Psychological Research
DOI 10.1007/s00426-011-0397-5

ORIGINAL ARTICLE

Gaze behaviour during space perception and spatial decision making

Jan M. Wiener · Christoph Hölscher · Simon Büchner · Lars Konieczny

Received: 15 March 2011 / Accepted: 17 November 2011

Abstract A series of four experiments investigating gaze behavior and decision making in the context of wayfinding is reported. Participants were presented with screenshots of choice points taken in large virtual environments. Each screenshot depicted alternative path options. In Experiment 1, participants had to decide between them to find an object hidden in the environment. In Experiment 2, participants were first informed about which path option to take as if following a guided route. Subsequently, they were presented with the same images in random order and had to indicate which path option they chose during initial exposure. In Experiment 1, we demonstrate (1) that participants have a tendency to choose the path option that featured the longer line of sight, and (2) a robust gaze bias towards the eventually chosen path option. In Experiment 2, systematic differences in gaze behavior towards the alternative path options between encoding and decoding were observed. Based on data from Experiments 1 and 2 and two control experiments ensuring that fixation patterns were specific to the spatial tasks, we develop a tentative model of gaze behavior during wayfinding decision making suggesting that particular attention was paid to image areas depicting changes in the local geometry of the environments such as corners, openings, and occlusions. Together, the results suggest that gaze during a wayfinding tasks is directed toward, and can be predicted by, a subset of environmental features and that gaze bias effects are a general phenomenon of visual decision making.

Keywords Visual attention · Wayfinding · Decision making · Gaze behavior · Spatial cognition · Space perception

Introduction

Where do we look when moving though the environment? In the context of navigation and wayfinding, eye-tracking studies have primarily investigated the role of gaze for the control of locomotory or steering behavior. Wayfinding, however, also comprises higher level processes such as encoding information into and retrieving it from spatial memory, path planning, and decision making at choice points. So far, very few eye-tracking studies investigated these processes to answer questions like: how does gaze behavior relate to spatial decision making? Is it possible to predict path choices by analyzing gaze behavior? And, which information are attended to when interpreting spatial situations and deciding between path alternatives? This work constitutes a first account to approach these questions.

Gaze behavior and navigation

Gaze behavior has been investigated in studies that required participants to move through space. Grasso, Prevost, Ivanenko, and Berthoz (1998) demonstrated anticipatory gaze behaviour when walking along curved paths. Similar results

J. M. Wiener (✉)
Department of Psychology, Bournemouth University,
Poole House, Talbot Campus, Fern Barrow,
Poole, Dorset BH12 5BB, UK
e-mail: jwiener@bournemouth.ac.uk

J. M. Wiener · C. Hölscher · S. Büchner · L. Konieczny
Center for Cognitive Science, Freiburg University,
Freiburg, Germany
e-mail: christoph.hoelscher@cognition.uni-freiburg.de

S. Büchner
e-mail: simon.buechner@cognition.uni-freiburg.de

L. Konieczny
e-mail: lars.konieczny@cognition.uni-freiburg.de

also come from steering experiments: when driving a car around a curve, drivers gaze at the tangent point on the inside of the curve (Land & Lee, 1994; Land & Tatler, 2001). Moreover, steering performance systematically decreases if gaze is fixed (Wilkie & Wann, 2003), and is biased in the direction of gaze (Readinger, Chatziastros, Cunningham, Bülthoff, & Cutting, 2002). Gaze behavior while moving through space has also been investigated in tasks involving navigation through cluttered spaces containing large objects (Rothkopf, Ballard, & Hayhoe, 2007). Fixation patterns were task specific, systematically differing when participants were asked to approach the objects and when asked to avoid them. In all these studies, participants solved prototypical locomotion tasks, such as steering, obstacle avoidance and approach, that require moving through the world in response to sensory–motor input of the local surrounding.

Montello (2001) conceptually separates navigation behavior into locomotion and wayfinding. Locomotion refers to navigation behavior in response to current sensory–motor input of the immediate surroundings and is therefore strongly related to the act of physically moving through space. Wayfinding tasks, in contrast, aim at reaching destinations beyond the current sensory horizon and involve some representation of the environment, decision making and/or planning processes. These processes, although crucial for successful wayfinding and navigation, are not necessarily tied to the act of physically moving through space. Typical wayfinding tasks are search, exploration, route learning, and route planning. Wiener, Büchner, and Hölscher (2009) analyze the level of information required for different wayfinding tasks and show that in many cases knowledge about key choice points and their connectivity is sufficient.

The question of how gaze behavior relates to wayfinding-related tasks such as memorizing and recalling path choices and movement decisions along a route, or deciding between path alternatives at junctions when exploring novel environments, has been addressed only in very few studies. Schuchard, Connell, and Griffiths (2006) collected eye-tracking data with dementia patients navigating the hallways of a nursing home. These patients showed a lack of ability to attend to wayfinding-critical cues (like signage) predominantly focussing on the lower parts of their visual field, presumably concentrating on motion control issues (i.e., the locomotion-component of navigation). Vembar et al., (2004) employed eye-tracking techniques to capture visual attention while wayfinding in a virtual reality (VR) maze. Their data analysis is limited to the relative amount of time spent looking at the 3D environment versus a supplementary survey map of the same environment. Spiers and Maguire (2008) present eye-tracking evidence on taxi drivers' pattern of visual attention in a VR driving simulation of London. Finally, Allen and Kirasic (2003), using a slide-show paradigm similar to the one applied in the present study, identified difference in visual focus between scenes with high versus low density of wayfinding-critical cues. Neither of these studies identified how eye movements relate to deciding between route options and analyzed which specific geometric features of the environment were attended when doing so.

The relation between decision making and gaze behavior has been studied in non-spatial contexts, and has been shown to be closely tied to preferences (e.g., Simion & Shimojo, 2007; Armel, Beaumel, & Rangel, 2008). Shimojo, Simion, Shimojo, and Scheier (2003) demonstrated that participants when asked to select the most attractive face, display a stable gaze bias towards the eventually chosen face in the last second before reporting the decision. The *gaze-cascade model* states that gaze is not merely reflecting preferences, but is involved in the formation of preferences. Essentially, it suggests that gaze orientation towards a stimulus and preference for that stimulus are linked in a positive feedback loop. Glaholt and Reingold (2009) recently extended these findings by demonstrating that the gaze bias effect is not specific for preference choices but constitutes a more general phenomenon of visual decision making. However, up to now it remains an open question whether similar effects can also be observed in spatial decision making when deciding between path alternatives. This question will be approached in the current study.

Control of visual attention

Currently, it is also unknown which characteristics of the environment viewers are attending to when deciding between path alternatives in the context of wayfinding. The features people attend to when inspecting images of scenes in a non-spatial context, in contrast, have been investigated in numerous studies revealing bottom-up as well as of top-down influences (for an overview see Henderson, 2003). *Bottom-up approaches* state that certain features of a visual scene attract attention and result in shifts of attention. The most widely used bottom-up approach is that of saliency maps, representations of the stimulus coding the strength of different features that are known to be extracted in early vision such as color, intensity and orientation (Itti & Koch, 2000, 2001; Parkhurst, Law, & Niebur, 2002; Foulsham & Underwood, 2008). The *Contextual Guidance Model* of attention (Torralba, Oliva, Castelhano, & Henderson, 2006) extends pure bottom-up approaches by acknowledging that both context and task modulate the selection of image regions that are attended to. Generally, *top-down approaches* focus on task-related influences on gaze control. For example, gaze patterns when inspecting the same image systematically differ when judging the ages of people depicted or when estimating their material circumstances (Yarbus, 1967), and fixation patterns differ when viewers search for an object in a scene or when they are memorizing that scene. (Henderson, Weeks, & Hollingworth, 1999; Castelhano, Mack, & Henderson, 2009). In general, top-down approaches state that humans use their knowledge about the world or the task at hand to guide their fixations when inspecting sceneries (Henderson, Malcolm, & Schandl, 2009).

In the context of a wayfinding task in which participants have to decide between path alternatives, top-down influences would suggest that attention is directed to stimulus areas that convey information about the spatial structure of the visible space. This opens up the intriguing question of whether it is possible to relate a wayfinder's visual attention to purely geometrical properties of the depicted space.

Synopsis

Two questions shall be considered in this study: (1) how does gaze behavior relate to spatial decision making? Are participants' choices reflected by gaze bias effects? (2) How does control of visual attention in spatial decision making relate to stimulus characteristics? Is it possible to predict gaze behaviour by analyzing geometric features of the spatial scenes? Question 1 is addressed with two psychophysical experiments in which participants were presented with images taken at choice points in large, complex virtual mazes, each depicting path alternatives. Their task in Experiment 1 was to select between the alternatives in order to search for an object that was hidden in the environment. In Experiment 2 participants were instructed about which path alternative to choose in an encoding phase. In the subsequent decoding phase, they had to recall these choices. In order to address how gaze behaviour relates to spatial decision making, we calculate the likelihood that the eventually chosen path option is inspected as a function of time.

We then address the question of whether gaze behaviour during wayfinding decision making can be predicted by purely geometric features of the stimuli. First, two control experiments are presented to ensure that the fixation patterns in Experiments 1 and 2 were in fact task specific. We then relate participants' fixation patterns to quantitative descriptions of the geometry of the depicted scenes and compare the predictive power of different geometrical characteristics for the recorded fixation patterns.

Note that the present study was not designed to investigate the role of sensori-motor information or of allocentric, metrically embedded survey representations for navigation. Authors like Zetzsche, Galbraith, Wolter, and Schill (2009) argue that weaker representation of environments such as route-level knowledge about landmarks and connections between them (cf. Siegel & White, 1975), are often sufficient for successful wayfinding (see also Gillner & Mallot, 1998). For the given tasks this is in line with a taxonomic account recently presented by the authors (Wiener et al., 2009). The static images of wayfinding decision points in this paper do not provide the participants with information about the Euclidean distance or angle between the snapshots, as would be obtained from actual movement through space and as would be necessary for path integration or determining novel routes/shortcuts between destinations. Yet the stimuli do provide information about the potential environmental structure beyond the current view (line of sight and number of occlusions suggesting further path options, cf. Peponis, Wineman, Rashid, Kim, & Bafna, 1997) as well as unique local layout (structural landmarks, see Stankiewicz & Kalia, 2007). The current study investigates the influence of these factors on wayfinding decision making. Untangling the impact of sensori-motor information and correct metric knowledge about relations between decision points will be an issue of future research.

Experiment 1

Methods

Participants

Twenty subjects (14 women, mean age 22.45 ± 2.83 years) participated in the experiment. They were mostly university students and were paid 8 Euros an hour.

Stimuli

The stimuli were 30 screenshots from within large virtual architectural environments (for examples, see Fig. 1). Each screenshot was taken at a decision point, depicting two path options. Pilot experiments suggested that stimuli as depicted in the left column of Fig. 1 could be comprehended without many gaze shifts. This observation is in line with literature suggesting that a coarse interpretation of a scene can be computed without the need for selective visual attention (Oliva & Torralba, 2006). However, one goal of this study was the investigation of which environmental features are crucial for the comprehension of the spatial structure of the depicted scenes (space perception) and for wayfinding decision making. We therefore manipulated the stimuli by selectively adjusting the colors of floor and ceiling to that of the walls (compare left and right column in Fig. 1) such that the geometry of the depicted scene could not be comprehended without overt shifts of attention. While this manipulation allowed us to monitor which

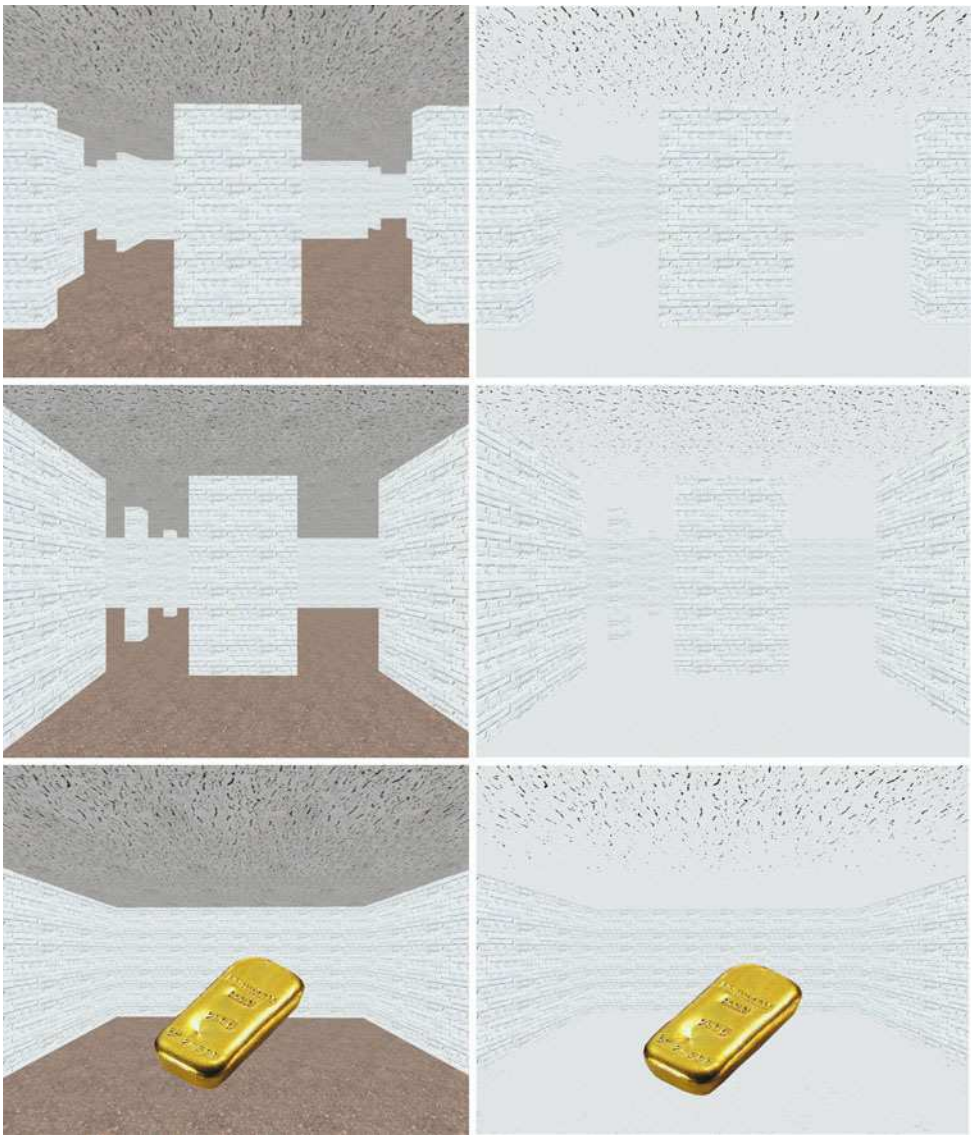

Fig. 1 *Left column* un-manipulated stimuli; *right column* manipulated, experimental stimuli used for the study; *upper and and middle row* two examples of decision points presented to participants; *lower row* the final stimulus presented at the end of each search, demonstrating that participants have found the hidden item (*the gold bar*)

environmental features participants attended to by recording eye-movements, one may argue that as a result of this manipulation it is unclear if, and how, findings from this study apply to more natural settings. In real life, however, we also make wayfinding decisions under bad lightning conditions and scenes cover much larger portions of the visual field than in our experiments. Under these conditions viewers will often not be able to comprehend the structure of the scene in a single glance only and will need to shift attention overtly. We assume that such attention shifts in real-world wayfinding will be directed to similar environmental properties as in the current study.

For the sake of clarity, further figures will depict the unmanipulated stimuli. Two versions of each stimulus were generated by mirroring the original stimulus along the vertical axis. Presentation of the original and the mirrored stimuli was balanced between participants.

Spatial analysis For analysing spatial characteristics of small-scale environments isovists have been suggested as objectively determinable basic elements (Benedikt, 1979). Isovists capture relevant spatial properties by describing the visible area from a given observation point with the use of viewshed polygons. The spatial characteristics of the scenes in this study were analyzed using a variant of isovist analysis: for each stimulus a depth profile was generated by contouring the edge between the floor and the walls (see Fig. 2 right). The height of the resulting contour describes the geometry of visible space by the distances of the vertical walls from the observer. Areas in the stimuli in which the profile is high relate to areas in space that are distant and vice versa. This analysis was chosen, because (1) the spatial situation is described from the perspective of the beholder capturing behaviorally relevant properties of space (Wiener et al., 2007; Franz & Wiener, 2008); (2) the visual system is

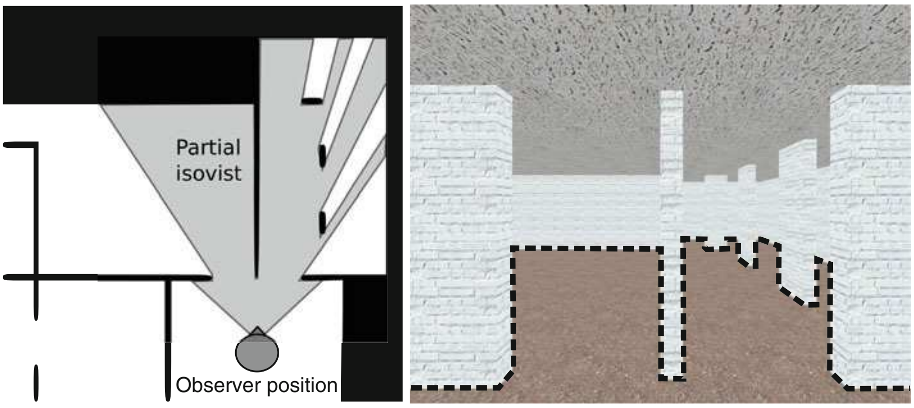

Fig. 2 *Left* Position in the maze from which one of the snapshots was taken. The *grey* area represents the isovist (depth profile) at this position; *right* corresponding view in the ego-perspective. The depth profile that is approximated by the *dashed line* is functionally equivalent to the partial isovist displayed on the *left*

able to use functionally equivalent information (i.e., angular declination) for distance judgments (Ooi, Wu, & He, 2001); (3) recent approaches in computational vision demonstrate that this information can also be computed directly from the image, even for complex natural outdoor scenes (Hoiem, Efros, & Hebert, 2007). The depth profiles were used to compare spatial properties of the left and right path alternative (left and right half of the stimulus). Earlier research suggests that line of sight (Conroy Dalton, 2003) and the local spatial complexity (Wiener et al., 2007) influence navigation and wayfinding behavior. Accordingly, we calculated longest line of sight for both path alternatives as well as the number of straight line segments to approximate their the spatial complexity. In order to verify the complexity measure, 12 volunteers, unaware of the experiment and its hypothesis, were asked to rate the stimuli according to their complexity on a scale from 1 to 7. The rated complexity for the different stimuli was highly correlated with the total number of straight line segments in the depth profiles ($r = 0.88$, $p < .001$).

Procedure

Participants first read a description of the experiment stating that their task was to search for an object (a gold bar) that was placed somewhere in the environment. They would be presented with a series of single choice points depicting alternative path options separated by a wall in the middle of the image. Their task was to decide between the left and the right path option. Participants were given no information about where to find the target object. Instead of actually walking through the environment they would then be presented with the next choice point that they would have encountered if actually navigating through the environment. In order to illustrate this procedure, participants were presented with a series of snapshots taken between two neighboring choice points, as well as an image of the target. By presenting static images, the current study separates decision making and memory-related processes that are crucial for successful wayfinding from processes related to experiencing continuous motion and controlling locomotion (cf. Montello, 2001; Wiener et al., 2009).

Before presenting a novel stimulus, participants fixated a small cross in the center of the screen and pressed the 'Space' bar. In order to indicate their path choice, they pressed the left or right cursor key. Each stimulus was presented for 5 s, irrespective of when participants responded. In contrast to the instructions participants received, their path choices (left or right) at single choice points had no influence on the image presented next. The images were presented in random order. The experiment was divided into 5 routes containing 4, 5, 6, 7, and 8 decisions. After the last decision, participants were presented with an image of a gold bar hovering in a small room (see Fig. 1).

Apparatus

The stimuli were displayed at a resolution of 1,024 × 768 pixels on a 20″ CRT monitor. The screen refresh rate was 100 Hz. Participants sat in front of the monitor at a distance of 60 cm, such that the resulting visual angle of the stimuli displayed was 37° (horizontally) × 28° (vertically). Eye movement were recorded using a SR Research Ltd. Eye-Link II eye-tracker, sampling pupil position at 500 Hz. The Eyelink II eye-tracker is a head-mounted system that does not require the participant's head to be constrained. The eye-tracker was calibrated using a 9-point grid. Fixations were defined using the detection algorithm supplied by SR Research.

Analysis

Behavioral data For each stimulus presented, the participants' decisions (left/right) as well as the corresponding response time was recorded.

Eye movement data For each stimulus three *interest areas*, vertically dividing the image in to a left part, a central

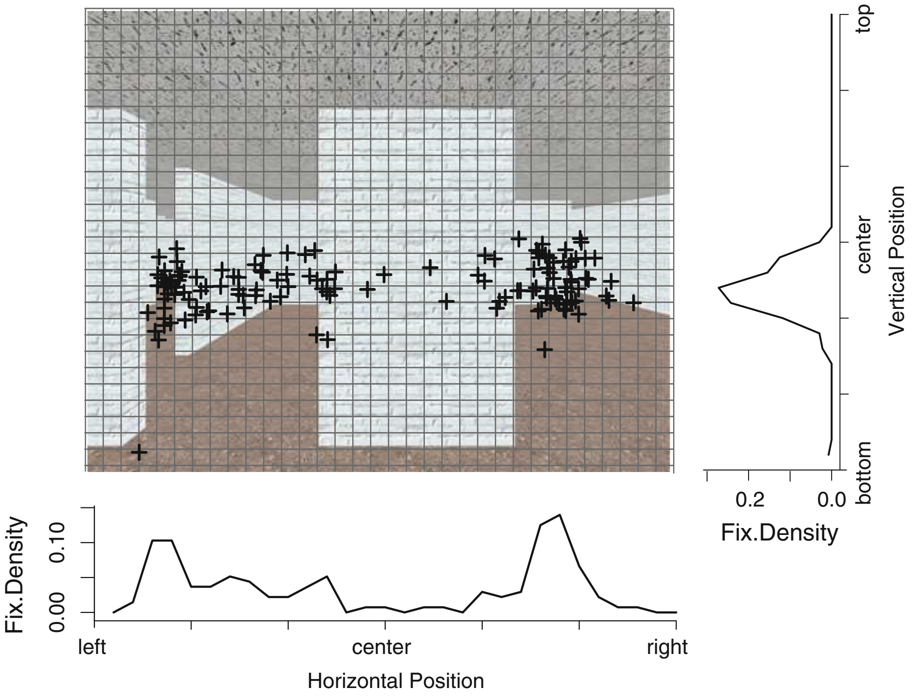

Fig. 3 Exemplary fixation pattern for one of the stimuli. Fixation densities are calculated separately for the horizontal and vertical image dimension

part, and a right part were defined. The width of the central interest area was adjusted such as to cover the central wall that varied in extension, the size of the left and the right interest areas were identical. Fixations were assigned to the different interest areas. If not stated otherwise, the initial fixations directed towards the central interest area were removed, as these fixations resulted from the requirement to look at the fixation cross before stimulus onset.

In order to compare fixation patterns between stimuli, a 30 × 30 grid was imposed on the stimuli and fixations were assigned to the corresponding grid-cells. Fixation density was calculated separately for the horizontal and vertical image dimension by summing up the number of fixations along the vertical or horizontal grid dimension. Each entry of the resulting two vectors with 30 entries each were then divided by the sum of the entire vector (see Fig. 3). Only fixations before participants reported their response (Experiments 1 and 2) entered this analysis.

Results

Behavioral data

Participants displayed a small yet significant tendency to choose the right path option [54.07 ± 3.70%: *t* test against chance level (50%): $t(19) = 2.28$, $p = .03$]. Their tendency to produce stereotypical responses was analyzed for each route. We evaluated which path option was chosen more often and divided the number of choices for this option by the total number of responses for that route, resulting in values between .5 when participants chose both path options equally often and 1 when they always chose the same path option. Note, that values are always larger than or equal to .5 and a value of exactly .5 is only possible for routes with an even number of responses. Chance level for this measure is .67, the participants' value was .66 ± 0.02. In 54.78% of the responses within one route, they switched from the left to the right or vice versa [*t* test against chance level (50%) : $t(19) = 1.30$, $p = .21$]. Together, these analyses suggest that participants were not using simple search strategies such as making right or left turns only.

In order to analyze which spatial features influenced participants' path choices, we calculated the 'difference in the length of the longest line of sight' between path alternatives as well as the 'difference in the number of vertical straight line segments in the depth profile' between path alternatives for each stimulus. The predictive power of these measures for decisions was then tested. As the two measures were correlated ($r = 0.43$), a step-wise linear regression for the mean decision score was calculated including both factors. Difference of line of sight significantly predicted mean decision scores, $\beta = 0.64$, $t(29) = 4.36$, $p < .001$ and explained a significant proportion of variance in mean decision scores, $R^2 = 0.41$, $F(1, 28) = 19.04$, $p < .001$. Difference of number of vertical straight line segments, in contrast, did not significantly contribute to this regression model, $\beta = -0.25$, $t(29) = -1.64$, $p > .1$. In a step-wise forward and backward regression, line of sight was included and number of vertical straight line segments was excluded (in both cases), suggesting that the number of vertical straight line segments (i.e., complexity) did not contribute to participants' decisions over an above the impact of line of sight. On average, participants

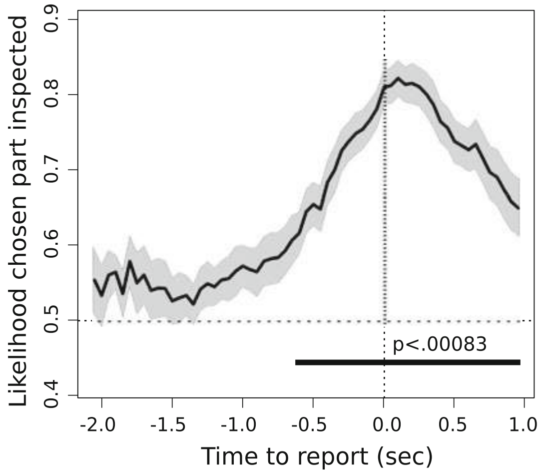

Fig. 4 The likelihood that the observer's gaze was directed towards the chosen part of the image (*left*/*right*) plotted against time (synchronized at time when decision was reported). The *black bar* indicates when the gaze bias was significantly different from chance level. The data represents the average across observers ($n = 20$) and trials ($n = 30$)

chose the path option featuring the longer line of sight in 60% of the cases.

Response times for the different stimuli ranged between 1,793 and 2654 ms (mean 2,277 ms). Marginally significant correlations were observed between response times and both, the total number of edges in the depth profiles ($r = 0.35$, $p = .06$), and the rated complexity of the stimuli ($r = 0.35$, $p = .06$)

Eye movement data

Gaze shifts After leaving the central interest area, participants on average made 6.20 fixations, crossing the mid-line 2.14 times, before reporting their decision. The first 6 fixations were characterized by an increase in eccentricity ($r = 0.87$, $p = .02$) suggesting that participants scanned the images from the center to the edges and inspected both sides of the images before reporting their decision.

Gaze bias The likelihood that observer's gaze was directed towards the eventually chosen part of the stimulus changed over the time course of the trials (see Fig. 4, left). Approximately 700 ms before participants reported their decisions, the likelihood of inspecting the chosen path option significantly increased above chance level, reaching a maximum of 82.18% around the time when decision was reported. Statistical significance was assessed by t test comparisons of gaze bias against chance level (50%) for each of 60 time points (50 ms intervals from 2 s before the response until 1 s after the response). In order to correct for multiple comparisons, alpha level was set at $p < 0.00083$ (Bonferroni adjustment).

Fixation patterns What parts of the scene did participants inspect while making their decisions? The right part of Fig. 5 summarizes fixation patterns for the horizontal and vertical stimulus location separately. Most noticeably the distribution of fixation density along the vertical image position was sharply tuned just below the horizontal center line of the images and there was very little variance in the fixation positions along the vertical position between stimuli. Accordingly, the average correlation between fixation densities along the vertical image dimension recorded for the 30 stimuli was very high ($r = 0.97$). The distribution of fixation density along the horizontal image position, in contrast, was rather broad and there were considerable differences between the stimuli (Fig. 5, right). Differences in fixation patterns between the different scenes were primarily due to differences in the horizontal dimension. Further analysis will therefore focus on the horizontal axis.

The averaged fixation density along the horizontal image location revealed two maxima, left and right of the vertical centerline of the images that relate to the two path options that participants inspected and compared while deciding between them (top right panel of Fig. 5). Figure 6 exemplarily displays the fixation densities along the horizontal position for three single stimuli. A qualitative analysis of fixation behavior for these stimuli suggests that participants paid close attention to the parts of the image in which the lines of sight were particularly long (left and middle example in Fig. 6). Furthermore, fixation densities of the right panel in Fig. 6, in which the longest lines of sight were identical for both choice alternatives, suggests that fixation density was also modulated by aspects of the local spatial complexity.

Taking these qualitative observations into account, a novel approach relating gaze behavior to geometrical properties of the depicted scenes is presented below. Predictions for gaze behavior will be derived by analyzing purely geometrical features of the scene depicted.

Discussion

Which path option do we explore when searching in unfamiliar environments? Participants showed a tendency to choose the option with the longer line of sight. Accordingly, a qualitative analysis of fixation patterns suggests that they paid attention to areas in the environment that featured long lines of sight (see Fig. 6; quantitative analyses of fixation patterns are reported below). This suggests that participants' gaze behavior reflects the significance of long lines of sight for their path choices. While related strategies have been demonstrated in other navigation and route selection studies (e.g., Golledge, 1995; Bailenson, Shum, & Uttal, 2000; Conroy Dalton, 2003), it remains unclear why participants chose the option with the longest line of sight.

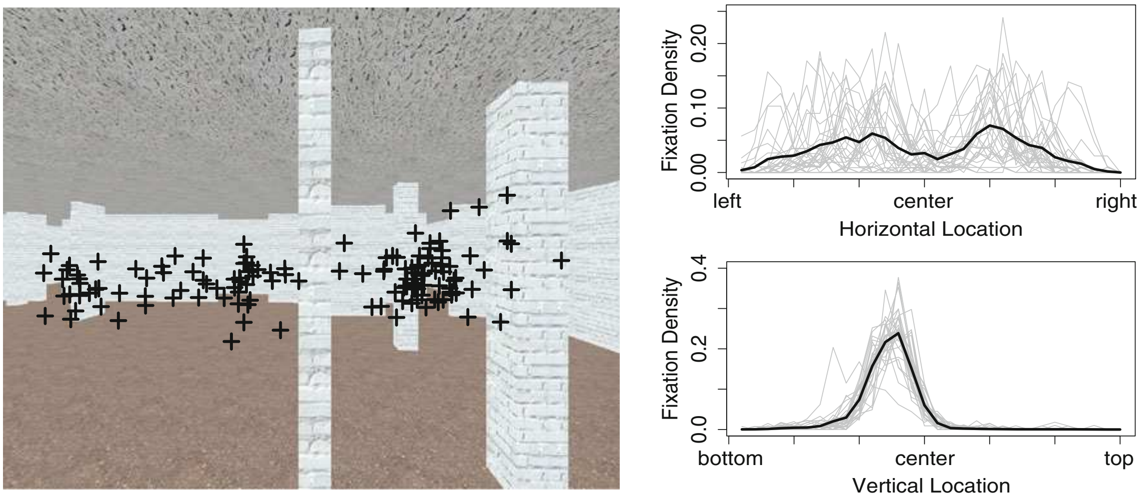

Fig. 5 *Left* Fixation pattern for one of the stimuli; *right* fixation densities for all stimuli for the horizontal (*top*) and vertical (*bottom*) image location. *Grey lines* depict fixation densities for the single stimuli (areas under curve sum up to 1); the *black lines* reflect the average over all 30 stimuli

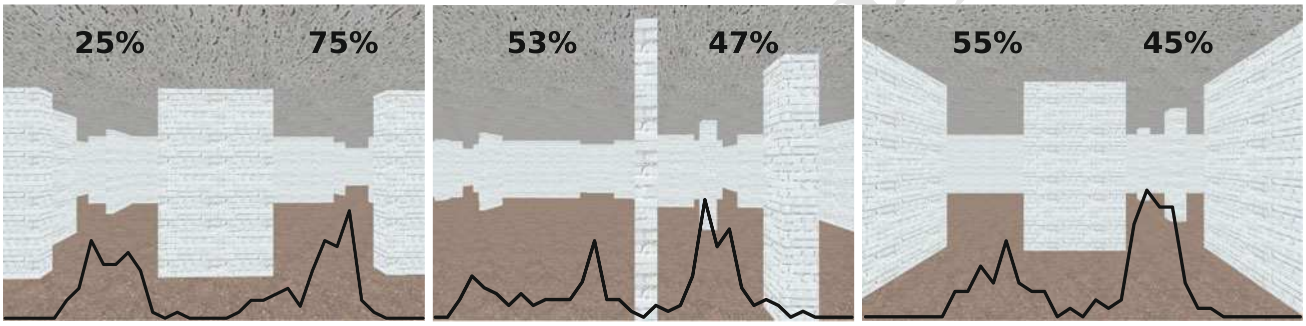

Fig. 6 Exemplary fixation densities plotted on top of three of the stimuli: Fixation densities (*black lines*) are plotted as a function of the horizontal position in the image. The *numbers* relate to the percentages the left or right path option was chosen

A possible explanation is that this path option promises greater information gain when traveling along than the alternative (Zetzsche et al., 2009). However, further research is needed to investigate this behavior.

The analysis of gaze behavior revealed a number of interesting results. First, gaze behavior reflected the spatial decision-making process: approximately 700 ms before observers reported their decisions, the likelihood that they inspected the eventually chosen path option significantly increased above chance level. While this is in line with earlier research on visual decision making (e.g., Shimojo et al., 2003; Simion & Shimojo, 2007), the results for the first time revealed a robust gaze bias effect while making wayfinding decisions. This demonstrates gaze bias effects when comparing alternative options within a single stimulus, rather than when comparing multiple separated stimuli and therefore puts decision making and gaze bias in a more realistic context. Moreover, the results suggest that gaze bias effects in visual decision making are rather broad phenomena that can be found in different tasks (cf. Glaholt & Reingold, 2009).

Experiment 2

Experiment 1 revealed a reliable gaze bias in a spatial search task in which participants were free to decide between options. In that sense, the task was similar to other visual decision-making tasks for which gaze bias effects have been reported (Glaholt & Reingold, 2009). Wayfinding, however, often requires the recognition of spatial situations encountered before and the encoding and decoding of earlier decisions or movements. One prototypical example is that of route navigation whereby the task is to memorize and later retrieve movement decisions at choice points (Trullier, Wiener, Berthoz, & Meyer, 1997). The aim of Experiment 2 was to investigate gaze behaviour in a task that was sought to resemble such route learning. Specific interest concerned whether gaze bias effects can also be observed in situations in which participants were instructed about their choice rather than making them and when recalling these earlier choices. Additionally, we were interested in the environmental features that participants attended to during encoding and retrieval.

Methods

Participants

Twenty subjects (mean age 24.71 ± 2.87 years) participated in the experiment. They were mostly university students and were paid 8 Euros an hour.

Stimuli

The stimuli used in Experiment 2 were identical to those of Experiment 1.

Procedure

Participants first read a description of the experiment along with a set of instructions stating that their task was to follow and remember a predefined route. The experiment consisted of several routes, each of which was subdivided into an encoding phase and a decoding phase.

In the *encoding phase*, participants were presented with a series of images depicting choice points. Prior to the presentation of each stimulus, participants received a text message on the screen instructing them to either choose the left or right path option. They were told to inspect the stimulus and press the corresponding button (left/right cursor key) as soon as they were convinced that they would remember the depicted scene along with the required response. Before a novel stimulus was presented, participants fixated a small cross in the center of the screen and pressed the 'Space' bar. Each stimulus was presented for 5 s, irrespective of when participants responded. Left and right responses during training were balanced across participants, routes, and stimuli.

Each encoding phase was followed by a *decoding phase* during which participants were presented the same stimuli they inspected during encoding, but in a random sequence. Their task was to repeat the path choices of the encoding phase. Again, each stimulus was presented for 5 s.

The experiment was divided into 5 routes containing 4, 5, 6, 7, and 8 decisions. After the last decision, participants were presented with an image of a gold bar hovering in a small room to indicate the end of the route (see Fig. 1).

Analysis

Behavioral data

For each stimulus presented, in both the encoding phase and the decoding phase, participants' behavioral responses (left/right) along with the corresponding response time was recorded. Correct and incorrect responses were assessed by comparing behavior in the encoding and the decoding phase.

Eye-movement data

The analysis of the eye-movement data was identical to Experiment 1.

Results

Behavioral data

Participants correctly recalled the required path option in 83.98% of the trials. An ANOVA did not reveal a main effect of the number of decisions to remember (ranging between 4 and 8) on memory performance [$F(4, 76.57) = 0.38$, $p = .83$]. Memory performance in the decoding phase could neither be predicted by measures of stimulus complexity [rated complexity ($r = -0.14$, $p = .45$), total number of straight line segments in the depth profiles ($r = -0.09$, $p = .63$), nor by the rated symmetry of the stimuli ($r = 0.23$, $p = .23$)].

Response times were shorter in the encoding phase than in the decoding phase (1,960 vs. 2,199 ms *t* test $t(19) = -2.49$, $p = .02$). Response times during the decoding phase was different for correct and incorrect responses (correct: 2,116 ms; incorrect: 2,693 ms; *t* test $t(18) = 4.50$, $p < .001$, one participant was removed from this analysis as he/she provided correct answers only).

Eye movement data

Gaze shifts

The average number of gaze shifts between the three interest areas (left, middle, and right part of the image) was smaller in the encoding than in the decoding phase (1.96 vs. 2.60: *t* test $t(19) = -3.46$, $p < .01$). Furthermore, during decoding participants shifted gaze more often during incorrect than during correct trials (correct: 2.46; incorrect: 3.49, *t* test $t(19) = -3.45$, $p < .01$).

After leaving the central interest area, participants made 5.44 fixations—crossing the mid-line 1.43 times in the encoding phase; and 6.26 fixations—crossing the mid-line 1.91 times in the decoding phase before reporting their decisions. The first 6 fixations after leaving the central interest area were characterized by an increase in eccentricity in both the encoding phase ($r = 0.82$, $p = .05$) and the decoding phase ($r = 0.89$, $p = .02$). As in Experiment 1, these results suggest that participants scanned the images from the center to the edges, inspecting both sides.

Gaze bias

In both experimental phases, participants displayed a significant gaze bias towards the chosen path option

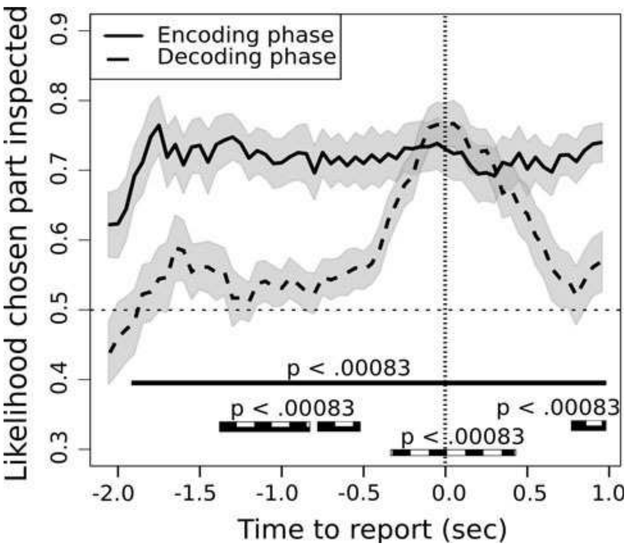

Fig. 7 The likelihood that the observer's gaze was directed towards the chosen part of the image (*left*/*right*) plotted against time (synchronized at time when decision was reported). The *solid black bar* and the *dashed black bar* indicate when the gaze bias in the corresponding conditions was significantly different from chance level. The combined *dashed/solid bar* indicates when the gaze bias effects differed significantly between the encoding and the decoding phase. The data represent the average across observers ($n = 20$) and trials ($n = 30$)

(see Fig. 7). The temporal dynamics of this bias, however, was different: in the encoding phase a significant bias was observed after 1,900 ms, i.e. immediately after stimulus onset (average response time 1,947 ms). This bias remained stable over the entire trial, even after the response was made. In the decoding phase participants initially distributed their gaze evenly across the two alternative path options. A significant gaze bias was observed only 350 ms before the response, reaching its maximum around the time of response and decaying afterwards. A comparison between encoding and decoding phase revealed significant differences in gaze bias both before and after the response was reported (between 1,450 and 500 ms before, and 800 ms after the response was reported; see Fig. 7). Statistical significance was assessed by *t* test comparisons of gaze bias [between conditions or against chance level (50%)] for each of 60 time points (50 ms intervals from 2 s before the response until 1 s after the response). In order to correct for multiple comparisons, alpha was set at $p < .00083$ (Bonferroni adjustment).

In addition, we compared gaze bias effects between experiments (Experiment 1, Experiment 2 encoding, Experiment 2 decoding). An ANOVA revealed a significant main effect of experiment ($F(2, 57) = 11.83$, $p < .001$). Overall, the gaze bias effect was strongest in the encoding phase of Experiment 2 (average bias over the 3 s period: .72), followed by Experiment 1 (.65) and the decoding phase of Experiment 2 (.59). Pairwise comparisons revealed significant differences between all experiments and conditions.

Fixation patterns

Fixation patterns in the encoding phase and decoding phase, as well as between Experiment 1 and Experiment 2 were strikingly similar. Figure 8 superimposes fixation densities along the horizontal dimension for four of the stimuli. The average correlation between the encoding and the decoding phase in Experiment 2 was $r = .92$ for the vertical position and $r = 0.81$ for the horizontal position. The average correlation between Experiment 1 and encoding phase of Experiment 2 was $r = 0.80$ (vertical) and $r = 0.78$ (horizontal). The average correlation between Experiment 1 and the decoding phase of Experiment 2 was $r = 0.86$ (vertical) and $r = 0.76$ (horizontal). These results demonstrate that, independent of the specific tasks and despite clear differences in the time course of gaze bias, visual attention was directed towards the same areas in the images.

Discussion

Performance in the route learning task was very good. Participants correctly recalled the required path choice in almost 84% of the cases. There was no significant effect of route length which was unexpected, as working memory load increases with more decision points to remember. A possible explanation is a floor effect—i.e. even with 8 decisions the memory task itself was fairly simple. While performance in the decoding phase could not be predicted by stimulus properties such as complexity (number of vertical straight line segments in depth profile or rated complexity) or symmetry, response times were correlated with memory performance. Response times in the decoding phase were shorter for correct trials and the number of gaze shifts in the decoding phase was significantly higher for incorrect trials. Most likely, these effects reflect participants' uncertainty in incorrect trials.

In addition to these performance-related effects, task-related influences on gaze behavior were observed. The gaze bias towards the chosen path option had a different temporal dynamic during encoding and decoding. Gaze bias effects did differ between all conditions in Experiments 1 and 2. For example, the gaze bias effect was stronger in Experiment 1 in which decisions were informed by evaluation of the path alternatives than in the decoding phase of Experiment 2 in which decisions were informed by memory retrieval. The temporal dynamic of the gaze bias effect also differed between the encoding and decoding phase of Experiment 2. During encoding, a significant bias was observed immediately after stimulus onset. During decoding, however, a

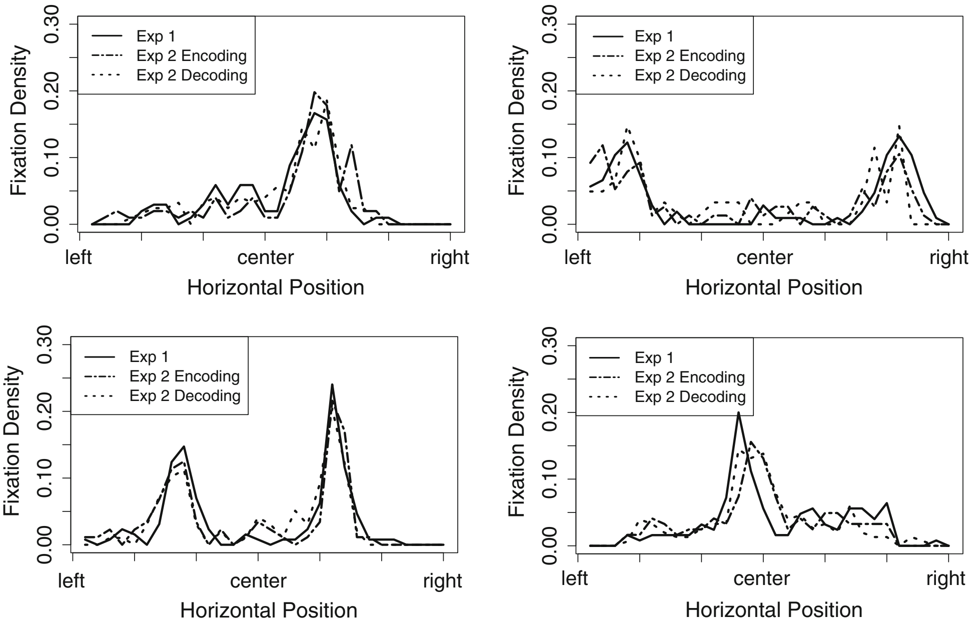

Fig. 8 Four examples of comparisons of fixation densities along the horizontal image location between Experiment 1 and the encoding phase and the decoding phase in Experiment 2

significant bias was observed only 350 ms before responding. In the context of wayfinding research, the early gaze bias during the encoding phase suggests that participants focused on the actual path when learning a guided route, rather than memorizing the entire scene and attaching directional information. The differences in gaze bias between encoding and decoding phase stand in contrast to earlier findings demonstrating similar scanpaths when encoding and later memorizing visual scenes (Humphrey & Underwood, 2008). The fact that participants in this study were not engaged in solving spatial tasks, but merely memorizing images, suggests that the spatial nature of the tasks in current study did in fact result in specific influences on gaze behavior.

The analysis of fixation patterns revealed very consistent patterns between the encoding and the decoding phase in Experiment 2: the fixation densities along both, the horizontal and the vertical image dimension were strongly correlated. This is surprising at first glance, as task-related differences in the time course of the gaze bias have been observed (see above). It can, however, be explained by the fact that the left/right responses were balanced across stimuli: 50% of the participants had to choose the left path option of a particular stimulus, thus focusing primarily on that side during encoding; the remaining participants had to choose the other path option. Fixation patterns were also strongly correlated between Experiment 1 and Experiment 2. While the task-related influences are evident primarily in the temporal characteristics of the gaze bias, the similarity in fixation patterns between experiments and conditions suggests that the image areas participants attended to most convey spatial information crucial to solving wayfinding decision tasks.

Experiment 3: free-viewing control

Motivation and procedure

Before analyzing in more detail which environmental features captured visual attention in Experiments 1 and 2, it is important to demonstrate that the recorded fixation patterns were, in fact, specific to the spatial tasks. To investigate this question a free-viewing control experiment was carried out. Sixteen participants were shown the same 30 stimuli used in Experiments 1 and 2 for 5 s each and their eye movements were recorded. Participants were not given a specific task other than inspecting the pictures they were shown (free-viewing task). Eye-movement data were analyzed as described in Experiment 1 (see Fig. 3), and fixation densities along the horizontal dimension were compared to Experiments 1 and 2. In a structured but open-ended interview after the experiment, participants were asked what they were thinking when inspecting the stimuli. While such subjective reports are sometimes considered unreliable, recent work demonstrates that individual fMRI time series and eye-tracking data recorded during wayfinding experiments closely matched events from retrospective verbal reports (Spiers & Maguire, 2006, 2007). This suggests that in the context of wayfinding, verbal reports can provide important and reliable information about the tasks that participants are engaged in.

Results and discussion

As in Experiments 1 and 2 fixations were sharply tuned along the vertical dimension to an area just below the horizon, but varied considerably along the horizontal image location. Average correlations for fixation densities along the horizontal dimension for the 30 stimuli between the control experiment and Experiments 1 and 2 were $r = 0.74$ demonstrating that the image areas inspected in this free-viewing condition were again similar to those in Experiments 1 and 2. As participants were not instructed to solve wayfinding decision-making tasks this suggests that the fixation patterns recorded in Experiments 1 and 2 were not specific to the spatial tasks. An analysis of the interviews after the experiments, however, offers an alternative explanation: 11 out of the 16 participants reported that they felt as if being in a maze when inspecting the stimuli. More importantly, 14 participants explicitly reported that they looked for path options and thought about how to navigate if they were actually given the chance to move (for examples of responses, see Appendix). In other words, 87.5% of the participants reported that they were spontaneously engaged in path choice tasks: they analyzed the spatial scene, inspected and compared the available path options. This is most likely due to the nature of the stimuli that afford a corresponding interpretation. Participants spontaneously assign a purpose and implicit task to the stimulus materials. These additional interpretations can hardly be conceptualized as low-level data-driven attentional processes, but rather appear to reflect high-level cognitive strategies that purposefully guide eye movements.

Experiment 4: free-viewing vertical control

Motivation and procedure

In Experiment 3, we demonstrated high correlations between the fixation patterns of Experiments 1 and 2 in which participants solved spatial tasks and the free-viewing condition of Experiment 3. While participants' retrospective reports suggest that they were engaged in spatial decision making in Experiment 3 even without being instructed to do so, an alternative explanation is that fixation patterns of Experiments 1, 2, and 3 were primarily driven by low-level image features and were not specific to the spatial tasks at all. In order to distinguish between these explanations we rotated the images by 90° in Experiment 4. This manipulation makes it less likely that participants interpret the images as spatial situations and should thus prevent them from spontaneously engaging in wayfinding tasks. At the same time the basic image features are not influenced by the manipulation. If participants' gaze behaviour was primarily driven by low-level image features, we expect similar fixation patterns as in Experiments 1 and 2, rotated by 90°. If, however, the fixation patterns in Experiments 1 and 2 were specific to the spatial tasks, we expect a clear reduction of correlation between fixation patterns of Experiments 1 and 2 and those in this control experiment. 18 participants were shown the same 30 stimuli used in Experiments 1 and 2 for 5 s each and their eye movements were recorded. As in the first control experiment, participants were not given a specific task other than inspecting the pictures they were shown (free-viewing task). For the analysis, eye-movement data was first rotated back by 90° and then analyzed as described in Experiment 1.

Results and discussion

Average correlations for fixation densities along the horizontal image dimension for the 30 stimuli between Experiment 4 and Experiments 1 and 2 were $r = 0.27$. This is a significant reduction as compared to the correlation between fixation patterns of Experiment 3 and Experiments 1 and 2 and [$r = 0.74$ vs. $r = 0.27$; $Z = 2.48$; $p = .01$ (two-tailed)]. These results demonstrate that rotating the experimental stimuli by 90° influenced gaze behavior, irrespective of the fact that the local image features were not manipulated, strongly suggesting that the fixation patterns of Experiments 1 and 2 do not solely result from low-level image features but were in fact specific to the spatial tasks that required analyzing the spatial scene depicted.

Wayfinding decision making and gaze behavior

While the coarse spatial layout of a scene can be computed without the need for selective visual attention (e.g., Oliva & Torralba, 2001, 2006), it remains an open question upon which environmental features such computation may rely. Or, in the context of this research: which features of architectural environments are being processed during space perception and wayfinding decision making? The stimuli used here are particularly suited to approach these questions since they predominantly depict spatial—i.e. geometrical—information. Moreover, due to the image manipulation the saliency of image areas depicting more distant parts of the environments, including openings and further path options, was selectively reduced. Fixations towards these areas are therefore unlikely to result from stimulus properties such as saliency. Rather these areas were inspected because they convey information crucial for comprehending the spatial situation and deciding between path alternatives. If that was true it should be possible to predict the gaze behavior recorded in Experiments 1 and 2 by analyzing purely geometrical properties of the scenes inspected.

Image saliency (Saliency Toolbox: Walther & Koch, 2006) did, in fact, correlate only weakly with fixation density: similar to the analysis presented in Experiment 1, the saliencies for the horizontal and vertical dimension were

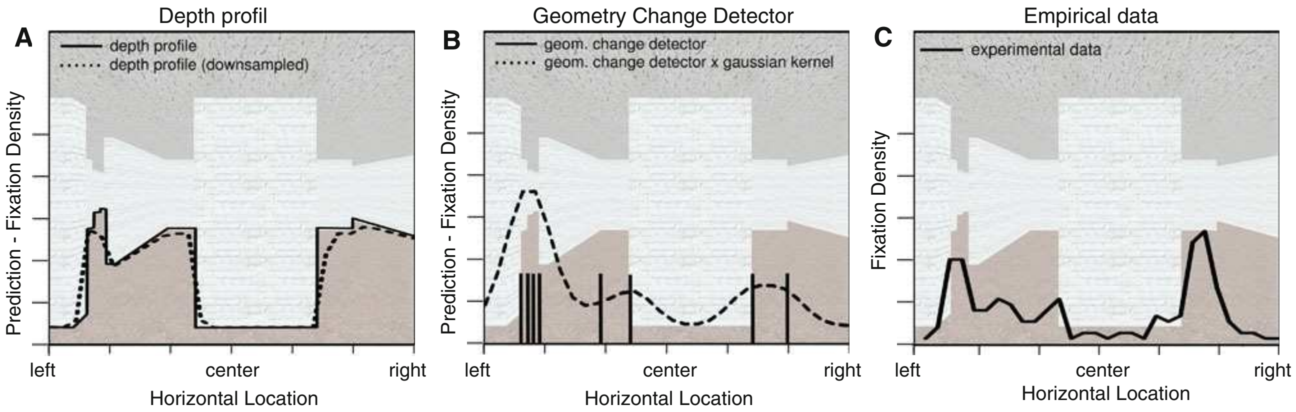

Fig. 9 Predictions for fixation densities of the two predictors (**a** Depth profile, **b** geometry change detector) for an exemplary stimulus and the corresponding empirical data (**c**)

calculated separately for each stimulus and were then correlated with the recorded fixation densities. The average correlation (after Fisher Z transformation) between the saliency and the horizontal fixation density was $r = -0.17$ and $r = 0.14$ for the vertical location, suggesting that saliency did not capture the image properties that guided visual attention in this particular setting. This, however, was not too surprising as saliency maps were designed for complex naturalistic scenes. Here, in contrast, we used screenshots of very simplistic VR sceneries that lack specific objects. Furthermore, by manipulating the images, we removed some stimulus features that saliency maps are particularly sensitive to, such as the high-contrast edges between walls and ceiling/floor.

What image areas did the participants inspect when making wayfinding decisions in the current setting? A qualitative analysis of gaze distributions suggests that participants were mostly attending to stimulus areas that allowed for long lines of sight (see Fig. 6). Given the wayfinding decision-making task in the experiments this appears to be a sensible viewing strategy, simply because these areas depict the available path options. There is, however, another explanation: rather than merely processing image areas depicting more distant parts of the environment, observers attend to areas featuring changes in the local geometry as generated by occlusions, openings, and further path options. Architectural theory suggests that these structural elements are of particular importance for wayfinding as they promise to reveal novel information about the environment (Peponis et al., 1997). In order to distinguish between these alternatives we developed quantitative measures that extract the according information directly from the stimuli. The first measure—the depth profile—codes egocentric distance information from the beholder to the surrounding walls, thus directly capturing the length of the lines of sight in different image areas (Fig. 2). The second measure—the geometry change detector—identifies image areas at which the local geometry of the depicted space changes, therefore highlighting areas with wall boundaries, openings and further path options. The predictive power of these two basic geometrical features for gaze behavior are investigated by correlating them to the gaze distributions measured in Experiments 1 and 2. We focus on fixation distributions along the horizontal image dimension, as fixation probabilities along the vertical dimension hardly varied between stimuli (Fig. 5).

Depth profile

For each of the 30 stimuli the depth profiles were extracted (see "Methods" and Fig. 2). Information about which parts of the visual scene allow for long lines of sight is directly coded in the height or elevation of the depth profile. In order to derive predictions for participants' gaze distribution, the resolution of the depth profile was reduced from 1,024 horizontal bins (the images were 1,024 × 768 pixel) to 30 horizontal bins by averaging the height of the depth profile over the respective bins (see Fig. 9a).

Geometry changes

Positions at which the geometry changes locally are captured by orientation changes in the depth profile (see Fig. 9). Image areas featuring many geometry changes mark locations with high spatial information density.[1] Note that not the point at which the orientation changes itself but this point together with its context—i.e. the immediate surrounding—carries information about how the geometry changes. Hence, a Gaussian kernel (with a half-amplitude width of $\sigma = 1.6$ bins, relating to 2° visual angle) was applied to the individual geometry changes. This half-amplitude

---

[1] The 2D-detector described by Zetzsche and Barth (1990) would highlight similar areas.

width was chosen as it roughly corresponds to the area of foveal acuity thus capturing the local geometry that could be perceived without further eye movements.[2]

As participants' fixation patterns were highly consistent between Experiments 1 and 2, empirical data were pooled. The average correlation over all 30 stimuli between the predictions (depth profile, geometry changes) and the empirical data (correlation coefficients were Fisher's Z transformed for averaging) was $r = 0.48$ for the depth profile and $r = 0.64$ for the geometry change detector. Given that earlier studies reported strong central fixation biases independent of viewing task and image features (e.g., Tatler, 2007), we added eccentricity as a predictor. We then fitted a mixed effects model in R, using the lmer function in the lme4 package (Bates, 2005), with the log odds of fixation probability in each horizontal bin as dependent variable, and with depth profile, geometry changes, and eccentricity (squared distance from center) as fixed factors. Stimulus was used as random factor. Removing random slopes was justified ($X^2(2) < 1$, ns.). All fixed factors were significant (all $ts > 10$). Not surprisingly, eccentricity turned out to be the strongest predictor (cf. Tatler, 2007), followed by geometry change and depth profile (eccentricity: estimate $-.006$, standard error .0003, $t = -20.38$; geometry change detector: estimate = 13.37, standard error = 0.72, $t = 18.61$; depth profile: estimate = 11.20, standard error = 1.01, $t = 11.15$).

Taken together, these results demonstrate that simple analyses of geometrical features allow predicting fixation patterns during space perception and wayfinding decision making for the stimuli in this study. Together with results from Experiment 4, this corroborates the notion that the gaze behavior in Experiments 1 and 2 is the result of the spatial nature of the wayfinding decision tasks and that it relates to the analysis of the geometry and the comparison of path alternatives. Moreover, the results suggest that participants were attending image areas depicting changes in the local geometry more than image areas allowing for long lines of sight. These changes are not only related to plain corners, but also to occlusions, openings, and further path options which have to be inspected and compared when making wayfinding decisions (cf. Peponis et al., 1997). Further studies in our lab will address to what extent these results generalize to wayfinding decision-making tasks that are presented with richer and more complex naturalistic scenes.

General discussion

The overall aim of this study was to investigate the relation of gaze behavior and decision making in the context of wayfinding. In two experiments, participants were presented with images of choice points displaying two path options and were either asked to decide between them in order to search for an object that was hidden in the environment (Experiment 1), or they were instructed how to decide at each decision point and were later asked to reproduce these decisions (Experiment 2). While the first task sought to resemble a search or exploration task in a novel environment, the latter task sought to resemble a route learning task, in which a navigator is first led along a route and is later asked to recall the path choices at different branching points. Both experiments revealed reliable gaze bias effects towards the chosen path option. In addition, the fixation patterns were very similar between experiments, suggesting that participants inspected the same environmental features while choosing between path options. In two control experiments we ensured that the fixation patterns recorded in Experiments 1 and 2 were in fact driven by the spatial tasks and not by other low-level image features.

The gaze bias effects observed in Experiments 1 and 2 extend earlier results on visual decision making in non-spatial domains (e.g., Shimojo et al., 2003) and demonstrate that gaze behavior reflects spatial decision making also in the context of wayfinding. While the experiments were not designed to investigate whether gaze bias merely reflects decisions or is also influencing them (cf. gaze-cascade model, Shimojo et al., 2003), the results provide additional evidence for the notion that gaze bias effects constitute a rather general phenomenon of decision making that are found in different tasks (Glaholt & Reingold, 2009). The time course of the gaze bias effects in the experiments was modulated by the task. For example, in the encoding phase of Experiment 1, we observed a strong gaze bias directly after stimulus onset that remained stable even after participants responded. In the decoding phase of Experiment 2, in contrast, a significant gaze bias effect was observed much later, only 350 ms before the participants responded. The differences in gaze bias between conditions of Experiment 2 are interesting, as they are in contrast to recent work reporting similar scan-paths between encoding and decoding of natural scenes (Humphrey & Underwood, 2008; Foulsham & Underwood, 2008). Participants in these studies, however, were not solving spatial tasks but were only instructed to memorize the scenes.

The task-related modulation of the gaze bias strongly suggests that the analysis of gaze behaviour can be a promising means to reveal novel insights into the information processing underlying specific wayfinding tasks. There is, for example, a broad consensus that route knowledge can be conceptualized as a sequence of stimulus–response pairs or recognition-triggered responses (e.g., Trullier et al., 1997; Kuipers, 2000; Mallot & Gillner, 2000). It is, however, an open question which information exactly is memorized to

---

[2] Shape information can be perceived even parafoveally. Therefore the small amplitude width can be considered conservative.

recognize the choice points along the route. As the current virtual environments were lacking any objects that could have served as landmarks, participants had to rely on geometrical information. One possibility is that they memorized the choice points by encoding the entire stimulus or spatial situation (cf. snapshot-like memories, Cartwright & Collett, 1982; Gillner, Weiss, & Mallot, 2008); a second possibility is that participants primarily focused on the path option belonging to the actual route. In the encoding phase of Experiment 1 we observed a strong gaze bias towards the path option that had to be chosen directly after stimulus onset. This suggests that rather than memorizing the entire scene and attaching directional information, participants memorized the visual appearance of the actual path option they had to choose. This encoding strategy does not require the explicit encoding of directional information and is similar to a beacon strategy in which the navigator moves towards a recognized landmark (Waller & Lippa, 2007).

Which information is depicted in the image areas that captured visual attention in Experiments 1 and 2? Behavioral results of Experiment 1 as well as qualitative analyses of fixation patterns in both experiments suggested that local geometric features, such as the line of sight and changes in the geometry had an influence on what people did and where they looked. We therefore analyzed the geometry of the scenes and related it to participants' fixation patterns. Inspired by isovist analysis (Benedikt, 1979), the visible geometry of the scenes was described by depth profiles from which also salient geometrical features were derived. Specifically, the areas in the depth profile that correspond to local geometry changes were detected. This geometry change detector highlights structures such as corners, openings, and occlusions and had a strong predictive power for where people were looking when inspecting the two-dimensional projections of three-dimensional complex architectural indoor scenes. These results are in line with architectural theory emphasizing the crucial role of these structural elements for wayfinding (Peponis et al., 1997). In real-world wayfinding, both geometric structure features and object and landmark features interact to inform spatial decision making. Recently, Frankenstein, Büchner, Tenbrink, and Hölscher (2010) demonstrated the impact of sematically rich landmarks alongside geometric information. Current research in our labs addresses how well the current results generalize to visual attention in more complex natural scenes.

It is noteworthy that results of Experiment 1 revealed how the line of sight was a stronger predictor for participants' movement decisions than the number of vertical straight lines in the depth profiles. For gaze behaviour, however, changes in the local geometry (points between the straight line segments in the depth profile) were a stronger predictor than the depth profile itself (coding the length of the line of sight). In other words, geometry changes are good predictors for fixation patterns, but weaker predictors for movement decisions, while line of sight is a good predictor for decisions, but a weaker predictor for fixation patterns. Concentrating future research on this double dissociation could further inform our understanding of the intricate relation between gaze behaviour and decisions.

Which parts of the scene did participants attend to when inspecting the stimuli and while deciding between alternative path options (Experiment 1) or while encoding or decoding path choices (Experiment 2)? For all stimuli, fixations were narrowly tuned along the vertical axis, i.e. viewers focused their attention around the horizon. This is a sensible viewing strategy in architectural indoor spaces in which the area around the horizon—as opposed to the floor and the ceiling—provides most information about the local geometry. Such a viewing strategy suggests that participants were not merely responding to areas with high visual complexity or saliency, but were actually interpreting the three-dimensional structure of the sceneries depicted. This is further corroborated by a comparison of the saliency of the 30 stimuli with the actual fixation densities. Fixation densities along the horizontal axis systematically differed between stimuli, suggesting that participants directed their attention to specific areas in the environment. At the same time, participants' fixations patterns were similar between Experiment 1, 2 and 3. In other words, participants inspected the same areas, irrespective of the specific spatial task. A likely explanation for this is that these image areas are crucial for the comprehension and interpretation of the geometry of the depicted space. This is further corroborated by Experiment 4 in which participants viewed the same scenes but rotated by 90°. These images could not be as easily decoded as depicting spatial scenes and accordingly, fixation patterns did not closely match those of Experiments 1 and 2.

Taken together, the experiments in this paper show that gaze behavior is directed towards environmental features that have been suggested as decisive for spatial learning and decision making in navigation-related tasks. The attentional patterns are guided both by features of the environment and by properties of the task at hand, be it route learning or spatial exploration. It is an important issue for spatial cognition research to untangle how environmental information, spatial abilities and cognitive strategies contribute to effective human navigation behavior. The paper underscores that spatial attention, measured via gaze behavior and fixation patterns, connects environmental features to spatial decision making. Overall, the results of this work suggest that the integrated analysis of navigation decision behavior and gaze behavior can play a key role in the investigation of the information processing and the cognitive strategies underlying human wayfinding.

Acknowledgments This work was supported by the Volkswagen Foundation (Tandem project: “Wayfinding Strategies in Behavior and Language”) and the SFB/TR8 “Spatial Cognition”. Special thanks to J. Wendler, J. Henschel, and A. Günther for their help in carrying out the experiment and analyzing the data.

Appendix

Control Experiment 1: structured interviews

Excerpts of participants’ answers in the interviews after the control experiments are reported below. Specifically, the interviewer asked participants what they thought while inspecting the images in the control experiments. Here parts were selected that relate to the notion that participants spontaneously engaged in a wayfinding task.

Participant 1: I noticed the possible routes and paths that could have been taken if I was in the environment. I focused on exploring these routes as this had more relevance and was of higher interest than the ceiling and walls.

Participant 2: Where is the way out?

Participant 3: Looked at it as a Labyrinth and thought about which way I would go/proceed to find my way. I was looking for ways out.

Participant 4: I was looking for doors, exits. It reminded me of displays of museums, where you walk around. How would I walk around to look at the displays that would be on the walls

Participant 5: I looked for openings, like a clear path. It reminded me on playing computer games when I was growing up egoshooter. I was looking as if I was inside the room, it seemed like a maze, I was looking for a clear path

Participant 6: It reminded me of a maze. I was looking for exits and entrances, sort of paths/ways to move on

Participant 7: I mentally walked through the corridors, thought about how to go on

Participant 8: First I attended to the center of the room, then I looked around

Participant 9: It was some kind of maze—find your way—remember where I am—I tried to remember what was the configuration of walls. Ive looked what the configuration of walls looks like!

Participant 10: I was looking how to go. It looked like you have to find your way in the labyrinth. I looked for ways to get around the obstacle in the middle! I analyzed the pattern of this place, I imagined standing in the labyrinth and thought about how to move in it. Where is the farest points in space?

Participant 11: I was looking for spaces to walk down/ it was like a maze, I realized there were two path options

Participant 12: I was looking for ways out. It reminded me of a maze, in a maze I look at everything , I looked for the layout. There were always two path choices in front and then more distant ones I inspected the distant path options, thinking about where I was gonna go.

Participant 13: I felt like I was in a maze, it looked familiar. When I saw there was a gap in the wall I wanted to see what was here. There were walls and spaces and I wanted to see behind them. I could have gone in different places

Participant 14: At first I was looking at the walls, then I thought if I was in that scene how would I escape I was looking for which way looked safest. There were differences between right and left side. I was seeing it as if it was a maze, of which I was in the middle and deciding which way to go!

Participant 15: (translated from German) I fixated the points that were furthest away—if somebody came around the corner here, I would detect that immediately. When something was directly ahead of me, I looked at the wall and surveyed the surroundings . This looked like a maze, checked how somebody might come into the room, not how I could escape [semi-professional gamer, 300 days of gaming in 4 years].

Participant 16: I looked mainly at the entrances/exits, it felt like a maze , I was looking for my way out.

References

Allen, G., & Kirasic, K. (2003). Visual attention during route learning: A look at selection and engagement. In *Cosit 2003* (pp. 390–398). Berlin: Springer-Verlag.

Armel, C., Beaumel, & A., Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgment and Decision Making, 3*(5), 396–403.

Bailenson, J. N., Shum, M. S., & Uttal, D. H. (2000). The initial segment strategy: A heuristic for route selection. *Memory Cognition, 28*(2), 306–318.

Bates, D. (2005). Fitting linear mixed models in r. *R News, 5*, 27–30.

Benedikt M. L. (1979) To take hold of space: Isovists and isovist fields. *Environment and Planning B, 6*, 47–65.

Cartwright, B.A., & Collett, T.S. (1982) How honey bees use landmarks to guide their return to a food source. *Nature, 295*, 560–564.

Castelhano, M., Mack, M., & Henderson, J. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision, 9*(3), 1–15.

Conroy Dalton, R. (2003). The secret is to follow your nose: Route path selection and angularity. *Environment & Behavior, 35*(1), 107–131.

Foulsham, T., Underwood, G. (2008) What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision, 8*, 1–17.

Frankenstein, J., Büchner, S., Tenbrink, T., & Hölscher, C. (2010) Influence of geometry and objects on local route choices during wayfinding. In Hölscher, C., Shipley, T., Belardinelli, M., Bateman, M., & Newcombe, N. (Eds.), *Spatial cognition vii* (Vol. 6222, pp. 41–53).

Franz, G., & Wiener, J. (2008). From space syntax to space semantics: a behaviorally and perceptually oriented methodology for the efficient description of the geometry and topology of environments. *Environment Planning B: Planning and Design, 35*(4), 574–592.

Gillner, S., & Mallot, H.A. (1998). Navigation and acquisition of spatial knowledge in a virtual maze. *Journal of Cognitive Neuroscience, 10*(4), 445–463.

Gillner, S., Weiss, A.M., & Mallot, H.A. (2008) Visual homing in the absence of feature-based landmark information. *Cognition, 109*, 105–122.

Glaholt, M. G., & Reingold, E. M. (2009) The time course of gaze bias in visual decision tasks. *Visual Cognition, 17*(8), 1228–1243.

Golledge, R. (1995). Path selection and route preference in human navigation: A progress report. In: Frank, A., Kuhn, W. (Eds.) *Spatial information theory: A theoretical basis for GIS (COSIT'95), Lecture Notes in Computer Science: 988* (pp. 207–222) Berlin: Springer.

Grasso, R., Prevost, P., Ivanenko, Y., & Berthoz, A. (1998). Eye-head coordination for the steering of locomotion in humans: an anticipatory synergy. *Neuroscience Letters, 253*, 115–118.

Henderson, J. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences, 7*(11), 498–504.

Henderson, J., Malcolm, G., & Schandl, C. (2009). Searching in the dark: cognitive relevance drives attention in real-world scenes. *Psychonomic Bulletin & Review, 16*, 850–856.

Henderson, J., Weeks, P., & Hollingworth, A. (1999). Effects of semantic consistency on eye movements during scene viewing. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 210–228.

Hoiem, D., Efros, A., & Hebert, M. (2007). Recovering surface layout from an image. *International Journal of Computer Vision, 75*(1), 151–172.

Humphrey, K., & Underwood, G. (2008). Fixation sequences in imagery and in recognition during the processing of pictures of real-world scenes. *Journal of Eye Movement Research, 2*(2), 1–15.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*, 1489–1506.

Itti, L., & Koch, C. (2001) Computational modelling of visual attention. *Nature Reviews Neuroscience, 2*, 194–203.

Kuipers, B. (2000). The spatial semantic hierarchy. *Artificial Intelligence, 19*(1–2), 191–233.

Land, M., & Lee, D. (1994) Where we look when we steer. *Nature, 369*(6483), 742–744.

Land, M., & Tatler, B. (2001). Steering with the head. the visual strategy of a racing driver. *Current Biology, 11*, 1215–1220.

Mallot, H. A., & Gillner, S. (2000). Route navigating without place recognition: What is recognised in recognition-triggered responses?. *Perception, 29*(1), 43–55.

Montello, D. R. (2001). Spatial cognition. In *International encyclopedia of the social & behavioral sciences* (pp. 14771–14775). Oxford: Pergamon Press.

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: a holistic representation of the spatial envelope. *International Journal of Computer Vision, 42*(3), 145–175.

Oliva, A., & Torralba, A. (2006). Building the gist of a scene: the role of global image features in recognition. *Progress in Brain Research, 155*, 23–36.

Ooi, T., Wu, B., & He, Z. (2001). Distance determined by the angular declination below the horizon. *Nature, 414*, 197–200.

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research, 42*, 107–123.

Peponis, J., Wineman, J., Rashid, M., Kim, S.H., & Bafna, S. (1997). On the description of shape and spatial configuration inside buildings: convex partitions and their local properties. *Environment and Planning B: Planning and Design, 24*(5), 761–781.

Readinger, W. O., & Chatziastros, A., Cunningham, D. W., Bülthoff, H. H., Cutting, J. E. (2002). Gaze-eccentricity effects on road position and steering. *Journal of Experimental Psychology: Applied, 8*(4), 247–258.

Rothkopf, C.A., Ballard, D.H., & Hayhoe, M.M. (2007). Task and context determine where you look. *Journal of Vision, 7*, 1–20.

Schuchard R., Connell B., & Griffiths P. (2006) An environmental investigation of wayfinding in a nursing home. In *Proceedings of the 2006 symposium on eye tracking research & applications, san diego ca* (p. 33).

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience, 6*, 1317–1322.

Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 10, pp. 9–55). New York: Academic Press.

Simion, C., & Shimojo, S. (2007): Interrupting the cascade: Orienting contributes to decision making even in the absence of visual stimulation. *Perception & Psychophysics, 69*(4), 591–595.

Spiers, H., & Maguire, E. (2006). Thoughts, behaviour, and brain dynamics during navigation in the real world. *Neuroimage, 31*(4), 1826–1840.

Spiers, H., & Maguire, E. (2007). Decoding human brain activity during real-world experiences. *Trends in Cognitive Sciences, 11*, 356–365.

Spiers, H., & Maguire, E. (2008). The dynamic nature of cognition during wayfinding. *Journal of Environmental Psychology, 28*, 232–249.

Stankiewicz, B. J., & Kalia A. A. (2007). Acquisition of structural versus object landmark knowledge. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 378–390.

Tatler, B. W. (2007). The central fixation bias in scene viewing: selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision, 7*, 1–17.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. *Psychological Review, 113*, 766–786.

Trullier, O., Wiener, S. I., Berthoz, A., & Meyer, J. (1997). Biologically based artificial navigation systems: Review and prospects. *Progress in Neurobiology, 51*, 483–544.

Vembar, D., Iyengar, N., Duchowski, A., Clark, K., Hewitt, J., & Pauls, K. (2004). Effect of visual cues on human performance in navigating through a virtual maze. In *Eurographics symposium on virtual environments*.

Waller, D., & Lippa, Y. (2007). Landmarks as beacons and associative cues: their role in route learning. *Memory Cognition, 35*, 910–924.

Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks, 19*, 1395–1407.

Wiener, J., Büchner, S., & Hölscher, C. (2009). Towards a taxonomy of wayfinding: A knowledge-based approach. *Spatial Cognition and Computation, 9*(2), 152–165.

Wiener, J., Franz, G., Rossmanith, N., Reichelt, A., Mallot, H., & Bülthoff, H. (2007). Isovist analysis captures properties of space relevant for locomotion and experience. *Perception, 36*(7), 1066–1083.

Wilkie, R., & Wann, J. (2003). Eye-movements aid the control of locomotion. *Journal of Vision, 3*, 677–684.

Yarbus, A. (1967). *Eye movements and vision*. New York: Plenum

Zetzsche C., & Barth E. (1990) Fundamental limits of linear filters in the visual processing of two-dimensional signals. *Vision Res, 30*, 1111–1117.

Zetzsche, C., Galbraith, C., Wolter, J., & Schill, K. (2009) Representation of space: Image-like or sensorimotor?. *Spatial Vision, 22*(5), 409–424.