Using Social Cues for Wayfinding: An Eye-Tracking Study of Attention Allocation and Evidence Accumulation in a Virtual Reality T-Intersection

Yu Qiao
School of Geography, University of Leeds, UK

Carlo Campagnoli
School of Psychology, University of Leeds, UK

Ed Manley[1]
School of Geography, University of Leeds, UK

Abstract

In synchronous weak social wayfinding, route decisions are shaped in real time by others' behaviour, yet the sampling and integration of social information during decision-making remain poorly understood. We investigated whether attention to pedestrians in a virtual-reality T-intersection reflects evidence accumulation preceding route commitment. Participants navigated scenes containing avatar crowds in which crowd size, proportion of credible others, directional majority, walking speed, and movement style were manipulated in an incomplete within-subjects orthogonal design. Continuous eye tracking measured attention to social cues, the destination landmark, and the surrounding scene. Results showed that gaze was directed predominantly toward social cues, whereas direct inspection of the landmark was rare. Among the manipulated factors, movement style and walking speed produced the clearest, though small, effects on fixation–saccade dynamics. Crucially, social gaze rose sharply early in the trial and then declined gradually before route commitment, suggesting that participants initially sampled social evidence and progressively reduced inspection as the decision unfolded. These findings indicate that social navigation involves accumulating and integrating others' behaviour over time rather than simply following them.

2012 ACM Subject Classification Information systems → Spatial-temporal systems; Human-centered computing → Empirical studies in HCI; Applied computing → Psychology

Keywords and phrases Social navigation, wayfinding, eye tracking, virtual reality, evidence accumulation, attention

Digital Object Identifier 10.4230/LIPIcs.COSIT.2026.28

Category Short Paper

1 Introduction

Wayfinding is not merely a geometric problem of choosing the shortest or most direct path. In naturalistic environments, route decisions are shaped by spatial layout, landmark structure, and the behaviour of other people [8, 3, 14, 2]. Dalton et al. [4] describe wayfinding as a social activity and distinguish weak synchronous social wayfinding, in which route choice is shaped in real time by others' behaviour without direct interaction, from stronger or temporally separated forms. Evidence from virtual environments likewise shows that crowd density and movement can systematically alter route choice and locomotion [11, 6, 7].

A key question in social wayfinding is how information from other people influences route decisions. For social information to affect behaviour, it must first be attended to. Attention does not merely accompany choice; it determines which information is sampled and how it is

[1] corresponding author

17th International Conference on Spatial Information Theory (COSIT 2026).
Editors: Sabine Timpf, Gabriele Filomena, Armand Kapaj, Rui Zhu, Nicholas Giudice, and Ed Manley;
Article No. 28; pp. 28:1–28:9
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

weighted during deliberation, shaping evidence accumulation and the eventual decision [9, 12]. Work on eye movements in decision making further shows that gaze can actively modulate information processing rather than passively reflect visual interest [10, 13]. This perspective suggests that gaze can provide insight into how social information is gathered and used during navigation. We therefore ask three related questions: how gaze is allocated between social cues and spatial cues, and among different social cues; how social gaze unfolds over time within a trial; and whether the temporal dynamics of social gaze are consistent with a process of evidence accumulation.

Eye tracking is therefore well suited to testing the process-based account of social navigation, especially in virtual reality, where route choice can be studied under controlled but behaviourally meaningful conditions. Prior work suggests that VR can reproduce key aspects of pedestrian route choice [5]. Yet challenges remain: fixation-saccade classification is less straightforward in dynamic 3D scenes [1], and gaze-to-object mappings are affected by geometric bias, because nearer objects are more likely to intersect gaze rays and may therefore appear more salient if uncorrected.

Against this background, the present study examined route choice in a virtual reality T-intersection. Participants viewed crowds of avatars moving toward one branch or the other under experimental control, while continuous eye tracking measured attention to social cues, the destination landmark, and the broader scene. The aim was to characterize how attention is allocated to social information during route choice and how social gaze evolves over time. In particular, we examined whether the temporal dynamics of social gaze are consistent with a process of evidence accumulation during navigation. This design captures both overall attention allocation and the moment-to-moment evolution of gaze prior to behavioural commitment.

This study makes three contributions. First, it extends social wayfinding research by treating gaze toward pedestrians as an evidentiary process unfolding before route commitment. Second, it improves methodology by combining sensitivity analysis for fixation-saccade classification with a geometric correction for distance-related hit probability. Third, it clarifies how social cues compete with spatial information in a controlled T-intersection and whether avatar motion modulates attention in a manner consistent with staged evidence accumulation. Together, these contributions position social gaze as a process-level bridge between perception and route decision in virtual navigation.

2 Methods

2.1 Experimental design and procedure

The experiment used an incomplete within-subjects orthogonal design based on an L36 array to manipulate five social-cue factors: pedestrian count, proportion of credible others, directional majority, walking speed, and movement style. Each participant completed nine unique conditions sampled from the full factorial set, with no repeated condition within participant. Trial order was randomized individually to reduce sequence effects.

The virtual environment was a minimal T-shaped intersection with identical geometry on both branches (Figure 1). The task was a route-decision task: at each trial, participants were instructed to imagine that they were on a city street, that they were trying to reach the stadium to watch a football match, and that they did not know which direction would lead to the destination. They were asked to use the available social and spatial cues to decide whether to go left or right. The destination landmark, operationalized as a stadium, was

Table 1 Experimental variables and levels.

Variable	Target path	Non-target path
Pedestrian count	Total pedestrians: 5, 20, 40	
Proportion of credible others	0	0, 20%, 75%
Directional majority	45%, 25%, 10%	55%, 75%, 90%
Walking speed	1.4 m/s	0.6, 1.4, 2.2 m/s
Movement style	Direct	Direct, Uncertain

placed behind one branch (target path) and counterbalanced across left and right sides across trials. This kept the destination available as a navigational reference without making one branch trivially dominant.

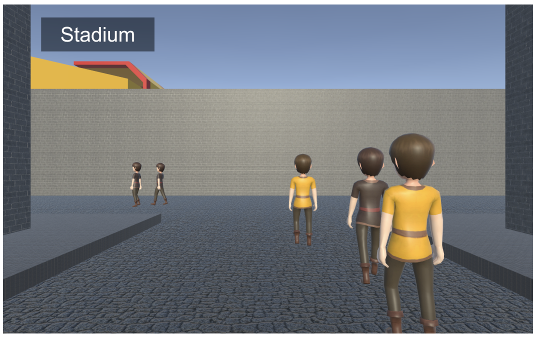

Figure 1 Virtual T-intersection environment.

Avatars were spawned along the approach corridor, reached the intersection at controlled timings, and then selected a branch according to the directional-majority manipulation before continuing with the assigned speed and movement style. The manipulations are summarized in Table 1. Pedestrian count varied across low, medium, and high levels. Credible others were operationalized as the percentage of avatars wearing football jerseys, which served as an association cue in this study. Directional majority indexed the extent to which the crowd favoured one branch. Walking speed varied across slow, intermediate, and fast levels. Movement style contrasted direct, goal-oriented motion with uncertain or wavering motion. The dependent variable was the participant's final route choice, recorded after 40 seconds when a response panel appeared and participants selected either the left or right branch.

A priori power analysis ($\alpha = .05$, power $= .80$) indicated that $N = 61$ was sufficient to detect medium-sized main effects (Cohen's $f \approx 0.25$) and typical interaction effects under the incomplete within-subjects L36 design. A total of 61 participants were recruited from the university participant pool and the local community. All participants had normal or corrected-to-normal vision and reported no severe motion or vestibular disorders. The study was approved by the Research Ethics Committee for Business, Environment, and Social Sciences at the University of Leeds, and all participants provided informed consent before participation. Participants received either a voucher or course credit as compensation.

2.2 Eye-movement classification and sensitivity analysis

Because the raw gaze log did not contain event-level labels, fixations and saccades were classified post hoc using an angular-velocity criterion. For each trial, gaze direction was represented as a unit vector at each sample. Angular displacement between consecutive samples was converted to angular velocity (deg/s), and samples exceeding a predefined threshold were labelled as saccades, with the remainder labelled as fixations. This yielded a trial-wise eye-state series aligned with areas of interest for subsequent analyses.

To determine an appropriate threshold, a sensitivity analysis was conducted over candidate values ranging from 2 to 100 deg/s. For each threshold, the full analysis pipeline was rerun and evaluated using four criteria: fixation–saccade balance, availability of valid time bins for social-attention summaries, magnitude of the balance index, and proportion of usable bins across factors. These were combined into a composite score favouring thresholds that preserved temporal variation while maintaining adequate coverage and moderate saccade rates. A threshold of 10 deg/s provided the optimal trade-off (saccade rate: 15.3%; coverage: 93.6%), and neighbouring values (7.5 and 12.5 deg/s) produced comparable results, indicating robustness. Accordingly, 10 deg/s was adopted for all analyses.

2.3 Geometric correction for distance-related hit probability

A further challenge in VR eye tracking is geometric sampling bias: nearer objects subtend larger visual angles and are therefore more likely to intersect a gaze ray independent of attentional relevance. To mitigate this effect, a condition-specific geometric baseline was estimated using Monte Carlo simulations of avatar trajectories.

For each condition, simulations reproduced the same spawning rules, movement parameters, and temporal structure as the VR task. At each time point, the probability that avatar i would be intersected by a gaze ray was approximated as:

$$p_i(t) = \begin{cases} 0, & z_i(t) \leq 0 \\ \min\left(1, \frac{A}{d_i(t)^2}\right), & z_i(t) > 0 \end{cases} \tag{1}$$

where $d_i(t)$ denotes the Euclidean distance to the observer, A is the effective projected frontal area of a pedestrian proxy, and $z_i(t)$ indicates depth relative to the camera plane. Probabilities were aggregated across avatars and averaged over time to obtain trial-level visibility, then averaged across simulation runs to derive condition-specific baselines. Empirical gaze measures were normalised against these baselines, allowing attention to be interpreted relative to expected geometric visibility.

Although this approach does not fully account for all visibility factors (e.g., occlusion), it substantially reduces distance-related bias and improves the validity of gaze-based attention measures.

3 Results

3.1 Overall allocation of attention

Across observations, gaze was directed mainly to social cues and other scene elements, whereas the spatial cue (the stadium) received minimal direct inspection. Social areas of interest accounted for a mean of 35.6% of gaze samples, other areas for 64.2%, and the spatial cue for only 0.2%. The median and upper quartile of spatial gaze were both 0.0%, indicating that direct fixations on the landmark were absent in most trials. In contrast, social gaze varied substantially across trials, suggesting that participants allocated more visual attention to pedestrian information than to the destination landmark within the present experimental setup.

This pattern was also evident at the condition level (Figure 2, left). Across the 36 experimental conditions, social gaze ranged from approximately 14% to 64%, with higher social allocation corresponding to lower gaze to other scene elements. Spatial gaze remained negligible in nearly all conditions and did not materially alter this ranking. Thus, the manipulations redistributed attention primarily between social and non-social scene elements rather than producing substantial inspection of the landmark. The participant-level plot showed substantial inter-individual variation, with some participants allocating more than 70% of gaze to social cues and others less than 20% (Figure 2, right). Even so, the same general structure held: social gaze and other gaze formed a near-complementary distribution, whereas spatial gaze remained marginal.

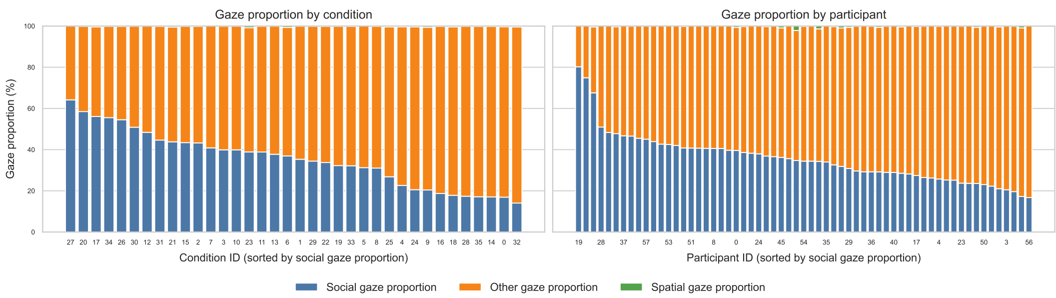

Figure 2 Gaze proportion by condition (left) and by participant (right).

3.2 Fixation–saccade dynamics and social cue properties

Using the 10 deg/s threshold, the gaze stream was strongly fixation-dominant: 84.7% of samples were classified as fixations and 15.3% as saccades. This indicates that participants spent most of the task in stable viewing states, with saccades functioning mainly as brief transitions between informative scene elements. The factor-level analyses therefore examined whether specific social cues altered this balance over time.

Across the three social factors, level-by-state contingency tests were significant in all cases, but effect sizes were small. Outfit showed a weak association with eye state ($\chi^2(1) = 58.98$, $p = 1.59 \times 10^{-14}$, Cramér's $V = 0.0068$), movement style showed the strongest association ($\chi^2(1) = 318.77$, $p = 2.68 \times 10^{-71}$, Cramér's $V = 0.0158$), and walking speed also reached significance ($\chi^2(2) = 104.31$, $p = 2.24 \times 10^{-23}$, Cramér's $V = 0.0090$). These results suggest that social cues biased visual sampling without producing large categorical changes in gaze behaviour. In other words, the cues did not determine whether participants looked, but rather how long they remained in a stable sampling state before shifting to another source of information.

The time-resolved balance maps provided a clearer picture of these effects. Outfit produced only limited differences between levels, with both conditions remaining fixation-dominant throughout most of the trial (Figure 3, top). Movement style showed the clearest temporal contrast: uncertain pedestrians were associated with a more fixation-weighted profile, whereas direct motion produced relatively more saccades, especially later in the trial (Figure 3, middle). Walking speed also modulated gaze allocation in a graded manner, with slower motion producing stronger fixation dominance and faster motion leading to earlier changes in balance (Figure 3, bottom). Overall, movement style had the strongest effect, speed a weaker but structured influence, and outfit the smallest impact.

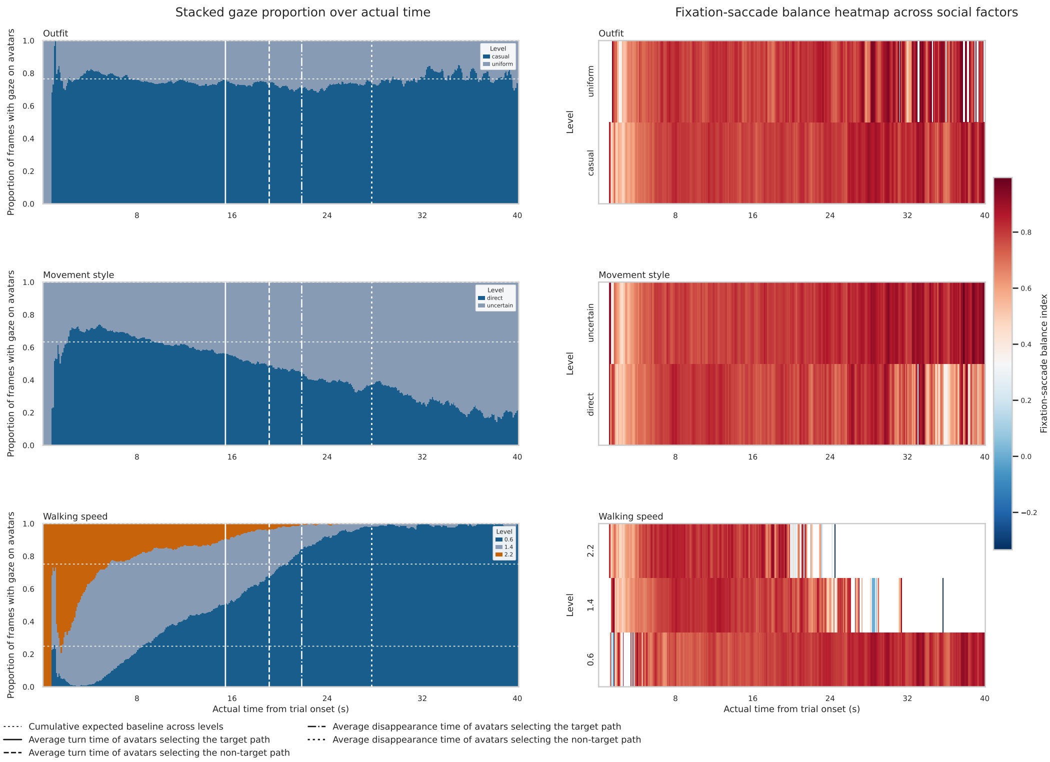

Figure 3 Left: Stacked gaze proportions over time across social factors (outfit, movement style, and walking speed; top to bottom). Right: Fixation–saccade balance heatmaps across social factors (outfit, movement style, and walking speed; top to bottom), where the balance index is defined as (Fixation – Saccade) / (Fixation + Saccade).

3.3 Temporal dynamics of social gaze as evidence accumulation

The temporal profile of social gaze suggests an evidence-accumulation process. In the raw data, mean social gaze rose sharply from near zero after trial onset, peaked at around 8–10 s, and then declined gradually for the rest of the trial (Figure 4, left). This early surge suggests rapid sampling of pedestrian information to form an initial directional hypothesis, followed by progressive redistribution of gaze as the decision advanced. Importantly, this peak occurred before the main behavioural commitments of the avatars, indicating that substantial evidence accumulation took place prior to turning.

After geometric correction, the same temporal structure remained but became more interpretable (Figure 4, right). The corrected curve preserved the early rise and peak, although its magnitude was reduced, indicating that part of the raw effect reflected distance-related hit probability rather than pure attentional preference. Once this bias was removed, the trajectory appeared less uniformly elevated and more clearly divided into stages of accumulation, consolidation, and stabilization. The turn of avatars selecting the target path occurred shortly after the peak of social attention, suggesting that participants likely reached sufficient evidence to initiate commitment while still monitoring social cues. After the relevant avatars left the scene, the corrected curve settled into a lower and more stable baseline with only small late-trial rebounds.

Overall, the temporal dynamics support a staged account of decision-making in which attention first accumulates social evidence, then tracks the commitment point, and finally stabilizes after the critical cues disappear. The gaze trajectory is therefore not only descriptive of where participants looked, but also informative about how route-relevant evidence was sampled, evaluated, and converted into a decision.

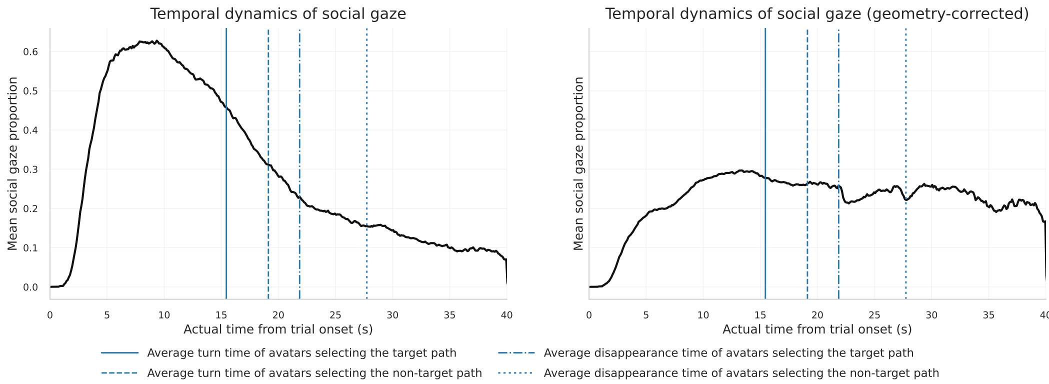

Figure 4 Temporal dynamics of social gaze before (left) and after (right) geometric correction.

4 Discussion

This study examined whether social information in wayfinding can be understood as an evidence-accumulation process. The results support this interpretation. Participants directed most of their gaze to social cues rather than to the destination landmark, indicating that, in this experimental context, the crowd was the main source of route decision-relevant information. However, this pattern should be interpreted within the present experimental context and does not imply that social cues are generally more important than spatial cues in naturalistic wayfinding. Social gaze also showed a clear temporal pattern: it rose rapidly at trial onset, peaked before or around the main turning window, and then declined as the scene unfolded. This trajectory is consistent with early evidence gathering followed by progressive commitment to a route.

Geometric correction further strengthened this account. In the raw data, social gaze was partly inflated by proximity-based hit probability, because nearby avatars were easier to intersect with a gaze ray. After correction, the same temporal structure remained but became more clearly linked to decision timing rather than spatial bias. This suggests that social gaze reflected not only visual accessibility, but also the informational value of pedestrian behaviour at different moments.

The fixation-saccade results are also consistent with a process account. Fixations dominated overall, indicating that participants spent most of the task in stable sampling states, while saccades functioned mainly as transitions between informative scene elements. Among the social cues, movement style was the most informative, with uncertain motion encouraging longer inspection. Walking speed showed a similar but weaker effect, suggesting that dynamic cues modulate the rate of evidence sampling rather than the mere presence of social attention. The small effect sizes should therefore not be taken as null effects; instead, they suggest that route choice was shaped by multiple weak cues integrated over time rather than by any single all-or-none trigger. This interpretation also points to a natural extension of the present work: a drift diffusion model with time-varying social evidence, in which gaze-gated avatar information is translated into a momentary evidence stream and accumulated over time to predict route choice more directly.

Methodologically, the study highlights the value of combining VR with careful gaze analysis. Threshold-based fixation classification is necessary in this setting, but threshold choice should be guided by sensitivity analysis rather than convention. Geometric correction is also essential when objects differ in distance and projected size, because apparent social attention can otherwise be overestimated. Future work should extend the correction model to include occlusion and line-of-sight constraints, which also shape visibility in crowd scenes.

The study has limitations. The environment was simplified to isolate social information, which improves internal validity but reduces ecological complexity. More importantly, route choice was reported only at the end of each 40-second trial, and explicit decision times were not collected, so some gaze samples may have been recorded after participants had already made their decision. In addition, gaze to avatars may reflect general attentional bias toward humans, motion saliency, or rendering-specific visual properties, rather than navigation-specific sampling of social cues. Future studies should record continuous choice timing, which would make it possible to test dynamic accumulation models more directly and to separate perceptual sampling from the final commitment to a route, and examine whether the proposed accumulation account generalizes to more complex and interactive urban environments.

5 Conclusion

This study shows that social information can guide wayfinding through evidence accumulation. Participants relied on pedestrians far more than on the destination landmark, and their gaze followed a clear temporal pattern: rapid early sampling, a peak before route commitment, and a later decline as the decision resolved. Movement style and walking speed had the clearest effects, but these were graded rather than dramatic. After correcting for geometric visibility, the temporal dynamics became especially compelling as evidence for a process account.

More broadly, social navigation appears to be a dynamic decision process in which people gather, compare, and integrate social evidence until it becomes sufficient for route choice. Future work could formalize this process with computational models of evidence accumulation, such as drift–diffusion models, combined with time-varying social cues derived from gaze behaviour. This would provide a more mechanistic account of social navigation and link observed gaze dynamics to latent decision processes.

References

1 Isayas Berhe Adhanom, Paul MacNeilage, and Eelke Folmer. Eye Tracking in Virtual Reality: A Broad Review of Applications and Challenges. *Virtual reality*, 27(2):1481–1505, 2023. doi:10.1007/s10055-022-00738-z.

2 Madeleine Alcock, Jan M. Wiener, and Doug Hardman. Searching in an unfamiliar environment: A phenomenologically informed experiment. *Phenomenology and the Cognitive Sciences*, 2024. doi:10.1007/s11097-024-09960-3.

3 Debjit Bhowmick, Stephan Winter, Mark Stevenson, and Peter Vortisch. The impact of urban road network morphology on pedestrian wayfinding behavior. *Journal of Spatial Information Science*, 21:203–228, 2020. doi:10.5311/JOSIS.2020.21.601.

4 Ruth C. Dalton, Christoph Hölscher, and Daniel R. Montello. Wayfinding as a Social Activity. *Frontiers in Psychology*, 10, 2019. doi:10.3389/fpsyg.2019.00142.

5 Yan Feng, Dorine C. Duives, and Serge P. Hoogendoorn. Using virtual reality to study pedestrian exit choice behaviour during evacuations. *Safety Science*, 137:105158, 2021. `doi:10.1016/j.ssci.2021.105158`.

6 Milad Haghani and Majid Sarvi. Pedestrian crowd tactical-level decision making during emergency evacuations. *Journal of Advanced Transportation*, 50(8):1870–1895, 2016. `doi:10.1002/atr.1434`.

7 Hengshan Li, Tyler Thrash, Christoph Hölscher, and Victor R. Schinazi. The effect of crowdedness on human wayfinding and locomotion in a multi-level virtual shopping mall. *Journal of Environmental Psychology*, 65:101320, 2019. `doi:10.1016/j.jenvp.2019.101320`.

8 Rui Li and Alexander Klippel. Wayfinding Behaviors in Complex Buildings: The Impact of Environmental Legibility and Familiarity. *Environment and Behavior*, 48(3):482–510, 2016. `doi:10.1177/0013916514550243`.

9 Felix Molter, Armin W. Thomas, Scott A. Huettel, Hauke R. Heekeren, and Peter N. C. Mohr. Gaze-dependent evidence accumulation predicts multi-alternative risky choice behaviour. *PLoS Computational Biology*, 18(7):e1010283, 2022. `doi:10.1371/journal.pcbi.1010283`.

10 Jacob L. Orquin and Simone Mueller Loose. Attention and choice: A review on eye movements in decision making. *Acta Psychologica*, 144(1):190–206, 2013. `doi:10.1016/j.actpsy.2013.06.003`.

11 Carlota Silva, Francisco Rebelo, Elisângela Vilar, and Paulo Noriega. Preliminary Study about Social Influence Over Wayfinding Decisions. *Procedia Manufacturing*, 3:5920–5926, 2015. `doi:10.1016/j.promfg.2015.07.905`.

12 Chih-Chung Ting and Sebastian Gluth. Unraveling information processes of decision-making with eye-tracking data. *Frontiers in Behavioral Economics*, 3, 2024. `doi:10.3389/frbhe.2024.1384713`.

13 Michel Wedel, Rik Pieters, and Ralf van der Lans. Modeling Eye Movements During Decision Making: A Review. *Psychometrika*, 88(2):697–729, 2023. `doi:10.1007/s11336-022-09876-4`.

14 Demet Yesiltepe, Ruth Conroy Dalton, and Ayse Ozbil Torun. Landmarks in wayfinding: A review of the existing literature. *Cognitive Processing*, 22(3):369–410, 2021. `doi:10.1007/s10339-021-01012-x`.