Eyes on the Street: An Outdoor Study of Pedestrian Navigation Support Using Augmented Reality Glasses

Maximilian Rose
NUILab - Computer Science Department
Colorado State University
Fort Collins, Colorado, USA
maximilian.rose@colostate.edu

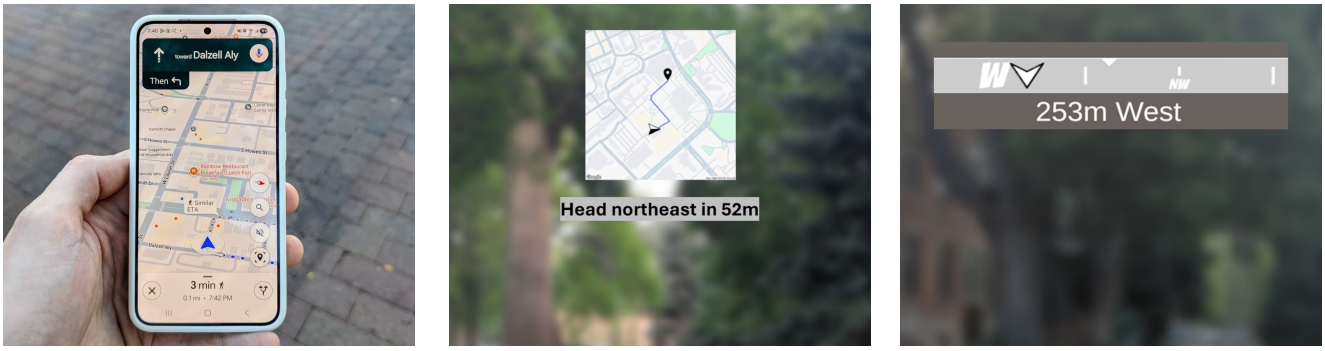
Zahra Borhani
Computer Science Department
Colorado State University
Fort Collins, Colorado, USA
faezeborhani@gmail.com

Dylan Schreiber
Computer Science
Colorado State University
Fort Collins, Colorado, USA
dylan.schreiber@colostate.edu

Nwadilioramma O Azuka-Onwuka
Faculty of Computer Science
Dalhousie University
Halifax, Nova Scotia, Canada
diliora.azukaonwuka@gmail.com

Mayra Donaji Barrera Machuca
Department of Computer Science
University of Calgary
Calgary, Alberta, Canada
mbarrera@ucalgary.ca

Francisco Raul Ortega
Computer Science
Colorado State University
Ft. Collins, Colorado, USA
fortega@colostate.edu



(a) Handheld (Google Maps) (b) AR Map (c) AR Compass

Figure 1: The three navigation interface conditions. (a) the Handheld baseline showing Google Maps with a prescriptive route; (b) the AR Map rendering a prescriptive route overlay in the AR glasses; (c) the AR Compass showing only a directional bearing and distance to the destination based on the user's current orientation. Map imagery in (a) and (b) © 2026 Google.

Abstract

Smartphone maps dominate pedestrian navigation. Yet, they demand repeated downward glances, diverting attention from the surrounding environment. Augmented reality (AR) glasses can place guidance in the user's field of view, but the most effective visualization for guiding users remains unclear. We present a between-subjects outdoor study (N = 59) comparing two AR interfaces against a smartphone baseline: an allocentric AR Map reproducing familiar turn-by-turn routing in a map, and a minimal egocentric AR Compass showing only a bearing using the head orientation and distance to the destination. Both AR interfaces yielded shorter total walked routes than the smartphone, though this difference arose almost entirely after participants reached the vicinity of the destination. All three conditions had comparable perceived workload and usability, but both AR conditions showed greater route-level speed variability. For spatial memory of the route, post-navigation sketch maps show that the AR Compass produced less useful route sketches than the route-based interfaces. Together, these findings suggest that the value of AR pedestrian navigation lies not in a single optimal display but in matching the interface to the user's goal. We distill these findings into design recommendations for minimal AR pedestrian navigation.

CCS Concepts

• **Human-centered computing** → **Mixed / augmented reality**; **Field studies**; **Usability testing**; **User studies**.

Keywords

Augmented reality, AR glasses, pedestrian navigation, wayfinding

This work is licensed under a Creative Commons Attribution 4.0 International License.

VRST '26, Sendai, Japan

© 2026 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-2811-2/2026/11
https://doi.org/10.1145/3822517.3848729

ACM Reference Format:
Maximilian Rose, Zahra Borhani, Dylan Schreiber, Nwadilioramma O Azuka-Onwuka, Mayra Donaji Barrera Machuca, and Francisco Raul Ortega. 2026. Eyes on the Street: An Outdoor Study of Pedestrian Navigation Support Using Augmented Reality Glasses. In *32nd ACM Symposium on Virtual Reality Software and Technology (VRST '26), November 16–18, 2026, Sendai, Japan.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3822517.3848729

1 Introduction

A new wave of consumer-oriented Augmented Reality (AR) glasses—such as the Meta Ray-Ban Display and the XREAL Air 2 Ultra—differs markedly from the previous generation of AR head-mounted displays (HMDs), including the Microsoft HoloLens 2 and Magic Leap 2. Whereas those devices are specialized for rendering stereoscopic content spatially registered to the physical world, consumer glasses are frequently used as a lightweight, head-stabilized display in the user's view. In this mode, they do not fully satisfy Azuma's classical definition of AR, which requires that a system combine real and virtual content, be interactive in real time, and register virtual content in three dimensions [5]. Yet, they present opportunities to consider AR use cases, such as outdoor pedestrian navigation. Our motivation is anchored in this shift toward everyday AR, and in how such devices might improve routine, real-world tasks.

Navigation is a fitting place to start, since the tool most people already rely on is more flawed than its ubiquity suggests [58]. The smartphone has become so central to how we find our way through unfamiliar cities that its shortcomings often go unexamined: walking with a phone angled toward them, pedestrians alternate their gaze between the screen and the pavement, repeatedly breaking contact with their surroundings. These limitations are twofold. The first is postural: consulting a handheld screen or other traditional paper maps forces pedestrians to look down repeatedly, and this frequency of glances correlates directly with navigational errors and walking stops [56]. Each downward glance also breaks visual contact with a continuously changing environment, degrading the perceptual foundation of situational awareness [15]. General phone use while walking is associated with an increased rate of pedestrian injuries [46], and experimental work links it directly to degraded situational awareness and altered gait when negotiating obstacles [32], adding a safety concern to handheld navigation. The second challenge is representational: a 2D top-down map must be mentally rotated and translated into 3D egocentric space [26] before it can guide movement, a transformation that significantly taxes working memory [64]. Users return to the screen precisely because this transformation is difficult to hold in mind — the down-glance is the visible symptom, and the mental rotation it serves is a substantial part of the underlying cause.

AR glasses offer a path to address both. Rendering navigation information in the user's field of view eliminates the down-glance without surrendering environmental contact, and AR's egocentric frame makes it possible to bypass the 2D-to-3D mental rotation altogether [66]. Prior empirical comparisons have shown the benefits of using AR navigation interfaces over non-AR alternatives for wayfinding [2, 29, 51]. However, in outdoor settings, these results remain inconsistent and underexplored [52]. Although outdoor environments have been used in AR navigation research [47, 63], most comparisons against non-AR alternatives have used handheld AR rather than AR glasses.

A small number of studies approach our comparison. Montuwy et al. [43] evaluated AR glasses against a smartwatch-based digital map for older pedestrians outdoors. Most relevant to our minimal AR Compass, Mazurkiewicz et al. [42] proposed *beeline* navigation (BeeAR), in which the destination is marked by a persistent AR landmark with no map or route, requiring users to choose their own path; compared against earlier free-choice and turn-by-turn navigation on the same outdoor routes, BeeAR users retraced their route and recalled more points of interest than free-choice users. Lakehal et al. [30, 31] likewise compared pedestrian spatial-knowledge acquisition between smartphones and AR glasses outdoors. Across these studies, the measures used vary considerably, making it important to examine multiple performance and experience dimensions, and the trade-offs among them, together.

This study presents a between-subjects study ($N = 59$) in which participants navigated a continuous outdoor route of five subroutes across a university campus using one of three interfaces: a) a *handheld* smartphone running Google Maps (baseline), see Figure 1a; b) an allocentric *AR Map*, a heads-up translation of the familiar top-down map, see Figure 1b; or c) an egocentric *AR Compass*, a minimal bearing-and-distance display, see Figure 1c. The two AR interfaces are designed to foreground different aspects of the navigation interface: the AR Map presents familiar top-down route-based guidance in a heads-up display, while the AR Compass additionally reduces the interface to a minimal egocentric direction-and-distance cue. We evaluated each interface along three complementary dimensions: navigation performance (path efficiency, walking speed, and route-level speed variability), spatial memory (assessed through post-navigation sketch maps), and user experience (workload, usability, and presence). The guiding research questions are:

- RQ_1**:** Does usage of AR glasses improve wayfinding performance and spatial memory for pedestrians compared to a traditional handheld device?
- RQ_2**:** How minimal can a navigation aid in a heads-up display be while still supporting efficient navigation, acceptable user experience, and spatial learning?
- RQ_3**:** What constitutes optimal usability and design for pedestrian navigation within the AR design space?

Both AR interfaces produced shorter walked routes than the smartphone, though an exploratory decomposition traced this difference almost entirely to behavior after reaching a destination rather than to more direct wayfinding, identifying arrival confirmation as a key weakness of handheld GPS navigation. This came at a shared cost: route-level speed variability rose in both AR conditions, while perceived workload, usability, and presence did not differ across any of the three. Finally, the minimal egocentric AR Compass matched the full map on navigation efficiency but produced less useful holistic route sketches. Landmark count and placement accuracy remained comparable across conditions, suggesting that the observed difference was specific to the overall route representation captured by the sketch rather than spatial memory broadly.

2 Related Work

2.1 AR vs. Non-AR Navigation

Two recent systematic reviews characterize AR wayfinding as an active but uneven research area [52, 70]. Prior research has compared AR navigation with paper maps [2, 18], digital and smartphone maps [34, 43, 51], voice guidance [23], and non-aided navigation [24], and frequently shows advantages in completion time, distance, and error rates [52]. Head-worn AR in particular has been associated with reduced mental and physical load and improved situational awareness and spatial learning relative to handheld AR [70].

This advantage is far from uniform, however, and the evidence is especially mixed for outdoor navigation. Several outdoor field studies found no significant difference between a handheld AR interface and a conventional digital map or voice guidance on wayfinding performance: Dong et al. [13] found comparable wayfinding performance between smartphone AR and a 2D electronic map; Dünser et al. [14] found no overall difference between an AR browser and a digital map, with the AR view less useful in some conditions; and Rehrl et al. [53] found that a digital map and voice guidance outperformed AR. The outdoor studies reporting a clear AR benefit have typically used weaker baselines such as paper maps or no guidance, or head-worn rather than handheld displays [2, 42, 43]. Two characteristics of this body of work limit its bearing on our question. First, the majority of evaluated systems are handheld or mobile rather than head-worn. Second, where outdoor comparisons against a strong digital-map baseline exist, they are largely confined to handheld AR and yield mixed results, leaving open the question of whether AR glasses offer a measurable advantage over the smartphone maps people actually use outdoors.

2.2 Information Design in AR Glasses

Beyond whether AR helps, a parallel line of work compares AR interfaces against one another to isolate which design choices matter. Two axes organize this space. The first is what information is shown: drawing on cognitive psychology, the displayed content is commonly framed in terms of landmark, route, and survey knowledge [57]. The second is how that information is anchored. Following the frame-of-reference classification of Billinghurst et al. [7], we distinguish world-stabilized AR, in which virtual elements are fixed to physical coordinates in the environment, from head-stabilized AR, in which elements remain fixed to the user's field of view regardless of head movement [9]. Our two AR interfaces sample this design space.

Several findings within this space motivate our design choices. Adding world-stabilized survey information does not reliably improve performance: Vasquez et al. [62] found that conventional north-up and track-up maps yielded fewer obstacle collisions and lower workload than a world-stabilized AR map, leading Zhao et al. [70] to caution that overlaying survey information onto the world is often a poor design choice. This became one of the considerations behind our decision to render the AR Map as a head-stabilized rather than world-stabilized overlay. On the question of reference frame, most directly relevant to our study, Xu et al. [66] compared egocentric and exocentric (allocentric [26]) AR cues on a head-mounted display and found the egocentric presentation faster—though, as with most of this work, the study was indoors. The form of the directional cue is likewise contested: Arntz et al. [4] found that a continuous navigation path outperformed discrete arrows. Liu et al. [39] showed that AR can actively mislead spatial memory, an effect partly mitigated by showing a route line rather than arrows alone.

These results expose a recurring tension between efficiency and spatial learning. Turn-by-turn guidance frees cognitive resources but can impair the spatial knowledge users retain. This effect is well documented for GPS-based navigation generally [12, 17, 25], and observed for AR specifically by Dong et al. [13], whose AR participants formed weaker route memory than those using a 2D map despite comparable wayfinding performance. A further constraint is the limited field of view of current AR glasses, which makes world-stabilized cues in the periphery easy to overlook [70] and argues for keeping persistent information minimal and centrally glanceable.

Few studies combine the threads applied to our setting. The closest precedent to our minimal AR Compass is the "beeline" bearing-to-target display of Mazurkiewicz et al. [42]: evaluated outdoors, BeeAR marks the destination with a persistent AR landmark and no map, and its users retraced their route and recalled more points of interest than free-choice navigators. Lakehal et al. [31] compared pedestrian spatial knowledge acquisition between smartphones and AR glasses outdoors, and found that AR glasses afforded significantly better memorization of both landmarks and survey knowledge. Their evaluation centered on spatial memory, however, and did not examine wayfinding efficiency; their two conditions also differed only in form factor: both delivered the same landmark-anchored, turn-by-turn guidance at decision points, leaving the effect of the underlying information design unexamined. What remains untested, therefore, is the specific comparison we make: *egocentric* and *allocentric* AR-glasses interfaces—varying the information design while holding the head-worn form factor constant—evaluated side by side outdoors against the everyday smartphone baseline, across performance, spatial memory, and experience.

2.3 Attentional Demands of Walking with AR

Walking, though often experienced as automatic, draws measurably on cognitive resources. A substantial dual-task literature shows that performing a concurrent cognitive task while walking degrades gait performance, most prominently by reducing walking speed—a phenomenon termed cognitive-motor interference (CMI) [1]. CMI is directly relevant to pedestrian navigation: consulting a navigation interface while walking is effectively a continuous dual-task in which spatial information must be read, interpreted, and translated into movement without interrupting locomotion. Beyond mean walking speed, such interference may surface as reduced route-level speed consistency. For head-mounted displays, this load may be compounded by novelty: guidance efficiency has been found to depend on a user's familiarity with the AR display [61].

3 UI Design

The design of our three conditions is organized around a central tension in pedestrian navigation: the trade-off between navigational efficiency and spatial learning. Prescriptive turn-by-turn

guidance can produce what has been described as "mindless" navigation [10], in which users reach a destination without forming a durable understanding of the space they traversed [44]. A minimal, non-prescriptive interface inverts this: by conveying only the direction toward a destination rather than a fixed route, it obliges users to interpret the environment and construct their own path. Whether this heightened engagement costs raw efficiency, or instead yields comparable performance while strengthening spatial learning, remains an open question.

The two AR interfaces evaluated in this study represent a subset of the broader AR navigation space, which spans different guidance techniques, including conventional arrows, in-situ signage, and world-anchored visual cues [35, 70]. The three conditions of this study vary in how guidance is presented and what information is provided. The Handheld condition serves as the baseline, providing prescriptive turn-by-turn navigation on a smartphone, reflecting current practice. The AR Map preserves the same broad guidance paradigm, while moving the information from a heads-down handheld device to a heads-up display. The AR Compass used the same head-worn hardware as the AR Map but replaces the prescriptive route map with a minimal egocentric direction-and-distance cue.

Smartphone Google Maps. Google Maps on a smartphone was selected as the baseline because it is by far the most widely used pedestrian navigation technique [19], lending the comparison ecological validity: any proposed heads-up system must justify itself against what people actually use, not against a convenient laboratory control. This is also the comparison most conspicuously missing from prior AR glasses research [52, 70], which has largely evaluated head-worn interfaces against one another or against paper maps rather than the smartphone standard. The application provides prescriptive turn-by-turn routing on a dynamically oriented map with combined text and visual instructions. Participants held and operated the phone as they normally would; position was logged unobtrusively by a background application.

Allocentric AR Map. The AR Map recreates the baseline smartphone navigation paradigm in heads-up form. It presents a top-down map rendered as a head-stabilized overlay in the user's field of view, preserving the prescriptive route line and street- and turn-based text instructions familiar from smartphone navigation (Figure 1b). The UI is primarily designed around one key difference: it uses a static north-up presentation rather than a dynamic track-up rotation, with the user's position and heading shown by a rotating marker at the map's center. This choice serves the study's core logic. The two AR conditions were designed to dissociate *why* down-glances impair handheld navigation. If the cost of a down-glance lies simply in diverting visual attention from the environment, then a heads-up map should resolve this regardless of its orientation. Prior work establishes that reading a north-up map requires a mental rotation to align the map's world-centered frame with the viewer's egocentric forward view, and that a track-up orientation removes part of this transformation [3, 64]. A north-up AR Map deliberately retains this allocentric-to-egocentric transformation, isolating it as the variable that distinguishes the two AR conditions. With the rotating position marker the user is supplied the same route information as a rotating map, while still requiring the mental rotation of aligning map to world. A static orientation also avoids the continuous compass-driven spin of a track-up display, which can introduce visual motion that becomes uncomfortable when rendered persistently in the field of view while walking. However, the AR Map also differs from Google Maps in more than posture alone. Beyond the heads-up form factor and the north-up orientation, the custom implementation also differs in its underlying rendering and update behavior, and its display of arrival. We therefore treat the Handheld-AR Map contrast as primarily reflecting display placement, while noting these co-varying factors.

Egocentric AR Compass. The AR Compass explores the minimal end of the design space (RQ_2). A directional arrow anchored to the user's heading points toward the destination's bearing, accompanied by the Euclidean distance to the target (Figure 1c). We render the Compass head-stabilized rather than world-anchored. While world-anchored targets can afford faster visual acquisition [41], world registration outdoors depends on GPS- and compass-derived pose whose jitter and latency make a stable world-locked overlay impractical; head-stabilization trades some acquisition speed for display stability. To make the arrow legible as a bearing, it sits against a rotating compass rose with a fixed heading marker, and the distance is annotated with the target's cardinal direction (e.g., "200 m west"). None of these elements convey route information: there is no map, route line, or turn instruction. The rose and cardinal labels act as signifiers [48] that communicate how to read the arrow rather than as navigational content in their own right, adding no information that the directional arrow doesn't already convey. The interface remains non-prescriptive, leaving all routing decisions to the user. Four considerations motivated this design. First, operating in the user's egocentric frame avoids the 2D-to-3D mental rotation required by top-down maps [64, 66]. Second, the minimal visual footprint keeps most of the visual field available, supporting the situational awareness that handheld navigation degrades [15]. Third, the design draws on the screen-edge compasses of commercial video games, which keep wayfinding support peripheral so that environmental engagement stays primary [49]. Fourth, the technique is inherently heads-up: on a handheld device a bearing display would force users to raise and reorient the phone after every heading change, reintroducing the very glance behavior it seeks to eliminate. Prior work indicates that content placed centrally or low in the view supports comprehension while walking better than top or peripheral placement [54]. Our persistent compass bar sits near the top of the field of view, a placement we revisit in light of participant feedback.

4 User Study

4.1 Methodology

Participants. We recruited 60 participants from a university campus. One participant was excluded because their data were not recorded correctly, leaving a final sample of 59 participants. This study was approved by the Institutional Review Board (IRB) of Colorado State University (Protocol #7278). Participants were assigned to one of three conditions using a counter-balanced Latin-square: AR Compass (n = 20), AR Map (n = 19), or Handheld smartphone (n = 20). All participants reported normal or corrected-to-normal vision and had no known mobility impairments. Participants with

corrected-to-normal vision were instructed to use contact lenses if possible, as the headset does not accommodate for glasses worn under it. Age ranged from 19 to 35 years (M = 23.7, SD = 4.2). Participants identified as male (n = 37), female (n = 19), or non-binary (n = 3). Participants received course credit as compensation for their time.

Apparatus. Both AR conditions used the commercially available XREAL Air 2 Ultra six-degree-of-freedom (6DoF) AR Glasses with a Samsung Galaxy S25 smartphone running a custom application developed in Unity (version 2022.3.21f1), with interaction logic implemented in C#. The application rendered interface elements as head-stabilized overlays, fixed to the user's field of view rather than anchored to world coordinates, consistent with the Rest-Frame Hypothesis [50]. GPS positioning was provided by the paired smartphone, logging position at five-second intervals, chosen to limit routing API requests over the course of the study. The user's heading, in contrast, was read directly from the device compass and updated several times per second. On the AR Map, the position marker therefore rotated responsively with the user's heading, though it was rendered in eight discrete 45° increments rather than continuously. Route geometry and turn-by-turn directions for the two AR conditions were obtained from the Google Maps Routes API (`computeRoutes` endpoint), queried with pedestrian travel mode for each origin–destination waypoint pair. A single route was requested per leg, with default routing preferences and no route modifiers. The same API-derived route lengths served as the reference (optimal) distances for the path-efficiency metric. Participants in the Handheld condition used a Samsung Galaxy S25 smartphone running Google Maps. GPS position was logged at 1 Hz using BasicAirData GPSLogger [6]. To ensure comparable sampling across conditions, these traces were downsampled to five-second intervals prior to analysis, matching the AR conditions' logging rate; all speed and distance measures were computed from the resampled data.

Procedure. Upon arrival, participants provided written informed consent and completed a brief demographics questionnaire. Participants assigned to an AR condition were then fitted with the head-mounted display paired with a Samsung Galaxy S25 smartphone. Those in the Handheld condition were only provided with the smartphone. All participants then completed a familiarization phase in which they explored their assigned interface.

The navigation task consisted of a continuous route consisting of five subroutes across Colorado State University campus, each directing the participant to a distinct GPS waypoint. Routes were designed with two goals beyond safety. First, to limit reliance on prior spatial knowledge, waypoints were placed at locations between buildings rather than at obvious, easily-named landmarks (e.g., a building's main entrance), so that participants could not identify the destination at a glance and disengage from the interface. Second, to reduce pedestrian congestion, routes avoided main thoroughfares, and sessions were scheduled during class periods (10–15 minutes past the hour) rather than the busier between-class intervals. Participants consequently encountered only sparse, intermittent pedestrian traffic. Participants self-determined arrival and verbally signaled the experimenter, who then manually advanced the system to the next subroute and thus only intervened between subroutes. This procedure was chosen to simulate a realistic user interaction for determining arrival. Each AR system did not signal arrival, while the baseline Google Maps displayed an arrival state upon being near the target. In all three conditions, however, the experimenter advanced to the next subroute only upon the participant's verbal confirmation, so the advancement procedure itself was identical. We mainly aimed to compare the visual channel, so any other output modality such as audio or haptic feedback was disabled. This applied throughout navigation, not only at arrival: the Handheld condition's turn-by-turn voice guidance was also disabled, so all conditions conveyed directions visually (via text and on-screen cues) rather than by audio. Participants were informed about this, so they wouldn't expect notifications. Upon completing all five subroutes, participants sat with the experimenter and completed three questionnaires on a tablet: the NASA-TLX [20], CSUQ [37], and IPQ [55]. They were then given a blank sheet of paper and asked to draw, from memory, the routes they had walked and any spatial features they could recall, following the sketch-map elicitation protocol established by Billinghurst and Weghorst [8]. This was followed by a brief semi-structured interview in which the experimenter asked the participant to describe their map-drawing strategy and their experience of the navigation interface. Finally, participants completed the Spatial Orientation Test (SOT) [22]. This order was chosen to conduct any measures related to the task immediately after its conclusion. The SOT was consequently placed at the end of the session, as the opening window was reserved for equipment setup and familiarization. The total session duration ranged from approximately 45 to 60 minutes.

Experimental Design. The study employed a between-subjects design with Navigation Interface (*AR Compass, AR Map, Handheld*) as the single independent variable. The subroutes were completed in a fixed order, with each start point coinciding with the previous endpoint, resulting in a single continuous route across the study area. Participants were instructed to navigate to each waypoint as efficiently as possible using only the assigned interface. They were not given additional verbal directions. To avoid differences in route and environment, we chose a between-subjects design. The route was designed to avoid moving cars, but to ensure safety further, the experimenter followed at a distance that allowed them to intervene in potentially dangerous situations.

4.2 Metrics

Spatial Ability. Baseline spatial reasoning ability was measured using the Spatial Orientation Test (SOT) [22], a validated psychometric instrument that requires participants to imagine being at one landmark in an environment and point toward another. Each trial is scored as the absolute angular error between the indicated and true directions (0–90°). Participants completed nine trials; trials not completed within the time limit were assigned a ceiling penalty of 90°, consistent with standard SOT scoring practice [22]. The SOT was administered at the end of the session. While administering a baseline measure post-task is atypical, randomization should distribute any task-induced effects evenly across conditions, and indeed no between-group difference was found ($F(2, 56) = 0.02$, $p = .977$), indicating the measure was not differentially biased.

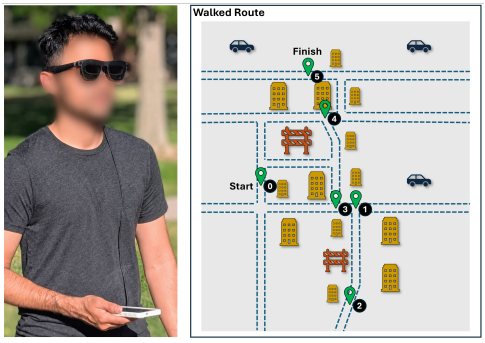


Figure 2: Overview of the study setup. Left: a participant navigating with the AR glasses. Right: the six waypoints defining the study route, numbered in the order visited. Participants walked from Start through waypoints 1—4 to Finish as one continuous route; because the AR Compass provided no prescribed path, no single walked path is shown. Icons © Microsoft.

Walking Speed. Active mean walking speed was computed via the five-second GPS samples over the active movement phase of each route, excluding stationary periods below 0.35 m/s at either end (Phone: $M = 1.28 \pm 0.12$ m/s; AR Compass: $M = 1.37 \pm 0.15$ m/s; AR Map: $M = 1.30 \pm 0.13$ m/s). Stationary periods at the beginning and end of each route —– arising from the time taken to read the interface before setting off, or to confirm arrival — were excluded to isolate locomotion-phase speed from interface interaction time.

Route-Level Speed Variability. Route-level speed variability was operationalized as the standard deviation of walking speed within the active phase of each route (SD in m/s). Prior work has shown that walking speed can decrease under cognitive dual-task conditions, reflecting the shared cortical resources demanded by concurrent cognitive and motor processing [1]. Speed standard deviation was therefore used as a complementary index of route-level speed consistency, with greater variability taken to reflect more frequent stop-and-start behavior during navigation. We emphasize that this is a coarse, route-level measure derived from position sampled every five seconds, not a stride-level gait metric: it captures fluctuations in walking pace rather than gait kinematics, and elevated values may reflect hesitation and attentional interruption but also turns, momentary stops, or positioning noise.

Path Efficiency. Path efficiency was defined as the ratio of the walkable reference distance to the participant's actual walked distance, a distance-based efficiency measure common in AR navigation evaluation [24, 52], here expressed so that 1.0 denotes an optimal path. We use the shortest walkable route (via Google Maps) rather than straight-line distance as the reference, so that efficiency reflects route choice rather than unavoidable detours around physical obstacles.

User Experience. User experience was captured with three standard questionnaires. Perceived workload was measured using the raw NASA Task Load Index (NASA-TLX; unweighted mean of six 20-point subscales, range 0–100), which shows sensitivity equivalent to the weighted version [20, 21]. Usability was assessed using the Computer System Usability Questionnaire (CSUQ) [37], a 19-item instrument (1–7, higher is better) selected over the SUS for its greater dimensionality [38]. Spatial presence was measured with the Igroup Presence Questionnaire (IPQ) [55]; as the IPQ was developed for enclosed VR, its applicability to outdoor MR is a known limitation [59], and it was included primarily for comparability with prior HMD research.

4.3 Data Analysis

Before analysis, normality was assessed via Shapiro-Wilk tests and Q-Q plot inspection. Active mean walking speed ($W = .973$, $p < .001$) and route-level speed variability ($W = .971$, $p < .001$) showed only mild departures from normality and were analyzed using linear mixed-effects models (LMMs) with Condition as a fixed effect and ParticipantID and Route as crossed random intercepts. Path efficiency ($W = .808$, $p < .001$) and all questionnaire measures were severely skewed and were analyzed using the Aligned Rank Transform (ART) [65]. ART models for route-level data included ParticipantID as a random intercept to account for five repeated routes per participant; participant-level models (overshoot decomposition, questionnaire scores) used one-way ART without random effects. Main effects are reported with Kenward-Roger degrees of freedom for LMMs. Post-hoc pairwise contrasts used Tukey correction throughout. Effect sizes are reported as marginal R^2 for LMMs [45] and partial η^2 for ART ANOVAs. We used mixed-effects models to accommodate the unequal group sizes ($n = 20/20/19$) arising from the excluded participant without bias [40]; all fixed-effect tests used Type III sums of squares with Kenward-Roger degrees of freedom.

A summary of all statistical tests and pairwise contrasts is provided in Table 1.

4.4 Objective Navigation Performance

Walking Speed. An LMM revealed a significant main effect of Condition, $\chi^2(2) = 6.73$, $p = .035$, $R^2_m = .074$. Tukey-corrected pairwise comparisons showed that participants in the AR Compass condition walked significantly faster than those in the Handheld condition ($\Delta = 0.090$ m/s, $t(56) = 2.50$, $p = .041$; Figure 3). The AR Map condition did not differ significantly from either the Handheld ($p = .809$) or AR Compass conditions ($p = .165$). The marginal R^2 of .074 indicates that Condition accounts for approximately 7% of variance in walking speed; the conditional $R^2 = .674$ reflects substantial between-participant and between-route variability in baseline walking pace.

Route-Level Speed Variability. Speed standard deviation, an index of route-level speed consistency capturing momentary hesitation and stop-start behavior during navigation, was analyzed with a second LMM (Phone: $M = 0.21 \pm 0.12$ m/s; AR Compass: $M = 0.44 \pm 0.16$ m/s; AR Map: $M = 0.44 \pm 0.16$ m/s). A highly significant main effect of Condition was observed, $\chi^2(2) = 250.93$, $p < .001$, $R^2_m = .460$, indicating that navigation interface alone accounted

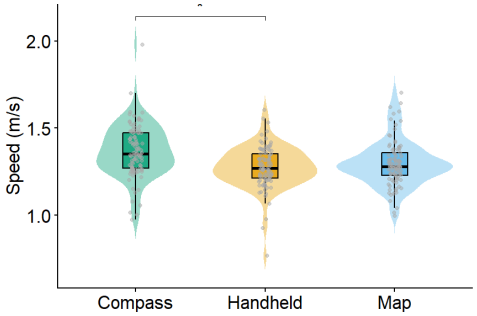


Figure 3: Active mean walking speed by condition. Only the AR Compass differed significantly from the Handheld baseline.

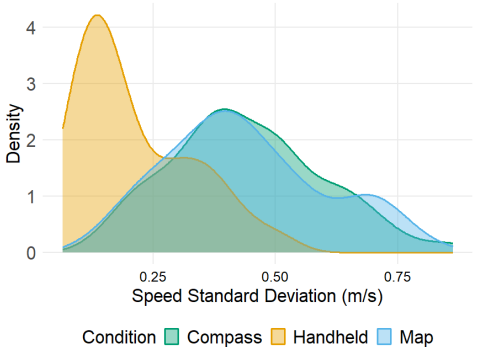


Figure 4: Route-level speed variability (SD) by condition. Both AR conditions show substantially greater variability than Handheld, and are near-identical to each other.

for 46% of variance in route-level speed variability[1]. Pairwise comparisons revealed that both AR conditions exhibited significantly greater speed variability than the Handheld condition (AR Compass vs. Phone: Δ = 0.229 m/s, $t(56) = 13.89$, $p < .001$; AR Map vs. Phone: Δ = 0.225 m/s, $t(56) = 13.48$, $p < .001$), while the two AR conditions were indistinguishable from one another ($\Delta = 0.004$, $p = .973$; Figure 4).

Path Efficiency. Path efficiency differed significantly across conditions (Phone: *Mdn* = .678 [.050]; AR Compass: *Mdn* = .710 [.059]; AR Map: *Mdn* = .709 [.065]), $F(2, 56) = 17.46$, $p < .001$, $\eta_p^2 = .38$. Both AR conditions achieved significantly higher path efficiency than the Handheld condition (AR Compass vs. Phone: $t(56) = 5.03$, $p < .001$; AR Map vs. Phone: $t(56) = 5.18$, $p < .001$), while efficiency was statistically equivalent between the two AR conditions ($p = .976$; Figure 5). The equivalence between the prescriptive AR Map and the non-prescriptive AR Compass is notable given their fundamentally different routing approaches.

Target Arrival Behavior. During analysis of path efficiency, we observed that the Handheld condition's longer total route distances did not appear to stem from less direct routes toward the targets,

[1]A boundary singular fit was returned for the Route random intercept, indicating negligible route-level variance relative to the dominant Condition effect. Point estimates and p-values are unaffected.

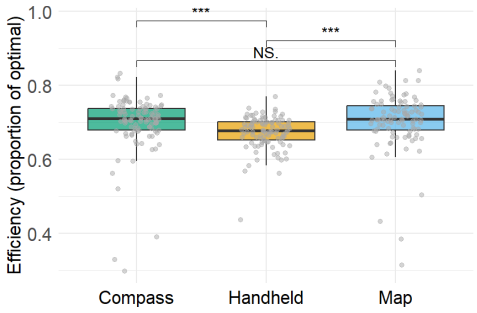


Figure 5: Path efficiency by condition. Both AR conditions were significantly more efficient than Handheld; the two AR interfaces did not differ.

but rather from additional distance accumulated near the destinations themselves. To investigate this pattern, we conducted an exploratory post hoc decomposition of total walked distance into two sequential components per route: (1) *Approach distance:* Cumulative walked distance at the moment the participant first entered a 15-meter radius of the target waypoint. This captures the efficiency of the wayfinding phase independent of post-arrival behavior. (2) *Post-arrival overshoot:* All additional walked distance accumulated after first entering the arrival zone. This captures the participant's ability to precisely confirm that they have reached the target and to cease navigation.

The 15-meter threshold corresponds to the spatial precision of consumer-grade GPS receivers under open-sky conditions [69]. Both components were averaged across the five routes per participant and analyzed using one-way ART ANOVAs, consistent with the participant-level analyses reported above. Given the post hoc nature of this analysis, the following results should be considered hypothesis-generating rather than confirmatory.

Approach Phase. An ART ANOVA on mean approach distances revealed a significant but modest effect of Condition, $F(2, 56) = 4.06$, $p = .023$, $\eta_p^2 = .13$ (Phone: *Mdn* = 165 m [5.7]; AR Compass: 166 m [9.4]; AR Map: 172 m [8.6]). Pairwise comparisons indicated that this effect was driven by a single contrast: Handheld participants had significantly shorter approach distances than AR Map participants ($t(56) = -2.85$, $p = .017$), while neither comparison involving the AR Compass reached significance ($p = .408$ vs. Phone; $p = .265$ vs. Map). Handheld users were thus at least as efficient — and marginally more direct — in navigating *to* each target zone.

Post-Arrival Overshoot. Post-arrival overshoot showed a different picture (Phone: *Mdn* = 20.4 m [2.2]; AR Compass: 7.8 m [4.1]; AR Map: 4.1 m [3.0]). An ART ANOVA returned a very large, highly significant effect of Condition, $F(2, 56) = 127.79$, $p < .001$, $\eta_p^2 = .82$. All three pairwise contrasts were significant (all $p < .001$): Handheld participants overshot targets by substantially greater distances than both AR Compass ($t(56) = 9.24$) and AR Map participants ($t(56) = 15.91$), and the AR Compass produced significantly more overshoot than the AR Map ($t(56) = 6.79$). The narrow IQR of 2.2 m for Handheld overshoot indicates that this pattern was highly consistent across participants, rather than attributable to a small number of extreme cases.

Table 1: Summary of statistical models and pairwise comparisons. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$. LMMs (Type III Wald χ^2) report marginal R^2 (variance explained by Condition alone). ART ANOVAs report partial η^2. All post-hoc contrasts are Tukey-corrected.

Metric	Model	Main Effect	Effect Size	Significant Pairwise Contrasts
Active Mean Speed	LMM	$\chi^2(2) = 6.73$, $p < .05$	$R^2_m = .074$	AR Comp > Phone*
Route-Level Speed Variability	LMM	$\chi^2(2) = 250.93$, $p < .001$	$R^2_m = .460$	AR Comp > Phone***; AR Map > Phone***
Path Efficiency	ART-MM	$F(2,\ 56) = 17.46$, $p < .001$	$\eta^2_p = .38$	AR Comp > Phone***; AR Map > Phone***
Approach Distance	ART	$F(2,\ 56) = 4.06$, $p = .023$	$\eta^2_p = .13$	Phone < AR Map*
Overshoot Penalty	ART	$F(2,\ 56) = 127.79$, $p < .001$	$\eta^2_p = .82$	Phone > AR Comp***; Phone > AR Map***; AR Comp > AR Map***
SOT Angular Error	ART	$F(2,\ 56) = 0.02$, $p = .977$	$\eta^2_p < .01$	
NASA-TLX Workload	ART	$F(2,\ 56) = 0.41$, $p = .668$	$\eta^2_p = .01$	
CSUQ Usability	ART	$F(2,\ 56) = 0.73$, $p = .486$	$\eta^2_p = .03$	
IPQ Spatial Presence	ART	$F(2,\ 56) = 0.41$, $p = .663$	$\eta^2_p = .01$	

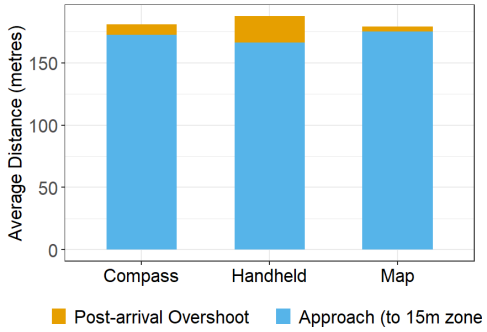


Figure 6: Decomposition of mean total route distance. Approach segments (blue) are comparable across conditions; post-arrival overshoot (orange) drives the longer Handheld routes.

The contrast in effect sizes is informative: overshoot ($\eta^2_p = .82$) substantially exceeded approach distance ($\eta^2_p = .13$), indicating that the Condition effect on total route length was driven almost entirely by post-arrival behavior.

Spatial Ability Analysis. We additionally explored whether baseline spatial ability moderated the effect of Condition, examining whether the best-performing interface differed for participants with stronger versus weaker spatial reasoning. Each participant's ability was summarized as their median SOT angular error, mean-centered, and entered together with its interaction with Condition into models for walking speed (linear mixed-effects) and path efficiency (participant-level rank-based regression [27]). No Condition × spatial-ability interaction was found for either walking speed ($F(2, 53) = 0.21$, $p = .809$) or path efficiency ($F = 0.71$, $p = .495$). However, the between-subjects design affords limited statistical power to detect interaction effects at this sample size.

4.5 Subjective User Experience

ART ANOVAs on self-reported outcomes yielded no significant differences across conditions for perceived workload (NASA-TLX: Phone = 29.0 [15.2], AR Compass = 28.5 [12.8], AR Map = 30.0 [10.0]; $F(2, 56) = 0.41$, $p = .668$, $\eta^2_p = .01$), system usability (CSUQ: Phone = 6.03 [1.20], AR Compass = 5.91 [1.27], AR Map = 5.88 [0.72]; $F(2, 56) = 0.73$, $p = .486$, $\eta^2_p = .03$), or spatial presence (IPQ: Phone = 3.75 [1.25], AR Compass = 3.82 [1.41], AR Map = 3.86 [0.82]; $F(2, 56) = 0.41$, $p = .663$, $\eta^2_p = .01$). Despite the substantial objective differences reported above, participants rated the cognitive demands, usability, and sense of presence of the three interfaces as equivalent.

4.6 Post-Navigation Sketch Map Analysis

Post-navigation sketch maps were scored blind to condition along three components adapted from Billinghurst and Weghorst [8]: *map usefulness* (a holistic 1–3 rating of how well the sketch would guide a stranger along the route), *landmark count*, and *landmark placement accuracy* (proportion of drawn landmarks correctly positioned). Scoring reliability was quantified using intraclass correlation coefficients (ICC(2,1); two-way, absolute agreement) [28] for landmark count and placement accuracy, and quadratic-weighted Cohen's κ [11] for the ordinal usefulness rating, interpreted following established benchmarks [28, 33]. Inter-rater reliability (20 maps, two raters) was good for landmark count (ICC = .85) and placement accuracy (ICC = .84, 95% agreement), and moderate for the holistic usefulness rating (weighted $\kappa = .44$, resolved by consensus). Each was analyzed with a one-way ART ANOVA. Only map usefulness differed across conditions, $F(2, 57) = 7.64$, $p = .001$, $\eta^2_p = .21$: contrary to the expectation that the minimal AR Compass might foster richer spatial learning, it produced the *least* useful sketches (Compass: *Mdn* = 2; Phone and AR Map: *Mdn* = 3; both contrasts $p < .01$, Compass vs. Phone and vs. Map). Landmark count and placement accuracy did not differ (both $p > .23$); participants recalled comparable numbers of landmarks and placed them near the ceiling regardless of the interface. The Compass deficit was thus specific to the route's holistic representation, not landmark memory. Finally,

of the three map measures, only placement accuracy was associated with baseline spatial ability ($\rho = -.26$, $p = .048$), with weaker spatial ability linked to less accurate placement.

4.7 Qualitative Feedback

Post-navigation interviews and open-ended survey responses were transcribed, reduced to discrete statements, and grouped into themes via affinity diagramming by two researchers. We summarize the themes most relevant to the quantitative findings.

Attention and the Novelty Curve. Both AR conditions were widely felt to keep users visually engaged with the environment—several contrasted glancing up at the display favorably against the smartphone down-glance. The novelty cut both ways. On the one hand, a subset reported that the interface drew attention away from their surroundings; on the other, several found it engaging, with one noting they felt *"pretty satisfied because the navigation was really cool and intuitive to use"* (P16, Compass). That the distraction reports tended to cluster early in the session, when *"learning the AR technique"* (P17, Compass) demanded keen attention, suggests the attentional cost is largely a novelty effect; many participants reported no interference at all. This learning-curve pattern corresponds to the elevated route-level speed variability seen in both AR conditions.

Prescriptive Routes Versus Reactive Bearings. The sharpest experiential divide fell between the two AR interfaces. Compass users described navigating reactively—*"After making a conscious decision about the first route, I just went on autopilot and stopped thinking about it too much."* (P16, Compass)—following a heading without an overarching plan, whereas the route-based interfaces gave users an explicit structure to follow. This offers an experiential correlate of the spatial-memory result: an explicit route appears to give users something to encode that a bare bearing does not.

Hardware and Requested Refinements. Participants also raised device-level frustrations—outdoor visibility in bright sunlight, the glasses' form factor, and the smartphone tether supplying the compass heading—that are properties of current hardware rather than of the navigation concepts under test. Their improvement suggestions, including stronger color contrast, an indicator for off-screen targets, explicit arrival confirmation, optional map rotation, and audio or haptic cues, inform the design recommendations in Section 5.

5 Discussion

Does usage of AR glasses improve wayfinding performance and spatial memory for pedestrians compared to a traditional handheld device? (RQ_1). Both AR interfaces produced shorter walked routes than the smartphone. Yet our decomposition of walked distance showed that this efficiency advantage did not arise where one might expect. The smartphone's efficiency deficit accrued almost entirely *after* participants reached the vicinity of a destination, as an overshoot. This points to a specific failure mode of GPS-based handheld navigation: it competently routes users to the approximate location of a target but provides weak confirmation of arrival, leaving users to overshoot and double back. The use of additional voice assistance could help address this issue, but it carries its own route-memory cost [67] and isn't always available, especially for pedestrians. This difference appears to stem from information availability. No condition issued an automatic arrival notification; participants judged arrival from the proximity information each interface displayed. That both AR conditions produced far less overshoot suggests their continuously visible distance readout supported more precise arrival judgments than the smartphone, whose distance information competes for space with the map and requires a down-glance to consult.

Both AR conditions, however, exhibited significantly more speed variability than smartphone users. This pattern is consistent with an attentional cost of operating an unfamiliar device and overlay, in line with cognitive-motor interference accounts of dual-task walking [1], though our measure cannot separate attentional interruption from other sources of speed fluctuation such as turns or stops. The qualitative data suggest this cost is a novelty effect rather than an intrinsic property of AR navigation. Whether this disruption would attenuate with sustained use is a direction for longitudinal work.

Crucially, these objective differences did not surface in participants' self-reports: perceived workload, usability, and presence did not differ across the three interfaces. We read this not as a null result but as an objective–subjective dissociation of the kind frequently observed in HCI [16, 68]: the novel AR interfaces achieved *subjective parity* with a mature, everyday technology despite measurable behavioral differences. It also cautions against relying on subjective instruments alone in this domain — behavioral measures revealed effects that questionnaires did not.

How minimal can a navigation aid in a heads-up display be while still supporting efficient navigation, acceptable user experience, and spatial learning? (RQ_2). The case for pursuing minimalism is not merely aesthetic. Most commercial AR headsets have a field of view under 60°, leaving room for little more than short messages [36], and roughly a third of observed everyday AR-glasses use involves several windows or applications at once [60]—so navigation must share a constrained display.

In this study, the AR Compass tested the minimal end of the design space: a single bearing and a distance readout, with no map, route line, or turn instruction. Path efficiency was statistically indistinguishable from the far richer AR Map, and the AR Compass condition showed the highest walking speed of the three. However, it produced significantly less useful route sketches than both the AR Map and the smartphone, even though participants in all conditions recalled a comparable number of landmarks and placed them with equal accuracy.

These results suggest that a minimal bearing-and-distance cue is sufficient for reaching a destination —– matching a full map in efficiency and speed —– but not for producing a usable representation of the route traveled. How minimal an aid can be depends on the goal: efficient wayfinding can tolerate minimalism, while producing a usable route representation appears to benefit from explicit route structure.

What constitutes optimal usability and design for pedestrian navigation within the AR design space? (RQ_3). By synthesizing our quantitative findings with participants' feedback, we want to offer the following design recommendations for AR pedestrian navigation.

- Interfaces should signal proximity and completion unambiguously rather than leaving users to infer arrival from a diminishing distance value.
- Systems intended to help users learn an environment should present the route as a connected structure; where display minimalism is a priority, a lightweight route trace may suffice.
- For outdoor navigation in bright conditions, high-contrast color-coded cues placed toward the periphery preserve both legibility and the view of the environment.
- Provide options to arrange the position and size of navigation information easily.
- When displaying maps, most users prefer the map rotating with their orientation rather than mentally aligning a north-up map with their heading. Nevertheless, a static north-up map remains efficient, so it should be left as an option for users struggling with motion sickness.

6 Limitations

Several limitations qualify these findings. The study was conducted in a single outdoor environment with a participant pool drawn from one university (ages 19–35), so generalization to other settings and age groups is untested.

Campus familiarity may have influenced participants' navigation and spatial-learning outcomes. To mitigate this concern, participants were randomly assigned to conditions, and routes were designed to connect non-obvious locations rather than easily identifiable landmarks, limiting the extent to which prior campus knowledge could reveal a destination or route. Nonetheless, participants' familiarity with campus was not directly measured. Future studies could explicitly assess environmental familiarity and account for it in the analysis.

The between-subjects design, while avoiding cross-interface learning effects, provides limited power to detect interactions, including the spatial-ability moderation we explored. Our spatial-ability measure was further constrained by ceiling penalties on incomplete SOT trials.

Our AR Map used a fixed north-up orientation to preserve the mental-rotation demand that distinguishes it from the egocentric Compass. This was a deliberate experimental control, but it departs from the track-up orientation common in deployed navigation and preferred by several of our participants, and we cannot separate the effect of the north-up presentation from that of the map format itself. A track-up AR Map might yield different efficiency or spatial-memory outcomes, and comparing orientations directly is a valuable direction for future work. The AR Map's heading marker was rendered in eight discrete 45° increments; although it updated responsively, this coarse angular resolution may have made aligning the map to the surrounding environment less precise than a continuously rotating marker would allow.

The overshoot decomposition was an exploratory, post hoc analysis and should be confirmed in a study designed around it; its 15-meter arrival threshold, though grounded in GPS accuracy, was an analyst's choice. Additionally, the Handheld condition used a commercial application whose localization, routing, and arrival logic are not observable to us; while this is precisely what gives the baseline its ecological validity, it means we cannot fully separate the contribution of the interface from that of the underlying navigation implementation, particularly for the overshoot result. The consistency of that effect across participants (IQR = 2.2 m) suggests a systematic difference rather than idiosyncratic application behavior, but a comparison against a custom-built handheld map would be needed to isolate the two. Our speed-variability measure also depends on the positioning pipeline. Although both conditions' speed was computed from position resampled to a common five-second interval, the underlying sources differed—the smartphone's commercial positioning versus the AR application's logging—and we cannot fully exclude a contribution of pipeline characteristics to the observed difference in variability. To isolate the visual guidance channel, we disabled the Handheld condition's audio and haptic feedback, including its arrival alert. In everyday use, smartphone navigation may provide such multimodal cues, which might in turn reduce the overshoot we observed; our results therefore speak to visual-only guidance rather than to the full multimodal experience of commercial navigation apps. Finally, the study captured a single session with unfamiliar hardware: the route-level speed variability cost we attribute to novelty, and the broader question of how these interfaces perform once users acclimate, can only be resolved by longitudinal investigation.

7 Conclusion

We presented a between-subjects outdoor study ($N = 59$) comparing everyday smartphone navigation with two heads-up navigation interfaces designed for modern AR glasses: an allocentric AR Map and a minimal egocentric AR Compass, and evaluated them across walking behavior, spatial memory, and user experience. Rather than identifying a single best interface, our results locate each modality at a different point along a shared set of trade-offs. Both AR interfaces produced shorter walked routes than the smartphone — though an exploratory decomposition traced this to arrival behavior rather than more direct wayfinding — and were rated equivalent in workload, usability, and presence. However, both AR conditions showed greater route-level speed variability, an attentional effect that the qualitative analysis suggests may stem from novelty. This decomposition isolates arrival confirmation, rather than wayfinding, as a distinct weakness of handheld GPS navigation. Finally, between the two AR designs the minimal AR Compass matched the full AR Map in efficiency and speed, but yielded less useful route sketches. These findings carry a broader message for AR pedestrian navigation: its value does not lie in a single optimal display, and we recommend matching the interface design to its goal. As consumer AR glasses move toward everyday life, we hope this field-study approach offers a useful starting point for studying these interfaces in real-world use.

Acknowledgments

ONR N00014-24-1-2214, National Science Foundation NSF 2444132, NSF 2439474, NSF 2327569, NSF 2238313, and NSF 2223432. We also thank Sophia XR (sophiaxr.com) for partial travel support; Sophia XR is an Agentic AI and Spatial Extended Reality company.

References

[1] Emad Al-Yahya, Helen Dawes, Lesley Smith, Andrea Dennis, Ken Howells, and Janet Cockburn. 2011. Cognitive motor interference while walking: a systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews* 35, 3 (2011), 715–728.

[2] B Anbaroğlu, İB Coşkun, and HH Gürler. 2020. Which way is 'Yildiz Amfi'? Augmented reality vs. paper map on pedestrian wayfinding. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 44 (2020), 53–60.

[3] Anthony J Aretz and Christopher D Wickens. 1992. The mental rotation of map displays. *Human Performance* 5, 4 (1992), 303–328.

[4] Alexander Arntz, Dustin Keßler, Nele Borgert, Nico Zengeler, Marc Jansen, Uwe Handmann, and Sabrina C. Eimler. 2020. Navigating a Heavy Industry Environment Using Augmented Reality - A Comparison of Two Indoor Navigation Designs. In *Virtual, Augmented and Mixed Reality. Industrial and Everyday Life Applications*, Jessie Y. C. Chen and Gino Fragomeni (Eds.). Springer International Publishing, Cham, 3–18. doi:10.1007/978-3-030-49698-2_1

[5] Ronald T Azuma. 1997. A survey of augmented reality. *Presence: Teleoperators & Virtual Environments* 6, 4 (1997), 355–385.

[6] BasicAirData. [n. d.]. GPS Logger. https://play.google.com/store/apps/details?id=eu.basicairdata.graziano.gpslogger. Android application. Accessed: February 2026.

[7] Mark Billinghurst and Hirokazu Kato. 1999. Collaborative mixed reality. In *Proceedings of the first international symposium on mixed reality*, Vol. 1. 261–284.

[8] M. Billinghurst and S. Weghorst. 1995. The use of sketch maps to measure cognitive maps of virtual environments. In *Proceedings Virtual Reality Annual International Symposium '95*. 40–47. doi:10.1109/VRAIS.1995.512478

[9] Zahra Borhani, Prashast Sharma, and Francisco R Ortega. 2023. Survey of annotations in extended reality systems. *IEEE Transactions on Visualization and Computer Graphics* 30, 8 (2023), 5074–5096.

[10] Jaewoo Chung, Francesco Pagnini, and Ellen Langer. 2016. Mindful navigation for pedestrians: Improving engagement with augmented reality. *Technology in Society* 45 (2016), 29–33.

[11] Jacob Cohen. 1968. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin* 70, 4 (1968), 213–220.

[12] Louisa Dahmani and Véronique D. Bohbot. 2020. Habitual use of GPS negatively impacts spatial memory during self-guided navigation. *Scientific Reports* 10, 1 (April 2020), 6310. doi:10.1038/s41598-020-62877-0

[13] Weihua Dong, Yulin Wu, Tong Qin, Xinran Bian, Yan Zhao, Yanrou He, Yawei Xu, and Cheng Yu. 2021. What is the difference between augmented reality and 2D navigation electronic maps in pedestrian wayfinding? *Cartography and Geographic Information Science* 48, 3 (May 2021), 225–240. doi:10.1080/15230406.2021.1871646

[14] Andreas Dünser, Mark Billinghurst, James Wen, Ville Lehtinen, and Antti Nurminen. 2012. Exploring the use of handheld AR for outdoor navigation. *Computers & Graphics* 36, 8 (Dec. 2012), 1084–1095. doi:10.1016/j.cag.2012.10.001

[15] Mica R Endsley. 1995. Measurement of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 37, 1 (1995), 65–84.

[16] Thomas Fritz, Andrew Begel, Sebastian C Müller, Serap Yigit-Elliott, and Manuela Züger. 2014. Using psycho-physiological measures to assess task difficulty in software development. In *Proceedings of the 36th international conference on software engineering*. 402–413.

[17] Aaron L. Gardony, Tad T. Brunyé, Caroline R. Mahoney, and Holly A. Taylor. 2013. How Navigational Aids Impair Spatial Memory: Evidence for Divided Attention. *Spatial Cognition & Computation* 13, 4 (Oct. 2013), 319–350. doi:10.1080/13875868.2013.792821

[18] Brian F Goldiez, Ali M Ahmad, and Peter A Hancock. 2007. Effects of augmented reality display settings on human wayfinding performance. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 37, 5 (2007), 839–845.

[19] Google. 2026. Google Maps Updates: Gemini Arrives, Immersive View Expands and More. https://blog.google/products-and-platforms/products/maps/gemini-google-maps-navigation-updates/. Google Maps serves more than 2 billion monthly users. Accessed: 06.13.2026.

[20] Sandra G Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage publications Sage CA: Los Angeles, CA, 904–908.

[21] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.

[22] Mary Hegarty and David Waller. 2004. A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence* 32, 2 (2004), 175–191.

[23] Haosheng Huang, Manuela Schmidt, and Georg Gartner. 2012. Spatial Knowledge Acquisition with Mobile Maps, Augmented Reality and Voice in the Context of GPS-based Pedestrian Navigation: Results from a Field Test. *Cartography and Geographic Information Science* 39, 2 (Jan. 2012), 107–116. doi:10.1559/15230406392107

[24] Jonathan Huang, Max Kinateder, Matt J Dunn, Wojciech Jarosz, Xing-Dong Yang, and Emily A Cooper. 2019. An augmented reality sign-reading assistant for users with reduced vision. *PloS one* 14, 1 (2019), e0210630.

[25] Toru Ishikawa, Hiromichi Fujiwara, Osamu Imai, and Atsuyuki Okabe. 2008. Wayfinding with a GPS-based mobile navigation system: A comparison with maps and direct experience. *Journal of Environmental Psychology* 28, 1 (March 2008), 74–82. doi:10.1016/j.jenvp.2007.09.002

[26] Roberta L Klatzky. 1998. Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In *Spatial cognition: An interdisciplinary approach to representing and processing spatial knowledge*. Springer, 1–17.

[27] John D Kloke and Joseph W McKean. 2012. Rfit: Rank-based estimation for linear models. *The R Journal* 4, 2 (2012), 57–64.

[28] Terry K Koo and Mae Y Li. 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine* 15, 2 (2016), 155–163.

[29] Yamato Kuwahara, Hung-ya Tsai, Yuya Ieiri, and Reiko Hishiyama. 2019. Evaluation of a campus navigation application using an AR character guide. In *International conference on collaboration and technology*. Springer, 242–250.

[30] Aymen Lakehal, Sophie Lepreux, Christos Efstratiou, Christophe Kolski, and Pavlos Nicolaou. 2021. Investigating Smartphones and AR Glasses for Pedestrian Navigation and their Effects in Spatial Knowledge Acquisition. In *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '20)*. Association for Computing Machinery, New York, NY, USA, 1–7. doi:10.1145/3406324.3410722

[31] Aymen Lakehal, Sophie Lepreux, Christos Efstratiou, Christophe Kolski, and Pavlos Nicolaou. 2023. Spatial Knowledge Acquisition for Pedestrian Navigation: A Comparative Study between Smartphones and AR Glasses. *Information* 14, 7 (July 2023), 353. doi:10.3390/info14070353

[32] Eric M Lamberg and Lisa M Muratori. 2012. Cell phones change the way we walk. *Gait & posture* 35, 4 (2012), 688–690.

[33] J Richard Landis and Gary G Koch. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33, 1 (1977), 159–174.

[34] Chun-I Lee. 2022. Benefit analysis of gamified augmented reality navigation system. *Applied Sciences* 12, 6 (2022), 2969.

[35] Jaewook Lee, Fanjie Jin, Younsoo Kim, and David Lindlbauer. 2022. User preference for navigation instructions in mixed reality. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 802–811.

[36] Lik-Hang Lee, Tristan Braud, Simo Hosio, and Pan Hui. 2021. Towards augmented reality driven human-city interaction: Current research on mobile headsets and future challenges. *ACM Computing Surveys (CSUR)* 54, 8 (2021), 1–38.

[37] James R Lewis. 1995. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction* 7, 1 (1995), 57–78.

[38] James R Lewis. 2002. Psychometric evaluation of the PSSUQ using data from five years of usability studies. *International Journal of Human-Computer Interaction* 14, 3-4 (2002), 463–488.

[39] Bing Liu, Linfang Ding, Shengkai Wang, and Liqiu Meng. 2024. Misleading effect and spatial learning in head-mounted mixed reality-based navigation. *Geospatial Information Science* 27, 2 (March 2024), 408–422. doi:10.1080/10095020.2022.2137063

[40] Steven G Luke. 2017. Evaluating significance in linear mixed-effects models in R. *Behavior research methods* 49, 4 (2017), 1494–1502.

[41] Pavel Manakhov, Ludwig Sidenmark, Ken Pfeuffer, and Hans Gellersen. 2024. Gaze on the go: Effect of spatial reference frame on visual target acquisition during physical locomotion in extended reality. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–16.

[42] Bartosz Mazurkiewicz, Marcelo de Lima Galvão, and Ioannis Giannopoulos. 2023. BeeAR: Augmented Reality Beeline Navigation for Spatial Knowledge Acquisition. *Proc. ACM Hum.-Comput. Interact.* 7, MHCI (Sept. 2023), 199:1–199:17. doi:10.1145/3604246

[43] Angélique Montuwy, Béatrice Cahour, and Aurélie Dommes. 2019. Using Sensory Wearable Devices to Navigate the City: Effectiveness and User Experience in Older Pedestrians. *Multimodal Technologies and Interaction* 3, 1 (March 2019), 17. doi:10.3390/mti3010017

[44] Stefan Münzer, Hubert D Zimmer, Maximilian Schwalm, Jörg Baus, and Ilhan Aslan. 2006. Computer-assisted navigation and the acquisition of route and survey knowledge. *Journal of environmental psychology* 26, 4 (2006), 300–308.

[45] Shinichi Nakagawa and Holger Schielzeth. 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in ecology and evolution* 4, 2 (2013), 133–142.

[46] Jack L Nasar and Derek Troyer. 2013. Pedestrian injuries due to mobile phone use in public places. *Accident Analysis & Prevention* 57 (2013), 91–95.

[47] Xin Hui Ng and Woan Ning Lim. 2020. Design of a mobile augmented reality-based indoor navigation system. In *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. IEEE, 1–6.

[48] Don Norman. 2013. *The design of everyday things: Revised and expanded edition*. Basic books.

[49] Margaree Peacocke, Robert J. Teather, Jacques Carette, I. Scott MacKenzie, and Victoria McArthur. 2018. An empirical comparison of first-person shooter information displays: HUDs, diegetic displays, and spatial representations. *Entertainment Computing* 26 (2018), 41–58.

[50] Jerrold Douglas Prothero. 1998. *The role of rest frames in vection, presence and motion sickness.* University of Washington.

[51] Xiaohe Qiu, Zhen Yang, Jinxing Yang, Qijun Wang, and Duming Wang. 2024. Impact of AR navigation display methods on wayfinding performance and spatial knowledge acquisition. *International Journal of Human–Computer Interaction* 40, 10 (2024), 2676–2696.

[52] Zhiwen Qiu, Armin Mostafavi, and Saleh Kalantari. 2025. Use of augmented reality in human wayfinding: a systematic review. *Virtual Reality* 29, 4 (Sept. 2025), 154. doi:10.1007/s10055-025-01226-w

[53] Karl Rehrl, Elisabeth Häusler, Sven Leitinger, and Daniel Bell. 2014. Pedestrian navigation with augmented reality, voice and digital map: final results from an in situ field study assessing performance and user experience. *Journal of Location Based Services* 8, 2 (April 2014), 75–96. doi:10.1080/17489725.2014.946975

[54] Rufat Rzayev, Paweł W Woźniak, Tilman Dingler, and Niels Henze. 2018. Reading on smart glasses: The effect of text position, presentation type and walking. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–9.

[55] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments* 10, 3 (2001), 266–281.

[56] Suguru Shimokihara, Takayuki Tabira, Yuriko Ikeda, Michio Maruta, Gwanghee Han, Taishiro Kamasaki, Yuma Hidaka, Yoshihiko Akasaki, Wataru Kukizaki, and Yusuke Kumura. 2025. Association between smartphone glances during application-based navigation and pedestrian navigation performance: a real-world experimental study. *mHealth* 11 (Oct. 2025), 63. doi:10.21037/mhealth-25-33

[57] Alexander W. Siegel and Sheldon H. White. 1975. The Development of Spatial Representations of Large-Scale Environments. In *Advances in Child Development and Behavior*, Hayne W. Reese (Ed.). Vol. 10. JAI, 9–55. doi:10.1016/S0065-2407(08)60007-5

[58] Phoebe O Toups Dugas, Bhavani Seetharaman, Rebecca J Fleming, Shano Liang, Don Samitha Elvitigala, Jennifer A Rode, and Nicolas LaLone. 2025. Mobile Maps Continue to Fail Pedestrians: Synthesised Reflective Auto-Aggro-Ethnographies of Walking. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–12.

[59] Tanh Quang Tran, Tobias Langlotz, and Holger Regenbrecht. 2024. A survey on measuring presence in mixed reality. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–38.

[60] Tram Thi Minh Tran, Shane Brown, Oliver Weidlich, Soojeong Yoo, and Callum Parker. 2025. Wearable AR in everyday contexts: Insights from a digital ethnography of youtube videos. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–18.

[61] Camille Truong-Allié, Alexis Paljic, Alexis Roux, and Martin Herbeth. 2021. User Behavior Adaptive AR Guidance for Wayfinding and Tasks Completion. *Multimodal Technologies and Interaction* 5, 11 (Nov. 2021), 65. doi:10.3390/mti5110065

[62] Holland M. Vasquez, Justin G. Hollands, Greg A. Jamieson, and Michael J. Agnew. 2022. A mirror in the sky: the effects of map format and user expertise on navigation performance and mental workload. *Ergonomics* 65, 4 (April 2022), 604–617. doi:10.1080/00140139.2021.1976848

[63] Prashant Verma, Kushal Agrawal, and V Sarasvathi. 2020. Indoor navigation using augmented reality. In *Proceedings of the 2020 4th International Conference on Virtual and Augmented Reality Simulations*. 58–63.

[64] Christopher D Wickens. 1999. Frames of reference for navigation. *Attention and Performance* 16 (1999), 113–144.

[65] Jacob O Wobbrock, Leah Findlater, Darren Gergle, and James J Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 143–146.

[66] Fang Xu, Tianyu Zhou, Hengxu You, and Jing Du. 2024. Improving indoor wayfinding with AR-enabled egocentric cues: A comparative study. *Advanced Engineering Informatics* 59 (2024), 102265.

[67] Yawei Xu, Tong Qin, Yulin Wu, Cheng Yu, and Weihua Dong. 2022. How do voice-assisted digital maps influence human wayfinding in pedestrian navigation? *Cartography and Geographic Information Science* 49, 3 (2022), 271–287.

[68] Yei-Yu Yeh and Christopher D Wickens. 1988. Dissociation of performance and subjective measures of workload. *Human factors* 30, 1 (1988), 111–120.

[69] Paul A Zandbergen. 2009. Accuracy of iPhone locations: A comparison of assisted GPS, WiFi and cellular positioning. *Transactions in GIS* 13 (2009), 5–25.

[70] Yu Zhao, Holly Gagnon, Jeanine Stefanucci, Sarah Creem-Regehr, and Bobby Bodenheimer. 2025. A Systematic Review of the Use of Augmented Reality in Pedestrian Navigation. *ACM Comput. Surv.* 58, 5 (Nov. 2025), 122:1–122:32. doi:10.1145/3770917