This is the version of the article before peer review or editing, as submitted by an author to The JOURNAL OF NEURAL ENGINEERING. IOP Publishing Ltd is not responsible for any errors or omissions in this version of the manuscript or any version derived from it.
The Version of Record is available online at https://iopscience.iop.org/article/10.1088/1741-2552/ad8b6f

Wide-angle simulated artificial vision enhances spatial navigation and object interaction in a naturalistic environment

Sandrine Hinrichs[1], Louise Placidet[2], Antonin Duret[2], Colas Authié[3], Angelo Arleo[2] and Diego Ghezzi[1,4,*]

[1] Medtronic Chair in Neuroengineering, Center for Neuroprosthetics and Institute of Bioengineering, School of Engineering, École Polytechnique Fédérale de Lausanne, Geneva, Switzerland

[2] Sorbonne Université, INSERM, CNRS, Institut de la Vision, Paris, France

[3] Streetlab, Paris, France

[4] Ophthalmic and Neural Technologies Laboratory, Department of Ophthalmology, University of Lausanne, Hôpital ophtalmique Jules-Gonin, Fondation Asile des Aveugles, Lausanne, Switzerland

* Corresponding author: info@ghezzilab.org

Abstract

Objective. Vision restoration approaches, such as prosthetics and optogenetics, provide visual perception to blind individuals in clinical settings. Yet their effectiveness in daily life remains a challenge. Stereotyped quantitative tests used in clinical trials often fail to translate into practical, everyday applications. On the one hand, assessing real-life benefits during clinical trials is complicated by environmental complexity, reproducibility issues, and safety concerns. On the other hand, predicting behavioral benefits of restorative therapies in naturalistic environments may be a crucial step before starting clinical trials to minimize patient discomfort and unmet expectations.

Approach. To address this, we leverage advancements in virtual reality technology to conduct a fully immersive and ecologically valid task within a physical artificial street environment. As a case study, we assess the impact of the visual field size in simulated artificial vision for common outdoor tasks.

Main Results. We show that a wide visual angle (45°) enhances participants' ability to navigate and solve tasks more effectively, safely, and efficiently. Moreover, it promotes their learning and generalization capability. Concurrently, it changes the visual exploration behavior and facilitates a more accurate mental representation of the environment. Further increasing the visual angle beyond this value does not yield significant additional improvements in most metrics.

Significance. We present a methodology combining augmented reality with a naturalistic environment, enabling participants to perceive the world as patients with retinal implants would and to interact physically with it. Combining augmented reality in naturalistic environments is a valuable framework for low vision and vision restoration research.

1. INTRODUCTION

Visual impairment and blindness are significant global health challenges, affecting millions worldwide [1]. These conditions severely limit mobility, spatial awareness, object interaction, and social engagement, placing a substantial burden on patients and their families. Efforts to combat blindness can be categorized into three main strategies: prevention, preservation, and restoration [2,3]. Early intervention aims to prevent the onset of blindness or at least slow its progression to preserve natural vision for as long as possible. These measures include preventive health and medicine, genetic therapies, and pharmacological treatments [2]. Yet, a prerequisite for their effectiveness is the early identification of the disease. However, some cases of blindness are unpredictable or unavoidable with current treatments. In such cases, restorative approaches are necessary. These can be divided into two categories: restoring natural tissue through regenerative medicine [4] and artificially stimulating the surviving visual neurons using methods like visual prostheses or optogenetics [5–7].

A crucial question in the field of vision restoration is determining the quality of restored vision needed to significantly improve daily functioning and quality of life. Identifying the key parameters that enhance a patient's ability to perform everyday activities is vital for developing and evaluating new treatments. Clinically, visual impairment and blindness are primarily assessed using two metrics: visual acuity, which measures the ability to discern fine details, and visual field size, which defines the area visible when fixing on a central point [8]. While most artificial vision approaches have focused on restoring visual acuity for tasks like object recognition and reading, maintaining a sufficient visual field is equally important for daily activities [9]. Peripheral visual field loss, such as tunnel vision, significantly impairs ability to navigate and move safely, such as in retinitis pigmentosa or glaucoma [10–13].

In this study, we explored how modulating the visual field size affects human behavior in daily outdoor activities for blind individuals with artificially restored vision. Our hybrid approach combined simulated artificial vision (SAV) using a head-mounted display (HMD) for augmented reality (AR) [14,15] with naturalistic behavioral assessments in an artificial street laboratory[16]. Most knowledge about artificial vision currently comes from retinal implants [17–19], while reports on other implants[20–23] and optogenetics [24] are limited. Hence, we simulated phosphene perception from the POLYRETINA device: a retinal implant designed to provide high-resolution and wide-field artificial vision [25–29]. But, we anticipate that our findings could apply to other forms of artificial vision and restorative approaches (e.g., optogenetics). Our results demonstrate that a wide visual angle (45°) significantly improves participants' ability to navigate and complete common outdoor tasks in a naturalistic environment more effectively, safely, and efficiently than a smaller visual angle (20°). Moreover, it promotes learning and generalization capability. This enhancement likely results from better visual exploration behaviors and a more accurate mental representation of the environment. Interestingly, increasing the visual angle beyond 45° did not yield substantial additional benefits. These findings highlight the importance of visual field size in artificial vision, especially in naturalistic settings. Understanding how visual field size impacts the effectiveness of artificial vision technologies is crucial for optimizing their design and improving user quality of life. Additionally, the use of SAV in an AR naturalistic environment offers a valuable framework for research in low vision and vision restoration.

2. METHODS

2.1 Ethical authorization. The study was approved by the Ethical Committee CPP Ile de France V (ID_RCB 2015-A01094-45, No. CPP: 16122 MSB).

2.2 Participants. 47 participants were recruited in a between-subjects study design from the SilverSight cohort, which comprised around 350 individuals enrolled at the Institut de la Vision and the Quinze-Vingts National Ophthalmology Hospital (Paris, France) [30]. The sample size was akin to prior studies on orientation and mobility in low or artificial vision [31]. Inclusion criteria were normal or corrected-to-normal vision, no limitations in physical mobility without assistance, less than 45 years of age, not affected by neurological disorder, and naïve to SAV. Participants with corrected-to-normal vision were asked to wear contact lenses. However, if not feasible, they were allowed to wear glasses within the HMD. Participation in the study was voluntary. A financial compensation in the form of a gift card was given for participation. Participants and experimenters were native French speakers, so the study was conducted in French. Participants were unaware of the specific research question.

2.3 Naturalistic environment. The study was conducted in an artificial street available at the Institut de la Vision (http://www.streetlab-vision.com, Paris, France) to replicate naturalistic outdoor conditions while capturing behaviorally relevant parameters and controlling light and sound [16]. The experimental area is a rectangular space sized 8.55 m × 4.30 m (**Figure 1a**). White fabric represented the sidewalks to ensure safety and minimize tripping risks. Lighting remained steady at 2300 lux, while background sounds mimicked a quiet street for a more immersive experience and to mask unintended auditory cues guiding exploration and spatial behavior. In the street, three real-life outdoor stations are represented: the home door, an ATM, and a post box (cyan insets in **Figure 1a**).

2.4 Simulated artificial vision. Participants wore a HMD (HTC VIVE Pro Eye) providing real-time vision of the street from the front camera input (**Figure 1b**). The HMD was operated with a backpack computer. Unity (version 2019.2.16f1) and Cg shaders converted the HMD camera input into SAV in real-time. SAV mimicked phosphene perceptions. The basic phosphenes were circles sized 80 μm and spaced 120 μm, corresponding approximately to 0.275° and 0.414° respectively. Parameters were based on the POLYRETINA prosthetic device [25–27,29]. A distortion mimicking unintended axon fiber activation was set to $\lambda = 2$ [14,15]. SAV operated at a 5 Hz with an 11 ms frame duration (1 frame on / 17 frames off) and was exclusively presented to the dominant eye. A random variability to the SAV was included at each frame: phosphene size varied randomly within ± 30% of the basic size, brightness varied randomly from 50% (gray) to 100% (white), and 10% of phosphenes were randomly not presented. SAV accounted for perceptual fading of phosphenes and a compensation strategy was included as previously described [15,28].

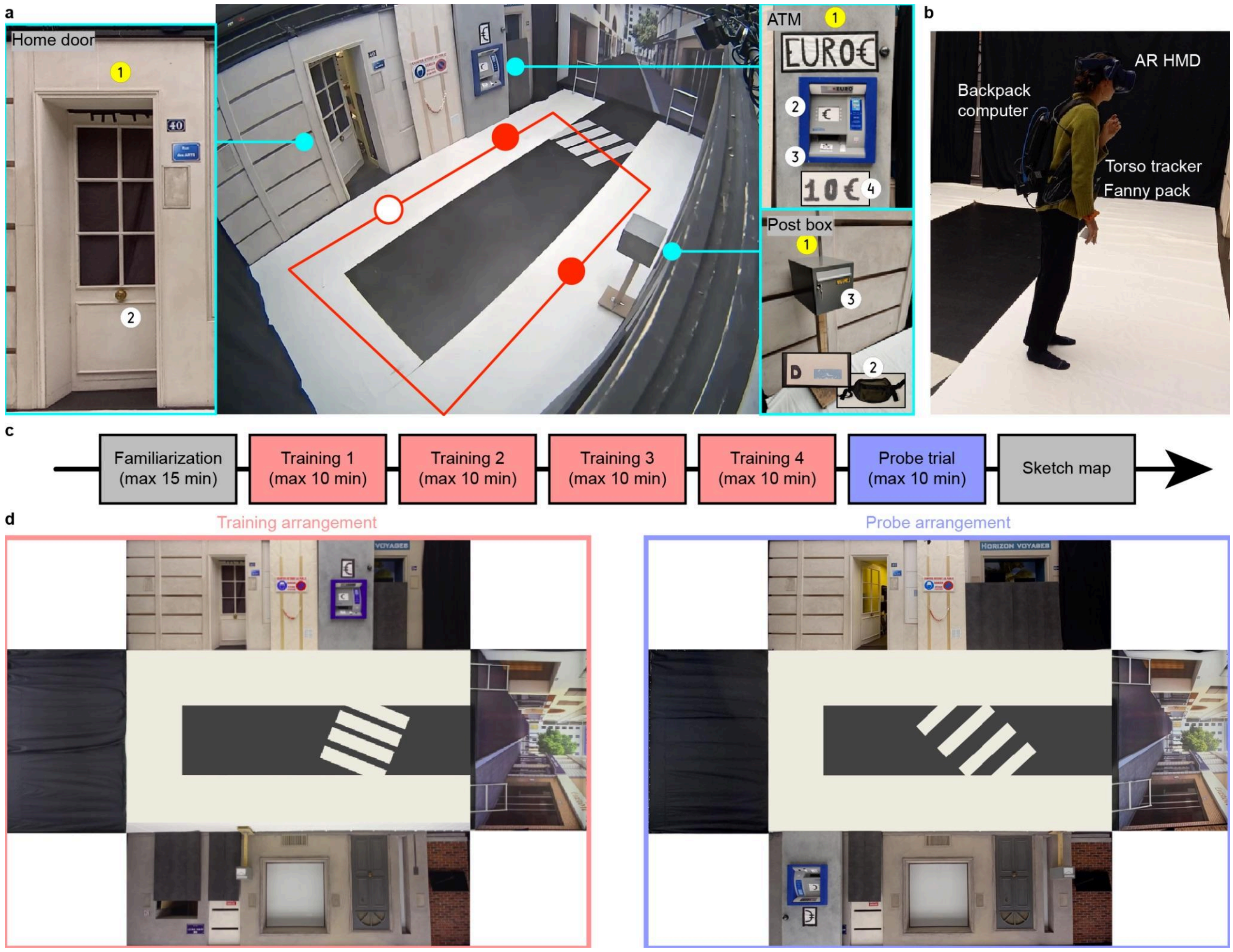


Figure 1*. Study overview.* ***a*** *Street overview including the three stations (home door, ATM, and post box; cyan insets) connected by a path (red line). The home door is the start/end point (white circle).* ***b*** *Participant wearing a HMD, a backpack computer, a torso tracker and a fanny pack.* ***c*** *Experimental protocol: familiarization phase, four training sessions, one probe trial, and sketch map drawing.* ***d*** *Street arrangements during the training sessions (left) and the probe trial (right).*

2.5 Study design. Participants performed a real-life outdoor activity in the artificial street wearing the AR HMD. After a familiarization phase, they performed five repetitions of the activity: four training sessions and one probe trial (**Figure 1c**). During the activity, they had to reach the 3 stations (task 1 in yellow), and solve the associated tasks: 3 at the ATM (tasks 2, 3 and 4 in white), 2 at the post box (tasks 2 and 3) and 1 when they returned home (task 2). The street arrangement changed between the training sessions and the probe trial to test the participants' generalization capability (**Figure 1d**). Last, participants drew a sketch map of the street seen during the probe trial.

Participants were randomly assigned to one of three testing groups (n = 14 per group) and one control group (n = 5), each performing the activity under a different viewing condition (**Table 1**). The testing groups perceived the street in SAV with three visual fields (**Figure 2**): circular of 20° in diameter (SAV 20°), circular of 45° in diameter (SAV 45°), and squared 98° x 98° corresponding to the HMD visible full field (SAV FF). In the control group, participants also wore the HMD, but the camera image was not altered. In this manner, subjects were exposed to comparable conditions (e.g., weight and resolution of the headset). The control group provides a performance reference in normal vision (NV), and it is not included in the statistical analysis.

	Testing groups			Control group
	SAV 20°	**SAV 45°**	**SAV FF**	**NV**
Number	14	14	14	5
Age (mean ± SD)	28 ± 4.71	28.64 ± 6.63	28.93 ± 4.38	34.4 ± 3.5
Age (min - max)	20 - 37	18 - 39	22 - 34	30 - 39
Sex	8 F / 5 M / 1 D	9 F / 4 M / 1 D	10 F / 4 M / 0 D	0 F / 5 M / 0 D

***Table 1**. Demographic data for study participants. Age is in years. F: female; M: male; D: diverse.*

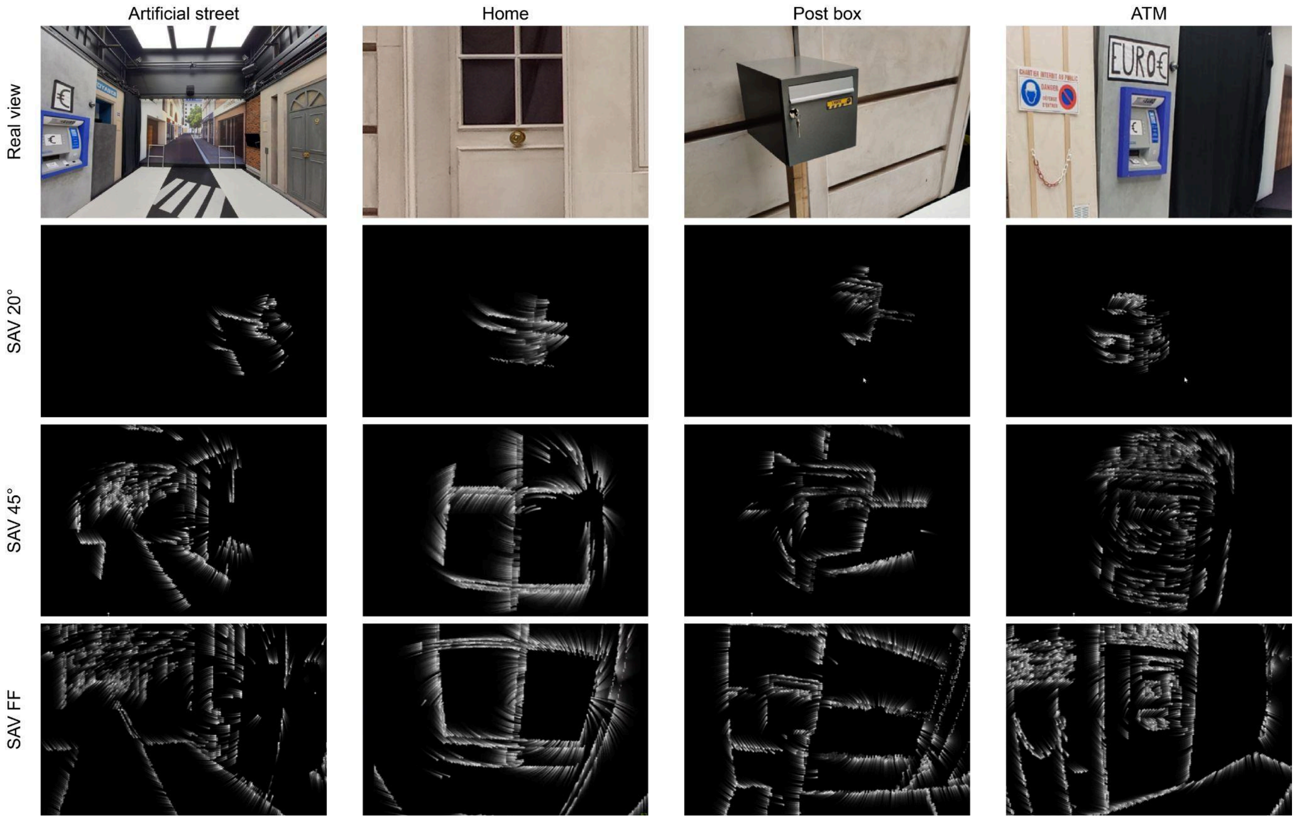


***Figure 2**. Comparison among viewing conditions. Images of the artificial street, the home door, the post box and the ATM are converted to SAV with different visual angles: restricted to 20° (SAV 20°), restricted to 45° (SAV 45°), and unrestricted HMD full field (SAV FF).*

2.6 Experimental protocol. The experimental protocol includes a preparation phase, the activity in the artificial street and the final procedures.

2.6.1 Preparation. Each participant was welcomed at the entrance hall of the Institut de la Vision where they received instructions about the study. Then, each participant was guided to the control room, accessible through the home door of the artificial street, to start the preparation and the familiarization phase. To prevent any prior exposure to the artificial street, participants were blindfolded while guided to the control room. Participants underwent eye dominance testing (Dolman method) and were equipped with the HMD, the backpack computer, the fanny pack, and the torso tracker. Once all equipment was arranged, participants familiarized themselves for 15 min with SAV. During familiarization, participants were exposed to their assigned condition (SAV 20°, SAV 45°, SAV FF, or NV) and had to solve a series of simple tasks (**Figure 3**): identifying the borders of the table in front of them, identifying position and names of five objects (three cups and two spoons), placing a cup on a coaster, and reading the number "35" and the word "CHAT" (French word for "CAT"). The tasks were chosen such that general skills required for solving the tasks in the artificial street were briefly trained while not being directly related to the tasks in the artificial street. Following familiarization, detailed instructions for the activity were provided (**Supplementary Material**) and the participant was guided to the starting point in the artificial street, facing the home door. The standard eye tracking calibration was conducted, and a countdown indicated the start of the trial.

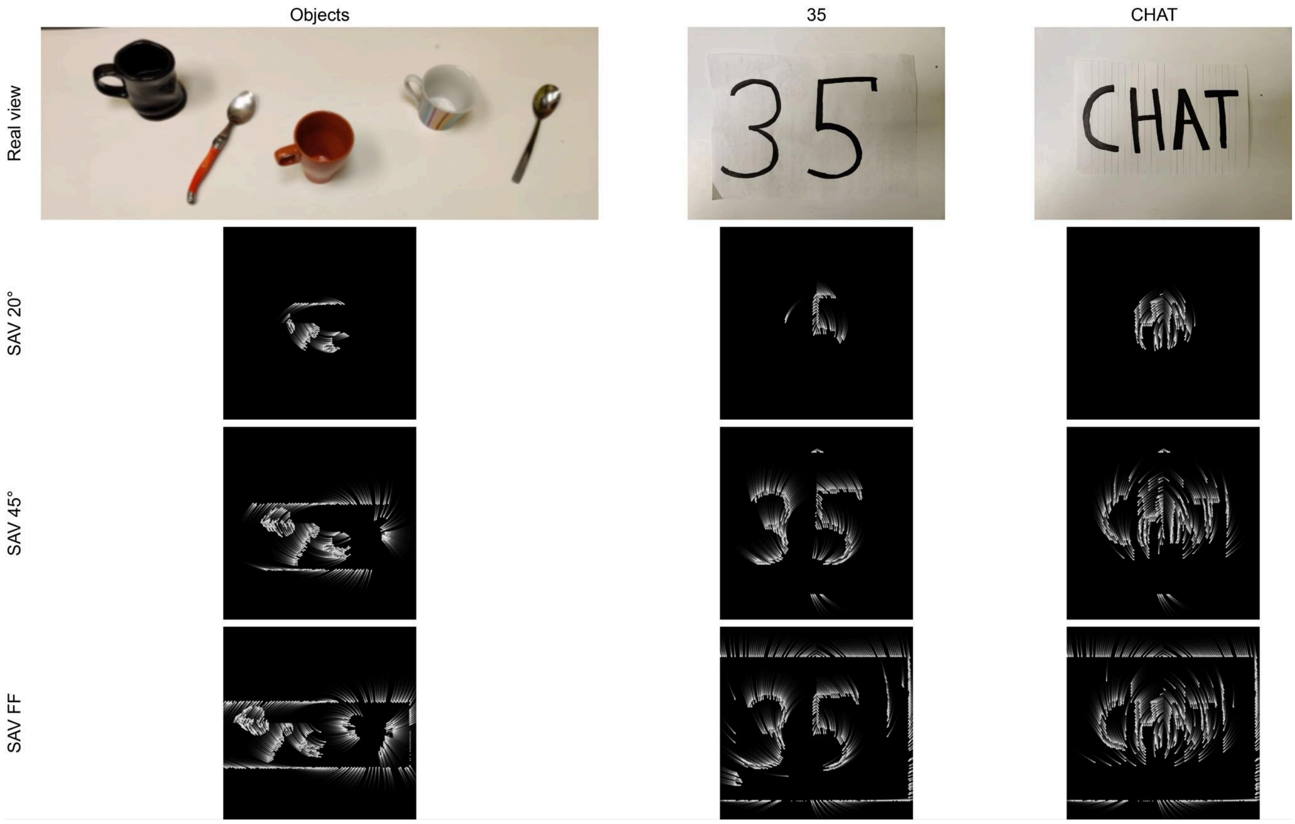


***Figure 3**. Tasks used during the familiarization phase: (1) identifying the borders of the table in front of them, (2) identifying position and names of five objects (three cups and two spoons) on the table, (3) placing a cup on a coaster, (4) reading the number "35" and (5) reading the word "CHAT".*

2.6.2 Activity. Participants were asked to imagine being in a hurry and to complete an outdoor activity in the artificial street as quickly as possible, consisting of: 1) retrieving money from an ATM, 2) posting a letter in the Post box and 3) going back home. Their starting point was in front of their home door, facing it, with the door also being their endpoint. Each station had tasks to be solved. For the ATM station, participants had to 1) correctly localize it, 2) touch the screen, 3) take the bill, and 4) identify its value. For the post box, they were instructed to 1) correctly localize it, 2) read the letter which indicated the destination on the envelope (e.g. "D" for Denmark), and 3) post it in the post box. Once both tasks were solved, the home task consisted in 1) correctly locating the home door and 2) touching the handle. A fanny pack was used to hold the envelope and the bill, so keeping participants' hands free. For all tasks, participants were instructed to rely on vision only and to not use their hands for scanning. The order of the station to reach (ATM first or Post box first) was deliberately left to the participants to assess their spontaneous exploration behavior. Importantly, participants were naïve to the arrangement of the artificial street and encountered it solely within their designated condition (SAV 20°, SAV 45°, SAV FF, or NV). Participants were instructed to remain safe and not step on the street. During the whole experiment, one experimenter was present within the artificial street to take notes on the participant's performance and ensure the participant's safety. The study was controlled via custom control software built in Unity (version 2019.2.16f1) and operated by the experimenter following the subject with a laptop computer.

Participants completed five repetitions of the activity to assess learning and generalization to unfamiliar arrangements. The first four repetitions served as training sessions, while the fifth repetition as the probe trial. Each repetition was constrained by a 10-min time limit. If a participant failed to complete all tasks within this time limit, they were stopped and guided back to the starting point. To ensure that participants relied on their vision rather than on their memory, participants were told that the arrangement of the street may or may not have changed between repetitions. In reality, the arrangement remained constant for the training sessions (training arrangement) and varied for the probe trial (probe arrangement). Both arrangements had comparable travel distances and were specifically designed to maximize the difference between the shortest and second-shortest paths. To this end, a customized Python algorithm was used to generate one million pairs of random arrangements, calculate the length of each path, and identify the 10 pairs with the largest differences. The two arrangements with the greatest path disparities were selected for the training and probe trials. The best routes for the training and probe arrangements were 20.4 m and 18.6 m respectively. The difference between the shortest and second-shortest paths was 6.8 m for the training arrangement and 6.2 m for the probe arrangement. This method aimed to maximize the sensitivity in assessing how directly participants navigate through the environment. The decision to impose a 10-min cutoff and conduct five repetitions was based on pilot experiments, revealing these numbers as a compromise between an appropriate number of training sessions, while keeping the participants' fatigue and frustration at a manageable level. At each repetition conclusion, marked by either task completion or the time limit, the HMD turned black, leaving the participant in the darkness. Participants were guided to the control room for a 5-minute resting period. Between repetitions, ambient loud music was employed to mask noises generated during the preparation of the street arrangement for the subsequent repetition. To give the participants no indication that the arrangement was not changed between the training trials, it was pretended that objects were also rearranged. For all repetitions, the preparation was the same. Participants were given the same instructions, eye-tracking was calibrated, trackers were checked for functionality, and participants were positioned at the starting point, waiting for the countdown to start.

2.6.3 Final procedures. After the probe trial, participants drew a sketch map of the last arrangement of the artificial street, capturing station-related and unrelated elements on both the floor and walls.

2.7 Data collection. The HMD recorded head position and rotation. A HTC VIVE tracker on the torso recorded torso movements. Eye movements were tracked using the *SRanipal* eye-tracking software (version 1.3.1.1). Head position and rotation, torso movements and eye movements were sampled at 100 Hz. Three HTC VIVE trackers were placed in the room corners for calibration purposes. Fixations were calculated in Python based on a combination of the PyGaze toolbox [32] and the Augmented Reality Eye Tracking Toolkit for HMDs [33]. Fixations were defined as sequences of eye movement samples where the angle between consecutive samples did not exceed 1.6° and each fixation lasted at least 250 ms [34]. The experimenter was not blinded to the trial conditions.

2.8 Statistical analysis. Preprocessing was performed in Python (version 3.9.1) and statistical analyses in R (version 4.1.1). Linear mixed effects models with viewing condition, trial and their interaction as fixed effects and participants as random effects were used for task-solving efficacy, navigation safety, and spatial navigation efficiency variables. For visual exploration behavior, fixation location as well as the respective interactions were added as additional fixed effects. The following procedure was applied to select the appropriate model. First, the standard linear mixed effects model (with the respective fixed and random effects), as well as generalized linear effects models which additionally account for zero-inflation, overdispersion and, if applicable, other distribution families, were specified using the *lmer* function from the R lme4 package and the *glmmTMB* package in R [35,36]. This step was necessary in cases of, for instance, half-integer count variables such as the reaching score, where a Tweedie distribution was more appropriate. Furthermore, it allowed for zero-inflation, which was the case in measures like the number of collisions. Subsequently, all specified models were compared based on their differences in Akaike's information criterion (dAIC) using the R *AICtab* function. The model with the lowest dAIC was tested for violation of model assumptions. Model diagnostics of correct distribution, dispersion, and homogeneity of variance of the residuals were checked using the *simulateResidulas* function from the DHARMa package in R (https://github.com/florianhartig/DHARMa) [37]. If the assumptions were not met, the assumptions of the model with the second lowest dAIC were tested. This procedure was iterated until a model finally met the assumptions. This model was then selected for further analysis. In some cases, it was not possible to identify a model which met all assumptions. In this case, the model with the fewest assumption violations was selected. It is worth noting that linear mixed models have been reported to be robust to violations of distributional assumptions [38]. All models and assumption tests are in **Supplementary Material**.

Once the appropriate model was identified, an analysis of variance was performed on this model to test for statistically significant main and interaction effects. Statistical significance was always set to $p < 0.05$. Only if the respective effects were significant, holm-adjusted post hoc comparisons were conducted with the *lsmeans* function[39] to investigate which viewing condition significantly differs during the probe test, shows learning effects between the first and the last training sessions and generalizes between the last training session and the probe trial. Learning was identified when the following two conditions apply. First, the post hoc test between the first and the last training session reveals a statistically significant change. Second, the sign of the test statistics indicates increased performance: positive test statistics when a performance increase is reflected by a reduction over time or negative test statistics otherwise. If statistically significant learning was identified, additional generalization was inferred when the post hoc comparison between the last training session and the probe trial reveals no significant difference in the opposite direction of the learning effect. For fixational exploration behavior, the same statistical approach was used. However, instead of learning we identified behavioral changes over the training sessions and generalization of these changes to the probe trial.

For the mental representation of the artificial street, the mean sketch map score and the spread of the rater ratings (operationalized as the averaged raters' standard deviations across all items) were analyzed. A linear mixed effects model with participant condition as fixed and rater as random effect was used for the mean sketch map score, and a linear regression of participant condition on rater standard deviation was used for the rater spread. For the former, the previously described procedure was followed to check for model assumptions and choose the appropriate model.

Extreme outliers were excluded from statistical analysis and labeled as empty circles in the plots (defined as values above Q3+3*IQR or below Q1-3*IQR, where Q1 and Q2 refer to the 25^{th} and 75^{th} percentiles respectively and IQR to the interquartile range). The number of participants included in each plot is specified in **Table 2**.

	Trial	**Efficacy**				**Safety**			**Efficiency**				**Exploration**			
		RS	**MS**	**TS**	**TT**	**C**	**DS**	**TS**	**PL**	**PT**	**HD**	**WS**	**IT**	**ST**	**NF**	**IF**
SAV 20°	**Train 1**	14	14	14	14	14	14	14	7	7	7	7	14	14	14	14
	Train 2	14	14	14	14	14	13	13	7	7	7	7	13	13	13	13
	Train 3	14	14	14	14	14	14	14	9	9	9	9	14	14	14	14
	Train 4	14	14	14	14	14	14	14	9	9	9	9	14	14	14	14
	Probe	14	14	14	14	14	13	13	10	10	10	10	13	13	13	13
SAV 45°	**Train 1**	14	14	14	14	14	13	13	9	9	9	9	13	13	13	13
	Train 2	14	14	14	14	14	13	13	11	11	11	11	13	13	13	13
	Train 3	14	14	14	14	14	13	13	12	12	12	12	12	13	13	13
	Train 4	14	14	14	14	14	13	13	11	11	11	11	13	13	13	13
	Probe	14	14	14	14	14	12	12	11	11	11	11	12	12	12	12
SAV FF	**Train 1**	14	14	14	14	14	14	14	13	13	13	13	14	14	14	14
	Train 2	14	14	14	14	14	13	13	13	13	13	13	13	13	13	13
	Train 3	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14
	Train 4	14	14	14	14	14	13	13	13	13	13	13	13	13	13	13
	Probe	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14
NV	**Train 1**	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
	Train 2	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
	Train 3	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
	Train 4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
	Probe	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5

Table 2. Number of participants per parameter included in each data plot.

3. RESULTS

The experiment tackles four questions. Does a wide visual angle in SAV provide a significant performance increase during the probe trial (SAV 45° vs SAV 20°)? Does an even larger visual angle in SAV provide further performance improvements during the probe trial (SAV 45° vs SAV FF)? Do the different viewing conditions differ in terms of learning during the training sessions? If yes, do they generalize this learning to the probe trial? To answer these questions, we compared the three testing groups across three performance variables: task-solving efficacy, navigation safety, and spatial navigation efficiency. For each performance variable, we assessed the effect of the viewing condition to determine if widening the visual angle provides a performance increase during the probe trial. Then, we assessed changes over repetitions to determine if the different viewing conditions differ in terms of learning during the training sessions and generalization to the probe trial. Details on statistical results are reported in the **Supplementary Materials**.

3.1 A wide visual angle increases efficacy, safety and efficiency

Task-solving efficacy is here defined as the ability to complete the activity as fast as possible with the highest success rate. To quantify efficacy, we computed four parameters. The reaching score (RS, **Figure 4a**) quantifies the number of correctly reached stations, awarding one point when participants reach a station and identify it correctly. Half a point is given if a station is reached correctly but via collisions or hand touch, zero otherwise (**Supplementary Table 1**). The mistake score (MS, **Figure 4b**) counts the number of wrongly reached stations: which is when participants reach a station (e.g., ATM) but identify it wrongly (e.g., they try to post a letter instead of retrieving money). The normalized task score (NS, **Figure 4d**) assesses the participants' ability to interact with the station and solve the associated tasks. It assigns one point for each performed task at each station (e.g., touching the ATM panel, taking the bill, and reading its value). If a task is completed after several attempts, or by hand exploration, it scores half a point (**Supplementary Table 1**). Because the number of tasks differ between stations, for each station the score is normalized to the number of respective tasks. Finally, the total time (TT, **Figure 4e**) is the time needed by participants to complete the activity. The complete quantification for the training sessions and the probe trial is reported in **Supplementary Figure 1**.

A significant 'viewing condition' main effect is present for RS ($\chi 2 = 23.61$, $p < 0.0001$, df = 2, two-tailed ANOVA on generalized linear mixed effects model with Tweedie distribution), NS ($\chi 2 = 15.55$, $p = 0.0004$, df = 2, two-tailed ANOVA on generalized linear mixed effects model with Tweedie distribution) and TT ($\chi 2 = 15.19$, $p = 0.0005$, df = 2, two-tailed ANOVA on linear mixed effects model). For these three parameters, during the probe trial, a wide visual angle (SAV 45°) resulted in a significant performance increase compared to SAV 20° (**Figure 4a,d,e**; RS: $z = -2.60$, $p = 0.0187$, two-tailed post hoc z-test; NS: $z = -2.98$, $p = 0.0059$, two-tailed post hoc z-test; TT: $t(68) = 3.77$, $p = 0.0007$, two-tailed post hoc t-test). The same applies to SAV FF compared to SAV 20° (**Figure 4a,d,e**; RS: $z = -3.59$, $p = 0.0010$, two-tailed post hoc z-test; NS: $z = -3.27$, $p = 0.0033$, two-tailed post hoc z-test; TT: $t(68) = 4.21$, $p = 0.0002$; two-tailed post hoc t-test). In contrast, SAV FF did not increase performance compared to SAV 45°

(**Figure 3a,d,e**; RS: z = -1.22, p = 0.2234, two-tailed post hoc z-test; NS: z = -0.29, p = 0.7752, two-tailed post hoc z-test; TT: t(64) = 0.45, p = 0.6523, two-tailed post hoc t-test). For MS, participants under SAV 20° made more mistakes than participants under SAV 45° and SAV FF during the probe trial (**Figure 4b**). However, the 'viewing condition' main effect is not statistically significant ($\chi 2$ = 0.25, p = 0.8812, df = 2, two-tailed ANOVA on generalized linear mixed effects model with Tweedie distribution). Nevertheless, the combination of RS and MS during the probe trial highlights that the wider the visual angle, the more stations the participants reached correctly and the fewer mistakes they made. (**Figure 4c**). Half of the participants under SAV 20° reached as many or more stations incorrectly than correctly (below the unity line). Only two participants under SAV 45° scored on the unity line, and none below. All participants under SAV FF scored above the unity line.

The summary plot recapitulates the results obtained for all task-solving efficacy variables and illustrates the gain provided by a wide visual angle (**Figure 4f**). Qualitatively, a large performance increase is observed when increasing the SAV from 20° to 45°, while only little incremental benefit is observed with an even wider SAV visual angle.

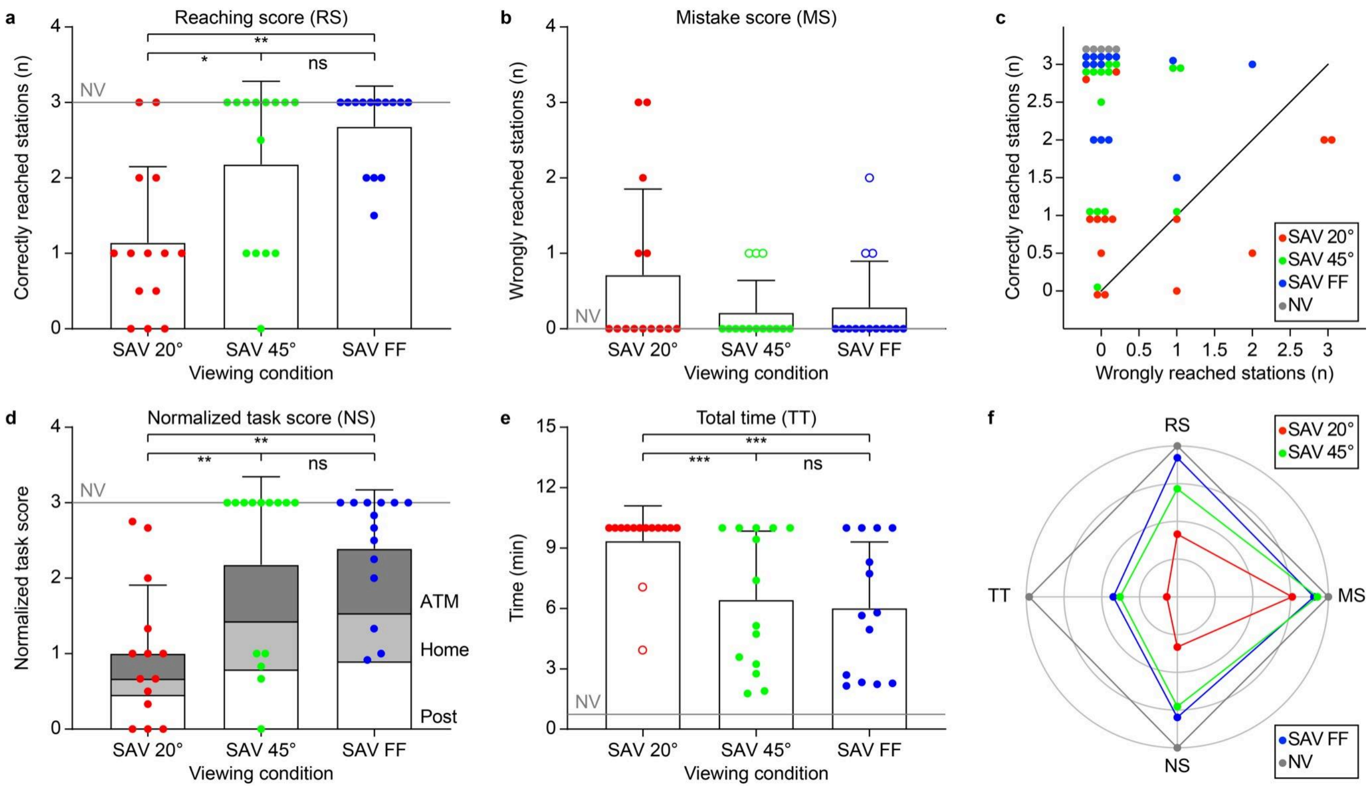


***Figure 4**. Task-solving efficacy during the probe trial as a function of the SAV viewing condition. **a,b** Quantification of reaching score (RS, **a**) and mistake score (MS, **b**). **c** Combination of RS and MS. The black line is the unity line. **d,e** Quantification of normalized task score (NS, **d**) and total time (TT, **e**). Bars are mean ± SD of all participants (n = 14 per viewing condition). Empty circles indicate outliers. The gray line is the mean performance of participants under NV (n = 5). Results from two-tailed post hoc comparisons based on the respective linear mixed effects models (nt: not tested; ns: p > 0.05: *: p <*

*0.05; **: p < 0.01; ***: p < 0.001).* ***f*** *Summary plot with mean RS, MS, NS, and TT normalized to the best performance among all viewing conditions. Gray circles correspond to 1, 0.75, 0.5, and 0.25.*

While efficacy in solving tasks is important, it is not the only relevant criterion in daily life. Another key factor is the ability to act and navigate safely, for example by avoiding collisions with objects and not walking in dangerous areas such as the street (red trajectories in **Figure 5a** and **Supplementary Figure 2**). Hence, we investigated the extent of safe behavior exhibited by participants. Evaluation of safe behavior encompasses three parameters: the total distance that participants traveled on the street which they were instructed not to step on (DS, **Figure 5b**), the total time that participants spent on the street (TS, **Figure 5c**), and the number of collisions with walls and objects (NC, **Figure 5d**). For the first two parameters (DS and TS), the number of participants is reduced to n = 13 under SAV 20° and n = 12 under SAV 45°, because data tracking was corrupted for three participants. For NC, the problem was not relevant since collisions were counted manually by the experimenter. The complete quantification for the training sessions and the probe trial is reported in **Supplementary Figure 3**.

A significant 'viewing condition' main effect is present for DS ($\chi 2 = 92.19$, $p < 0.0001$, df = 2, two-tailed ANOVA on generalized linear mixed effects model with Tweedie distribution), TS ($\chi 2 = 109.20$, $p < 0.0001$, df = 2, two-tailed ANOVA on generalized linear mixed effects model with Tweedie distribution) and NC ($\chi 2 = 9.86$, $p = 0.0072$, df = 2, two-tailed ANOVA on generalized linear mixed effects model with Poisson distribution). For the three parameters, during the probe trial, a wide visual angle (SAV 45°) provided a significant increase in navigation safety compared to SAV 20° (DS: $z = 2.26$, $p = 0.0477$; TS: $z = 3.56$, $p = 0.0007$; NC: $z = 3.85$, $p = 0.0004$, two-tailed post hoc z-tests). The same applies to SAV FF compared to SAV 20° (DS: $z = 4.83$, $p < 0.0001$; TS: $z = 6.31$, $p < 0.0001$; NC: $z = 3.02$, $p = 0.0051$; two-tailed post hoc z-tests). In contrast, SAV FF did not increase safety compared to SAV 45° for DS and NC (DS: $z = 1.94$, $p = 0.0529$; NC: $z = -1.43$, $p = 0.1524$, two-tailed post hoc z-tests) but it did for TS ($z = 3.00$, $p < 0.0027$, two-tailed post hoc z-test).

The summary plot illustrates the gain of a wide visual angle for all navigation safety variables (**Figure 5e**). As for task-solving efficacy, an increase in safety is observed when increasing the SAV visual angle from 20° to 45°, while little benefit is observed for a further increase of the visual angle.

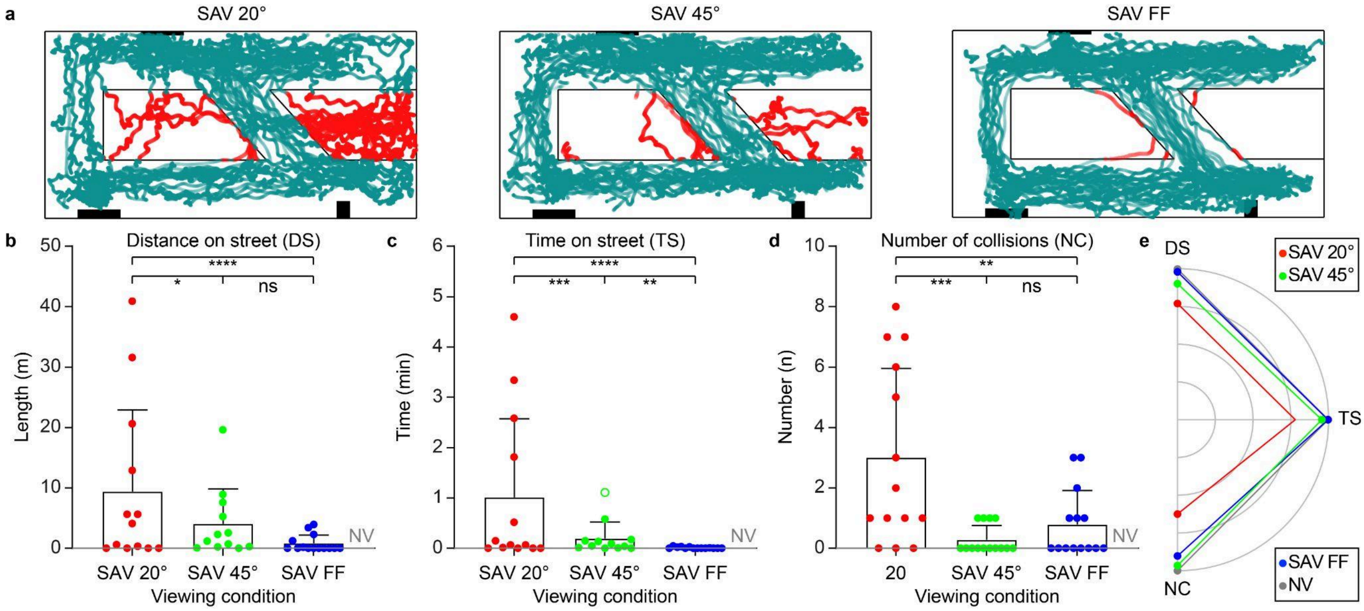


***Figure 5**. Navigation safety during the probe trial as a function of the SAV viewing condition. **a** Overlay of all participants' trajectories for SAV 20°, SAV 45°, and SAV FF. Dark green shows when participants are on the sidewalk and crosswalk, whereas red indicates when they are on the street. Black rectangles represent the three stations. **b-d** Quantification of distance on street (DS, **b**), time on street (TS, **c**), and number of collisions (NC, **d**). Each bar plot is the mean ± SD of all participants (for DS and TS: n = 13 for SAV 20°, n = 12 for SAV 45°, and n = 14 for SAV FF; for NC: n = 14 participants per viewing condition). Empty circles indicate outliers. The gray line shows the mean performance of participants under NV (n = 5). Results from two-tailed post hoc comparisons based on respective linear mixed effects models (ns: p > 0.05; *: p < 0.05; **: p < 0.01; ***: p < 0.001: **** p < 0.0001). **f** Summary plot with mean DS, TS, and NC normalized to the best performance among all viewing conditions. Gray circles correspond to 1, 0.75, 0.5, and 0.25.*

The third performance variable is spatial navigation efficiency. Navigation trajectories differ between participants and repetitions (**Figure 6a**). While participants under NV followed a direct trajectory between the stations, this ability is reduced under SAV with a restricted visual angle. To quantify spatial navigation efficiency, we excluded from the analysis the periods in which participants solve the tasks at the stations (orange and cyan in **Figure 6a**), which occurs when a participant reaches a station (RS is not zero), is in the interaction zone (gray semicircle in **Figure 6a**). The remaining trajectories were divided into path segments between two successfully reached stations (blue in **Figure 6a**).

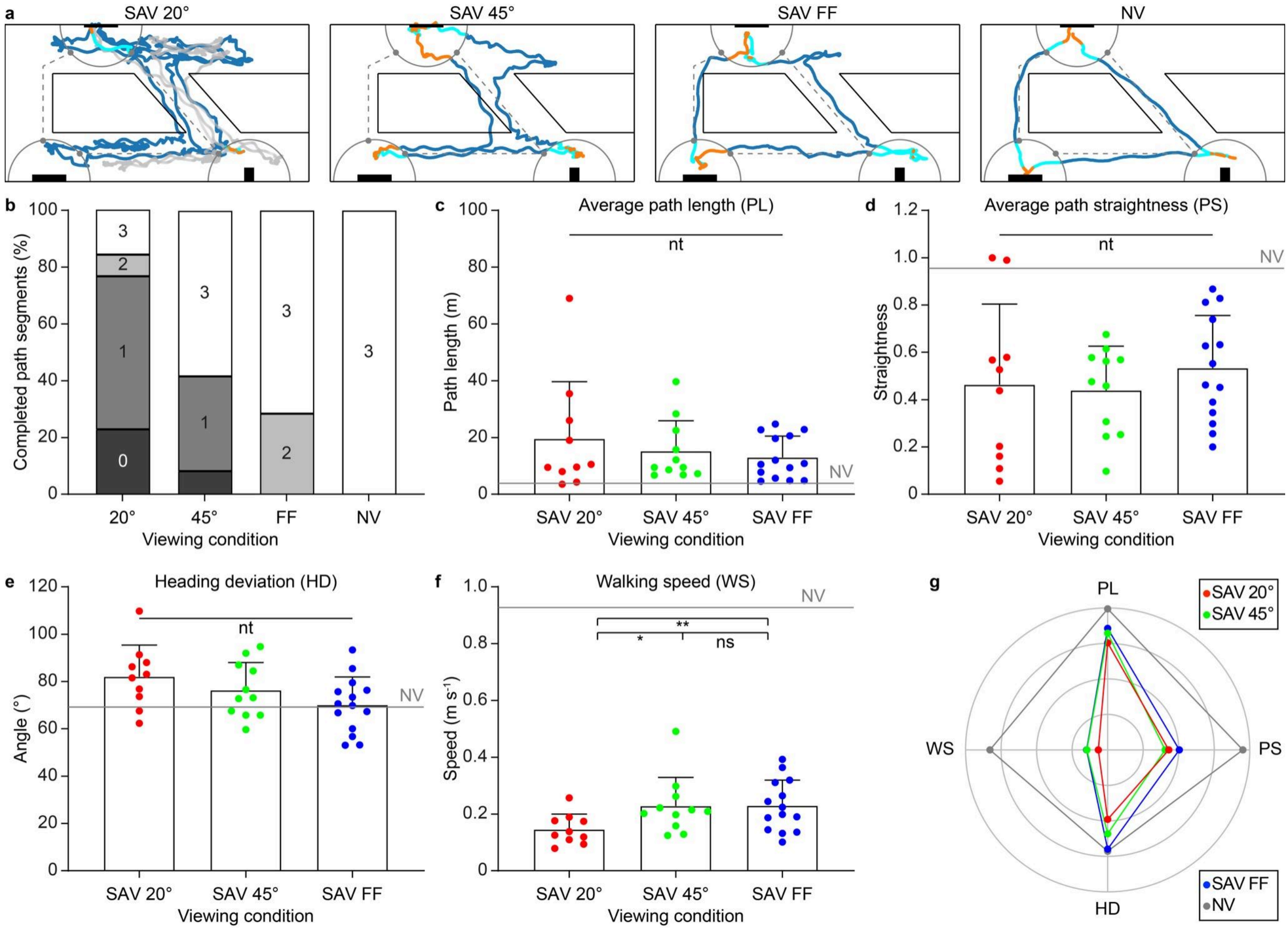


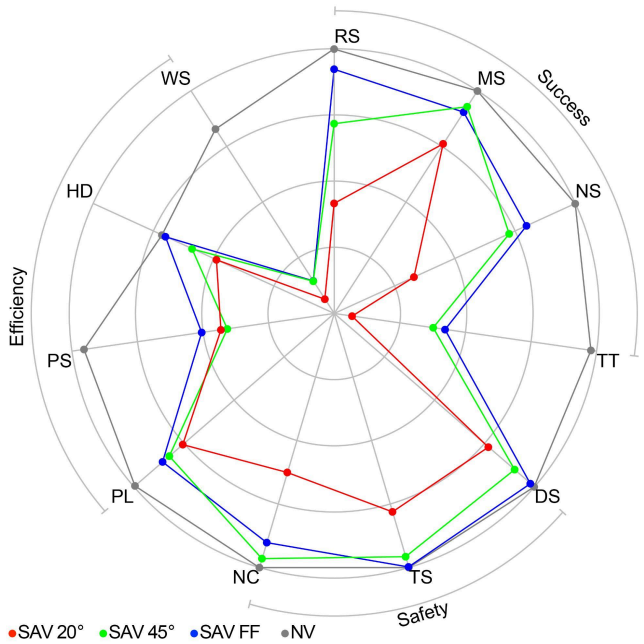
***Figure 6**. Spatial navigation efficiency during the probe trial as a function of the SAV viewing condition. **a** Trajectories from representative participants under SAV 20°, SAV 45°, SAV FF, and NV during the probe trial. The gray semicircles are the interaction zones (1.2 m radius) around each station (black rectangles). Gray dashed lines are optimal trajectories connecting stations. Blue shows navigation segments between stations, orange indicates approaching a station and interaction, cyan marks when a participant is leaving a station to the next station, and gray indicates the last navigation segment in case the next station is not reached. **b** Percentage of completed path segments during the probe trial across all participants for each viewing condition. The color code corresponds to the total number of completed segments out of 3 (black: 0; dark gray: 1; light gray: 2; white: 3). **c-f** Quantification of average path length (PL, **c**), average path straightness (PS, **d**), average heading deviation (HD, **e**), and average walking speed (WS, **f**). Each bar plot shows the mean ± SD of all participants who completed at least one path segment (n = 10 / 13 for SAV 20°, n = 11 / 12 for SAV 45°, and n = 14 / 14 for SAV FF). Empty circles indicate outliers. The gray line shows the mean performance of participants under NV (n = 5). Results from two-tailed post hoc comparisons based on the respective linear mixed effects models (nt: not tested; ns: p > 0.05; *: p < 0.05; **: p < 0.01). **f** Summary plot with mean PL, PS, HD and WS normalized to the best performance among all viewing conditions. Gray circles correspond to 1, 0.75, 0.5, and 0.25.*

Efficiency was first gauged through the number of completed path segments (**Figure 6b**), which is when a participant successfully reaches the next station (RS is not zero). The smaller the visual angle, the less likely participants are to complete all three segments (white box), and the more likely they do not complete any segment (black box). As for navigation safety, the total number of participants is n = 13 under SAV 20° and n = 12 under SAV 45° since data tracking was corrupted for three participants.

Subsequently, we only considered completed path segments and assessed efficiency using four parameters: average path length (PL, **Figure 6c**), average path straightness (PS, **Figure 6d**), average heading deviation (HD, **Figure 6e**), and average walking speed (WS, **Figure 6f**). Each parameter is calculated separately for each path segment within a trial and, for each trial, all paths are then averaged. The choice of computing efficiency measures only on completed path segments instead of the total trajectories is due to the large variability in participants' trajectories. This variability is mainly caused by two reasons. First, participants chose their paths and station order. Second, participants did not always reach at least one station (number of participants reaching at least one station: n = 10 / 13 under SAV 20°, n = 11 / 12 under SAV 45°, n = 14 / 14 under SAV FF and n = 5 / 5 under NV). PL and WS are derived from motion trajectories. PS is the ratio of optimal trajectory length (gray dashed lines in **Figure 6a**) to the PL[30]. HD is the mean angular disparity of momentary velocity vectors along the trajectory relative to the direction of the station the participant is heading for. The complete quantification for training sessions and the probe trial is reported in **Supplementary Figure 4**.

While a general improvement in spatial navigation is qualitatively observed in participants with an increased visual angle (**Figure 6a**), the models for PL, PS and HD do not reveal statistical significance for the 'viewing condition' main effect (PL: $\chi 2 = 4.63$, $p = 0.0987$, df = 2; PS: $\chi 2 = 1.78$, $p = 0.4115$, df = 2; HD: $\chi 2 = 5.26$, $p = 0.0720$, df = 2; two-tailed ANOVA on linear mixed effects models). Nevertheless, SAV visual angle influences the speed at which participants navigate between stations. WS reveals a significant 'viewing condition' main effect ($\chi 2 = 26.34$, $p < 0.0001$, df = 2, two-tailed ANOVA on linear mixed effects model). In the probe trial, a wide visual angle (SAV 45°) allows for a significantly faster WS compared to SAV 20° ($t(136) = -2.27$, $p = 0.0493$, two-tailed post hoc t-test). The same applies to SAV FF compared to SAV 20° ($t(136) = -3.15$, $p = 0.0060$, two-tailed post hoc t-test). In contrast, SAV FF did not increase WS compared to SAV 45° ($t(136) = -0.48$, $p = 0.6348$; two-tailed post hoc t-test). The lack of a statistically significant 'viewing condition' main effect for PL, PS and HD should be considered against the fact that the analysis accounts only for completed path segments in those participants who managed to reach at least one station. Under SAV 20°, 3 out of 13 participants did not reach any station and, thus, have been excluded. For 7 more, only 1 out of 3 path segments was completed. As a result, this analysis is intrinsically biased towards the better performing participants from which there are fewer in the SAV 20° viewing condition (**Figure 6b**). And, for these successful cases, once the participant managed to complete a path segment, results suggest that an increase in SAV visual angle does not influence the extent to which they perform direct paths towards the stations.

The summary plot shows the gain of the wider visual angle for the efficiency performance variable (**Figure 6g**). Compared to success and safety, the overall performance increase is less striking. Improvement is visible for HD and WS, but less so for PL and PS. It should, again, be noted that this analysis is biased towards the better performing participants, of which there are fewer in the SAV 20° viewing condition. Among those, participants travel similar distances to find a station regardless of the viewing condition, but a wider visual angle allows them to do so in less time.



A final summary plot (**Fig. 7**) resumes these findings, bringing together the three performance variables (task-solving efficacy, navigation safety, and spatial navigation efficiency). The SAV visual field size impacts participants' performance. Overall, enlarging the SAV visual angle provides a performance increase across all analyzed variables, in particular from SAV 20° to SAV 45°.

Figure 7*. Aggregated summary plot. Mean parameters for each performance variable normalized to the best performance among all viewing conditions. Gray circles correspond to 1, 0.75, 0.5, and 0.25.*

3.2 A wide visual angle changes the visual exploration behavior and facilitates a more accurate mental representation of the environment

Now, we examine possible factors explaining the enhanced performance observed with an increased SAV visual angle. First, we evaluated how the visual angle influences the participants' visual exploration behavior. Then, we probed how it influences the participants' mental representation of the artificial street.

Spatial navigation efficiency (**Figure 6**) suggests that participants under all SAV viewing conditions travel similar distances until they find the next station. One can thus infer that, in order to find their direction, participants explore the environment to a similar extent. However, the finding that a wider visual angle enables them to walk faster suggests that participants exploited different exploration behaviors between the different viewing conditions. First, we analyzed the time participants took to inspect the street at the start of the trial (Inspection Time, IT; **Figure 8a**). More specifically, IT is the time from the moment participants stop looking at the home door at trial start until the time they leave the interaction zone (gray semicircle; **Figure 6a**). The total number of participants is reduced to n = 13 under SAV 20° and n = 12 under SAV 45°. IT reveals a significant 'viewing condition' main effect ($\chi 2 = 17.07$, $p = 0.0002$, $df = 2$, two-tailed ANOVA on linear mixed effects model). However, during the probe trial, IT under SAV 20° do not differ significantly from IT under SAV 45° or SAV FF (SAV 20° vs SAV 45°: $t(163) = -0.35$, p = 0.7294; SAV 20° vs SAV FF: $t(163) = 2.11$, p = 0.0733; two-tailed post hoc t-tests). Only participants under SAV FF spend significantly less time inspecting the room at the start than participants under SAV 45° (t(163)

= 2.51, p = 0.0396, two-tailed post hoc t-test). The complete quantification for training sessions and the probe trial is reported in **Supplementary Figure 5**.

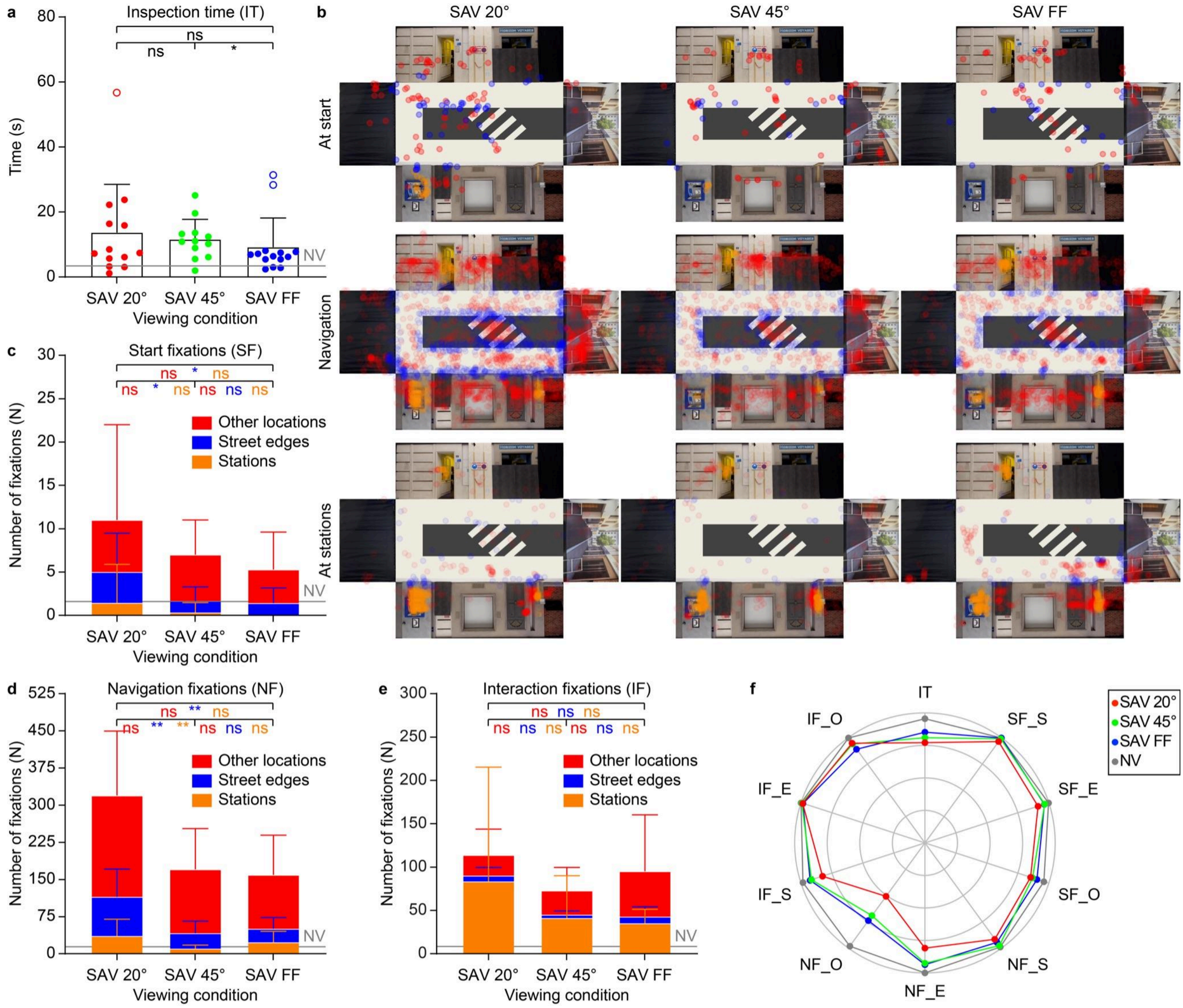


***Figure 8**. Visual exploration behavior during the probe trial as a function of the SAV viewing condition. **a** Quantification of inspection time (IT). Bar plot shows the mean ± SD of all participants (n = 13 for SAV 20°, n = 12 for SAV 45°, and n = 14 for SAV FF). Empty circles indicate outliers. The gray line represents the average performance of participants under NV (n = 5). **b** Gaze fixations of all participants during the probe trial divided into fixations at trial start (SF, top row), during navigation (NF, middle row), and during station interaction (IF, bottom row). Orange dots show fixations falling on stations, blue dots show fixations falling on street edges, and red dots show remaining fixations. **c-e** Quantification of start fixations at the start of the probe trial (SF, **c**), navigation fixations (NF, **d**) during navigation between stations, and interaction fixations (IF, **e**) during interaction at each station. Each bar plot represents the mean ± SD of all participants (n = 13 for SAV 20°, n = 12 for SAV 45°, and n = 14 for SAV FF). Fixations are divided into fixations on street edges (blue), stations (orange), and other locations (red). The gray line shows the mean total number of fixations under NV (n = 5). Results from two-tailed post hoc*

*comparisons based on respective generalized linear mixed effects models (ns: p > 0.05; *: p < 0.05; **: p < 0.01).* ***f*** *Summary plot with mean IT, SF, NF and IF normalized to the best performance among all viewing conditions. Fixations are divided based on the fixation point: stations (SF_S, NF_S and IF_S), street edges (SF_E, NF_E and IF_E), and other locations (SF_O, NF_O and IF_O).*

To further investigate the participants' exploration behavior, we analyzed fixational gaze behavior during IT (Start Fixations, SF; **Figure 8b** top row and **Figure 8c**), during navigation between stations (Navigation Fixations, NF; **Figure 8b** middle row and **Figure 8d**), and during interaction at each station (Interaction Fixations, IF; **Figure 8b** bottom row and **Figure 8e**). Moreover, fixations are divided into three groups based on their location: street edges (_E), stations (_S), and other locations (_O). The rationale for this division is that street edges help navigation[40] and stations are the important landmarks in the artificial street.

SF shows a significant 'viewing condition' main effect (χ2 = 10.45, p = 0.0054, df = 2, two-tailed ANOVA on generalized linear mixed effects model with Poisson distribution), a significant 'fixation location' main effect (χ2 = 345.81, p < 0.0001, df = 4, two-tailed ANOVA on generalized linear mixed effects model with Poisson distribution) and a significant 'viewing condition x fixation location' interaction effect (χ2 = 14.44, p = 0.0060, df = 4, two-tailed ANOVA on generalized linear mixed effects model with Poisson distribution), indicating that the viewing condition has an impact on what participants looked at during IT. At the start of the probe trial, participants under SAV 20° focused significantly more on street edges compared to participants with wider visual angles (SAV 20° vs SAV 45°: z = 2.78, p = 0.0163; SAV 20° vs SAV FF: z = 2.62, p = 0.0175; two-tailed post hoc z-tests) while no significant difference is observed between SAV 45° and SAV FF (z = -0.27, p = 0.7856, two-tailed post hoc z-tests). No significant difference between viewing conditions during the probe trial is present for fixations on stations (SAV 20° vs SAV 45°: z = 0.00, p = 1; SAV 20° vs SAV FF: z = 0.00, p = 1; SAV 45° vs SAV FF: z = 0.00, p = 1; two-tailed post hoc z-tests) or other locations (SAV 20° vs SAV 45°: z = -0.40, p = 0.8389; SAV 20° vs SAV FF: z = 0.81, p = 0.8389; SAV 45° vs SAV FF: z = 1.27, p = 0.6134; two-tailed post hoc z-tests).

NF shows a significant 'viewing condition' main effect (χ2 = 15.64, p = 0.0004, df = 2, two-tailed ANOVA on generalized linear mixed effects model with negative binomial distribution), a significant 'fixation location' main effect (χ2 = 1094.83, p < 0.0001, df = 2, two-tailed ANOVA on generalized linear mixed effects model with negative binomial distribution), and a significant 'viewing condition x fixation location' interaction effect (χ2 = 23.84, p < 0.0001, df = 4, two-tailed ANOVA on generalized linear mixed effects model with negative binomial distribution), indicating that the viewing condition has also an impact on where participants look during navigating. During the probe trial and while walking, participants under SAV 20° focus significantly more on street edges compared to participants with wider visual angles (SAV 20° vs SAV 45°: z = 3.06, p = 0.0066; SAV 20° vs SAV FF: z = 3.05, p = 0.0066; two-tailed post hoc z-tests) while no significant difference is present between the SAV 45° and SAV FF (z = -0.01, p = 0.9937, two-tailed post hoc z-test). SAV 45° further leads to significantly fewer fixations on stations compared to SAV 20° (z = 3.24, p = 0.0036, two-tailed post hoc z-test) while no significant difference is present between SAV 20° and SAV FF (z = 1.68, p = 0.1714, two-tailed post hoc z-test) and between SAV 45° vs SAV FF (z = -1.72, p = 0.1714, two-tailed post hoc z-test). No significant difference is present for fixations on other locations (SAV 20° vs SAV 45°: z = 2.08, p = 0.0745; SAV 20° vs SAV FF: z = 2.35, p = 0.0567; SAV 45° vs SAV FF: z = 0.28, p = 0.7828; two-tailed post hoc z-tests).

In contrast to SF and NF, the 'viewing condition' main effect is not statistically significant for IF ($\chi2$ = 4.85, p = 0.0884, df = 2, two-tailed ANOVA on generalized linear mixed effects model with negative binomial distribution) but the 'fixation location' is ($\chi2$ = 215.02, p < 0.0001, df = 2, two-tailed ANOVA on generalized linear mixed effects model with negative binomial distribution) together with the 'viewing condition x fixation location' interaction ($\chi2$ = 56.06, p < 0.0001, df = 4, two-tailed ANOVA on generalized linear mixed effects model with negative binomial distribution). However, the probe test does not reveal significant differences between viewing conditions during interaction at stations, neither for fixations on street edges (SAV 20° vs SAV 45°: z = 0.22, p = 1; SAV 20° vs SAV FF: z = -0.29, p = 1; SAV 45° vs SAV FF: z = -0.59, p = 1; two-tailed post hoc z-tests) nor for fixations on stations (SAV 20° vs SAV 45°: z = 0.93, p = 0.7020; SAV 20° vs SAV FF: z = -0.27, p = 0.7866; SAV 45° vs SAV FF: z = -1.37, p = 0.5124; two-tailed post hoc z-tests) nor for fixations on other locations (SAV 20° vs SAV 45°: z = -0.10, p = 1; SAV 20° vs SAV FF: z = -0.36, p = 1; SAV 45° vs SAV FF: z = -0.29, p = 1; two-tailed post hoc z-tests).

The summary plot during the probe trial highlights that the viewing condition influences the participant's visual exploration behavior (**Figure 8f**). In particular, at the trial start and during navigation, participants exposed to a wider visual angle rely significantly less on street edges for orientation. Also, during navigation, they perform significantly fewer fixations on stations.

These findings suggest that the visual field size under SAV has an impact on participants' performance (task-solving efficacy, navigation safety, and spatial navigation efficiency) and visual exploration behavior while performing daily activities. We hypothesize that a wider visual angle might facilitate the mental representation of the street. To test this hypothesis, participants drew a sketch map of the artificial street arrangement seen during the probe trial (**Figure 9a**). Instructions for sketch map creation emphasized the inclusion of environmental features without artistic concerns (**Supplementary Materials**) [41,42]. While some participants accurately depicted all relevant elements and additional details in correct proportions (high-score map), others provided a rough approximation but omitted important parts (medium-score map), and still others failed to reproduce even the general structure of the street (low-score map).

We assessed the accuracy of the mental representation of the artificial street using the sketch map evaluation metric inspired by Lynch's work on urban navigation[43,44] based on landmarks representation and orientation, route segments and structures, presence of additional landmarks, and orientation (**Supplementary Table 2**) [42]. 13 independent and naïve raters (7 female and 6 male), aged 23-37 years (min-max) and 29.15 ± 3.90 years (mean ± SD) evaluated the sketch maps.

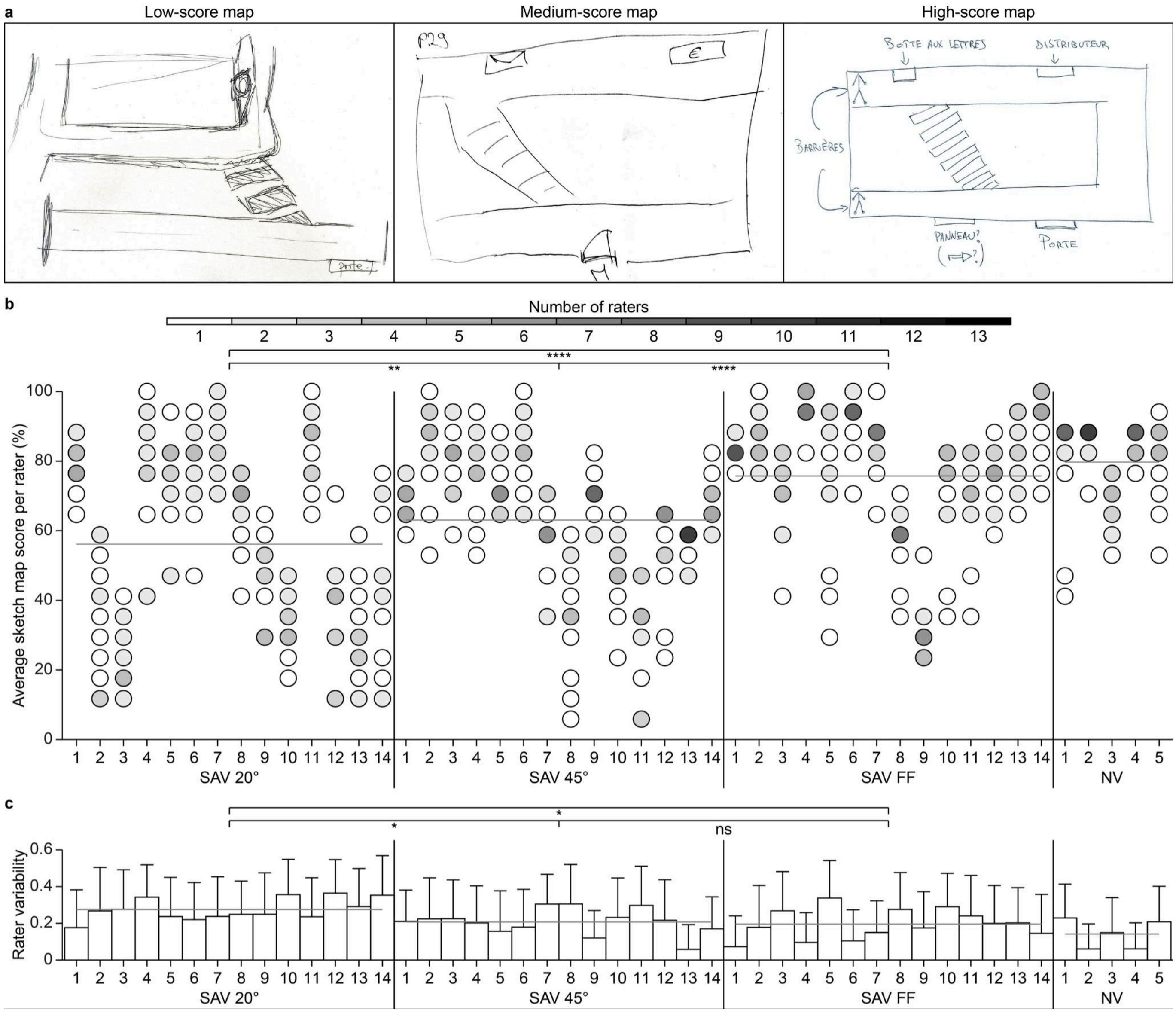


***Figure 9**. Mental representation after the probe trial. **a** Representative sketch maps illustrating the participants' mental representations of the probe arrangement. **b** Mean score assigned to each individual sketch map (one per participant on the x-axis). The symbol color corresponds to the number of raters assigning the respective score. The horizontal gray lines show the average score among all participants per viewing condition. **c** Mean rating SD per question of the sketch map evaluation metric. Bar plots show the mean ± SD. The horizontal gray lines are the average variabilities among all participants per viewing condition. Result from two-tailed post hoc comparisons based on a linear mixed effects model for **b** and a linear model for **c** (ns: $p > 0.05$; *: $p < 0.05$; **: $p < 0.01$; ****: $p < 0.0001$).*

We used two measures to quantify the difference between viewing conditions: the average score among all raters (**Figure 9b**) and the variability between raters (**Figure 9c**). For the average score, the ANOVA on a generalized linear model with the 'viewing condition' as fixed effect and the 'rater' as random effect revealed a statistically significant 'viewing condition' effect ($\chi 2 = 80.78$, $p < 0.0001$, $df = 2$). A significant difference appeared for all viewing conditions (**Figure 9b**; SAV 20° vs SAV 45°: $t(539) = -3.00$, $p = 0.0029$; SAV 20° vs SAV FF: $t(539) = -8.45$, $p < 0.0001$; SAV 45° vs SAV FF: $t(539) = -6.42$, $p < 0.0001$;

two-tailed post hoc t-tests) indicating that an increased visual angle leads to a more accurate representation of the artificial street. We choose SD to measure variability among raters. SD was first calculated for every question of the sketch map evaluation metric, and then averaged within participants to obtain the rater variability (**Figure 9c**). A significant 'viewing condition' effect is also present in rater variability ($F_{(2,39)}$ = 5.39, p = 0.0086, two-tailed one-way ANOVA on linear regression with rater SD as dependent and viewing condition as independent measure), with ratings for SAV 45° and SAV FF being less variable than the ones for SAV 20° (**Figure 9c**; SAV 20° vs SAV 45°: p = 0.0190; SAV 20° vs SAV FF: p = 0.0170; SAV 45° vs SAV FF: p = 0.6800; two-tailed post hoc t-tests).

The smaller visual angle resulted in a greater score variability among raters. This result indicates increased difficulty in achieving a consensus rating, probably because rating a poor sketch map is more difficult than one that matches the artificial street well. Overall, sketch maps of participants with larger visual angles reach higher and less variable scores from independent raters. This result suggests that a larger visual angle allows for a better mental representation of the artificial street, which, in turn, allows participants to interact more successfully with the environment.

3.3 A wide angle fosters learning during training sessions which generalizes to the probe trial.

We have found that a wider SAV visual angle (SAV 45° compared to SAV 20°) enhances participants' ability to navigate and solve daily tasks in a naturalistic environment more effectively, safely, and efficiently. Additionally, a wider visual angle triggers a change in visual exploration behavior and facilitates a more accurate mental representation of the environment. Now, we address the questions whether or not the different viewing conditions differ in terms of learning during the training sessions and, if yes, whether or not these learnings generalize to the probe trial. To answer these questions, we assessed the 'repetition' main effect for each parameter in each performance variable.

For task-solving efficacy, we found a significant 'repetition' main effect for RS ($\chi2$ = 53.12, p < 0.0001, df = 4, two-tailed ANOVA on generalized linear mixed effects model with Tweedie distribution), NS ($\chi2$ = 40.58, p < 0.0001, df = 4, two-tailed ANOVA on generalized linear mixed effects model with Tweedie distribution) and TT ($\chi2$ = 69.19, p < 0.0001, df = 4, two-tailed ANOVA on linear mixed effects model), but not for MS ($\chi2$ = 2.55, p = 0.6352, df = 4, two-tailed ANOVA on generalized linear mixed effects model with Tweedie distribution). Participants reached significantly more stations at the last training session compared to the first one under all viewing conditions (RS; **Figure 10a**; SAV 20°: z = -3.33, p = 0.0017; SAV 45°: z = -4.42, p < 0.0001; SAV FF: z = -3.40, p = 0.0014; two-tailed post hoc z-tests) and completed significantly more tasks (NS; **Figure 10c**; SAV 20°: z = -2.98, p = 0.0058; SAV 45°: z = -3.26, p = 0.0022; SAV FF: z = -2.86, p = 0.0084; two-tailed post hoc z-tests), indicating learning for both RS and NS. The learning generalizes to the probe trial, which does not show any statistically significant difference compared to the last training session for both RS (**Figure 10a**; SAV 20°: z = 1.05, p = 0.2950; SAV 45°: z = 0.51, p = 0.6082; SAV FF: z = 0.83, p = 0.4068; two-tailed post hoc z-tests) and NS (**Figure 10c**; SAV 20°: z = 0.57, p = 0.5658; SAV 45°: z = -0.35, p = 0.7262; SAV FF: z = -0.43, p = 0.6672; two-tailed post hoc z-tests). Participants completed the trial significantly faster (TT) in the last training session compared to the first one under SAV 45° and SAV FF (**Figure 10d**; SAV 45°: $t(142) = 4.53$, $p < 0.0001$; SAV FF: $t(142) = 7.11$, $p < 0.0001$; two-tailed post hoc t-tests) but not under SAV 20° ($t(144) = 0.57$, $p = 1$, two-tailed post hoc t-test). The learning under SAV 45° and SAV FF generalizes to the probe trial (**Figure 10d**; SAV 45°: $t(141) = 0.28$, $p = 0.7778$; SAV FF: $t(141) = -1.12$, $p = 0.2633$; two-tailed post hoc t-tests).

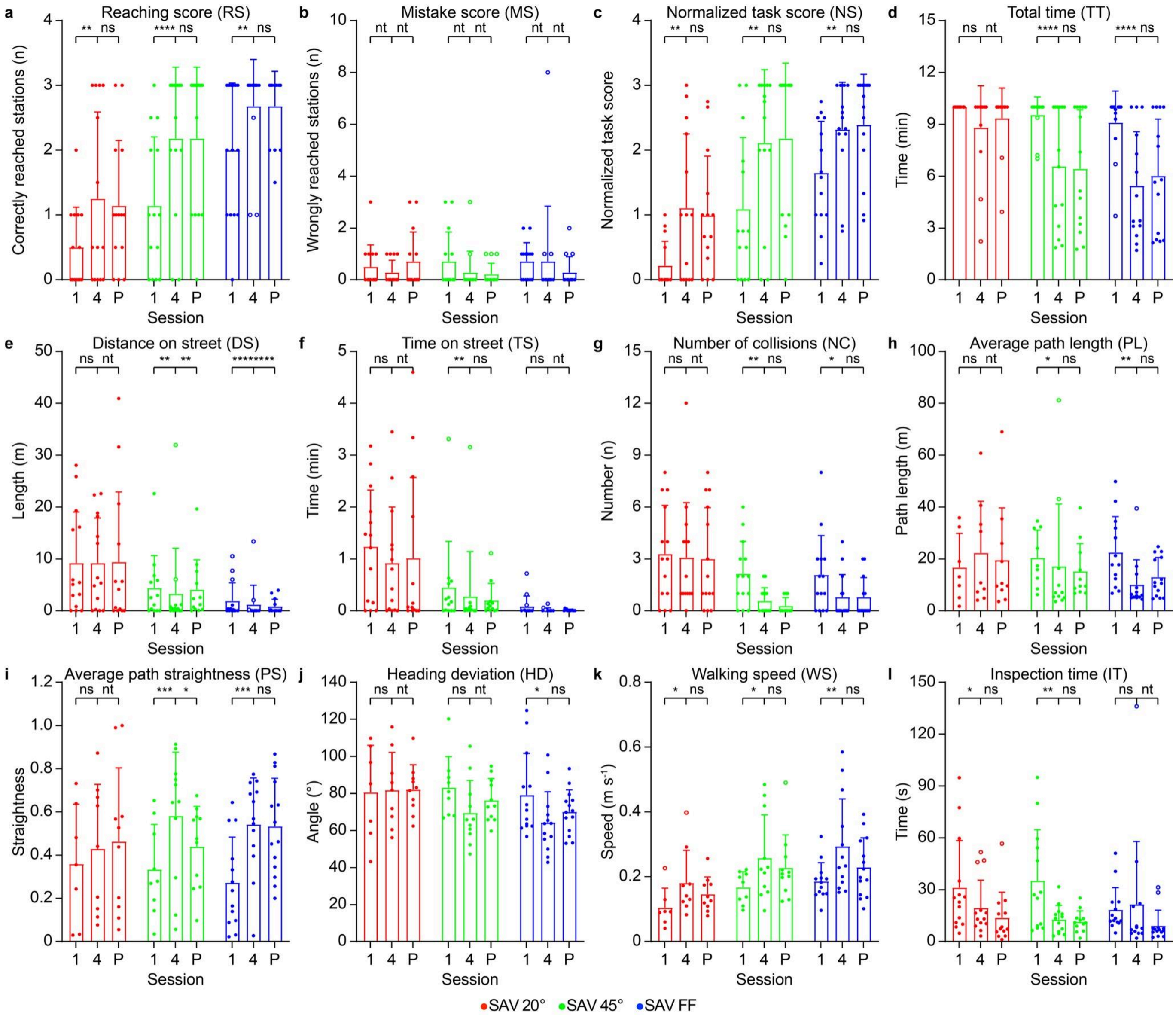


Figure 10. Learning and generalization as a function of the SAV viewing condition. *Quantification of learning and generalization for the reaching score (RS,* ***a****), mistake score (MS,* ***b****) normalized task score (NS,* ***c****), total time (TT,* ***d****), distance on street (DS,* ***e****), time on street (TS,* ***f****), number of collisions (NC,* ***g****), average path length (PL,* ***h****), average path straightness (PS,* ***i****), average heading deviation (HD,* ***j****), average walking speed (WS,* ***k****), and inspection time (IT,* ***l****). Bar plots show mean ± SD of all participants. Empty circles are outliers. 1: training session 1; 4: training session 4; P: probe trial. Results from two-tailed post hoc comparisons based on the respective linear mixed effects models (nt: not tested; ns: p > 0.05: *: p < 0.05; **: p < 0.01; ***: p < 0.001; ****: p < 0.0001).*

For navigation safety, we identified a significant ‘repetition’ main effect in DS (χ2 = 3500.19, p < 0.0001, df = 4, two-tailed ANOVA on generalized linear mixed effects model with Tweedie distribution), TS (χ2 = 20.36, p = 0.0004, df = 4, two-tailed ANOVA on generalized linear mixed effects model with Tweedie distribution), and NC (χ2 = 15.23, p = 0.0042, df = 4, two-tailed ANOVA on generalized linear mixed effects model with Poisson distribution). We found that participants travel significantly shorter distances on the street (DS) at the last training session compared to the first one under SAV 45° and SAV FF (**Figure 10e**; SAV 45°: z = 3.38, p = 0.0014; SAV FF: z = 24.91, p < 0.0001; two-tailed post hoc z-tests) while do not under SAV 20° (z = 0.16, p = 1, two-tailed post hoc z-test). However, both learning effects under SAV 45° and SAV FF do not generalize to the probe trial (**Figure 10e**; SAV 45°: z = -3.14, p = 0.0017; SAV FF: z = -30.63, p < 0.0001; two-tailed post hoc z-tests). Also, participants spend significantly less time on the street (TS) at the last training session compared to the first one under SAV 45° (**Figure 10f**; z = 3.44, p = 0.0011, two-tailed post hoc z-test) while do not under SAV 20° and SAV FF (SAV 20°: z = 1.21, p = 0.4544, SAV FF: z = 1.56, p = 0.1182; two-tailed post hoc z-tests). Learning under SAV 45° generalizes to the probe trial (**Figure 10f**; z = -1.66, p = 0.0965, two-tailed post hoc z-test). Finally, participants have significantly fewer collisions (NC) at the last training session compared to the first one under SAV 45° and SAV FF (**Figure 10g**; SAV 45°: z = 3.32, p = 0.0018; SAV FF: z = 2.74, p = 0.0124; two-tailed post hoc z-tests) and do not under SAV 20° (z = 0.32, p = 1, two-tailed post hoc z-test). Learning under SAV 45° and SAV FF generalizes to the probe trial (**Figure 10g**; SAV 45°: z = 1.13, p = 0.2577; SAV FF: z = 0.00, p = 1; two-tailed post hoc z-tests).

For spatial navigation efficiency, the ‘repetition’ main effect is statistically significant for all measures (PL: χ2 = 17.47, p = 0.0016, df = 4; PS: χ2 = 41.13, p < 0.0001, df = 4; HD: χ2 = 10.26, p = 0.0363, df = 4; WS: 2 = 54.79, p < 0.0001, df = 4; two-tailed ANOVA on linear mixed effects models). Participants traveled significantly shorter PL at the last training session compared to the first one under SAV 45° and SAV FF (**Figure 10h**; SAV 45°: t(114) = 2.67, p = 0.0176; SAV FF: t(114) = 3.56, p = 0.0011; two-tailed post hoc t-tests) but not under SAV 20° (t(119) = -1.00, p = 0.6391, two-tailed post hoc t-test). The learning effect under SAV 45° and SAV FF generalizes to the probe trial (**Figure 10h**; SAV 45°: t(119) = -1.64, p = 0.1034; SAV FF: t(113) = -1.22, p = 0.2260; two-tailed post hoc t-tests). Participants also perform significantly straighter paths (PS) in the last training session compared to the first one under SAV 45° and SAV FF (**Figure 10i**; SAV 45°: t(144) = -3.59, p = 0.0009; SAV FF: t(144) = -3.97, p = 0.0002; two-tailed post hoc t-tests) while no significant difference is observed under SAV 20° (SAV 20°: t(144) = -0.57, P = 1; two-tailed post hoc t-test). In this case, only the learning effect under SAV FF generalizes to the probe trial (**Figure 10i**; SAV 45°: t(144) = 2.03, p = 0.0439; SAV FF: t(144) = -0.12, p = 0.9075, two-tailed post hoc t-tests). HD is significantly lower in the last training session compared to the first one under SAV FF (**Figure 10j**; t(115) = 2.36, p = 0.0395; two-tailed post hoc t-test) but not significantly different under SAV 20° and SAV 45° (SAV 20°: t(120) = -0.23, p = 1; SAV 45°: t(116) = 2.06, p = 0.0841; two-tailed post hoc t-tests). The learning effect under SAV FF generalizes to the probe trial (**Figure 10j**; t(114) = -0.82, p = 0.4118; two-tailed post hoc t-test). Finally, WS is significantly higher in the last training session compared to the first one under all viewing conditions (**Figure 10k**; SAV 20°: t(136)

= -2.29, p = 0.0473; SAV 45°: t(136) = -2.61, p = 0.0200; SAV FF: t(136) = -3.58, p = 0.0010; two-tailed post hoc t-tests). In all viewing conditions, this learning effect generalizes to the probe trial (**Figure 10k**; SAV 20°: t(136) = 0.04, p = 0.9657; SAV 45°: t(136) = 0.88, p = 0.3795; SAV FF: t(136) = 1.92, p = 0.0572; two-tailed post hoc t-tests).

Last, IT also shows a significant 'repetition' main effect ($\chi 2$ = 36.44, p < 0.0001, df = 4, two-tailed ANOVA on linear mixed effects model). Participants took significantly less time to inspect the room at the last training session compared to the first one under SAV 20° and SAV 45° (**Figure 10l**; SAV 20°: t(163) = 2.45, p = 0.0305; SAV 45: t(163) = 2.88, p = 0.0091; two-tailed post hoc t-tests) while no significant difference is observed under SAV FF (t(163) = 1.52, p = 0.1315, two-tailed post hoc t-test). Both learning effects under SAV 20° and SAV 45° generalize to the probe trial (**Figure 10l**; SAV 20°: t(163) = 0.30, p = 0.7628; SAV 45°: t(163) = 0.34, p = 0.7374; two-tailed post hoc t-tests).

In summary, learning and generalization is observed in 4 parameters out of 12 under SAV 20° (RS, NS, WS, and IT), 8 parameters under SAV 45° (RS, NS, TT, TS, NC, PL, WS, and IT) and 8 parameters under SAV FF (RS, NS, TT, NC, PL, PS, HD, and WS). Changes in fixational gaze behavior over the training sessions and generalization of these changes to the probe trial are reported in **Supplementary Figures 6** and **7** and **Supplementary Material**.

DISCUSSION

This study documents that a wide visual angle (45°) under SAV enhances participants' ability to navigate and solve tasks in a naturalistic environment more effectively, safely, and efficiently than a small visual angle (20°). Moreover, it promotes their learning and generalization capability. Notably, further increasing the visual angle beyond 45° does not yield significant additional improvements in most metrics. These results may be attributed to the different visual exploratory behaviors adopted as a function of the visual angle. Additionally, a wider visual angle enables participants to construct a more precise mental representation of the environment.

We evaluated the influence of the visual field size on the ability of simulated blind individuals with restored vision to engage in daily activities within a naturalistic environment. Previous research has predominantly focused on the impact of resolution [8,9] rather than visual angle, primarily due to the technological limitations of earlier approaches (such as prostheses), which required a trade-off between resolution and visual angle [45]. Recent advancements in vision restoration technologies might overcome this challenge, enabling the combination of high resolution with wide visual angle, and now raising the question of the appropriate visual angle for efficient vision restoration. Previous studies identified a minimal visual angle between 20° to 35° as critical for navigation and daily tasks under NV [11] and SAV [14]. However, using stereotyped tasks and environments questions whether this critical value holds for a more complex and naturalistic context, particularly under artificial vision. Similarly, prior experiments on patients with artificial vision have mainly been conducted in controlled environments using simple and stereotyped tasks [17,23,24,46,47]. Hence, the predictive validity of these results for the patients' real-world performance remains uncertain [48]. Alternatively, a growing body of research used virtual reality (VR) to offer a more immersive approach [14,15,49,50]. However, existing VR studies primarily focus on virtual perception only, rather than interaction with the physical environment; thus, neglecting the importance of sensorimotor integration in naturalistic real-life activities. In this study, we capitalize on the advances of AR/VR technology to design a fully immersive and ecologically valid task within a physical artificial street. In contrast to previous studies, by incorporating SAV in AR, participants are not only tasked with perceiving but also physically interacting with their surroundings. This approach provides a unique opportunity to bridge the gap between laboratory experiments and real-world experience in low-vision and vision restoration research.

Our results highlight the importance of a wide visual angle in helping individuals with SAV to effectively solve tasks in a naturalistic setting. We found that crossing the 45° angle led to a significant improvement in performance compared to the 20° angle. However, increasing the angle beyond 45° did not show much extra benefit for most tasks. This finding suggests that there is a point between 20° and 45° where further increases in visual angle do not add much to how well people can see and interact with their surroundings. Yet, these results do not necessarily imply that a wide angle in artificial vision yields equivalent performance compared to normal vision. Indeed, the differences between SAV and NV conditions varied greatly across measures. Although participants' ability to complete tasks and maintain safety with wider visual angles neared NV levels, their speed differed considerably from NV performance. This finding suggests that, even with a large visual angle, artificial vision still faces challenges, particularly in quickly mastering daily tasks. Therefore, increasing the visual angle may be a necessary, but not sufficient, element for improvement. Specifically for this simulation, as it emulates an instance of an epiretinal implant, one factor is the elongated shape of the phosphenes caused by the activation of many axons, which has been previously identified as impactful in artificial vision [14,51]. Slow temporal perception characteristic of prosthetic vision and perception fading are other important limiting factors [15,52,53]. Some participants reported struggling with the low temporal resolution. Additionally, since SAV was displayed in only one eye, as prosthetics and optogenetics are delivered monocularly or unilaterally, many participants encountered difficulties with depth perception, citing it as the main reason for collisions.

The brain is plastic and holds the capacity to adapt to new signals over time. This property was reported to also alter perception in the case of auditory [54] and limb [55] prostheses. Even though patients with artificial vision reported improvements through training, the quality of their perception did not change over time [48]. Rather, the improvements appear to be task-specific, stemming from the patients' enhanced ability to interpret the given signal more effectively [56]. Our results suggest that, for most measures, learning played an important role. Participants performed more successfully, safely, and efficiently in the last training session than they did in the first one and reported an easier recognition of the visual input over time. However, the extent to which this result applies varies with the visual angle. Participants with a larger angle demonstrated learning across multiple measures, whereas participants limited to 20° showed improvement in only a few measures. This might explain limited changes over time in previous clinical tests, as the implants evaluated in prior clinical trials were constrained to visual angles up to 20° [8,9]. Consequently, our findings suggest that a visual angle greater than 20° is crucial for enhancing patients' performance over time. It is important to note that, while clinical studies spanned several years, our study was restricted to 5 consecutive repetitions within a few hours on the same day. Different learning patterns may emerge over longer periods. Additionally, the learning effect does not merely reflect memory performance since we informed participants that the arrangement might change between repetitions. Yet, it is possible that participants still recognized the same arrangement during learning sessions and relied on memory. However, two observations contradict this possibility. First, many participants explored different routes during the sessions. Second, even if some did not, their ability to generalize the learned behavior to the probe trial highlights the reliance on vision as opposed to memory.

Understanding the underlying mechanisms behind performance variations under SAV is crucial. Efficiency measures suggest that participants under different SAV viewing conditions explore and travel similar distances to reach stations. However, the visual exploration behavior differs, as indicated by faster walking speeds among participants with a wider visual angle. Analysis of gaze data supports this observation, revealing that participants with a 20° visual angle focus more on street edges compared to those with larger visual angles. The focus on street edges aligns with literature indicating that individuals with low vision rely more on such cues for navigation [40,57]. Consequently, the emphasis on street edges among participants with a 20° visual angle may cause them to overlook other relevant features on walls, potentially explaining their lower performance despite the fact that they traveled similar distances to find a station. Moreover, the fact that straight street edges are easier to recognize than the more complex shapes on the walls may contribute to the discrepancy in performance. In summary, while all groups covered similar distances exploring the street, participants with wider visual angles were better equipped to identify station-relevant details, leading to improved performance. Conversely, participants with smaller visual angles may have struggled to form an accurate mental representation of the artificial street, as reflected in their sketch maps. At the same time, while participants with an ultra-wide visual angle (SAV FF) draw significantly improved sketch maps compared to those with under 45°, this enhancement did not translate into improved efficacy, safety, and efficiency. To explain this result, it is important to consider that this study did not involve complex navigation tasks requiring a robust mental representation for effective performance. Thus, while an ultra-wide visual angle leads to a more accurate mental representation, its impact on navigation and interaction with the environment may vary depending on the task complexity.

It is reasonable to question how accurately the SAV reflects a real patient's perception of artificial vision. We strived to ground the simulation in known data and phenomena to the best of our ability. To this end, we chose to model the parameters of the POLYRETINA implant [14,15,25–29]. This decision is rooted in the majority of evidence and patient reports stemming from retinal implants [58,59], as well as in the ability of POLYRETINA specifically to offer high resolution and wide visual angle. In addition, the simulation accounts for anatomical, physiological, and phenomenological aspects reported by patients with previous retinal implants. For example, we incorporated findings regarding the variability in the shape of phosphenes experienced by patients with epiretinal implants [59]. We also accounted for perceptual fading and slow time resolution. Yet, a patient's perception may still vary from the simulation, although implementing the mentioned aspects should bring it closer to reality. Importantly, we argue that this approach offers significant benefits as the simulation enables testing complementary effects of vision restoration approaches compared to what clinical trials can achieve. This aspect might greatly aid in evaluating expected utility before clinical assessment, potentially avoiding unnecessary patients' discomfort [48]. Using the POLYRETINA implant to emulate phosphene perception in the simulation might be seen as another limitation. Nonetheless, we argue that the results and implications extend beyond the scope of POLYRETINA. Ultimately, this study focuses on the general visual properties crucial to enable a useful perception and interaction with the environment. Identifying these parameters is essential for all vision restoration therapies. Additionally, the broader application of using SAV to explore the utility of artificial vision in naturalistic settings extends beyond the constraints of POLYRETINA parameters. Also, our study emphasizes the importance of visual angle in artificial vision, but it can be adapted to explore any other parameter crucial for further improving vision restoration therapies.

The versatility of this approach allows simulating and testing in naturalistic scenarios virtually every restoration therapy, thus opening up a portfolio of opportunities for both fundamental and applied research. Yet, it is important to acknowledge that conditions in an artificial street still differ from real-world situations. Specifically, real-world situations involve maintaining visual attention with various intra- and inter-modality distractors, along with the additional psychological burden of navigating potentially hazardous elements. Despite these limitations, our study serves as a foundational step towards understanding the utility of artificial vision in daily life. Traditionally, vision restoration focuses on applying a specific technique to patients and assessing their capabilities through clinical tests in controlled settings with stereotyped tests. This approach opens up the possibility to revert this process by first assessing if the restored properties are sufficient for daily activities, and only after testing them in patients if appropriate. To achieve this goal, we advocate for a holistic, naturalistic approach that reflects real-world interactions to truly understand the utility of artificial vision.

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

DATA AVAILABILITY

The authors declare that the data supporting the findings of this study are available in the paper. Any additional requests for information can be directed to the corresponding author.

CODE AVAILABILITY

The SAV code is accessible online (https://github.com/lne-lab/polyretina_ar). The dataset and the analysis code to replicate the study will be available on Zenodo before publication.

REFERENCES

[1] Burton M J, Ramke J, Marques A P, Bourne R R A, Congdon N, Jones I, Tong B A M A, Arunga S, Bachani D, Bascaran C, Bastawrous A, Blanchet K, Braithwaite T, Buchan J C, Cairns J, Cama A, Chagunda M, Chuluunkhuu C, Cooper A, Crofts-Lawrence J, Dean W H, Denniston A K, Ehrlich J R, Emerson P M, Evans J R, Frick K D, Friedman D S, Furtado J M, Gichangi M M, Gichuhi S, Gilbert S S, Gurung R, Habtamu E, Holland P, Jonas J B, Keane P A, Keay L, Khanna R C, Khaw P T, Kuper H, Kyari F, Lansingh V C, Mactaggart I, Mafwiri M M, Mathenge W, McCormick I, Morjaria P, Mowatt L, Muirhead D, Murthy G V S, Mwangi N, Patel D B, Peto T, Qureshi B M, Salomão S R, Sarah V, Shilio B R, Solomon A W, Swenor B K, Taylor H R, Wang N, Webson A, West S K, Wong T Y, Wormald R, Yasmin S, Yusufu M, Silva J C, Resnikoff S, Ravilla T, Gilbert C E, Foster A and Faal H B 2021 The *Lancet Global Heal*th Commission on Global Eye Health: vision beyond 2020 Lancet Global Heal **9** e489–551

[2] Jacobson S G and Cideciyan A V 2010 Treatment Possibilities for Retinitis Pigmentosa *New Engl J Medicine* **363** 1669–71

[3] Kerschensteiner D 2023 Losing, preserving, and restoring vision from neurodegeneration in the eye *Curr. Biol.* **33** R1019–36

[4] Gelder R N V, Chiang M F, Dyer M A, Greenwell T N, Levin L A, Wong R O and Svendsen C N 2022 Regenerative and restorative medicine for eye disease *Nat Med* **28** 1149–56

[5] Ghezzi D 2015 Retinal prostheses: progress toward the next generation implants *Front Neurosci-switz* **9** 290

[6] Cehajic-Kapetanovic J, Singh M S, Zrenner E and MacLaren R E 2022 Bioengineering strategies for restoring vision *Nat Biomed Eng* 1–18

[7] Yue L, Weiland J D, Roska B and Humayun M S 2016 Retinal stimulation strategies to restore vision: Fundamentals and systems *Prog Retin Eye Res* **53** 21–47

[8] Borda E and Ghezzi D 2022 Advances in visual prostheses: engineering and biological challenges *Prog Biomed Eng* **4** 032003

[9] Ghezzi D 2023 The role of the visual field size in artificial vision *J Neural Eng* **20** 023001

[10] HAYMES S, GUEST D, HEYES A and JOHNSTON A **19**96 Mobility of People with Retinitis Pigmentosa as a Function of Vision and Psychological Variables *Optometry Vision Sci* **73** 621–37

[11] Hassan S E, Hicks J C, Lei H and Turano K A 2007 What is the minimum field of view required for efficient navigation? *Vision Res* **47** 2115–23

[12] Barhorst-Cates E M, Rand K M and Creem-Regehr S H 2016 The Effects of Restricted Peripheral Field-of-View on Spatial Learning while Navigating *Plos One* **11** e0**16**3785

[13] Barhorst-Cates E M, Rand K M and Creem-Regehr S H 2019 Navigating with peripheral field loss in a museum: learning impairments due to environmental complexity *Cognitive Res Princ Implic* **4** 41

[**14**] Thorn J T, Migliorini E and Ghezzi D 2020 Virtual reality simulation of epiretinal stimulation highlights the relevance of the visual angle in prosthetic vision J Neural Eng **17** 056019

[15] Thorn J T, Chenais N A L, Hinrichs S, Chatelain M and Ghezzi D 2022 Virtual reality validation of naturalistic modulation strategies to counteract fading in retinal stimulation J Neural Eng 19 026016

[16] Sahel J-A, Grieve K, Pagot C, Authié C, Mohand-Said S, Paques M, Audo I, Becker K, Chaumet-Riffaud A-E, Azoulay L, Gutman E, Léveillard T, Zeitz C, Picaud S, Dalkara D and Marazova K 2021 Assessing Photoreceptor Status in Retinal Dystrophies: From High-Resolution Imaging to Functional Vision *Am. J. Ophthalmol.* **230** 12–47

[17] Stingl K, Bartz-Schmidt K U, Besch D, Braun A, Bruckmann A, Gekeler F, Greppmaier U, Hipp S, Hörtdörfer G, Kernstock C, Koitschev A, Kusnyerik A, Sachs H, Schatz A, Stingl K T, Peters T, Wilhelm B and Zrenner E 2013 Artificial vision with wirelessly powered subretinal electronic implant alpha-IMS *Proc Royal Soc B Biological Sci* **280** 20130077

[**18**] Cruz L da, Dorn J D, Humayun M S, Dagnelie G, Handa J, Barale P-O, Sahel J-A, Stanga P E, Hafezi F, Safran A B, Salzmann J, Santos A, Birch D, Spencer R, Cideciyan A V, Juan E de, Duncan J L,

Eliott D, Fawzi A, Koo L C O de, Ho A C, Brown G, Haller J, Regillo C, Priore L V D, Arditi A, Greenberg R J and Group A I S 2016 Five-Year Safety and Performance Results from the Argus II Retinal Prosthesis System Clinical Trial *Ophthalmology* **123** 2248–54

[19] Edwards T L, Cottriall C L, Xue K, Simunovic M P, Ramsden J D, Zrenner E and MacLaren R E 2018 Assessment of the Electronic Retinal Implant Alpha AMS in Restoring Vision to Blind Patients with End-Stage Retinitis Pigmentosa Ophthalmology **125** 432–43

[20] Brindley G S and Lewin W S 1968 The sensations produced by electrical stimulation of the visual cortex *J Physiology* **196** 479–93

[21] Dobelle W H, Mladejovsky M G and Girvin J P 1974 Artificial Vision for the Blind: Electrical Stimulation of Visual Cortex Offers Hope for a Functional Prosthesis *Science* **183** 440–4

[22] DOBELLE W H, MLADEJOVSKY M G, EVANS J R, ROBERTS T S and GIRVIN J P 1976 'Braille' reading by a blind volunteer by visual cortex stimulation *Nature* **259** 111–2

[23] Fernández E, Alfaro A, Soto-Sánchez C, González-López P, Ortega A M L, Peña S, Grima M D, Rodil A, Gómez B, Chen X, Roelfsema P R, Rolston J D, Davis T S and Normann R A 2021 Visual percepts evoked with an Intracortical 96-channel microelectrode array inserted in human occipital cortex *J. Clin. Investig.* **131**

[24] Sahel J-A, Boulanger-Scemama E, Pagot C, Arleo A, Galluppi F, Martel J N, Esposti S D, Delaux A, Aubert J-B de S, Montleau C de, Gutman E, Audo I, Duebel J, Picaud S, Dalkara D, Blouin L, Taiel M and Roska B 2021 Partial recovery of visual function in a blind patient after optogenetic therapy Nat Med **27** 1223–9

[25] Ferlauto L, Leccardi M J I A, Chenais N A L, Gilliéron S C A, Vagni P, Bevilacqua M, Wolfensberger T J, Sivula K and Ghezzi D 2018 Design and validation of a foldable and photovoltaic wide-field epiretinal prosthesis *Nat Commun* **9** 992

[26] Leccardi M J I A, Chenais N A L, Ferlauto L, Kawecki M, Zollinger E G and Ghezzi D 2020 Photovoltaic organic interface for neuronal stimulation in the near-infrared *Commun Mater* **1 2**1

[27] Chenais N A L, Leccardi M J I A and Ghezzi D 2021 Photovoltaic retinal prosthesis restores high-resolution responses to single-pixel stimulation in blind retinas Commun Mater 2 28

[28] Chenais N A L, Leccardi M J I A and Ghezzi D 2021 Naturalistic spatiotemporal modulation of epiretinal stimulation increases the response persistence of retinal ganglion cell J Neural Eng 18 016016

[29] Vagni P, Leccardi M J I A, Vila C-H, Zollinger E G, Sherafatipour G, Wolfensberger T J and Ghezzi D 2022 POLYRETINA restores light responses in vivo in blind Göttingen minipigs Nat Commun **13** 3678

[30] Bécu M, Sheynikhovich D, Tatur G, Agathos C P, Bologna L L, Sahel J-A and Arleo A 2020 Age-related preference for geometric spatial cues during real-world navigation *Nat. Hum. Behav.* **4** 88–99

[31] Chang K J, Dillon L L, Deverell L, Boon M Y and Keay L 2020 Orientation and mobility outcome measures *Clin. Exp. Optom.* **103** 434–48

[32] Dalmaijer E S, Mathôt S and Stigchel S V der 2014 PyGaze: An open-source, cross-platform toolbox for minimal-effort programming of eyetracking experiments *Behav. Res. Methods* **46** 913–21

[33] Kapp S, Barz M, Mukhametov S, Sonntag D and Kuhn J 2021 ARETT: Augmented Reality Eye Tracking Toolkit for Head Mounted Displays *Sensors* **21** 2234

[34] Llanes-Jurado J, Marín-Morales J, Guixeres J and Alcañiz M 2020 Development and Calibration of an Eye-Tracking Fixation Identification Algorithm for Immersive Virtual Reality *Sens. (Basel, Switz.)* **20** 4956

[35] Bates D, Mächler M, Bolker B and Walker S 2015 Fitting Linear Mixed-Effects Models Using lme4 *J. Stat. Softw.* **67**

[36] Brooks M E, Kristensen K, Benthem K J ,van, Magnusson A, Berg C W, Nielsen A, Skaug H J, Mächler M and Bolker B M 2017 glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling *R J.* **9** 378

[37] Hartig F 2022 *DHARMa - Residual diagnostics for Hieraerchical models*

[38] Schielzeth H, Dingemanse N J, Nakagawa S, Westneat D F, Allegue H, Teplitsky C, Réale D,

Dochtermann N A, Garamszegi L Z and Araya-Ajoy Y G 2020 Robustness of linear mixed-effects models to violations of distributional assumptions *Methods Ecol. Evol.* **11** 1141–52

[39] Lenth R V 2016 Least-Squares Means: The R Package lsmeans J. Stat. Softw. **69**

[40] Matsuda Y, Kawauchi A and Motooka N 2021 Gazing behavior exhibited by people with low vision while navigating streets *J. Asian Arch. Build. Eng.* **20** 414–27

[41] Blajenkova O, Motes M A and Kozhevnikov M 2005 Individual differences in the representations of novel environments *J. Environ. Psychol.* **25** 97–109

[42] Zhong J Y and Kozhevnikov M 2016 Relating allocentric and egocentric survey-based representations to the self-reported use of a navigation strategy of egocentric spatial updating J. Environ. Psychol. **46** 154–75

[43] HUYNH N T, HALL G B, DOHERTY S and SMITH W W 2022 Interpreting urban space through cognitive map sketching and sequence analysis *Can. Geogr. Le Géographe Can.* **52** 222–40

[44] McCunn L J and Gifford R 2018 Spatial navigation and place imageability in sense of place *Cities* **74** 208–18

[45] He Y, Huang N T, Caspi A, Roy A and Montezuma S R 2019 Trade-Off Between Field-of-View and Resolution in the Thermal-Integrated Argus II System *Transl Vis Sci Technology* **8** 29

[46] Humayun M S, Dorn J D, Cruz L da, Dagnelie G, Sahel J-A, Stanga P E, Cideciyan A V, Duncan J L, Eliott D, Filley E, Ho A C, Santos A, Safran A B, Arditi A, Priore L V D, Greenberg R J and Group A I S 2012 Interim Results from the International Trial of Second Sight's Visual Prosthesis Ophthalmology **119** 779–88

[47] Palanker D, Mer Y L, Mohand-Said S, Muqit M and Sahel J A 2020 Photovoltaic Restoration of Central Vision in Atrophic Age-Related Macular Degeneration Ophthalmology **127** 1097–104

[48] Erickson-Davis C and Korzybska H 2021 What do blind people "see" with retinal prostheses? Observations and qualitative reports of epiretinal implant users Plos One 16 e0229189

[49] Hayes J S, Yin V T, Piyathaisere D, Weiland J D, Humayun M S and Dagnelie G 2003 Visually Guided Performance of Simple Tasks Using Simulated Prosthetic Vision *Artif Organs* **27** 1016–28

[50] Zapf M P, Matteucci P B, Lovell N H, Zheng S and Suaning G J 2014 Towards Photorealistic and Immersive Virtual-reality Environments for Simulated Prosthetic Vision: Integrating Recent Breakthroughs in Consumer Hardware and Software *2014 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2014** 2597–600

[51] Guo H, Wang Y, Yang Y, Tong S, Zhu Y and Qiu Y 2010 Object Recognition Under Distorted Prosthetic Vision *Artif. Organs* **34** 846–56

[52] Avraham D, Jung J-H, Yitzhaky Y and Peli E 2021 Retinal prosthetic vision simulation: temporal aspects *J. Neural Eng.* **18** 0460d9

[53] Fornos A P, Sommerhalder J, Cruz L da, Sahel J A, Mohand-Said S, Hafezi F and Pelizzone M 2012 Temporal Properties of Visual Perception on Electrical Stimulation of the Retina *Investigative Opthalmology Vis Sci* **53** 2720

[54] Purdy S C and Kelly A S 2016 Change in Speech Perception and Auditory Evoked Potentials over Time after Unilateral Cochlear Implantation in Postlingually Deaf Adults *Semin. Hear.* **37** 062–73

[55] Cuberovic I, Gill A, Resnik L J, Tyler D J and Graczyk E L 2019 Learning of Artificial Sensation Through Long-Term Home Use of a Sensory-Enabled Prosthesis *Front. Neurosci.* **13** 853

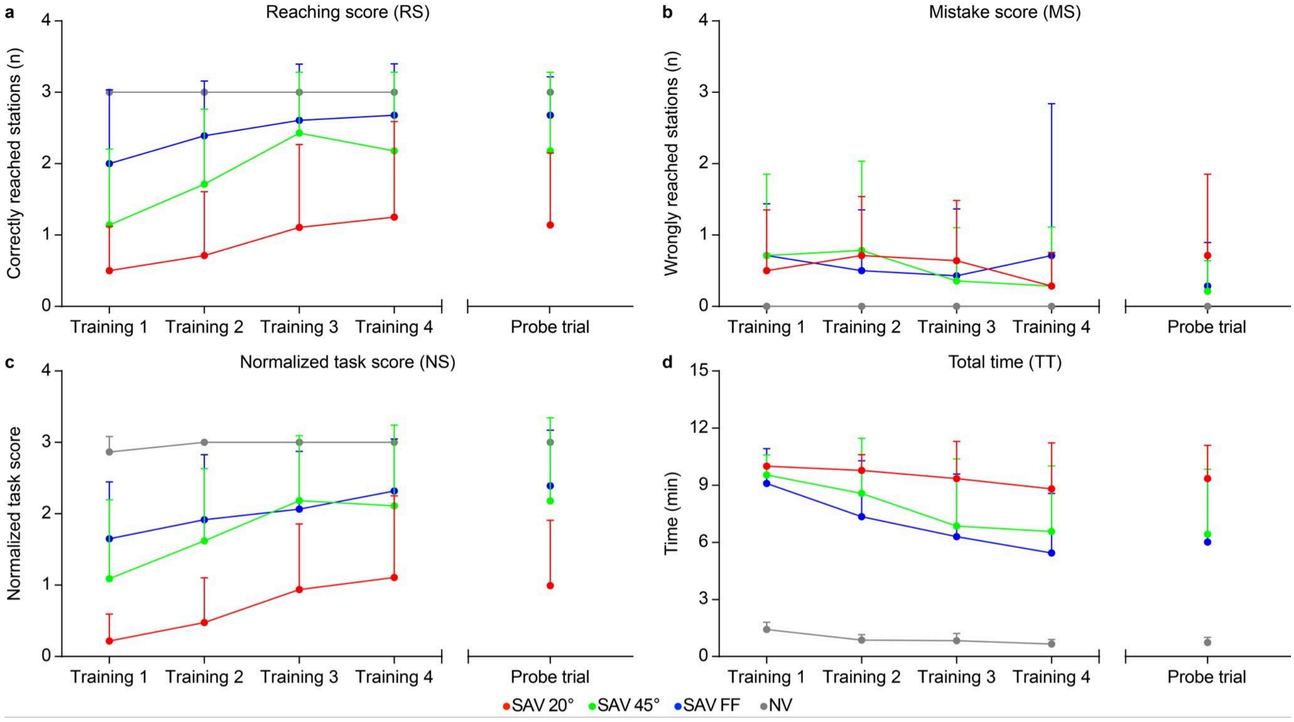
[56] Beyeler M, Rokem A, Boynton G M and Fine I 2017 Learning to see again: biological constraints on cortical plasticity and the implications for sight restoration technologies J Neural Eng 14 051003

[57] Authié C N, Berthoz A, Sahel J-A and Safran A B 2017 Adaptive Gaze Strategies for Locomotion with Constricted Visual Field *Front Hum Neurosci* **11** 387

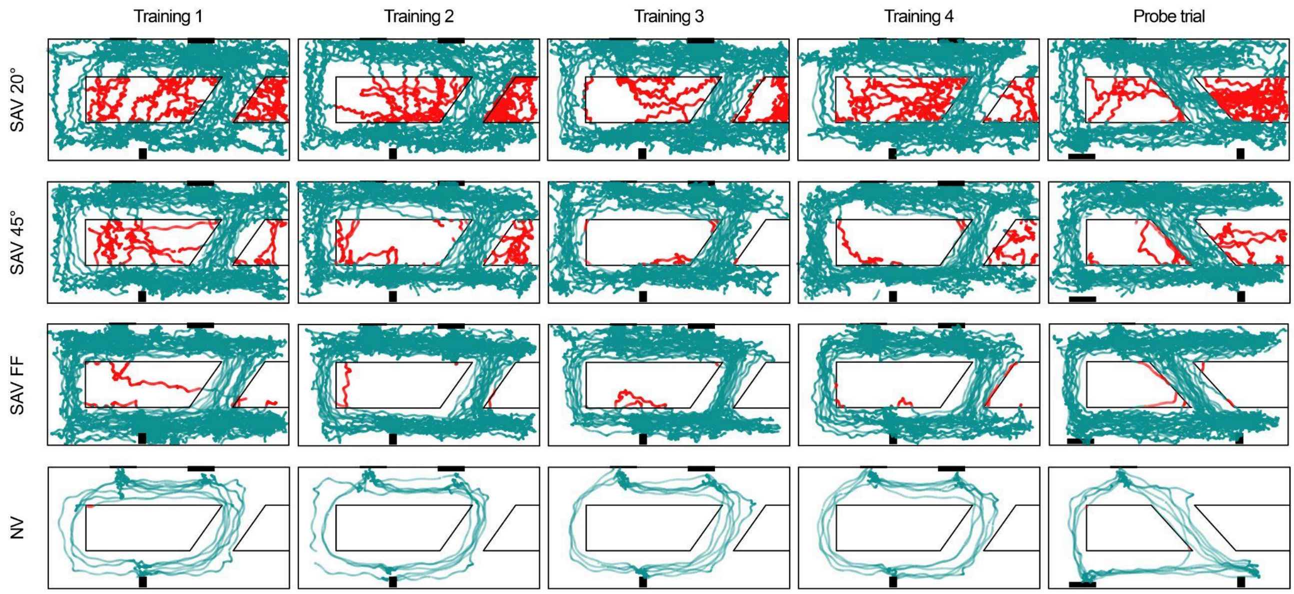
[58] Beyeler M, Rokem A, Boynton G and Fine I 2017 Modeling the perceptual experience of retinal prosthesis patients *J Vision* **17** 573

[59] Beyeler M, Nanduri D, Weiland J D, Rokem A, Boynton G M and Fine I 2019 A model of ganglion axon pathways accounts for percepts elicited by retinal implants *Sci Rep-uk* **9** 9199

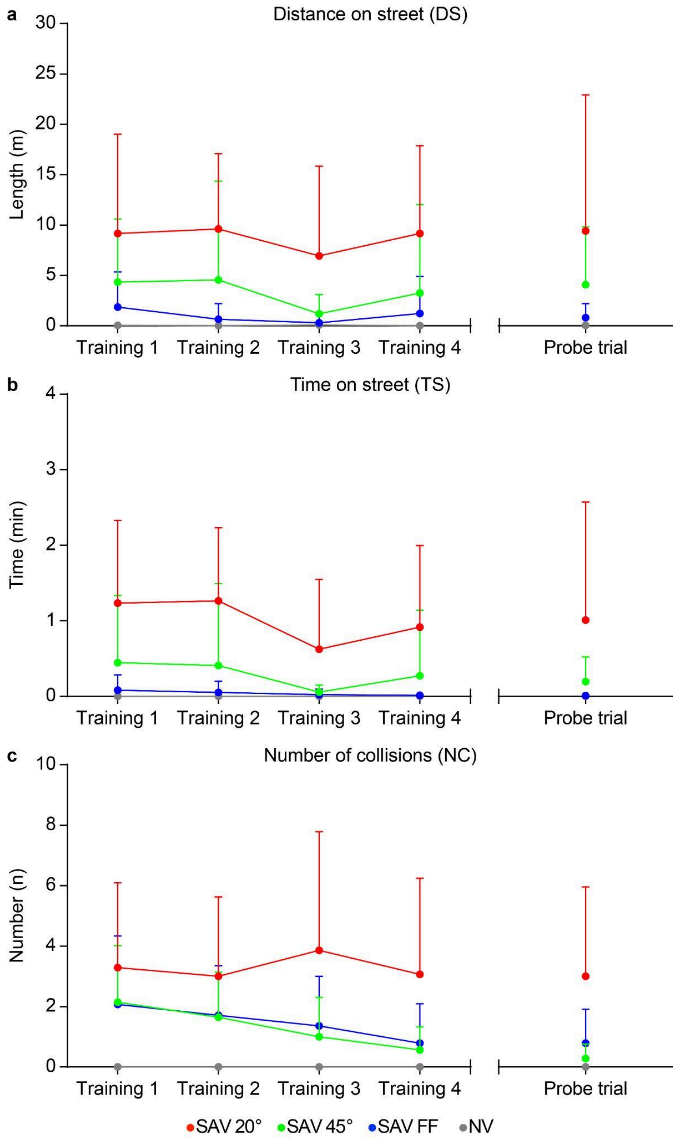
SUPPLEMENTARY FIGURES



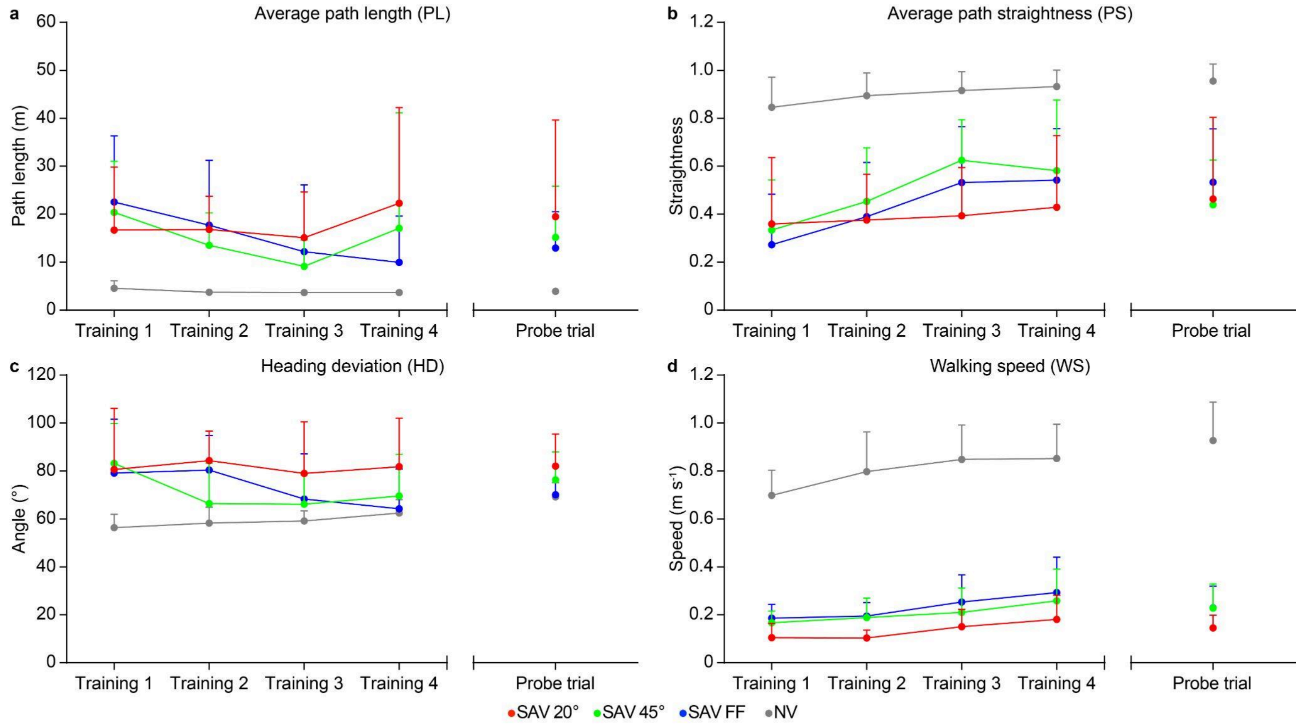
Supplementary Figure 1. Full quantification of task-solving efficacy as a function of the viewing condition. **a** Mean (± SD) reaching score (RS) during the training sessions and probe trial. **b** Mean (± SD) mistake score (MS) during the training sessions and probe trial. **c** Mean (± SD) normalized task score (NS) during the training sessions and probe trial. **d** Mean (± SD) total time (TT) during the training sessions and probe trial.



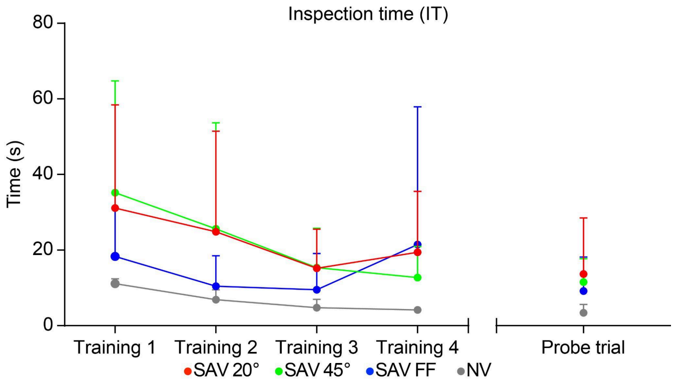
Supplementary Figure 2. Overlay of all participants' trajectories during training sessions and the probe trial for all viewing conditions. Trajectories are dark green when participants are on the sidewalk and crosswalk and red when they are on the street. Black rectangles indicate the stations (home, ATM, post box).



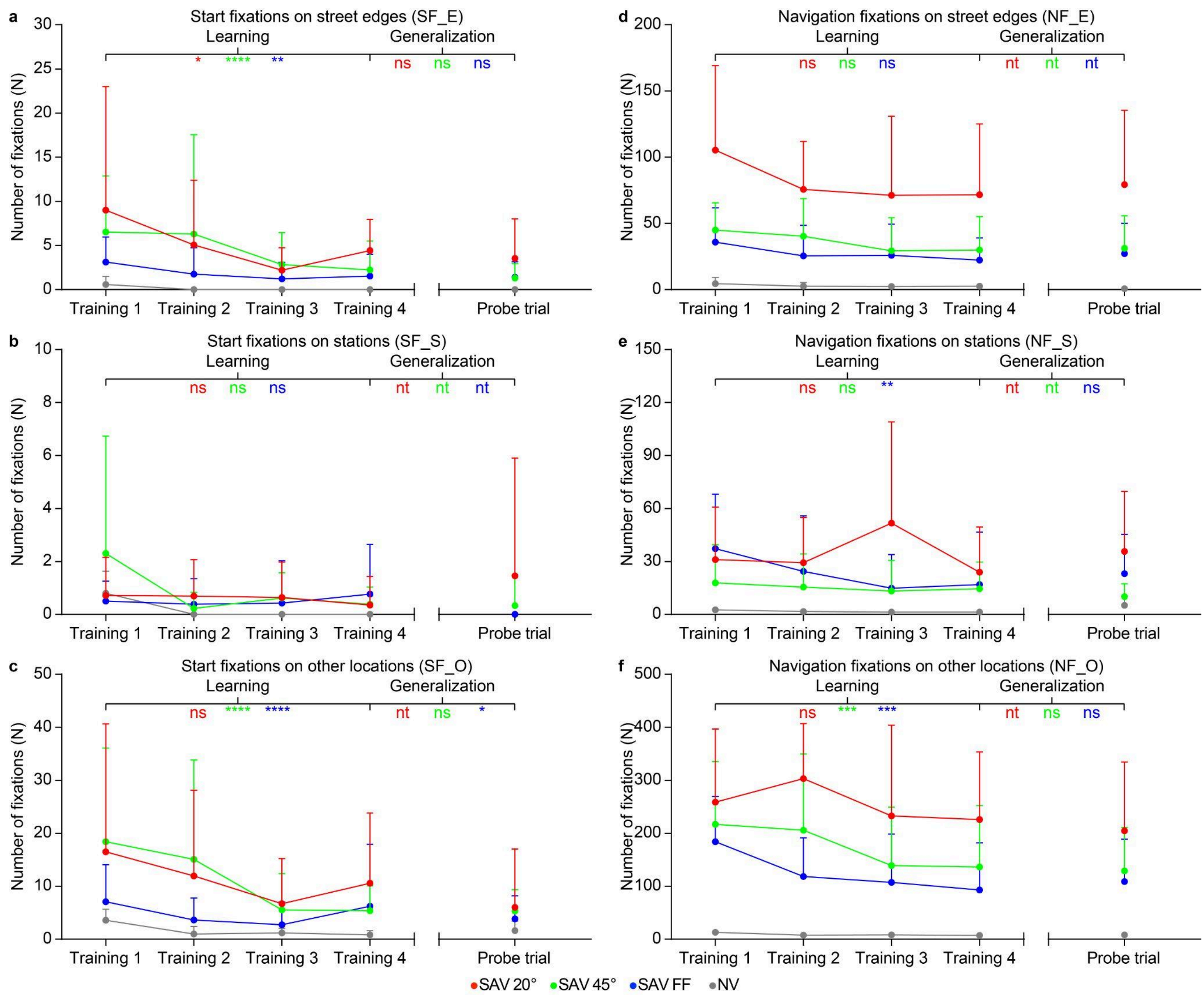
Supplementary Figure 3. Full quantification of navigation safety as a function of the viewing condition. **a** Mean (± SD) distance on street (DS) during the training sessions and probe trial. **b** Mean (± SD) time on street (TS) during the training sessions and probe trial. **c** Mean (± SD) number of collisions (NC) during the training sessions and probe trial.



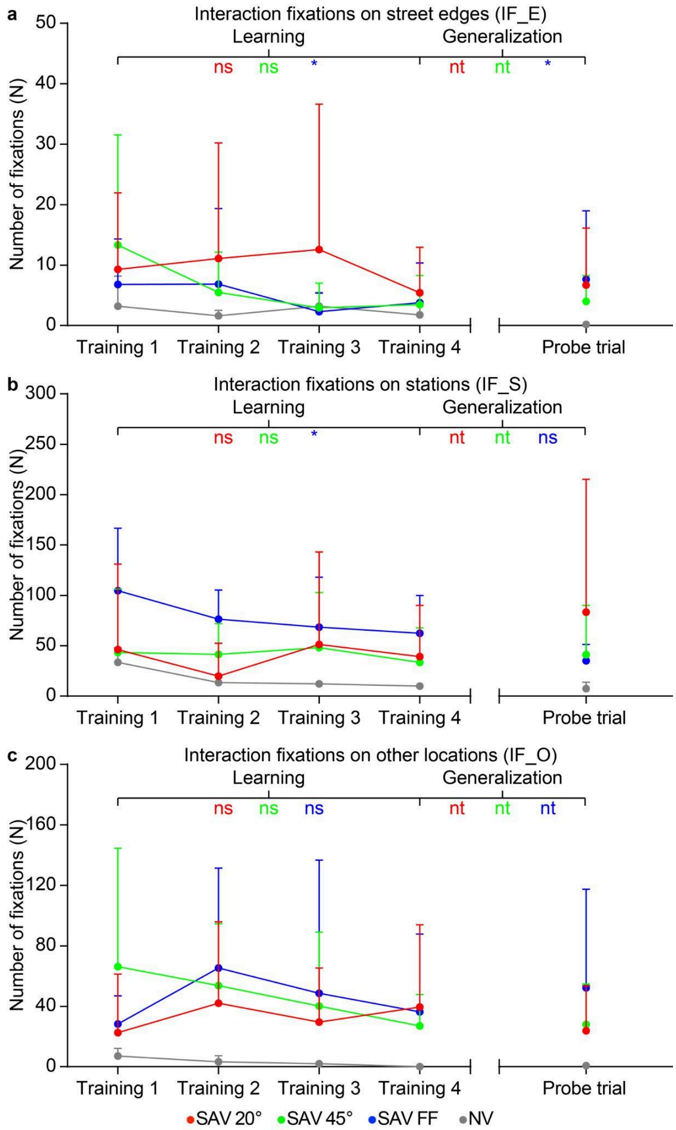
Supplementary Figure 4. Full quantification of visual navigation efficiency as a function of the viewing condition. **a** Mean (± SD) average path length (PL) during the training sessions and probe trial. **b** Mean (± SD) average path straightness (PS) during the training sessions and probe trial. **c** Mean (± SD) average heading deviation (HD) during the training sessions and probe trial. **d** Mean (± SD) average walking speed (WS) during the training sessions and probe trial.



Supplementary Figure 5. Full quantification of inspection time as a function of the viewing condition. Mean (± SD) average inspection time (IT) during the training sessions and probe trial.



Supplementary Figure 6. Quantification of fixational gaze behavior as a function of the viewing condition. **a** Mean (± SD) start fixations on street edges (SF_E) at the start of the training sessions and probe trial. **b** Mean (± SD) start fixations on stations (SF_S) at the start of the training sessions and probe trial. **c** Mean (± SD) start fixations on other locations (SF_O) at the start of the training sessions and probe trial. **d** Mean (± SD) navigation fixations on street edges (NF_E) during the training sessions and probe trial. **e** Mean (± SD) navigation fixations on stations (NF_S) during the training sessions and probe trial. **f** Mean (± SD) navigation fixations on other locations (NF_O) during the training sessions and probe trial. Results from two-tailed post hoc comparisons based on respective linear mixed effects models (nt: not tested; ns: $p > 0.05$; *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ****: $p < 0.0001$).



Supplementary Figure 7. Quantification of fixational gaze behavior as a function of the viewing condition. **a** Mean (± SD) interaction fixations on street edges (IF_E) at the start of the training sessions and probe trial. **b** Mean (± SD) interaction fixations on stations (IF_S) at the start of the training sessions and probe trial. **c** Mean (± SD) interaction fixations on other locations (IF_O) at the start of the training sessions and probe trial. Results from two-tailed post hoc comparisons based on respective linear mixed effects models (nt: not tested; ns: p > 0.05; *: p < 0.05).

SUPPLEMENTARY TABLES

Supplementary Table 1. Grading form for reaching score (RS) and normalized task score (NS).

Station		Task	Points
ATM			
Reaching score (RS)	The ATM was found	The ATM was not found	0
		The ATM was found by touch or because the participant bumped into it accidentally	0.5
		The ATM was found by visual search	1
Normalized task score (NS)	The ATM screen was touched	The ATM screen was not touched	0
		The participant touched a rectangular shape that was not the screen	0.5
		The ATM screen was identified visually and touched	1
	The ATM bill was taken	The bill was not taken	0
		The bill was taken by touched and not visual identification	0.5
		The bill was taken using only vision	1
	The value of the bill taken from the ATM was correctly read	The value of the bill was not read	0
		One of the two digits on the bill was correctly read	0.5
		The entire value of the bill was correctly read	1
Post box			
Reaching score (RS)	The post box was found	The post box was not found	0
		The post box was found by touch or because the participant bumped into it accidentally	0.5
		The post box was found by vision	1
Normalized task score (NS)	The country on the envelope was correctly read	The country on the envelope was not read	0
		The correct side of the envelope where the country was written was identified	0.5
		The country on the envelope was correctly read	1
	The envelope was correctly placed inside the mailbox	The envelope was not posted	0
		The envelope was placed inside the mailbox relying on touch	0.5
		The envelope was placed inside the mailbox using only vision	1
Home			
Reaching score (RS)	The home was found	The home was not found	0
		The participant went back to the street's quadrant in which the home door is located	0.5
		The home was found	1
Normalized task score (NS)	The handle of the door was touched	The handle of the door was not touched	0
		The handle of the door was found by touch	0.5
		The handle of the door was found by vision	1

Supplementary Table 2. Sketch Map Evaluation Grid

Evaluation Criteria	Score
Task-related landmarks (representation)	
The ATM is represented	0: no; 1: yes
The post box is represented	0: no; 1: yes
The home door is represented	0: no; 1: yes
The crosswalk is represented	0: no; 1: yes
Task-related landmarks (correct location)	
The ATM is in the correct location	0: no; 1: yes
The post box is in the correct location	0: no; 1: yes
The home door is in the correct location	0: no; 1: yes
Route segments and structure	
The sidewalk is represented as 3 segments	0: no; 1: yes
The 3 segments of the sidewalk are correctly placed (forming a “U” laying on its left side)	0: no; 1: yes
The segments of the sidewalks are approximately proportional in length with those of the actual street	0: no; 1: yes
The crosswalk is in the correct location (roughly in the center of the black part of the street)	0: no; 1: yes
The crosswalk is oriented in the right direction (top left to bottom right, see illustration)	0: no; 1: yes
The entire street is represented as a rectangle (with the home door edge being longer than the one of the fences)	0: no; 1: yes
Additional landmarks	
Is the other door represented	0: no; 1: yes
Are the fences are represented	0: no; 1: yes
Is the construction panel is represented	0: no; 1: yes
Is there any element represented that is not existing (such as the pole in the example sketch map)?	1: no; 0: yes
Orientation	
The sketch map of the total artificial street is drawn from the perspective of the home door	0: no; 1: yes

SUPPLEMENTARY NOTES

Changes in fixational gaze behavior over the training sessions and generalization of these changes to the probe trial

SF reveals a significant ‘repetition’ main effect ($\chi2$ = 273.01, p < 0.0001, df = 6; two-tailed ANOVA on generalized linear mixed effects model with Poisson distribution), a significant ‘viewing condition x repetition’ interaction effect ($\chi2$ = 63.16, p < 0.0001, df = 8, two-tailed ANOVA on generalized linear mixed effects model with Poisson distribution) as well as a significant ‘fixation location x repetition’ interaction effect ($\chi2$ = 17.82, p = 0.0226, df = 8, two-tailed ANOVA on generalized linear mixed effects model with Poisson distribution), indicating a change of fixation behavior over the course of trials. Only the three-way interaction effect ‘viewing condition x fixation location x repetition’ is not significant ($\chi2$ = 11.07, p = 0.8051, df = 16, two-tailed ANOVA on generalized linear mixed effects model with Poisson distribution). At the start of the trial, participants in all SAV viewing conditions perform fewer fixations on street edges in the last training session compared to the first one (**Supp. Fig. 6a**; SAV 20°: z = 2.48, p = 0.0261; SAV 45°: z = 4.60, p < 0.0001; SAV FF: z = 3.36, p = 0.0015; two-tailed post hoc z-tests). In all viewing conditions, this behavioral change generalizes to the probe trial (SAV 20°: z = -0.56, p = 0.5734; SAV 45°: z = 1.93, p = 0.0536; SAV FF: z = -1.24, p = 0.2153; two-tailed post hoc z-tests). The number of fixations on stations, in contrast, does not significantly change between the first and the last training session in all viewing conditions (**Supp. Fig. 6b**; SAV 20°: z = 0.00, p = 1; SAV 45°: z = 0.76, p = 0.8999; SAV FF: z = 0.00, p = 1; two-tailed post hoc z-tests). For the other fixations, only participants under SAV 45° and SAV FF change significantly between the first and the last training session (**Supp. Fig. 6c**; SAV 20°: z = 0.63, p = 0.5295; SAV 45°: z = 9.56, p < 0.0001; SAV FF: z = 4.51, p < 0.0001; two-tailed post hoc z-tests). Though only the change in the SAV 45° condition generalizes to the probe trial (SAV 45°: z = -0.11, p = 0.9102; SAV FF: z = -2.22, p = 0.0263; two-tailed post hoc z-tests).

NF also reveals a significant ‘repetition’ main effect ($\chi2$ = 63.31, p < 0.0001, df = 4, two-tailed ANOVA on generalized linear mixed effects model with negative binomial distribution), but no significant interaction effects (‘viewing condition x repetition’: $\chi2$ = 12.93, p = 0.1142, df = 8; ‘fixation location x repetition’: $\chi2$ = 6.98, p = 0.5385, df = 8; ‘viewing condition x fixation location x repetition’: $\chi2$ = 9.03, p = 0.9121, df = 16; two-tailed ANOVA on generalized linear mixed effects model with negative binomial distribution). The number of fixations on street edges don’t change significantly between the first and the last training session in all SAV viewing conditions (**Supp. Fig. 6d**; SAV 20°: z = 1.31, p = 0.3813; SAV 45°: z = 1.64, p = 0.2031; SAV FF: z = 1.16, p = 0.4939; two-tailed post hoc z-tests). The number of fixations on task stations only significantly decreases under SAV FF (**Supp. Fig. 6e**; SAV 20°: z = 1.03, p = 0.3047; SAV 45°: z = 0.16, p = 1; SAV FF: z = 2.99, p = 0.0055; two-tailed post hoc z-tests), with this SAV FF behavioral change generalizing to the probe trial (z = -1.85, p = 0.0640, two-tailed post hoc z-test). Only participants under SAV 45° and SAV FF show a significant decrease of remaining fixations between the first and the last training session during navigation (**Supp. Fig. 6f**; SAV 20°: z = 0.90, p = 0.6849; SAV 45°: z = 3.96, p = 0.0002; SAV FF: z = 4.00, p = 0.0001; two-tailed post hoc z-tests). Both changes generalize to the probe trial (SAV 45°: z = 0.32, p = 0.7511; SAV FF: z = -0.98, p = 0.3256; two-tailed post hoc z-tests).

IF reveals a significant 'repetition' main effect ($\chi 2$ = 23.08, p = 0.0001, df = 4, two-tailed ANOVA on generalized linear mixed effects model with negative binomial distribution) and a significant 'fixation location x repetition' interaction ($\chi 2$ = 20.99, p = 0.0072, df = 8, two-tailed ANOVA on generalized linear mixed effects model with negative binomial distribution). The remaining interactions are not significant ('viewing condition x repetition': $\chi 2$ = 10.71, p = 0.2189, df = 8; 'viewing condition x fixation location x repetition': $\chi 2$ = 18.73, p = 0.2832, df = 16; two-tailed ANOVA on generalized linear mixed effects model with negative binomial distribution). A significant change in gaze behavior over the sessions during station interaction is observed under SAV FF for fixations on street edges (**Supp. Fig. 7a**; z = 2.64, p = 0.0166, two-tailed post hoc z-test) and stations (**Supp. Fig. 7b**; z = 2.28, p = 0.0451, two-tailed post hoc z-test). However, only the latter change generalizes to the probe trial (street edges: z = -2.50, p = 0.0166; stations: z = 1.28, p = 0.2013, two-tailed post hoc z-tests). No further change behaviors are observed for fixations on street edges (**Supp. Fig. 7a**; SAV 20°: z = 1.91, p = 0.1136; SAV 45°: z = 2.12, p = 0.0680; two-tailed post hoc z-tests), stations (**Supp. Fig. 7b**; SAV 20°: z = -0.29, p = 0.7734; SAV 45°: z = -0.53, p = 1; two-tailed post hoc z-tests), and other locations (**Supp. Fig. 7c**; SAV 20°: z = -0.01, p = 1; SAV 45°: z = 1.99, p = 0.0928; SAV FF: z = 2.18, p = 0.0557; two-tailed post hoc z-tests).

Instructions before the familiarization phase

- La vision se base sur la détection des bords : est-ce que tu arrives à détecter des bords et identifier la géométrie de la table ? (à faire pointer du doigt)
- N'hésite pas à bouger légèrement la tête pour continuer à voir l'image et ne pas rester trop dans le noir. Essaye de ne pas rester statique et de faire des mouvements !
- N'hésite pas à bien t'avancer et aussi prendre du recul. Quand tu prends recul, tu pourras voir la forme générale et les bords, ce qui te permettra de détecter les limites et voir les éléments dans leur entièreté. Et en avançant, tu vas pouvoir percevoir les détails de ces éléments plus précisément. N'hésite pas à utiliser cette stratégie quand tu navigues ou quand tu regardes des objets pour mieux comprendre ce que tu vois.
- Si jamais tu te sens complètement perdu pendant la navigation, n'hésite pas à retrouver tes mains dans ton champ de vision, ça peut t'aider à mieux gérer ta vision.

Instructions given to the participants during the study before the start of each trial

- Rappelle-toi, tu es pressé parce que tu as un rendez-vous urgent qui t'attend.
- Tu dois donc faire les deux tâches dans un ordre optimal qui te permet de gagner le plus de temps possible. Il s'agit d'une rue artificielle, donc ne fais pas attention au plafond, il ne te donnera aucune indication pour réaliser les tâches, si ce n'est te retarder pour ton rendez-vous.
- Pour la tâche de la boîte aux lettres, tu dois d'abord trouver la boîte aux lettres, lire le pays indiqué sur la carte que tu veux envoyer, puis mettre la lettre dans la boîte à travers la fente. Tu dois placer la lettre dans la fente en utilisant seulement ta vision, et ne pas la toucher.
- Pour la tâche du distributeur, tu dois le trouver dans la pièce, puis identifier l'écran qui contient un symbole, le toucher. Tu pourras alors prendre le billet, identifier sa valeur puis le placer dans ta banane.
- Pour ces tâches, nous sommes intéressés par ta recherche visuelle. Il faut donc que tu ne touches que quand tu es sûr de reconnaître l'objet d'intérêt.
- Une fois que tu as réalisé ces deux tâches, tu dois rentrer chez toi rapidement et toucher la poignée de chez toi pour terminer cet essai. Entre chaque essai, la configuration peut ou pas avoir été modifiée.
- Tu dois respecter le code de la route et ne pas marcher sur la route sauf s'il y a des moyens mis en place pour traverser. Un côté de la route est limité par des petites barrières pour ne pas que tu ailles plus loin.
- N'hésite pas à utiliser les bords du trottoir comme point de repère pour naviguer, et aussi de ne pas passer trop de temps dans un seul endroit si tu n'es pas sûre, le temps presse !
- Nous serons présents dans la salle, mais nous ne communiquerons pas avec vous lors des tâches. Bon courage !

Instructions for the creation of sketch maps of the artificial street

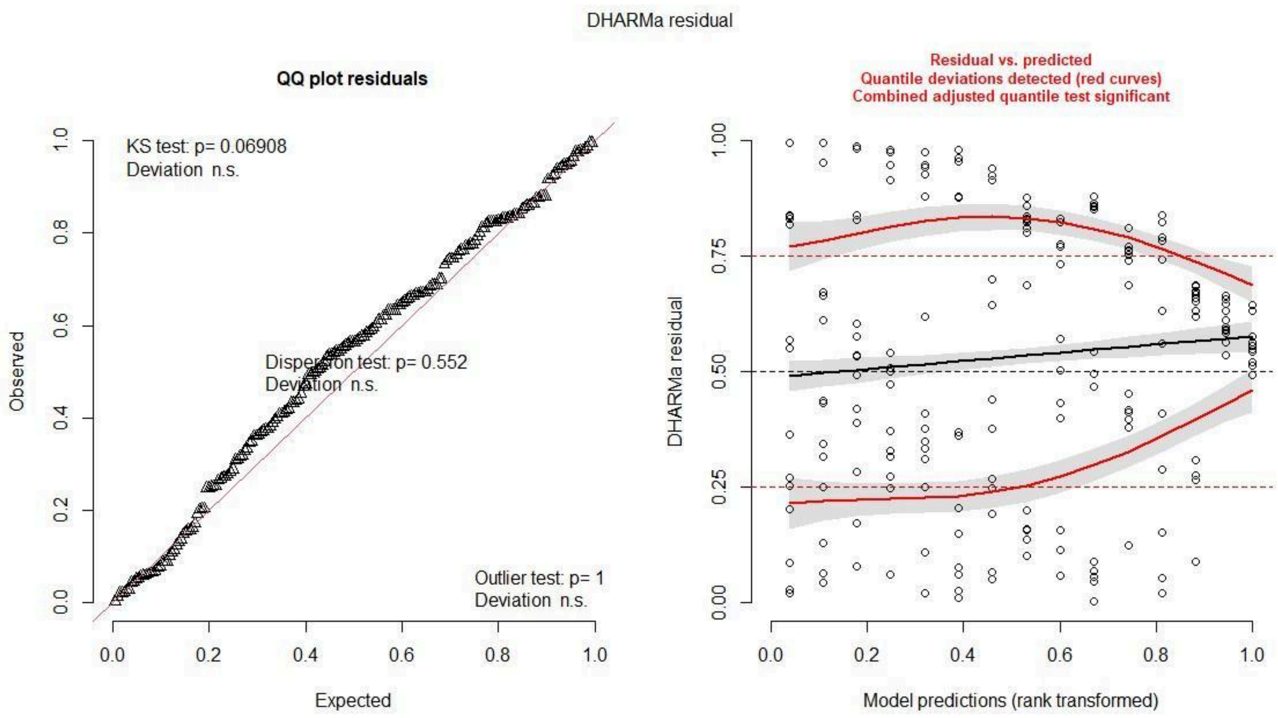
- Merci de dessiner une carte du dernier arrangement de la rue dans laquelle tu as navigué avec le plus de précision possible.
- Il faut y inclure le plus grand nombre possible de caractéristiques environnementales et topographiques.
- Les capacités artistiques ne sont pas importantes, merci d'illustrer la carte au mieux de tes capacités.

Model specifics, assumptions and statistics

- Reaching score (RS)

Best fitting model: Generalized linear mixed effects model with tweedie distribution and zero-inflation and dispersion correction. RS as dependent variable, viewing condition, trial and their interaction as fixed effects and, since the design is a repeated measures design, participant as random effect.

Model assumptions: assumption violations are highlighted in red.



Statistical results:

ANOVA(m): Analysis of Deviance Table (Type II Wald χ2-tests)			
	χ2	df	Pr(>χ2)
Viewing Condition	23.6131	2	7.46E-06
Repetition	53.1226	4	8.03E-11
Viewing Condition x Repetition	6.9217	8	0.5451

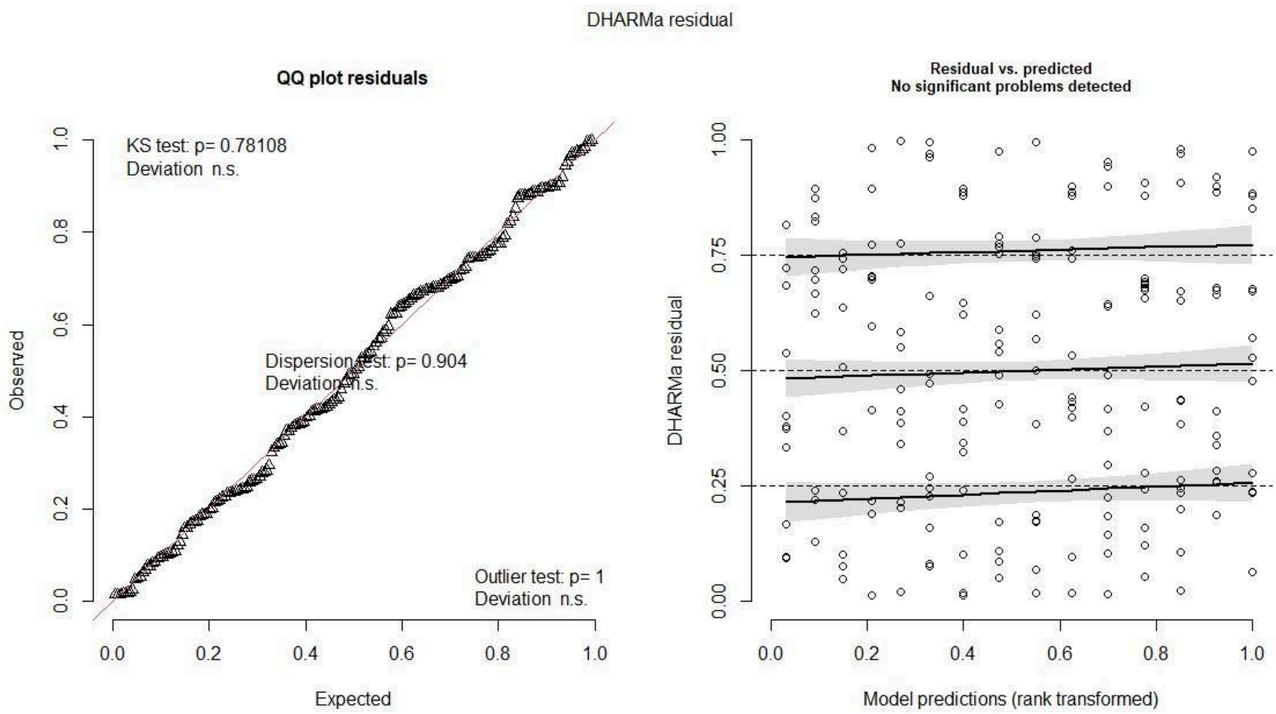
Post hoc pairwise comparisons for probe trial					
contrast	estimate	SE	df	z-ratio	p-value
SAV 20° vs SAV 45°	-0.474	0.182	Inf	-2.599	0.0187
SAV 20° vs SAV FF	-0.641	0.178	Inf	-3.594	0.001
SAV 45° vs SAV FF	-0.168	0.138	Inf	-1.217	0.2234
Results are given on the log (not the response) scale					
p value adjustment: holm method for 3 tests					

Post hoc pairwise comparisons for learning and generalization					
Viewing Condition = SAV 20°					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	-0.8116	0.2439	Inf	-3.328	0.0017
train4 vs probe	0.1879	0.1794	Inf	1.047	0.295
Viewing Condition = SAV 45°					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	-0.5196	0.1176	Inf	-4.417	<.0001
train4_v_probe	0.0504	0.0983	Inf	0.513	0.6082
Viewing Condition = SAV FF					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	-0.3117	0.0917	Inf	-3.399	0.0014
train4 vs probe	0.0711	0.0857	Inf	0.83	0.4068
Results are given on the log (not the response) scale					
p value adjustment: holm method for 2 tests					

- Mistake score (MS)

Best fitting model: Generalized linear mixed effects model with tweedie distribution and dispersion correction. MS as dependent variable, viewing condition, trial and their interaction as fixed effects and, since the design is a repeated measures design, participant as random effect.

Model assumptions:



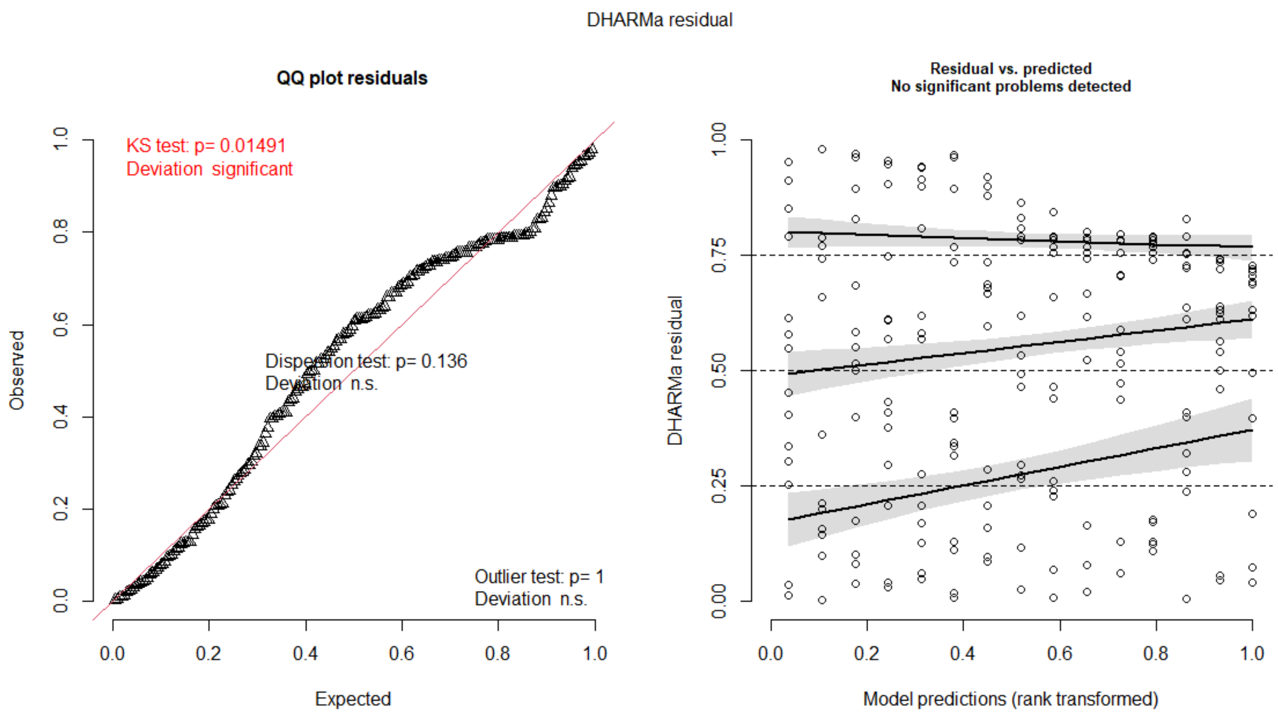
Statistical results:

ANOVA(m): Analysis of Deviance Table (Type II Wald χ2-tests)			
	χ2	df	Pr(>χ2)
Viewing Condition	0.2529	2	0.8812
Repetition	2.5529	4	0.6352
Viewing Condition x Repetition	1.065	8	0.9978

- Normalized task score (NS)

Best fitting model: Generalized linear mixed effects model with tweedie distribution and zero-inflation and dispersion correction. NS as dependent variable, viewing condition, trial and their interaction as fixed effects and, since the design is a repeated measures design, participant as random effect.

Model assumptions: assumption violations are highlighted in red.



Statistical results:

ANOVA(m): Analysis of Deviance Table (Type II Wald χ2-tests)			
	χ2	df	Pr(>χ2)
Viewing Condition	15.552	2	0.0004198
Repetition	40.578	4	3.29E-08
Viewing Condition x Repetition	12.475	8	0.1312181

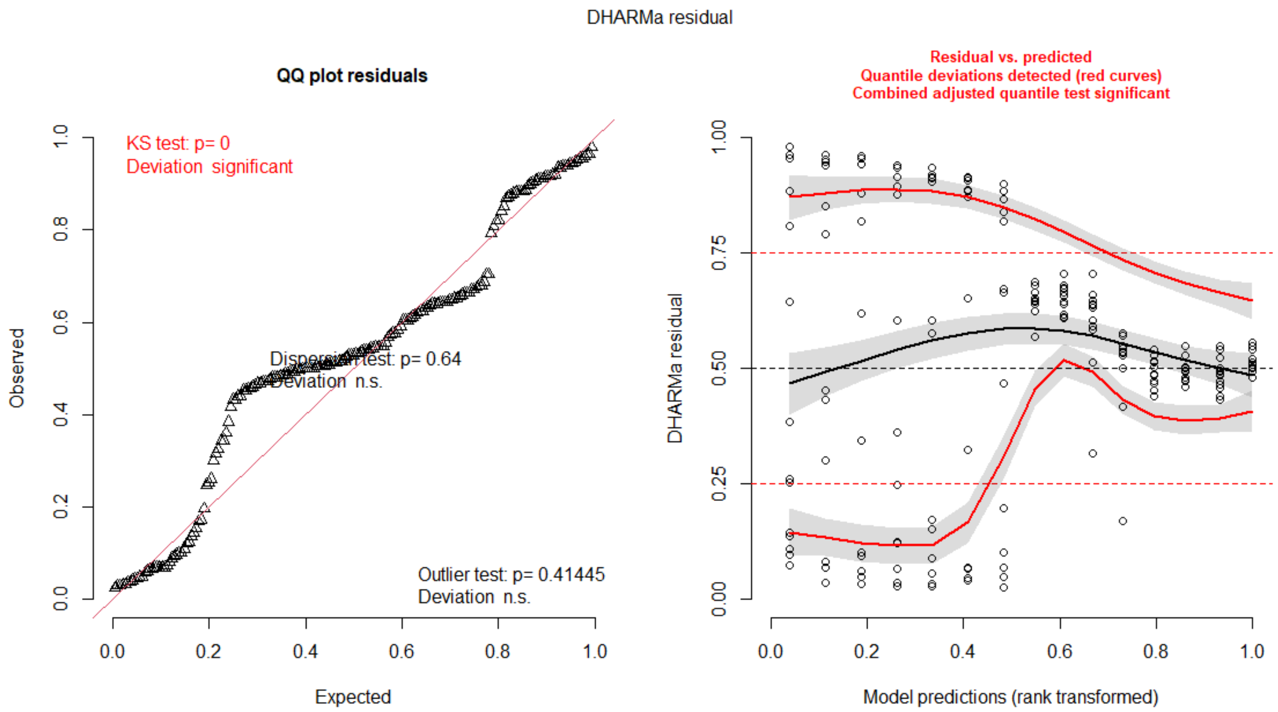
Post hoc pairwise comparisons for probe trial					
contrast	estimate	SE	df	z-ratio	p-value
SAV 20° vs SAV 45°	-0.6952	0.234	Inf	-2.975	0.0059
SAV 20° vs SAV FF	-0.7524	0.23	Inf	-3.265	0.0033
SAV 45° vs SAV FF	-0.0572	0.2	Inf	-0.286	0.7752
Results are given on the log (not the response) scale					
p value adjustment: holm method for 3 tests					

Post hoc pairwise comparisons for learning and generalization					
Viewing Condition = SAV 20°					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	-1.6156	0.5429	Inf	-2.976	0.0058
train4 vs probe	0.1124	0.1958	Inf	0.574	0.5658
Viewing Condition = SAV 45°					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	-0.6311	0.1935	Inf	-3.261	0.0022
train4 vs probe	-0.0344	0.0982	Inf	-0.35	0.7262
Viewing Condition = SAV FF					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	-0.3424	0.1196	Inf	-2.862	0.0084
train4 vs probe	-0.0303	0.0705	Inf	-0.43	0.6672
Results are given on the log (not the response) scale					
p value adjustment: holm method for 2 tests					

- Total time (TT)

Best fitting model: Linear mixed effects model. TT as dependent variable, viewing condition, trial and their interaction as fixed effects and, since the design is a repeated measures design, participant as random effect.

Model assumptions: assumption violations are highlighted in red.



Statistical results:

ANOVA(m): Analysis of Deviance Table (Type II Wald χ2-tests)			
	χ2	df	Pr(>χ2)
Viewing Condition	15.191	2	5.03E-04
Repetition	69.187	4	3.37E-14
Viewing Condition x Repetition	37.924	8	7.78E-06

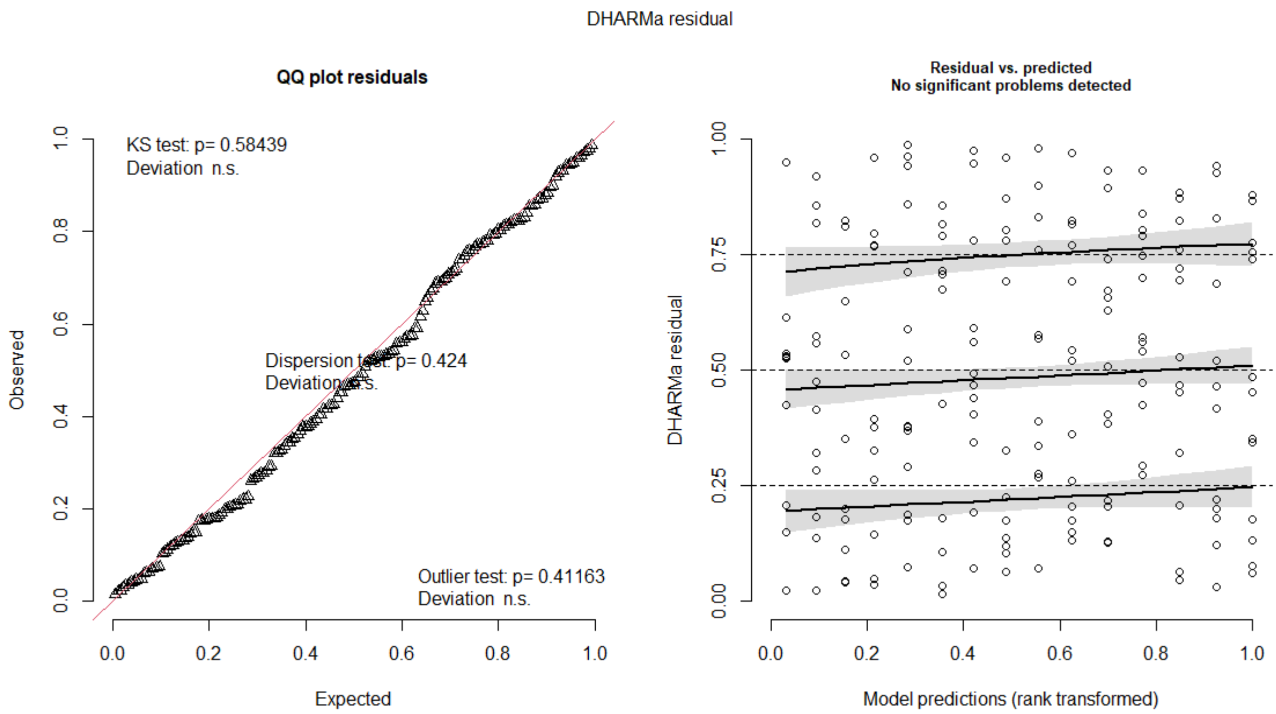
Post hoc pairwise comparisons for probe trial					
contrast	estimate	SE	df	t-ratio	p-value
SAV 20° vs SAV 45°	3.533	0.938	68.2	3.768	0.0007
SAV 20° vs SAV FF	3.95	0.938	68.2	4.212	0.0002
SAV 45° vs SAV FF	0.417	0.921	64.4	0.453	0.6523
Degrees-of-freedom method: kenward-roger					
p value adjustment: holm method for 3 tests					

Post hoc pairwise comparisons for learning and generalization					
Viewing Condition = SAV 20°					
contrast	estimate	SE	df	t-ratio	p-value
train1 vs train4	0.302	0.535	144	0.565	1
Viewing Condition = SAV 45°					
contrast	estimate	SE	df	t-ratio	p-value
train1 vs train4	2.481	0.548	142	4.528	<.0001
train4 vs probe	0.143	0.505	141	0.283	0.7778
Viewing Condition = SAV FF					
contrast	estimate	SE	df	t-ratio	p-value
train1 vs train4	3.773	0.53	142	7.112	<.0001
train4 vs probe	-0.567	0.505	141	-1.123	0.2633
Degrees-of-freedom method: kenward-roger					
p value adjustment: holm method for 2 tests					

- Distance on street (DS)

Best fitting model: Generalized linear mixed effects model with tweedie distribution and zero-inflation and dispersion correction. DS as dependent variable, viewing condition, trial and their interaction as fixed effects and, since the design is a repeated measures design, participant as random effect.

Model assumptions:



Statistical results:

ANOVA(m): Analysis of Deviance Table (Type II Wald χ2-tests)			
	χ2	df	Pr(>χ2)
Viewing Condition	92.185	2	< 2.2e-16
Repetition	3500.188	4	<2.2e-16
Viewing Condition x Repetition	88.764	8	8.29E-16

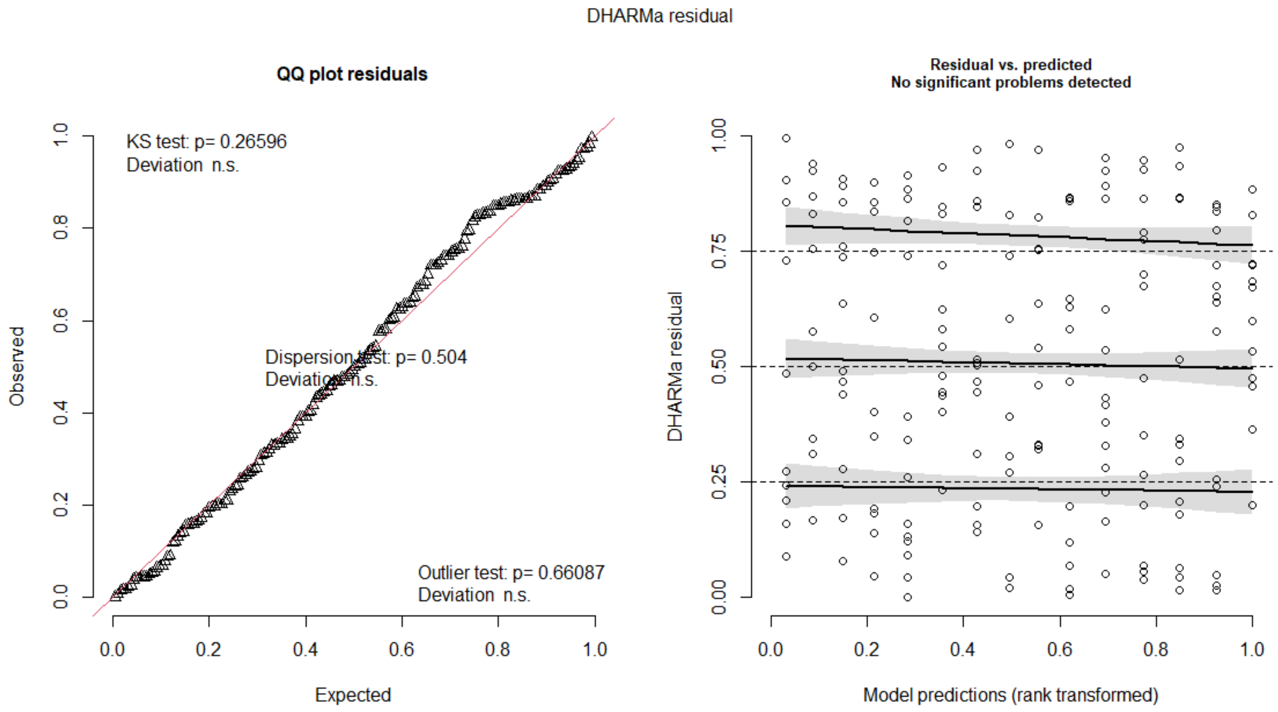
Post hoc pairwise comparisons for probe trial					
contrast	estimate	SE	df	z-ratio	p-value
SAV 20° vs SAV 45°	1.214	0.54	Inf	2.26	0.0477
SAV 20° vs SAV FF	2.147	0.44	Inf	4.825	<.0001
SAV 45° vs SAV FF	0.933	0.48	Inf	1.936	0.0529
Results are given on the log (not the response) scale					
p value adjustment: holm method for 3 tests					

Post hoc pairwise comparisons for learning and generalization					
Viewing Condition = SAV 20°					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	0.0527	0.3358	Inf	0.157	1
Viewing Condition = SAV 45°					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	2.239	0.6624	Inf	3.38	0.0014
train4 vs probe	-1.9853	0.6332	Inf	-3.135	0.0017
Viewing Condition = SAV FF					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	2.0198	0.0811	Inf	24.907	<.0001
train4 vs probe	-2.7966	0.0913	Inf	-30.634	<.0001
Results are given on the log (not the response) scale					
p value adjustment: holm method for 2 tests					

- Time on street (TS)

Best fitting model: Generalized linear mixed effects model with tweedie distribution dispersion correction. TS was taken as the dependent variable, viewing condition, trial and their interaction as the fixed effects and, since the design is a repeated measures design, participant as random effect.

Model assumptions:



Statistical results:

ANOVA(m): Analysis of Deviance Table (Type II Wald χ2-tests)			
	χ2	df	Pr(>χ2)
Viewing Condition	109.203	2	< 2.2e-16
Repetition	20.363	4	0.0004234
Viewing Condition x Repetition	15.325	8	0.0531312

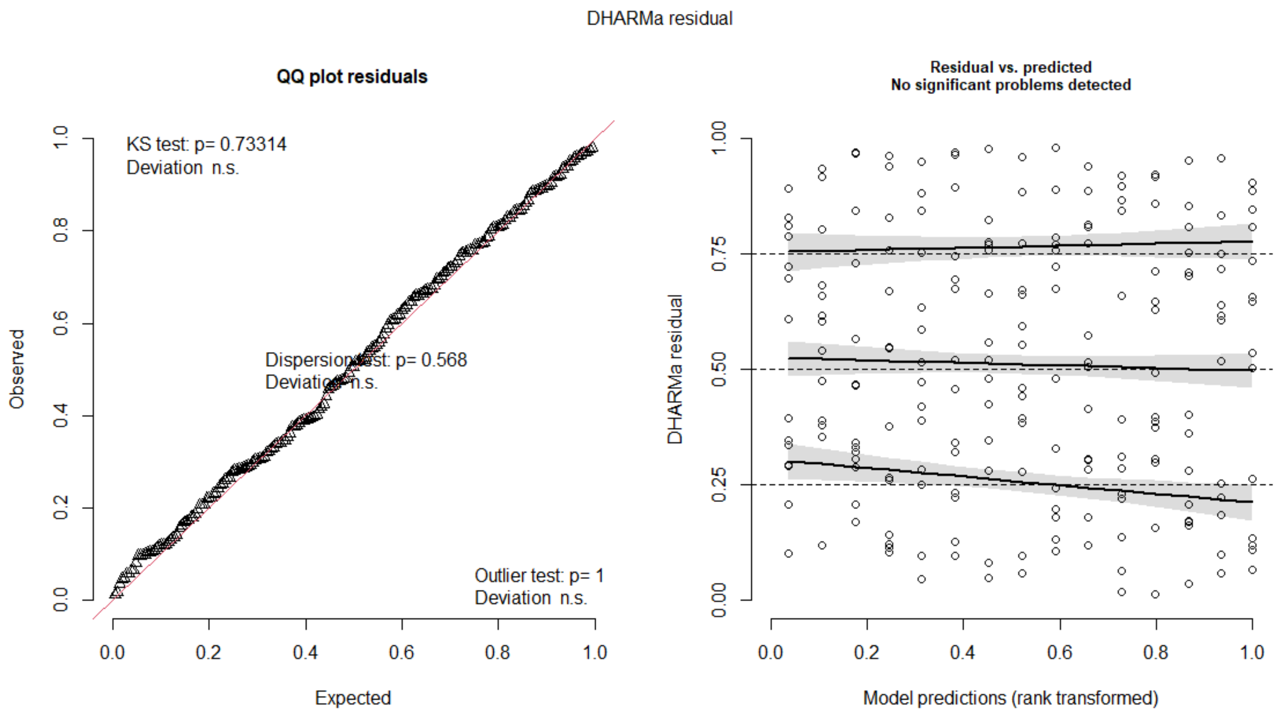
Post hoc pairwise comparisons for probe trial					
contrast	estimate	SE	df	z-ratio	p-value
SAV 20° vs SAV 45°	2.11	5.92E-01	Inf	3.558	0.0007
SAV 20° vs SAV FF	4.24	6.72E-01	Inf	6.305	<.0001
SAV 45° vs SAV FF	2.13	7.11E-01	Inf	2.998	0.0027
Results are given on the log (not the response) scale					
p value adjustment: holm method for 3 tests					

Post hoc pairwise comparisons for learning and generalization					
Viewing Condition = SAV 20°					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	0.4572	0.379	Inf	1.208	0.4544
Viewing Condition = SAV 45°					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	1.8849	0.547	Inf	3.444	0.0011
train4 vs probe	-0.9932	0.597	Inf	-1.662	0.0965
Viewing Condition = SAV FF					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	2.234	1.43	Inf	1.562	0.1182
Results are given on the log (not the response) scale					
p value adjustment: holm method for 2 tests					

- Number of collisions (NC)

Best fitting model: Generalized linear mixed effects model with poisson distribution. NC as dependent variable, viewing condition, trial and their interaction as fixed effects and, since the design is a repeated measures design, participant as random effect.

Model assumptions:



Statistical results:

ANOVA(m): Analysis of Deviance Table (Type II Wald χ2-tests)			
	χ2	df	Pr(>χ2)
Viewing Condition	9.8624	2	0.007218
Repetition	15.2349	4	0.004238
Viewing Condition x Repetition	21.8433	8	0.005215

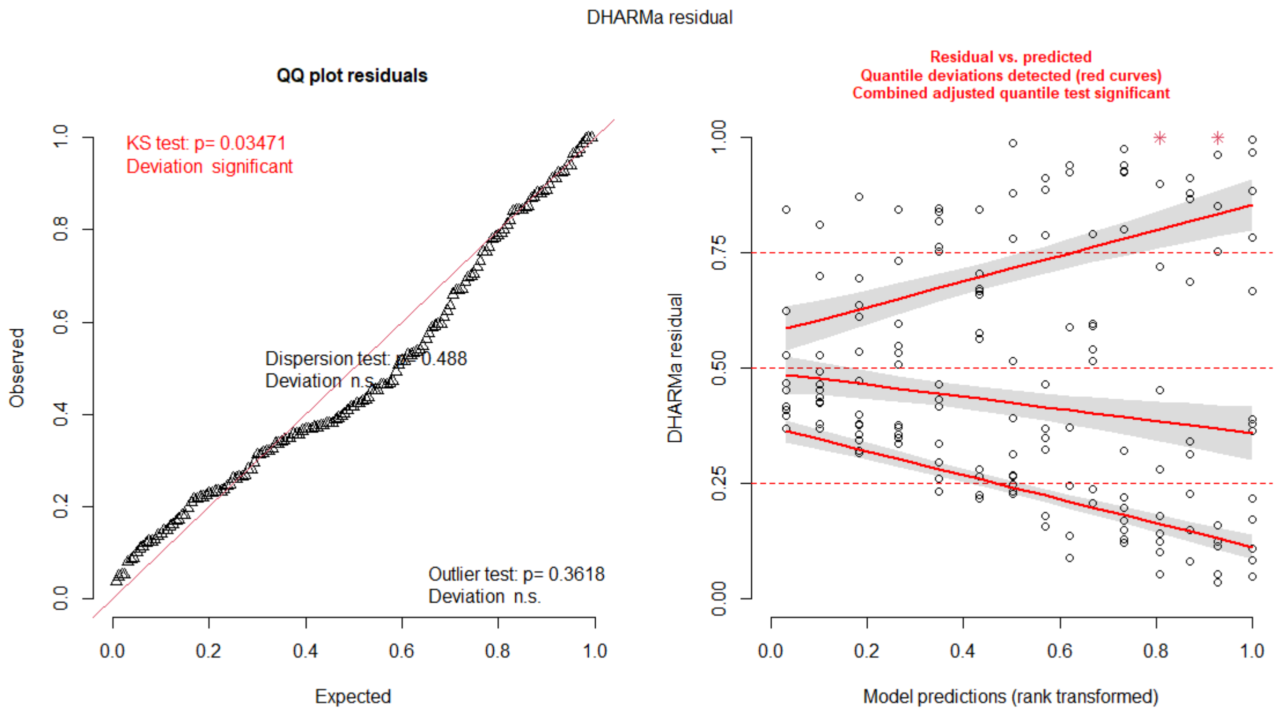
Post hoc pairwise comparisons for probe trial					
contrast	estimate	SE	df	z-ratio	p-value
SAV 20° vs SAV 45°	2.3576	0.612	Inf	3.852	0.0004
SAV 20° vs SAV FF	1.4035	0.465	Inf	3.018	0.0051
SAV 45° vs SAV FF	-0.954	0.667	Inf	-1.431	0.1524
Results are given on the log (not the response) scale					
p value adjustment: holm method for 3 tests					

Post hoc pairwise comparisons for learning and generalization					
Viewing Condition = SAV 20°					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	6.74E-02	0.212	Inf	0.318	1
Viewing Condition = SAV 45°					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	1.32E+00	0.398	Inf	3.322	0.0018
train4 vs probe	6.93E-01	0.612	Inf	1.132	0.2577
Viewing Condition = SAV FF					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	9.69E-01	0.354	Inf	2.738	0.0124
train4 vs probe	-1.17E-05	0.426	Inf	0	1
Results are given on the log (not the response) scale					
p value adjustment: holm method for 2 tests					

- Average path length (PL)

Best fitting model: Linear mixed effects model. PL as dependent variable, viewing condition, trial and their interaction as fixed effects and, since the design is a repeated measures design, participant as random effect.

Model assumptions: assumption violations are highlighted in red



Statistical results:

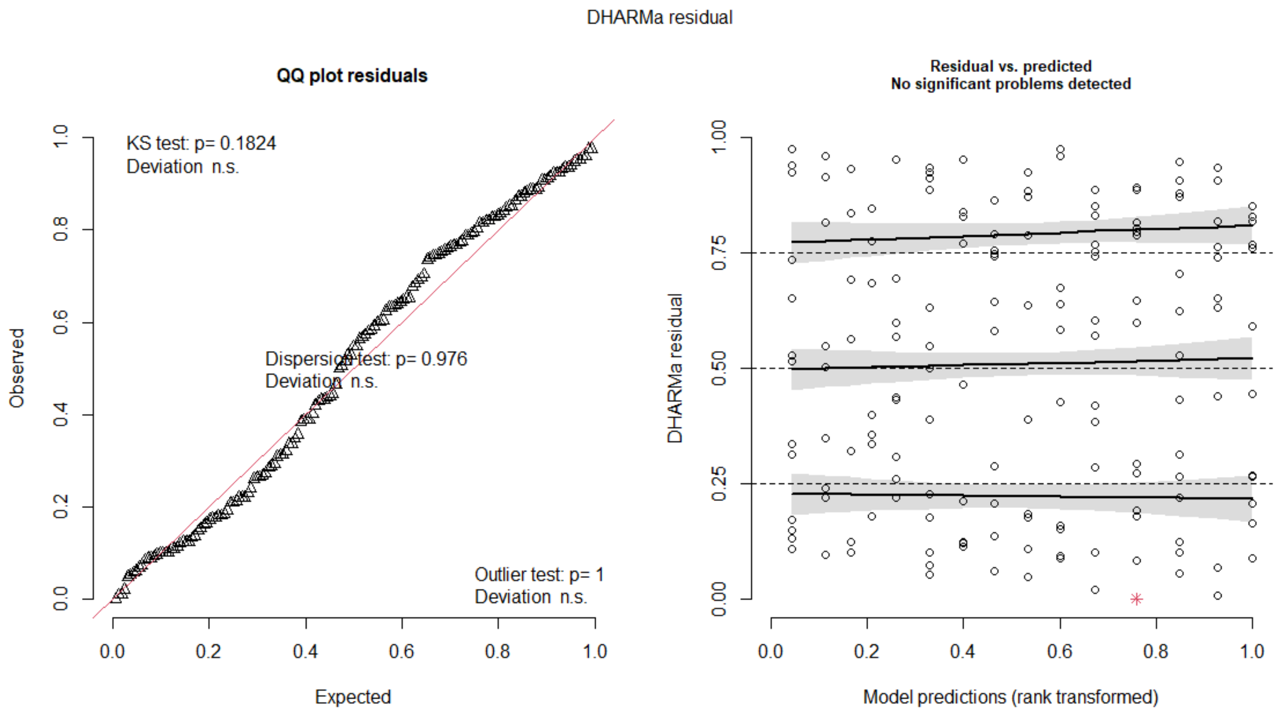
ANOVA(m): Analysis of Deviance Table (Type II Wald χ2-tests)			
	χ2	df	Pr(>χ2)
Viewing Condition	4.6316	2	0.098688
Repetition	17.4727	4	0.001564
Viewing Condition x Repetition	12.0613	8	0.148489

Post hoc pairwise comparisons for learning and generalization					
Viewing Condition = SAV 20°					
contrast	estimate	SE	df	t-ratio	p-value
train1 vs train4	-5.31	5.31	119	-1	0.6391
Viewing Condition = SAV 45°					
contrast	estimate	SE	df	t-ratio	p-value
train1 vs train4	13.14	4.93	114	2.665	0.0176
train4 vs probe	-7.76	4.73	119	-1.641	0.1034
Viewing Condition = SAV FF					
contrast	estimate	SE	df	t-ratio	p-value
train1 vs train4	14.9	4.19	114	3.558	0.0011
train4 vs probe	-5	4.11	113	-1.217	0.226
Degrees-of-freedom method: kenward-roger					
p value adjustment: holm method for 2 tests					

- Average path straightness (PS)

Best fitting model: Generalized linear mixed effects model with dispersion correction. PS as dependent variable, viewing condition, trial and their interaction as fixed effects and, since the design is a repeated measures design, participant as random effect.

Model assumptions:



Statistical results:

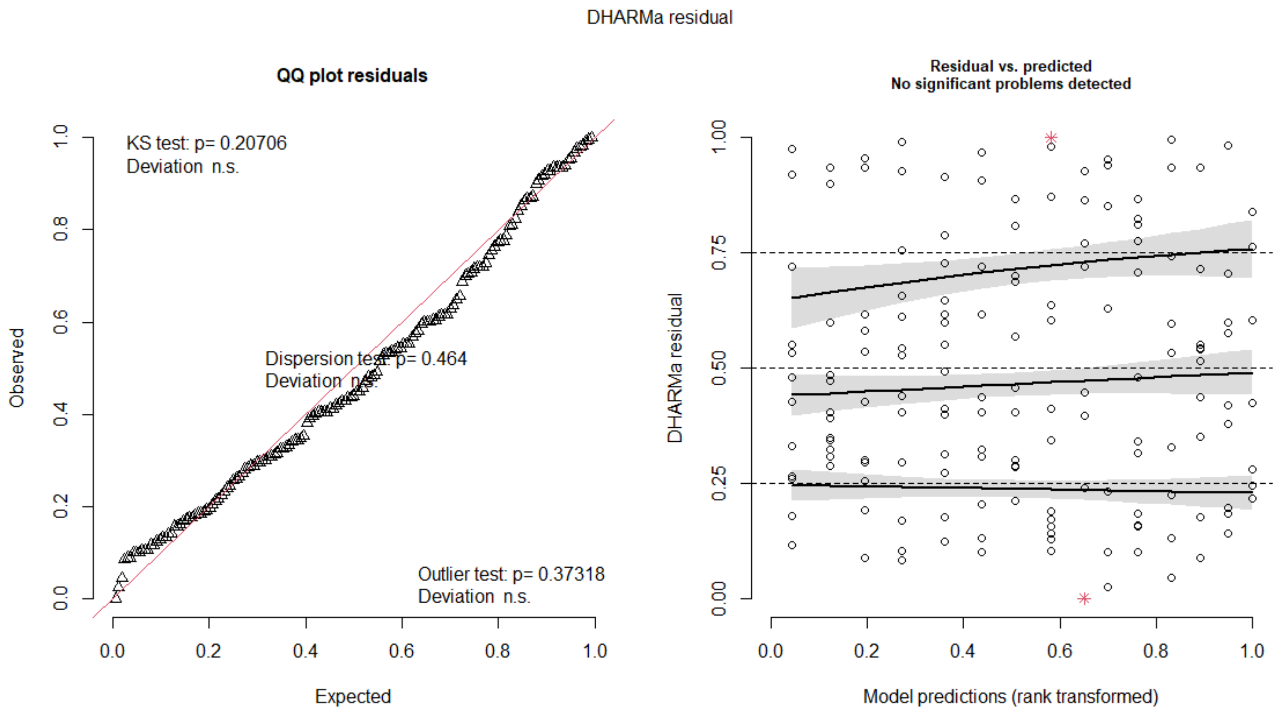
ANOVA(m): Analysis of Deviance Table (Type II Wald χ2-tests)			
	χ2	df	Pr(>χ2)
Viewing Condition	1.7758	2	0.4115
Repetition	41.1326	4	2.52E-08
Viewing Condition x Repetition	9.221	8	0.324

Post hoc pairwise comparisons for learning and generalization					
Viewing Condition = SAV 20°					
contrast	estimate	SE	df	t-ratio	p-value
train1 vs train4	-0.07278	0.1276	144	-0.57	1
Viewing Condition = SAV 45°					
contrast	estimate	SE	df	t-ratio	p-value
train1 vs train4	-0.27458	0.0765	144	-3.588	0.0009
train4_v_probe	0.1495	0.0735	144	2.033	0.0439
Viewing Condition = SAV FF					
contrast	estimate	SE	df	t-ratio	p-value
train1 vs train4	-0.26537	0.0669	144	-3.966	0.0002
train4 vs probe	-0.00763	0.0655	144	-0.116	0.9075
p value adjustment: holm method for 2 tests					

- Heading deviation (HD)

Best fitting model: Linear mixed effects model. HD as dependent variable, viewing condition, trial and their interaction as fixed effects and, since the design is a repeated measures design, participant as random effect.

Model assumptions:



Statistical results:

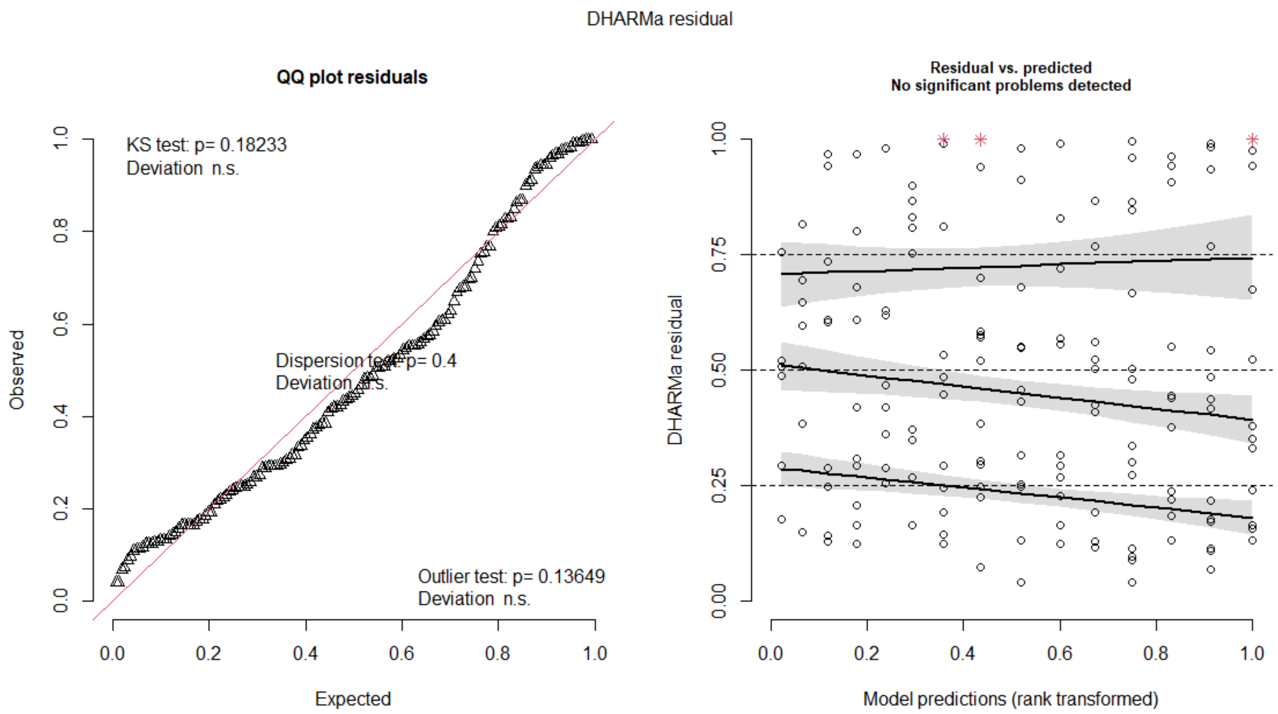
ANOVA(m): Analysis of Deviance Table (Type II Wald χ2-tests)			
	χ2	df	Pr(>χ2)
Viewing Condition	5.2621	2	0.072
Repetition	10.2592	4	3.63E-02
Viewing Condition x Repetition	10.3715	8	0.23991

Post hoc pairwise comparisons for learning and generalization					
Viewing Condition = SAV 20°					
contrast	estimate	SE	df	t-ratio	p-value
train1 vs train4	-1.71	7.6	120	-0.225	1
Viewing Condition = SAV 45°					
contrast	estimate	SE	df	t-ratio	p-value
train1 vs train4	13.77	6.7	116	2.056	0.0841
Viewing Condition = SAV FF					
contrast	estimate	SE	df	t-ratio	p-value
train1 vs train4	13.79	5.83	115	2.364	0.0395
train4 vs probe	-4.7	5.71	114	-0.824	0.4118
Degrees-of-freedom method: kenward-roger					
p value adjustment: holm method for 2 tests					

- Walking speed (WS)

Best fitting model: Generalized linear mixed effects model with zero-inflation and dispersion correction. WS was taken as dependent variable, viewing condition, trial and their interaction as fixed effects and, since the design is a repeated measures design, participant as random effect.

Model assumptions:



Statistical results:

ANOVA(m): Analysis of Deviance Table (Type II Wald χ2-tests)			
	χ2	df	Pr(>χ2)
Viewing Condition	26.3368	2	1.91E-06
Repetition	54.7853	4	3.60E-11
Viewing Condition x Repetition	3.3282	8	0.9121

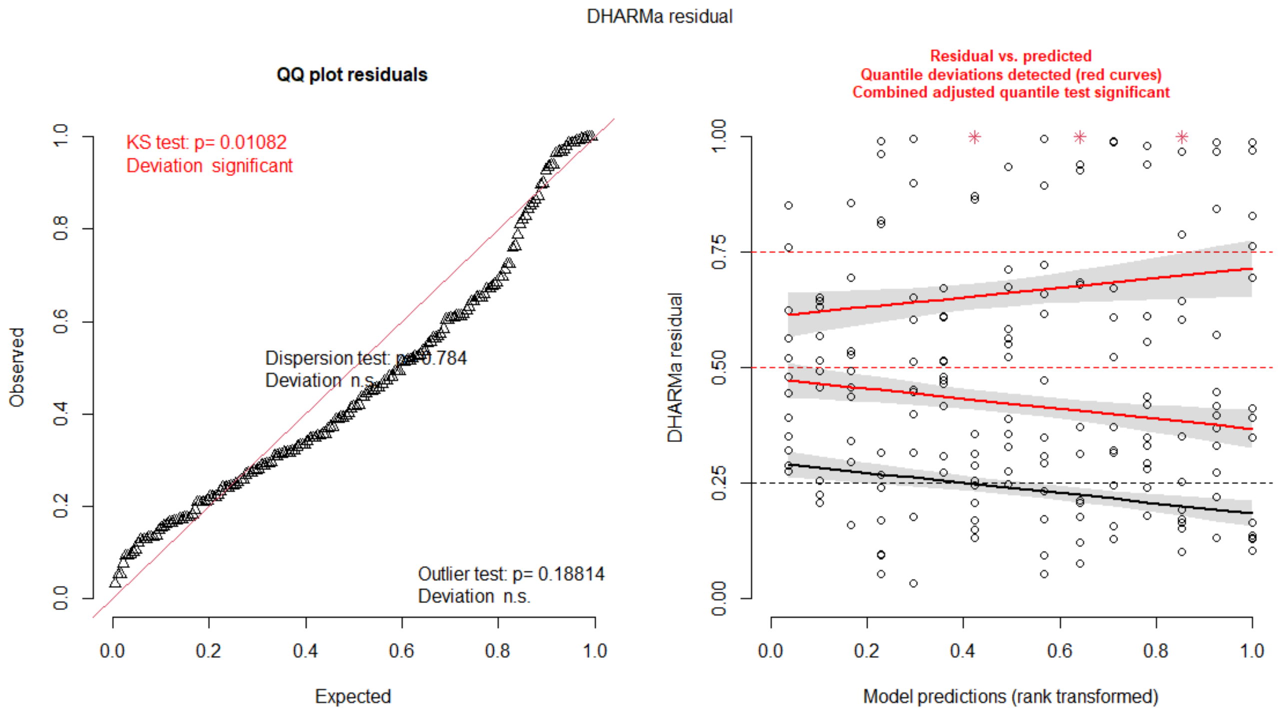
Post hoc pairwise comparisons for probe trial					
contrast	estimate	SE	df	t-ratio	p-value
SAV 20° vs SAV 45°	-0.0654	0.0288	136	-2.272	0.0493
SAV 20° vs SAV FF	-0.0795	0.0252	136	-3.152	0.006
SAV 45° vs SAV FF	-0.0141	0.0296	136	-0.476	0.6348
p value adjustment: holm method for 3 tests					

Post hoc pairwise comparisons for learning and generalization					
Viewing Condition = SAV 20°					
contrast	estimate	SE	df	t-ratio	p-value
train1 vs train4	-0.06277	0.0274	136	-2.288	0.0473
train4 vs probe	0.00121	0.0281	136	0.043	0.9657
Viewing Condition = SAV 45°					
contrast	estimate	SE	df	t-ratio	p-value
train1 vs train4	-0.09953	0.0381	136	-2.613	0.02
train4 vs probe	0.03661	0.0415	136	0.882	0.3795
Viewing Condition = SAV FF					
contrast	estimate	SE	df	t-ratio	p-value
train1 vs train4	-0.10979	0.0307	136	-3.581	0.001
train4 vs probe	0.06259	0.0326	136	1.918	0.0572
p value adjustment: holm method for 2 tests					

- Inspection time (IT)

Best fitting model: Generalized linear mixed effects model with zero-inflation and dispersion correction. IT as dependent variable, viewing condition, trial and their interaction as fixed effects and, since the design is a repeated measures design, participant as random effect.

Model assumptions: assumption violations are highlighted in red.



Statistical results:

ANOVA(m): Analysis of Deviance Table (Type II Wald χ2-tests)			
	χ2	df	Pr(>χ2)
Viewing Condition	17.069	2	1.97E-04
Repetition	36.442	4	2.35E-07
Viewing Condition x Repetition	12.899	8	0.1153596

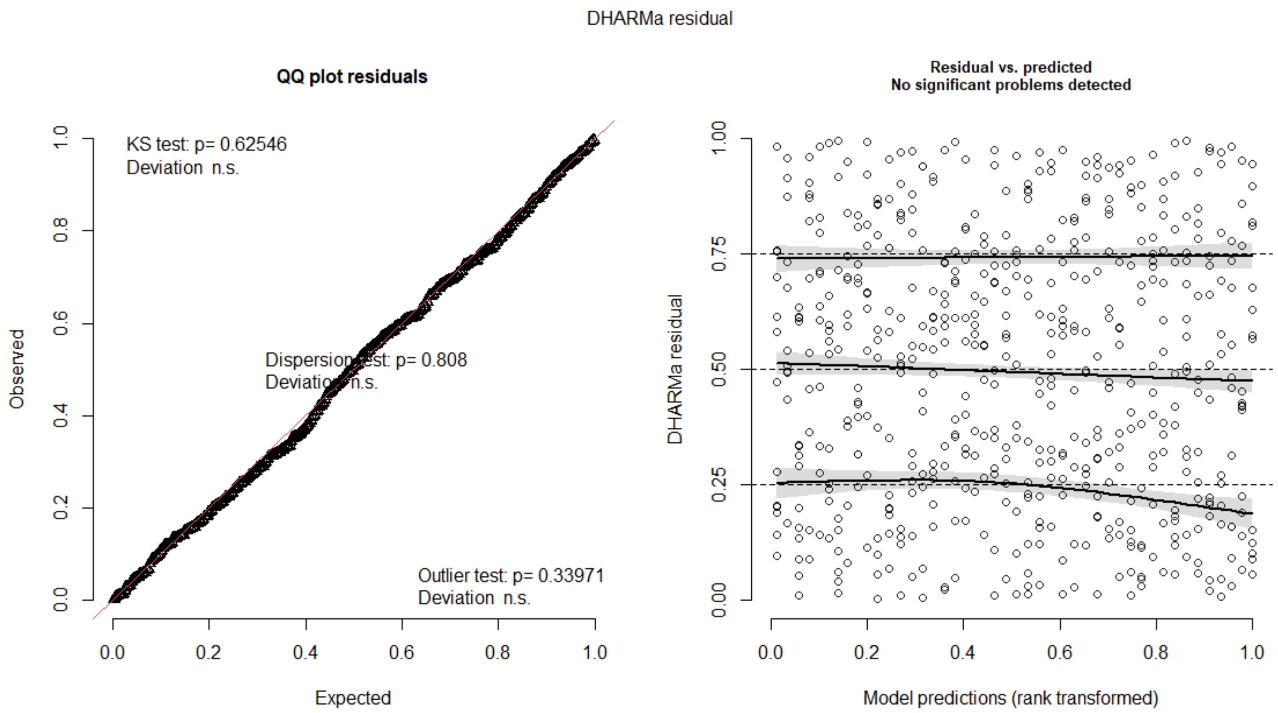
Post hoc pairwise comparisons for probe trial					
contrast	estimate	SE	df	t-ratio	p-value
SAV 20° vs SAV 45°	-0.8249	2.38	163	-0.347	0.7294
SAV 20° vs SAV FF	4.4653	2.12	163	2.107	0.0733
SAV 45° vs SAV FF	5.2902	2.11	163	2.506	0.0396
p value adjustment: holm method for 3 tests					

Post hoc pairwise comparisons for learning and generalization					
Viewing Condition = SAV 20°					
contrast	estimate	SE	df	t-ratio	p-value
train1 vs train4	19.23	7.84	163	2.452	0.0305
train4 vs probe	1.22	4.04	163	0.302	0.7628
Viewing Condition = SAV 45°					
contrast	estimate	SE	df	t-ratio	p-value
train1 vs train4	22.42	7.79	163	2.878	0.0091
train4 vs probe	1.24	3.68	163	0.336	0.7374
Viewing Condition = SAV FF					
contrast	estimate	SE	df	t-ratio	p-value
train1 vs train4	6.39	4.21	163	1.516	0.1315
p value adjustment: holm method for 2 tests					

- Start fixations (SF)

Best fitting model: Generalized linear mixed effects model with poisson distribution and zero-inflation correction. SF as dependent variable, viewing condition, trial, fixation location and their interaction as fixed effects and, since the design is a repeated measures design, participant as random effect.

Model assumptions:



Statistical results:

ANOVA(m): Analysis of Deviance Table (Type II Wald χ2-tests)			
	χ2	df	Pr(>χ2)
Fixation Location	345.808	4	< 2.2e-16
Viewing Condition	10.454	2	5.37E-03
Repetition	273.088	6	< 2.2e-16
Fixation Location x Viewing condition	14.441	4	6.01E-03
Fixation Location x Repetition	17.822	8	0.022599
Viewing Condition x Repetition	63.162	8	1.11E-10
Fixation Location x Viewing Condition x Repetition	11.071	16	0.80507

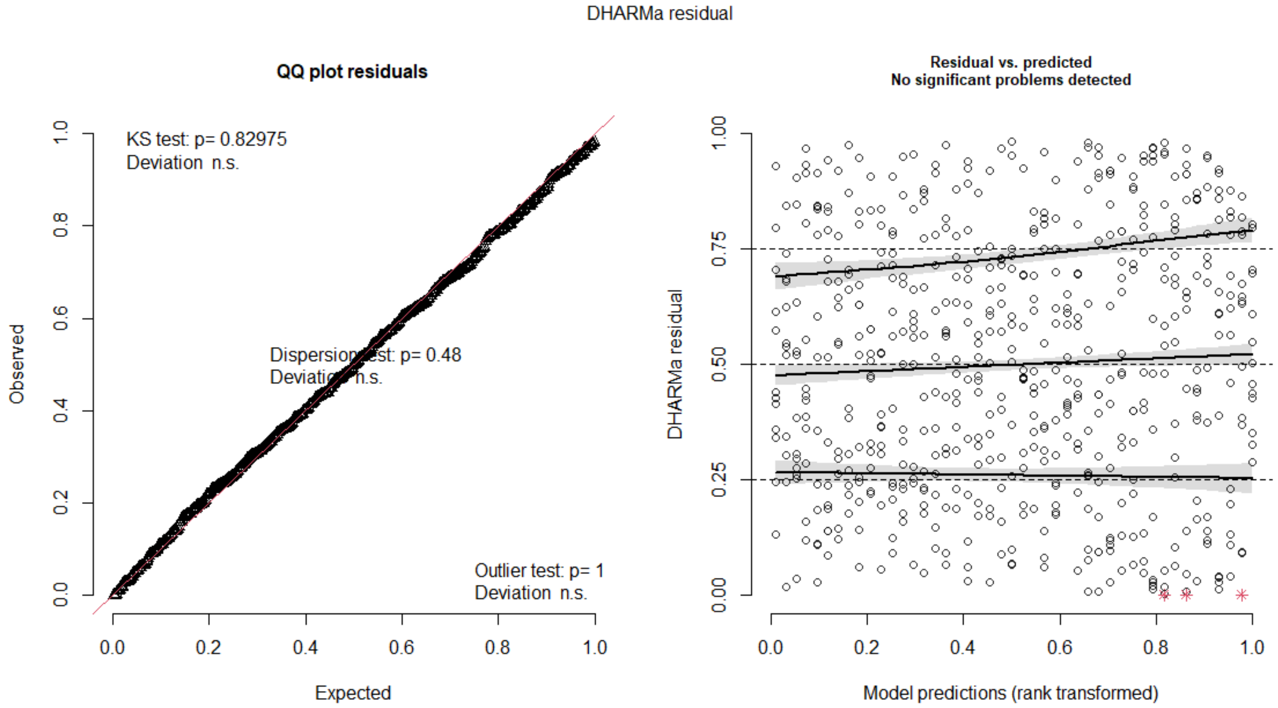
Post hoc pairwise comparisons for probe trial					
Fixation Location = Other Locations					
contrast	estimate	SE	df	z-ratio	p-value
SAV 20° vs SAV 45°	-0.1403	4.00E-01	Inf	-0.399	0.8389
SAV 20° vs SAV FF	0.2931	4.00E-01	Inf	0.807	0.8389
SAV 45° vs SAV FF	0.4334	3.00E-01	Inf	1.269	0.6134
Fixation Location = Stations					
contrast	estimate	SE	df	z-ratio	p-value
SAV 20° vs SAV 45°	5.7192	2.21E+05	Inf	0	1
SAV 20° vs SAV FF	7.4486	4.78E+05	Inf	0	1
SAV 45° vs SAV FF	1.7294	5.26E+05	Inf	0	1
Fixation Location = Street Edges					
contrast	estimate	SE	df	z-ratio	p-value
SAV 20° vs SAV 45°	1.1479	4.00E-01	Inf	2.781	0.0163
SAV 20° vs SAV FF	1.0264	4.00E-01	Inf	2.622	0.0175
SAV 45° vs SAV FF	-0.1215	4.00E-01	Inf	-0.272	0.7856
Results are given on the log (not the response) scale					
p value adjustment: holm method for 3 tests					

Post hoc pairwise comparisons for change over the course of trials and generalization					
Viewing Condition = SAV 20°, Fixation Location = Others Locations					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	0.0932	0.1	Inf	0.629	0.5295
Viewing Condition = SAV 45°, Fixation Location = Others Locations					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	1.308	0.1	Inf	9.561	<.0001
train4 vs probe	-0.0198	0.2	Inf	-0.113	0.9102
Viewing Condition = SAV FF, Fixation Location = Others Locations					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	1.6412	0.4	Inf	4.51	<.0001
train4 vs probe	-0.8529	0.4	Inf	-2.222	0.0263
Viewing Condition = SAV 20°, Fixation Location = Stations					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	18.4909	7695.5	Inf	0.002	1
Viewing Condition = SAV 45°, Fixation Location = Stations					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	0.4751	0.6	Inf	0.755	0.8999
Viewing Condition = SAV FF, Fixation Location = Stations					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	28.8232	743868.6	Inf	0	1
Viewing Condition = SAV 20°, Fixation Location = Street Edges					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	0.4266	0.2	Inf	2.482	0.0261
train4 vs probe	-0.1157	0.2	Inf	-0.563	0.5734
Viewing Condition = SAV 45°, Fixation Location = Street Edges					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	1.0654	0.2	Inf	4.596	<.0001
train4 vs probe	0.6446	0.3	Inf	1.93	0.0536
Viewing Condition = SAV FF, Fixation Location = Street Edges					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	1.1878	0.4	Inf	3.362	0.0015
train4 vs probe	-0.4917	0.4	Inf	-1.239	0.2153
Results are given on the log (not the response) scale					
p value adjustment: holm method for 2 tests					

- Navigation fixations (NF)

Best fitting model: Generalized linear mixed effects model with negative binomial distribution and dispersion correction. NF as dependent variable, viewing condition, trial, fixation location and their interaction as fixed effects and, since the design is a repeated measures design, participant as random effect.

Model assumptions:



Statistical results:

ANOVA(m): Analysis of Deviance Table (Type II Wald χ2-tests)			
	χ2	df	Pr(>χ2)
Fixation Location	1094.8337	2	< 2.2e-16
Viewing Condition	15.6446	2	4.01E-04
Repetition	63.312	4	5.83E-13
Fixation Location x Viewing condition	23.8435	4	8.59E-05
Fixation Location x Repetition	6.9823	8	0.5385438
Viewing Condition x Repetition	12.933	8	0.1141757
Fixation Location x Viewing Condition x Repetition	9.0306	16	0.9121467

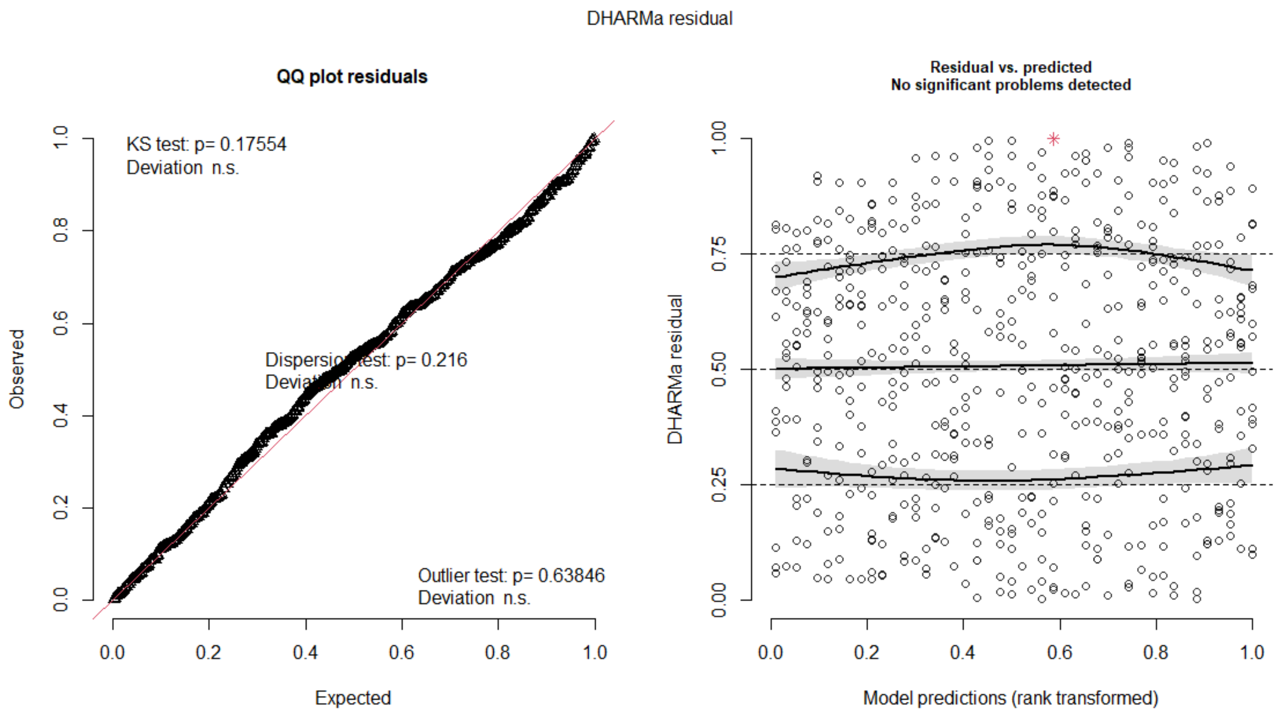
Post hoc pairwise comparisons for probe trial					
Fixation Location = Others Locations					
contrast	estimate	SE	df	z-ratio	p-value
SAV 20° vs SAV 45°	0.53532	0.257	Inf	2.083	0.0745
SAV 20° vs SAV FF	0.60445	0.257	Inf	2.347	0.0567
SAV 45° vs SAV FF	0.06913	0.251	Inf	0.276	0.7828
Fixation Location = Stations					
contrast	estimate	SE	df	z-ratio	p-value
SAV 20° vs SAV 45°	1.26205	0.389	Inf	3.243	0.0036
SAV 20° vs SAV FF	0.60323	0.36	Inf	1.677	0.1714
SAV 45° vs SAV FF	-0.65882	0.383	Inf	-1.718	0.1714
Fixation Location = Street Edges					
contrast	estimate	SE	df	z-ratio	p-value
SAV 20° vs SAV 45°	0.98615	0.322	Inf	3.063	0.0066
SAV 20° vs SAV FF	0.98358	0.322	Inf	3.051	0.0066
SAV 45° vs SAV FF	-0.00257	0.327	Inf	-0.008	0.9937
Results are given on the log (not the response) scale					
p value adjustment: holm method for 3 tests					

Post hoc pairwise comparisons for change over the course of trials and generalization					
Viewing Condition = SAV 20°, Fixation Location = Others Locations					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	0.1351	0.15	Inf	0.9	0.6849
Viewing Condition = SAV 45°, Fixation Location = Others Locations					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	0.4996	0.126	Inf	3.955	0.0002
train4 vs probe	0.0447	0.141	Inf	0.317	0.7511
Viewing Condition = SAV FF, Fixation Location = Others Locations					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	0.6347	0.159	Inf	3.999	0.0001
train4 vs probe	-0.1667	0.17	Inf	-0.983	0.3256
Viewing Condition = SAV 20°, Fixation Location = Stations					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	0.336	0.327	Inf	1.026	0.3047
Viewing Condition = SAV 45°, Fixation Location = Stations					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	0.054	0.343	Inf	0.157	1
Viewing Condition = SAV FF, Fixation Location = Stations					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	0.9793	0.327	Inf	2.994	0.0055
train4 vs probe	-0.63	0.34	Inf	-1.852	0.064
Viewing Condition = SAV 20°, Fixation Location = Street Edges					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	0.2954	0.226	Inf	1.309	0.3813
Viewing Condition = SAV 45°, Fixation Location = Street Edges					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	0.3737	0.228	Inf	1.637	0.2031
Viewing Condition = SAV FF, Fixation Location = Street Edges					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	0.3133	0.271	Inf	1.158	0.4939
Results are given on the log (not the response) scale					
p value adjustment: holm method for 2 tests					

- Interaction fixations (IF)

Best fitting model: Generalized linear mixed effects model with negative binomial distribution and zero-inflation and dispersion correction. IF as dependent variable, viewing condition, trial, fixation location and their interaction as fixed effects and, since the design is a repeated measures design, participant as random effect.

Model assumptions:



Statistical results:

ANOVA(m): Analysis of Deviance Table (Type II Wald χ2-tests)			
	χ2	df	Pr(>χ2)
Fixation Location	215.0191	2	< 2.2e-16
Viewing Condition	4.8519	2	8.84E-02
Repetition	23.0838	4	1.22E-04
Fixation Location x Viewing condition	56.0613	4	1.95E-11
Fixation Location x Repetition	20.9949	8	0.0071611
Viewing Condition x Repetition	10.7059	8	0.2189247
Fixation Location x Viewing Condition x Repetition	18.7253	16	0.2831662

Post hoc pairwise comparisons for probe trial

contrast	estimate	SE	df	z-ratio	p-value
Fixation Location = Others Locations					
SAV 20° vs SAV 45°	-0.0463	0.462	Inf	-0.1	1
SAV 20° vs SAV FF	-0.1628	0.455	Inf	-0.358	1
SAV 45° vs SAV FF	-0.1165	0.409	Inf	-0.285	1
Fixation Location = Stations					
SAV 20° vs SAV 45°	0.4304	0.461	Inf	0.933	0.702
SAV 20° vs SAV FF	-0.1157	0.427	Inf	-0.271	0.7866
SAV 45° vs SAV FF	-0.546	0.399	Inf	-1.37	0.5124
Fixation Location = Street Edges					
SAV 20° vs SAV 45°	0.1381	0.634	Inf	0.218	1
SAV 20° vs SAV FF	-0.1759	0.604	Inf	-0.291	1
SAV 45° vs SAV FF	-0.314	0.535	Inf	-0.587	1

Results are given on the log (not the response) scale

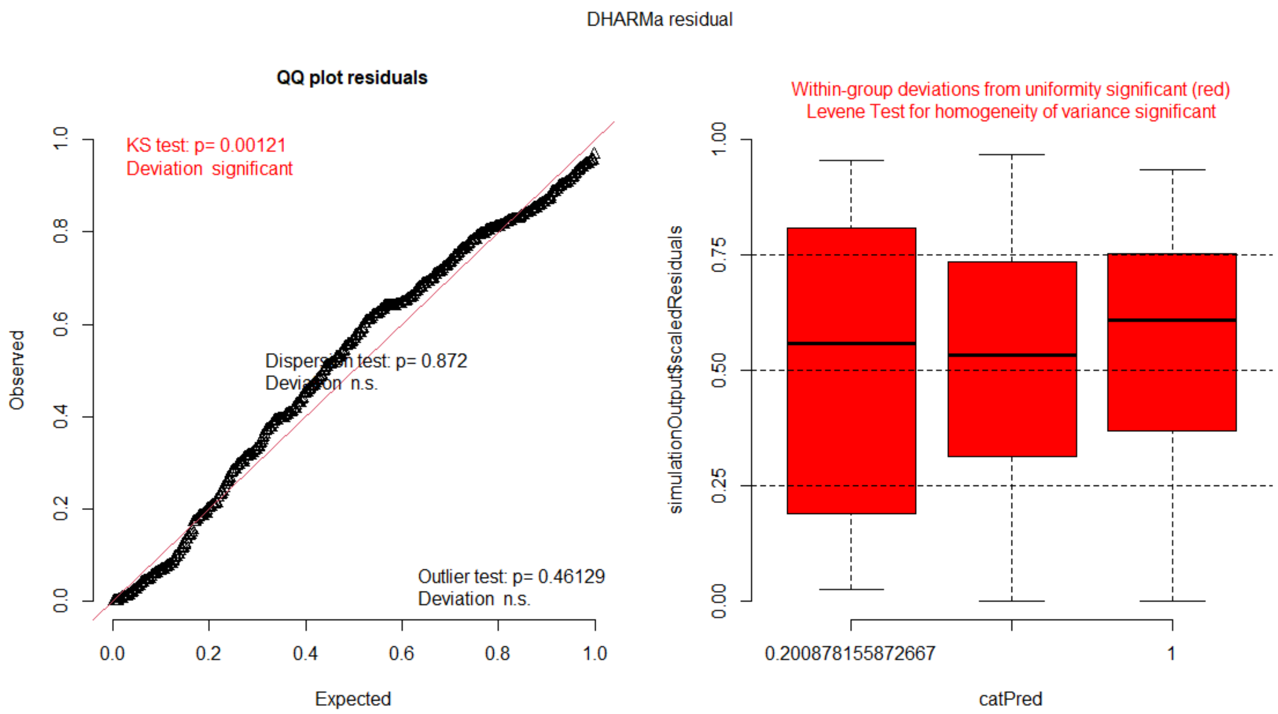
p value adjustment: holm method for 3 tests

Post hoc pairwise comparisons for change over the course of trials and generalization					
Viewing Condition = SAV 20°, Fixation Location = Others Locations					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	-0.0032	0.457	Inf	-0.007	1
train4 vs probe	-0.2306	0.407	Inf	-0.567	1
Viewing Condition = SAV 45°, Fixation Location = Others Locations					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	0.5865	0.294	Inf	1.992	0.0928
Viewing Condition = SAV FF, Fixation Location = Others Locations					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	0.79	0.362	Inf	2.184	0.0557
Viewing Condition = SAV 20°, Fixation Location = Stations					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	-0.1326	0.46	Inf	-0.288	0.7734
Viewing Condition = SAV 45°, Fixation Location = Stations					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	-0.1894	0.361	Inf	-0.525	1
Viewing Condition = SAV FF, Fixation Location = Stations					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	0.4768	0.209	Inf	2.281	0.0451
train4 vs probe	0.2989	0.234	Inf	1.278	0.2013
Viewing Condition = SAV 20°, Fixation Location = Street Edges					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	0.9618	0.505	Inf	1.905	0.1136
Viewing Condition = SAV 45°, Fixation Location = Street Edges					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	1.0705	0.505	Inf	2.12	0.068
Viewing Condition = SAV FF, Fixation Location = Street Edges					
contrast	estimate	SE	df	z-ratio	p-value
train1 vs train4	1.5236	0.577	Inf	2.639	0.0166
train4 vs probe	-1.4043	0.561	Inf	-2.502	0.0166
Results are given on the log (not the response) scale					
p value adjustment: holm method for 2 tests					

- Average sketch map score

Best fitting model: Generalized linear mixed effects model with dispersion correction. Sketch map score as the dependent variable, viewing condition as fixed effect and rater as random effect.

Model assumptions: assumption violations are highlighted in red.



Statistical results:

ANOVA(m): Analysis of Deviance Table (Type II Wald χ2-tests)			
	χ2	df	Pr(>χ2)
Viewing Condition	80.782	2	< 2.2e-16

Post hoc pairwise comparisons for probe trial					
contrast	estimate	SE	df	t-ratio	p-value
SAV 20° vs SAV 45°	-6.92	2.31	539	-2.996	0.0029
SAV 20° vs SAV FF	-19.59	2.32	539	-8.445	<.0001
SAV 45° vs SAV FF	-12.67	1.97	539	-6.418	<.0001
p value adjustment: holm method for 3 tests					

- Sketch map rater variability

Model: Linear model with rater SD as dependent and viewing condition as independent variable.

Model assumptions:

Shapiro test to test for normal distribution for each group		
Viewing Condition	estimate	p-value
SAV 20°	0.915	0.189
SAV 45°	0.945	0.482
SAV FF	0.972	0.899
Normality: parametric test can be performed		

Statistical results:

ANOVA(m): Analysis of Variance Table					
	df	Sum Sq	Mean Sq	F-value	Pr(>*F*)
Viewing Condition	2	0.052158	0.0260789	5.3932	0.008552
Residuals	39	0.188583	0.0048355		

Post hoc pairwise t-tests	
contrast	p-value
SAV 20° vs SAV 45°	0.019
SAV 20° vs SAV FF	0.017
SAV 45° vs SAV FF	0.68
p value adjustment: holm method for 3 tests	