voice system under design scrutiny. Paper maps are not an interactive software system in the SUS sense, so parallel SUS scoring was not performed. Nevertheless, future work could simultaneously incorporate SUS or another complementary instrument adapted for maps to contextualize subjective perceptions across modalities. Last, our findings reflect a single navigation app in a Japanese urban context, and replication is needed across different platforms and locations.

Conclusion

For healthy adults aged 50–69 navigating short urban routes, voice-only guidance matched paper maps performance while making the experience more relaxing. This greater relaxation may reflect reduced perceived cognitive effort when visual search and map-to-world transformations are offloaded to auditory cues. However, because the present study did not include objective workload metrics, this mechanism should be regarded as hypothetical and verified in future research. Furthermore, the system's usability was only moderate. Participant feedback provides a clear roadmap for improvement, which focuses on distance wording, prompt timing, intersection clarity, landmark anchoring, and noise robustness. These results provide a foundation for designing better navigation tools and for conducting rigorous evaluations in older and more cognitively vulnerable populations.

Data availability

The data that support the findings of this study are available on request from the corresponding author.

References

1. Webber, S. C., Porter, M. M. & Menec, V. H. Mobility in older adults: A comprehensive framework. *Gerontologist* **50**(4), 443–450 (2010).
2. Rosso, A. L., Taylor, J. A., Tabb, L. P. & Michael, Y. L. Mobility, disability, and social engagement in older adults. *J. Aging Health* **25**(4), 617–637 (2013).
3. Weisberg, S. M., Ebner, N. C. & Seidler, R. D. Getting LOST: A conceptual framework for supporting and enhancing spatial navigation in aging. *Wiley Interdiscip. Rev. Cogn. Sci.* **15**(2), e1669 (2024).
4. Chen, C. & Nakagawa, S. Physical activity for cognitive health promotion: An overview of the underlying neurobiological mechanisms. *Ageing Res. Rev.* **86**, 101868 (2023).
5. Izquierdo, M., Duque, G. & Morley, J. E. Physical activity guidelines for older people: Knowledge gaps and future directions. *Lancet Healthy Longev.* **2**(6), e380–e383 (2021).
6. De Silva, N. A. et al. Examining the association between life-space mobility and cognitive function in older adults: A systematic review. *Hindawi J. Aging Res.* **2019**, 1–9 (2019).
7. Moffat, S. D. & Resnick, S. M. Effects of age on virtual environment place navigation and allocentric cognitive mapping. *Behav. Neurosci.* **116**(5), 851 (2002).
8. Harris, M. A. & Wolbers, T. Ageing effects on path integration and landmark navigation. *Hippocampus* **22**(8), 1770–1780 (2012).
9. Yamamoto, N. & DeGirolamo, G. J. Differential effects of aging on spatial learning through exploratory navigation and map reading. *Front. Aging Neurosci.* **4**, 14 (2012).
10. Hu, X., Song, A., Wei, Z. & Zeng, H. Stereopilot: A wearable target location system for blind and visually impaired using spatial audio rendering. *IEEE Trans. Neural Syst. Rehabil. Eng.* **30**, 1621–1630 (2022).
11. El-Taher, F. E. Z. et al. A survey on outdoor navigation applications for people with visual impairments. *IEEE Access* **11**, 14647–14666 (2023).
12. Nagata, C. et al. Exploring voice navigation system usage in healthy individuals: towards understanding adaptation for patients with dementia. *Med. Sci. Innov.* **71**(3–4), 65–73 (2024).
13. Pillette, L. et al. A systematic review of navigation assistance systems for people with dementia. *IEEE Trans. Visual. Comput. Graph.* **29**(4), 2146–2165 (2022).
14. Peng, W. et al. Mediating role of homebound status between depressive symptoms and cognitive impairment among community-dwelling older adults in the USA: A cross-sectional analysis of a cohort study. *BMJ Open* **12**(10), e065536 (2022).
15. Xu, Y., Qin, T., Wu, Y., Yu, C. & Dong, W. How do voice-assisted digital maps influence human wayfinding in pedestrian navigation?. *Cartogr. Geogr. Inf. Sci.* **49**(3), 271–287 (2022).
16. Rehrl, K., Häusler, E., Leitinger, S. & Bell, D. Pedestrian navigation with augmented reality, voice and digital map: final results from an in situ field study assessing performance and user experience. *J. Locat. Based Serv.* **8**(2), 75–96 (2014).
17. Huang, J., Wang, H., Ding, S. & Wang, S. DuIVA: An intelligent voice assistant for hands-free and eyes-free voice interaction with the Baidu maps App. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining.* 3040–3050 (2022)
18. Ministry of Internal Affairs and Communications, Government of Japan. Summary of the 2023 Communication Usage Trend Survey Results (Outline). Available at https://www.soumu.go.jp/main_content/000950622.pdf?form=MG0AV3. Accessed 3, Feb 2025. (in Japanese).
19. Yamashita, R. et al. The mood-improving effect of viewing images of nature and its neural substrate. *Int. J. Environ. Res. Public Health* **18**(10), 5500 (2021).
20. Chen, C., Hagiwara, K. & Nakagawa, S. Introducing the Chen-HAgiwara Mood Test (CHAMT): A novel, brief scale developed in Japanese populations for assessing mood variations. *Asian J. Psychiatr.* **93**, 103941 (2024).
21. Kubo, T. et al. Characteristic patterns of fatigue feelings on four simulated consecutive night shifts by "Jikaku-sho Shirabe". *Sangyo Eiseigaku Zasshi* **50**(5), 133–144 (2008) (**in Japanese**).
22. Bangor, A., Kortum, P. & Miller, J. Determining what individual SUS scores mean: Adding an adjective rating scale. *J. Usability Stud.* **4**(3), 114–123 (2009).
23. Hatakeyama, Y., Sasaki, R. & Ikeda, N. Preliminary investigation of visual memory in Alzheimer's disease patients and elderly depressed patients using a touch panel task. *J. Geriatr. Psychiatry* **17**(6), 655–664 (2006) (**in Japanese**).
24. Ministry of Health, Labour and Welfare, Government of Japan. Healthy Life Expectancy Award, Care Prevention and Elderly Life Support Field, Minister of Health, Labor and Welfare Award, Company Category, Excellence Award. https://www.smartlife.mhlw.go.jp/pdf/award/award_8th_kaigo.pdf, 8, (2008) (in Japanese).
25. Takeuchi, Y. Sense of direction and its relationship with geographical orientation, personality traits and mental ability. *Jpn. J. Educ. Psychol.* **40**, 47–53 (1992) (**in Japanese**).
26. Kato, Y. & Takeuchi, Y. Individual differences in wayfinding strategies. *J. Environ. Psychol.* **23**(2), 171–188 (2003).

Author contributions

Conception: C.N., Methodology: C.N., T.Y., C.C., Investigation: C.N., T.O., Y. Sumida., Y.I., Y. Shimizu, Formal analysis: C.N., T.Y., C.C., Funding acquisition: C.N., Writing—original draft: C.N., C.C., Writing—reviewing & editing: all.

Funding

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-30590-5.

Correspondence and requests for materials should be addressed to C.N. or C.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.