The role of motion parallax and landmark information in place recognition

Bachelorarbeit
der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen

vorgelegt von

Nina Walker

Tübingen, Oktober 2015

Erklärung

Hiermit erkläre ich,

- dass ich diese Arbeit selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.
- Dass die eingereichte Arbeit weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahrens gewesen ist.

Tübingen, den

Danksagung

Hiermit möchte ich mich an erster Stelle bei Marc Halfmann und seiner immerwährenden Geduld und Beratung bedanken. Zudem möchte ich mich auch bei Herrn Professor Mallot bedanken, der durch seine Ideenfindung und Beratung diese Arbeit erst möglich gemacht hat. Außerdem geht mein Dank an die gesamte Arbeitsgruppe der Kognitiven Neurowissenschaft, welche einen durch ihre freundliche und verständnisvolle Atmosphäre willkommen hieß.

Des Weiteren möchte ich mich bei meiner Familie bedanken, die mich immer unterstützt und beraten hat. Und natürlich bei meinen Freunden, welche mich durch mein gesamtes Studium begleitet, verstanden und immer wieder aufgebaut haben. Es war eine wundervolle und unvergessliche Zeit!

Table of content

Zusammenfassung

In dieser Arbeit wurden Probanden in einem sogenannten „homing task" getestet, bei welchem ihnen nur Tiefeninformationen zu Verfügung standen um sich in einem virtuellen Raum zu orientieren. Dabei wurden Tiefenwahrnehmungs-cues in zwei Experimenten getestet. Im ersten Experiment wurde getestet, ob die Probanden den zuvor gezeigten Zielpunkt wieder fanden, wenn ihnen nur eine monokulare Ansicht zur Verfügung stand. Dabei wollten wir den Einfluss der Bewegungsparallaxe testen. Im zweiten Experiment wollten wir den Einfluss von Landmarkeninformation bzw. das Fehlen dieser testen. Hierbei war den Probanden eine binokulare Ansicht gestattet. Beide Experimente wurden an einem Stereoskop durchgeführt, welches einen virtuellen drachenförmigen Raum zeigte. Dieser drachenförmige Raum wurde dabei auf zwei verschiedene Arten dem Probanden präsentiert: einmal mit einer texturierten Tapete welche aus großen weißen Punkten auf schwarzem Hintergrund bestand und damit Farbgradienten also Hilfestellung bot, und zum anderen mit sogenannten „random dots", flackernden weißen kleinen Punkten auf schwarzem Hintergrund, sodass nur noch Tiefeninformationen als Hilfsmittel zur Orientierung dienten. Unsere Experimente zeigten, dass Bewegungsparallaxe alleine nicht ausreichend war um den „homing task" befriedigend zu absolvieren, demzufolge schränkte das Fehlen der binokularen Informationen die Performance stark ein. Im Gegensatz dazu, schränkte das Fehlen jeglicher Landmarkeninformationen die Probanden nicht ausschlaggebend ein. Somit zeigen unsere Ergebnisse, dass Landmarkeninformationen nicht notwendig sind während eine binokulare Sicht erforderlich ist.

Abstract

In this study the participants were tested on their ability to fulfil homing tasks when only depth cues are given. We tested the influence of two factors in two experiments. In the first experiment we wanted to test, whether the participants were still able to find the previously shown target point while only having monocular cues. Therefore we tried to test the influence of motion parallax. In the second experiment we wanted to test the influence of landmark information that is the absence of such information. During this second experiment the participants were allowed a binocular view. Both experiments were performed at a single-mirror stereoscope which showed a kite-shaped room. The virtual kite-shaped room was presented in two conditions: one with textured wallpaper, so that intensity information was also available, the second one with so-called random dot wallpaper, where no intensity information was available.

Our experiments indicate that the participants needed more information than motion parallax cues to fulfil the task and the absence of binocular cues limited their ability greatly. Contrary to this the absence of any landmark information did not limit participants' ability greatly. Therefore landmark information is not necessarily needed for homing tasks.

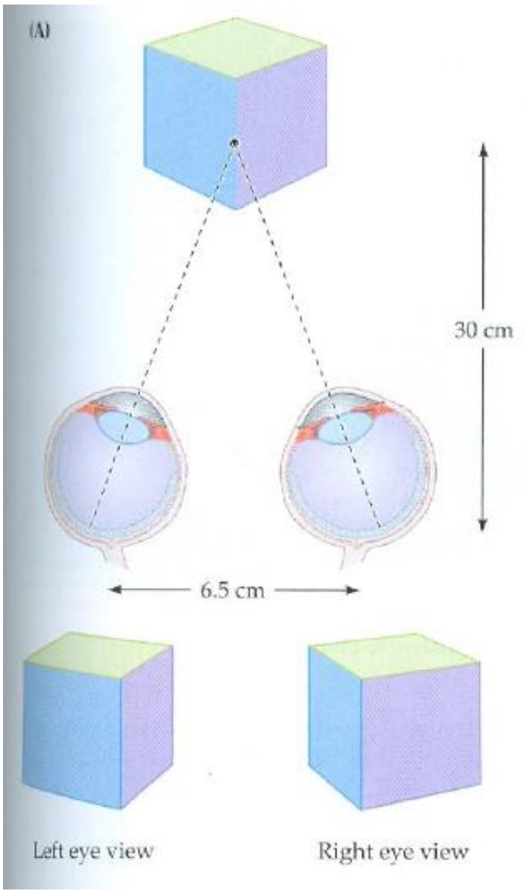
Introduction

Place recognition plays an important role in spatial cognition and during most navigational tasks. Nearly all species have to be able to navigate through their environment for different reasons including to find food sources, mating spots or their lair. Squirrels for example make use of recognizing a previously visited spot when trying to find hidden food during winter months. Place recognition is based on one's ability to remember certain features at the place's location. These features are usually so called landmarks. Most of the research of place recognition and navigation tasks is based on landmark recognition (e.g. Cartwright & Collett, 1983; Graham & Collett, 2002). Landmarks are defined in many different ways. Some have described landmarks as actual physical objects (Lynch, 1960), while others say that landmarks are not necessarily single objects but the configuration of several objects (Siegel & White, 1975). Some even go as far as calling depth estimates from an egocentric view landmark information (Hermer & Spelke, 1994). There has also been a discussion whether landmarks have to be consciously recognized by the observer. But the most important characteristics of landmarks are that they have to be recognizable, memorable and immobile. Additionally they should be relevant while making navigational decisions (Gillner, Weiß, & Mallot, 2008).

Another cue that can be used for place recognition is snapshot memory. Snapshot memory is best known to play a big role in insects' navigation. Honey bees for example use snapshot memory for homing tasks (Cartwright & Collett, 1983) as well as wood ants which use snapshot memory when finding previous food sites, even recalling those snapshots before reaching the destination (Durier, Graham, & Collett, 2003). But it was also proven to be used by humans during place recognition tasks (Gillner, Weiß, & Mallot, 2008). Snapshots are memorized pictures of the surrounding environment that can include several landmarks and the space surrounding them. These will later be compared to the currently viewed picture. Insects for example move around until the memorized and the current picture match (Durier, Graham, & Collett, 2003). This was also tested in the study by Cartwright and Collett. They trained honey bees by putting them into a room with only one black landmark close to a food source. After the training period they discovered the honey bees looked for the food source at the same place close to the landmark. But when changing the size of the landmark between training and testing the honey bees searched for the food source not at the same distance as in the first experiment but in a place where the landmark appeared to be the same size as during the training period. They therefore concluded that the honey bees learn the size and appearance of the landmark from the position of the food source. While searching for this, they move around until the retinal image matches the image they had remembered during the training period (Cartwright & Collett, 1983).

Our study is based on a previous study by Prozmann, which has shown that human place recognition is possible with only depth information. In their study the participants were tested in a virtual environment where all landmarks, and therefore any possible snapshot information, were removed. They only had depth information as a recognizable cue but were still able to succeed in a place recognition task (Prozmann, 2015).

Human depth perception is based on multiple factors of which some can be perceived monocularly and some require binocular vision.
One eye can perceive depth by either using occlusion, the size difference of closer and farther images, or motion parallax. Motion parallax is a result of seeing the environment from spatially different viewing points at different moments in time (Rogers & Graham, 1982). It can best be explained by the example occurring while sitting in a moving train and watching the outside world passing by. Objects that are farther away from the train like mountains move by relatively slowly while objects that are closer to the observer, like



trees, pass by very fast.

Figure 1: example of the two different images the two eyes perceive an image in caused by the two different angles (Purves et al., 2008)

Two eyes can additionally make use of stereopsis. Stereopsis is caused by the two images received by the two eyes being slightly different from one another. This again is a result of the two different visual angles of the eyes as seen in figure 1 by Purves et al., the difference of the resulting pictures is called binocular disparity (Purves, et al., 2008). Stereopsis is known to be used as a depth cue in many different species, of course in vertebrates but in insects as well (Rossel, 1983).

In the study by Prozmann participants were tested in a virtual environment that was presented on a single mirror stereoscope (see 3) providing them with a three-dimensional presentation of the stimulus.

A stereoscope in general is a gadgetry with the purpose of simulating a three-dimensional vision of two slightly different pictures.

The first originally invented stereoscope is the one invented by Sir Charles Wheatstone, the so called “Wheatstone stereoscope”. Sir Charles Wheatstone was the first scientist who discovered that the two images perceived by the two eyes make up a three-dimensional image by differing slightly from one another and by then being connected by the brain. Therefore two images in general which are being shifted horizontally from one another must simulate a virtual image of a three-dimensional environment. Using this knowledge Wheatstone created a device which was perfected by the invention of photographs (drawings always differed from one another in more than the horizontal position). The Wheatstone stereoscope was constructed as seen in figure 2: Two images (E’ and E) were positioned to the left and right side of the participant and reflected onto 2 mirrors (A’ and A) right in front of the observer. The two images and the mirrors are positioned in such way that the observer sees two images (one with each eye) that merge into one image.

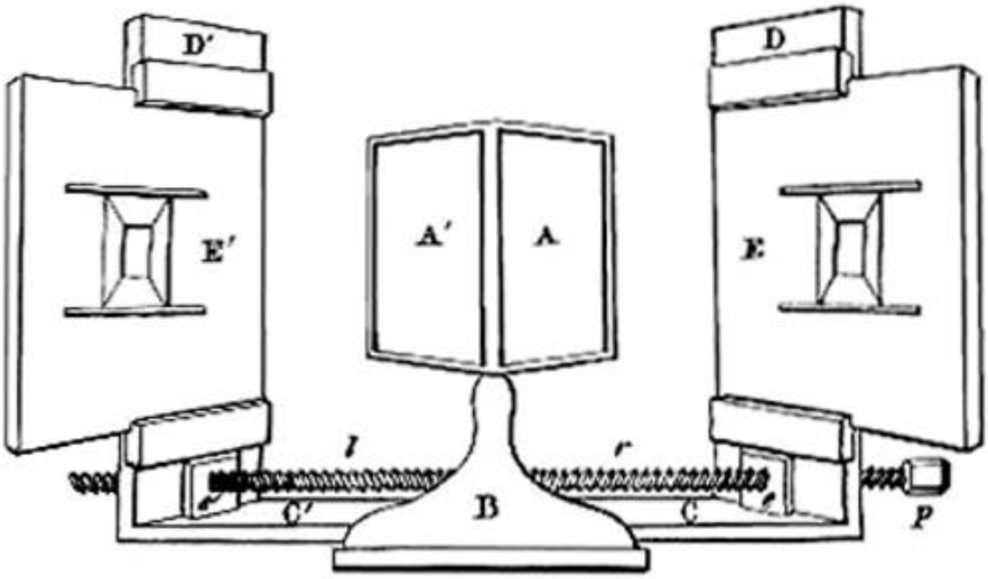


Figure 2: setup of the first Wheatstone stereoscope with the two images E' and E and the two mirrors A' and A reflecting these images

In the previous study by Prozmann the participants had to complete a place recognition task, a so-called "return-to-cued-location" task. They had to learn a goal position and then, after being displaced somewhere else in the kite-shaped room, had to return to this goal position as accurately and as fast as possible. The kite-shaped room was presented in two different conditions (one with several different features like intensity and snapshot information, the other one being featureless). Their results showed that the homing task was possible during both conditions but the participants' performance was impaired during the featureless condition where only depth cues were available. Additionally they had to complete a preliminary test where the participants' ability for stereo vision was tested but no correlation was found between their stereo ability and their performance in the main task (Prozmann, 2015). These results raise the following questions:

While the performance in the main task is not dependent on the participant's ability for stereo vision, motion parallax may play a bigger role in depth perception. Therefore we raise the question: Will the removing of the ability for stereo vision (thus only testing motion parallax) decrease the performance during the homing task?

Additionally we wanted to remove the possibility that the sharp corners of the kite-shaped room might still have provided the participants with landmark information. Consequently we rounded off the corners of the virtual room making them less pronounced. This raises our second question: Will the elimination of any landmark information decrease the performance of the participants?

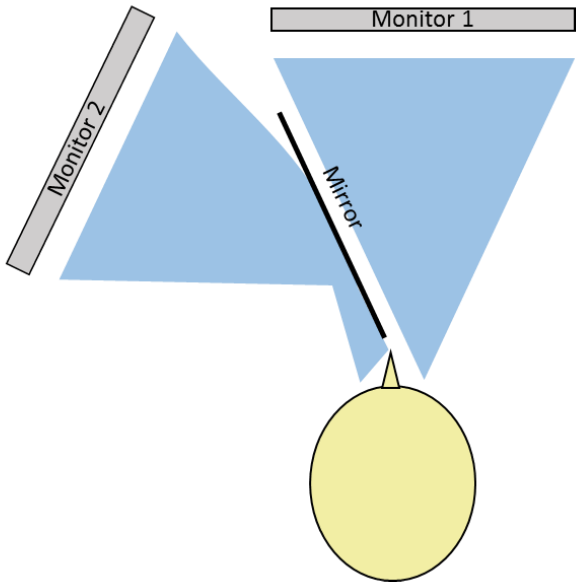
The experiments in this study were both based on the past experiments by Prozmann in which the participants were also tested at the stereoscope in the same kite-shaped virtual environment having the same task as in our experiment. The task was the same: being able to orient themselves where they were in the kite-shaped room and then being put somewhere else in the room and asked to return to the original position which was also the goal position. The past experiments were different as the participants always had both eyes uncovered being able to use binocular cues at all times which our participants weren't able to do in experiment 1. That way we wanted to test whether only motion parallax (without stereopsis) could give enough indication to perceive depth properly and be able to complete the homing tasks. Also there was one difference in the appearance of the virtual room presented to the participants in experiment 2 compared to the previous study in which the corners were sharper and not rounded. The results in the previous study suggested that the sharp corners were still used as landmarks. To avoid this we eliminated the landmarks by rounding off the corners making them less sharp and less noticeable in our second experiment. This way we wanted to make sure that only the depth perception was a cue being used for the homing task. Additionally we wanted to test whether only motion parallax (without stereopsis) could give enough indication to perceive depth properly and even be able to complete the homing tasks. We made sure of this by giving the participants an eye patch.

Material and Methods

In this experiment we test the different effects of monocular vision versus binocular vision which is based on disparity.

Setup

The stereoscope we used was constructed as follows: with the right eye the participants looked directly into monitor 1 and, in experiment 2, with the left eye the participants looked at the reflection in the mirror of monitor 2 as seen in figure 2. Therefore this stereoscope was a single mirror stereoscope as described by Kollin & Hollander (Kollin & Hollander, 2007). The monitors were both connected to the same computer while displaying two different images as in every stereoscope. These two images were slightly shifted horizontally from one another. Monitor 1 was positioned 80 cm from the participant; monitor 2 was positioned diagonally 62cm from the mirror (measuring from the center of the monitor to the center of the mirror). The angles were aligned in such a way, that the 2 images perceived by the observer merged into one three-dimensional image. During the experiment, the room in which the participant sat in front of the



stereoscope was completely darkened.

During both parts of the experiment participants were presented a kite-shaped room which was 21 meters long and 10 meters wide. There were 3 target points with the following coordinates: (3|13), (1|4) and (-2|16). In later references we speak of the "blue" ≙ (3|13), "red" ≙ (1|4) and "green" ≙ (-2|16) target points which doesn't mean they were colored but meant for an easier understanding as shown in figure 4.

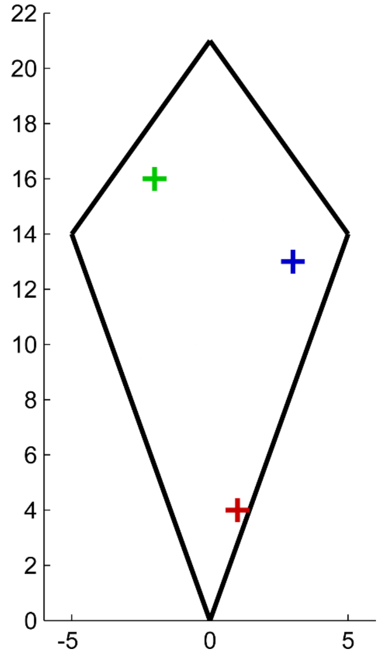


Figure 4: kite-shaped room with the blue (3|13), red (1|4) and green (-2|16) target points

Procedure

16 people were tested on 2 different days, 8 in the monocular experiment and 8 in the binocular experiment with the "rounded corners". The procedure of the main experiment was identical in both experiments.

The experiment consisted of 2 parts. During the first part the kite-shaped room was presented as a textured wallpaper consisting of black background and big white dots as seen in figure 5 (texture condition).

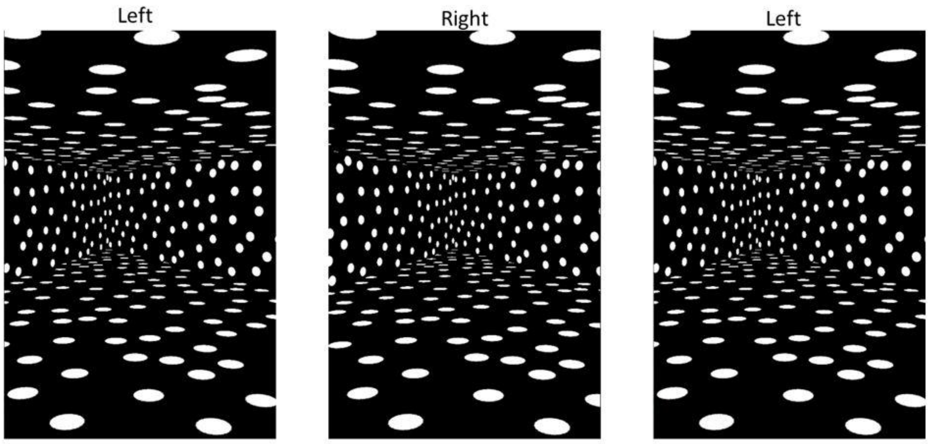


Figure 5: Participant's view in the first part of the experiments with the textured wallpaper; arranged for free fusion, with either crossed (left and right pictures) or un-crossed views (right and left pictures)

Each of the 14 trials (of which the first two were for familiarization and not to be included in the results) was started by the participants by pushing the "space" button. During the first part of each trial the participants could only move in a space of 1.5 meters diameter. The participants were instructed to look around and move around, within the given space, to find out where they were placed in the kite-shaped room. This was the goal position they should remember for this trial. If the participants thought they knew where they were they had to press "space" again and were placed at the starting position. The task was to walk back to the goal position and to confirm the position by pushing the "space" button again. The participants were immediately given a feed-back how far they were from the actual goal position by being driven back to the correct location.

After this the next 13 trials were running the exact same way. The participants were allowed a break between the trials but not during a trial as they were instructed to find the goal position as fast and as accurately as possible.

The second part of the experiment went exactly the same way as the first part with the exception of the design of the walls in the kite-shaped room. Here there were no big white dots but little flickering white dots which had a life span of 100-200ms and there was always another dot created somewhere else as soon as one dot "died" so that there were always 2000 dots, as seen in figure 6.

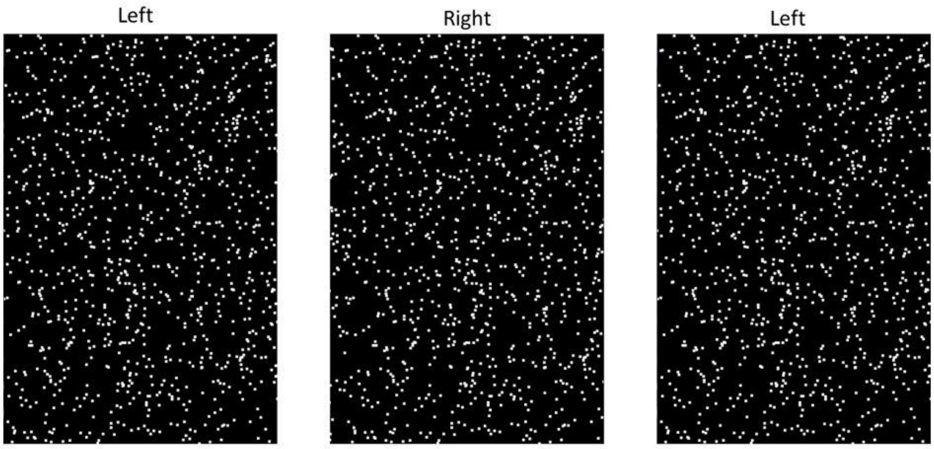


Figure 6: Participant's view in the second part of the experiments with the random dots; arranged for free fusion, with either crossed (left and right pictures) or un-crossed views (right and left pictures)

The two different conditions (texture and random dot) were performed on 2 different days by the participants, so they wouldn't run out of motivation. The experiments with the textured condition took about 15 minutes while the part of each of the experiments with the random dot condition took about 20 minutes, although both times varied strongly in between participants.

Before the experiment was started each participant was handed out a written form which explained the procedure of the test and what they were supposed to do. They could also take a first look of how the room was shaped so they would know what they should be seeing (a kite-shaped room). Additionally they had to sign a consent form.

Experiment 1

The participants had the left eye covered by an eye patch and could therefore only look with the right eye onto the right monitor. Because of the fact that participants could only look with one eye they couldn't make use of disparity and thus could only use motion parallax as a depth indicator. The procedure was as described above and displayed in figure 7.

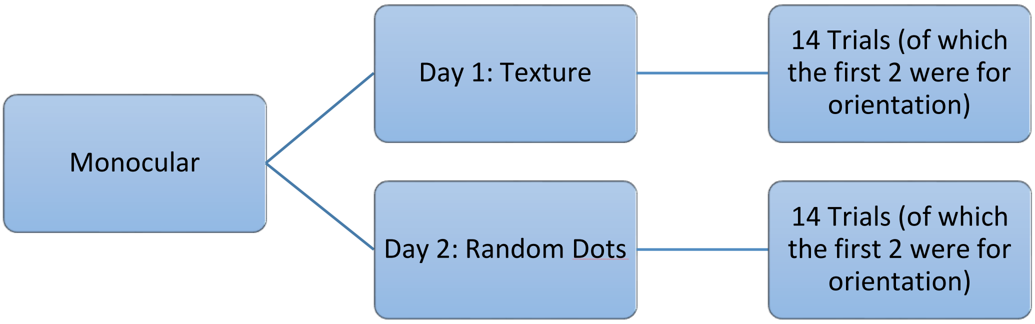


Figure 7: Procedure of experiment 1 in which the participants performed the two different conditions on 2 different days, each condition consisting of 14 trials

Experiment 2

This experiment was performed by the participants with both eyes uncovered. Therefore they were looking with the left eye into the mirror reflecting the image of the left monitor and with the right eye directly onto the right monitor.

Prior to the experiment the participants had to perform an experiment which tested their ability to see depth in the stereoscope. This pre-test was performed before the random dot condition but after the textured condition. The pre-test consisted of 24 trials in which the participant was presented 2 squares which were made up of random dots just as in the main experiment. They had to decide which one of the squares was closer to them. The squares had 3 different distances from one another (0.25m, 1m and 5m) which came randomized while each distance was presented 8 times. If they thought they knew whether the left or right square was closer to them then they had to confirm this by either clicking left or right on the computer mouse.

In the main experiment (during both texture and random dot condition) the kite-shaped room was altered in such a way that each corner was not sharp but rounded, all other aspects were exactly as in experiment 1 as shown in figure 8.

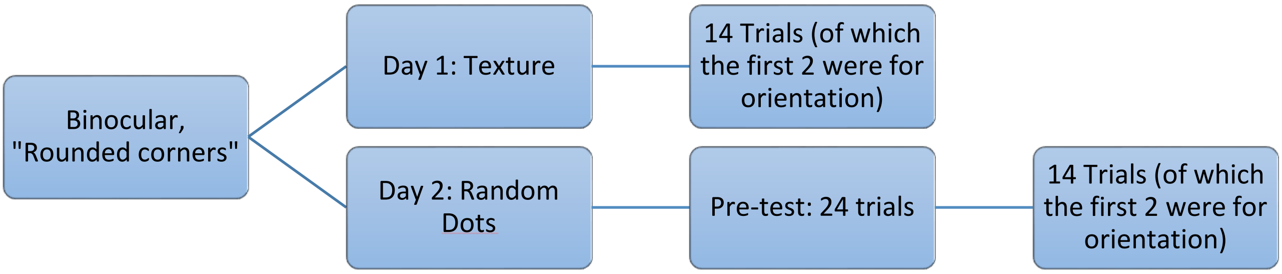


Figure 8: Procedure of experiment 2 in which the participants performed the two conditions on two different days with a pre-test before the random dot consisting of 24 trials

Error definition

To define errors in this study we changed the definition used in previous studies where every point that was more than 4 meters away from the correct position was defined as an error. The problem with this definition was that it didn't consider the geometry of the kite-shaped room. To define an error in a better way we separated the room according to Voronoi. We connected the three target points with one another so that a triangle was formed. Then we took the perpendicular bisectors of each side of the triangle and put a straight line through the perpendicular bisectors.

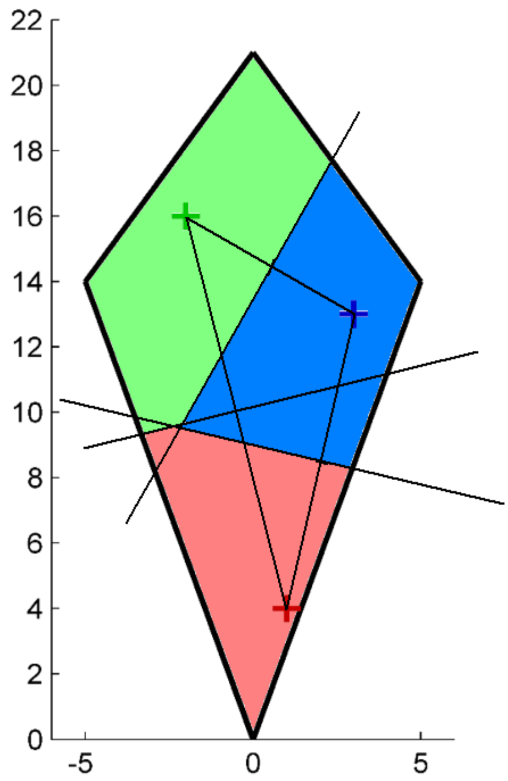


Figure 9: Separating the kite-shaped room into 3 areas according to Voronoi concludes into our error definition which says that every point closer to a different goal position than the correct is defined as an error

By dividing the kite-shaped room with the straight lines (which went through the perpendicular bisectors) three segments were generated. Each of them surrounded one of the goal points. Every point in each segment was now closer to this goal position than it was to one of the other two. Each answer landing in one of the resulting segments was defined as an error as long as it wasn't in the segment of the respective target point. In conclusion the chance level was therefore 1/3 as it was defined by area

Results

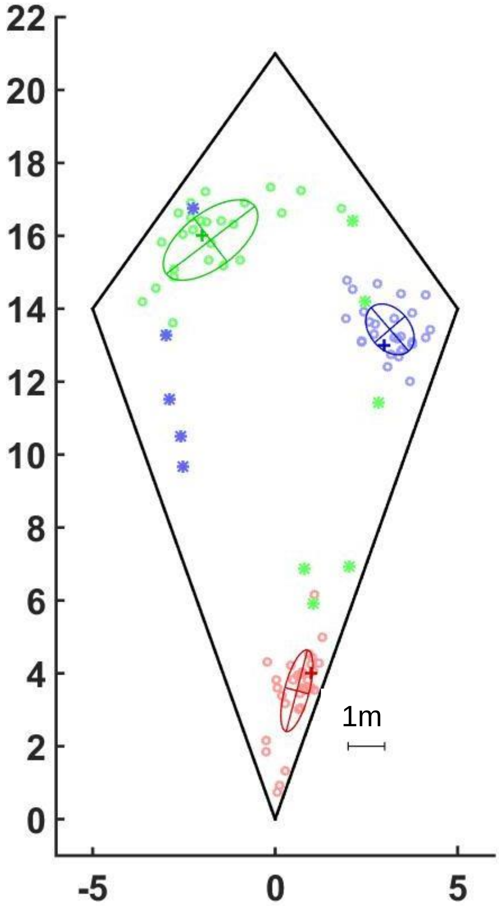
Distribution and error ellipses

Overall the results are as expected. The monocular cues to depth solely are not enough to perceive depth adequately, to complete orientation tasks satisfyingly. Also the results show that orientation tasks are indeed manageable without any intensity information.

Experiment 1

(a)

(b)



1m

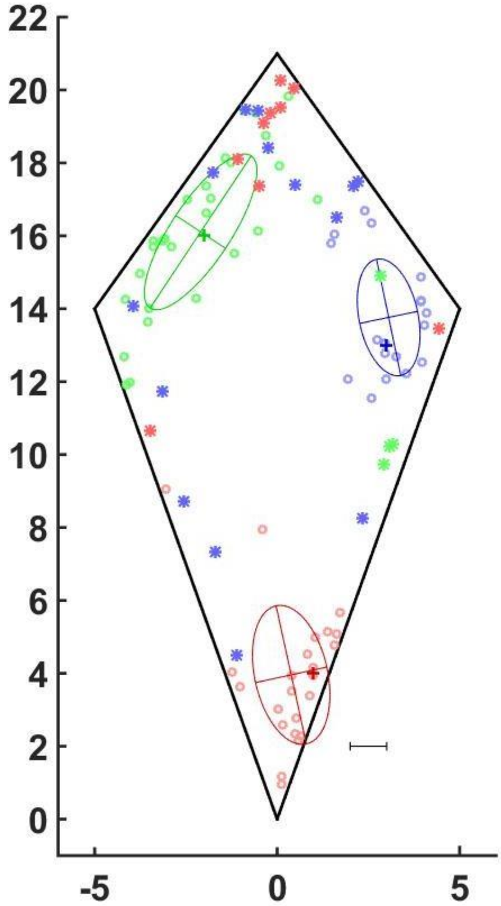


Figure 10: All answers (correct trials as circles and errors as asterisks) and error ellipses in experiment 1; (a) during the texture condition and (b) during the random dot condition

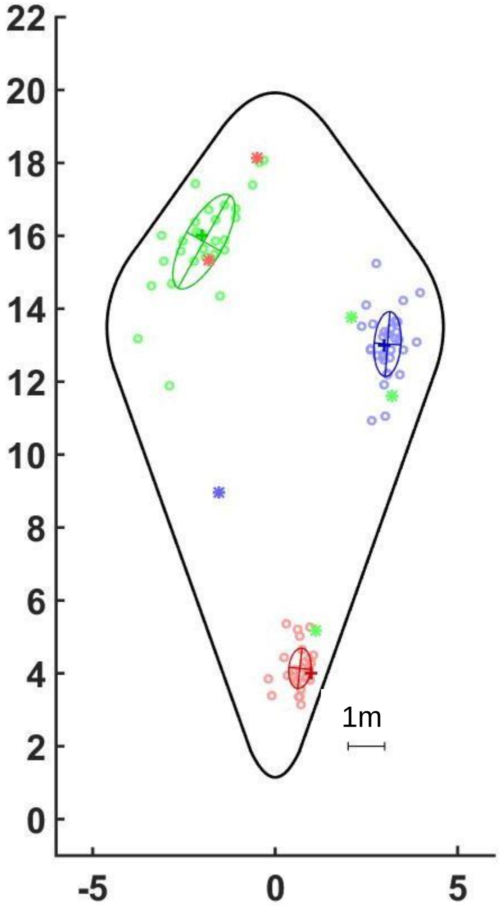
In the monocular task with the texture condition the 8 participants had 5 errors finding the blue target, didn`t have any errors finding the red target and 6 errors in finding the green target, resulting in 11 errors altogether as shown in figure 10 (a). In contrast the participants had 14 errors in finding the blue target, 9 errors finding the red target and 4 in finding the green target, combined resulting in 27 errors of the 96 answers given by all participants in the random dot condition. The error ellipses, which show the variance of the correct answers are much wider spread in the random dot condition than in the texture condition as seen in figure 10 (b). This can also be verified when looking at the actual size of the particular target point as shown in table 1.

Table 1: areas of the error ellipse for each target point per condition in m^2 in experiment 1

	Texture condition	Random dot condition
Blue target point	1.36 m^2	4.15 m^2
Red target point	1.23 m^2	6.09 m^2
Green target point	3.60 m^2	6.44 m^2

Experiment 2

(a) **(b)**



1m

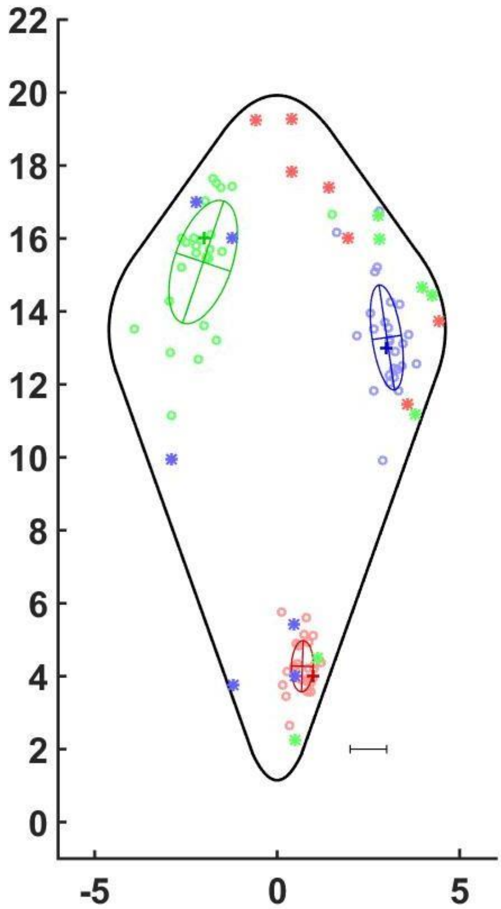


Figure 11: All answers (correct trials as circles and errors as asterisks) and error ellipses in experiment 2; (a) during the texture condition and (b) during the random dot condition

In experiment 2 the participants had 6 errors in the texture condition (1 in finding the blue target, 2 in finding the red target and 3 in finding the green target) as seen in figure 11 (a). In contrast to this we can see in figure 11 (b) that the participants had 20 errors altogether (6 within finding the blue target and 7 each in finding the red and green targets). The ellipses were again bigger in the random dot condition in comparison to the texture condition.

In addition the areas of the error ellipses were generally much bigger in experiment 1 than in experiment 2 (with the exception of the green target point during the texture condition) as seen in table 2.

Table 2: areas of the error eclipse for each target point per condition in m² in experiment 2

	Texture condition	Random dot condition
Blue target point	1.02 m²	1.84 m²
Red target point	0.52 m²	0.66 m²
Green target point	2.47 m²	4.39 m²

Comparison of distances of correct trials to target points

Furthermore we examined the distances, of all answers considered as being “correct trials”, to the respective target point in figures 10 and 11. These were actually statistical errors as they were never exactly on the actual target point.
Looking at the mean of every distance in figure 12 we can see that the participants were the farthest away from the correct target point during the monocular task with the random dot condition besides also being the farthest away looking solely at the texture condition. Comparing the texture with the random dot conditions the distances differed significantly in experiment 1 (Wilcoxon ranksum, $p<0.001$) and in the previous results ($p<0.001$), while it only showed a trend in experiment 2 ($p=0.0821$). Contemplating the distances during the random dot conditions of all 3 experiments they differed significantly among each other (monocular versus rounded corners: $p<0.001$; monocular versus previous results: $p<0.001$; rounded corners versus previous results: $p=0.023$).

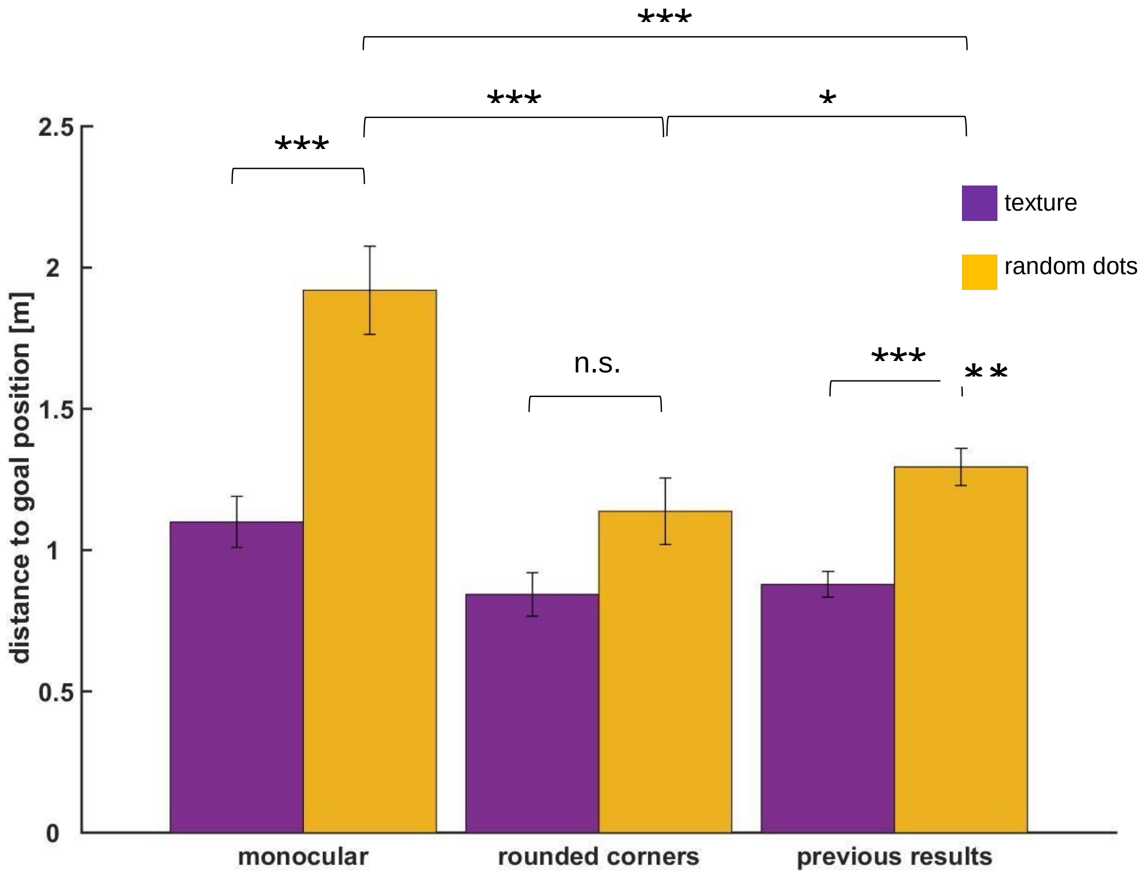


Figure 12: distances of the statistical errors to the actual target point in each condition as well as in previous results

Histograms of all distances to the target points

Comparing the histograms of the distances from the actual target point to the decision points given by the participants it seems obvious that in most conditions the participants appear to be closest to the respective target point while looking for the red position.

Experiment 1

Looking closer into the results of the first experiment with the texture condition the participants confirmed positions closest to the correct target point within looking for the red position (cf. figure 14) and were farthest away looking for the green position with even 3 confirmed positions being 9 to 11 meters away from the actual target point (cf. figure 15). Overall most of them were within 4 meters away from the target point.

In the random dot condition the participants confirmed the most (by number) farthest positions while finding the blue target with no single peak visible. Most confirmed positions were within a distance of 8 meters as seen in figure 13 (b). Searching for the red position there was a bigger peak with values of 0 to 2 meters and then dispersed single answers until 15 meters and 5 given answers at 15 to 17 meters distance to the actual target point (cf. figure 14 (b)). In finding the green target point the participants were mostly within 5 meters to the goal position target point with 3 exceptions between 7 and 8 meters (cf. figure 15 (b)).

(a) **(b)**

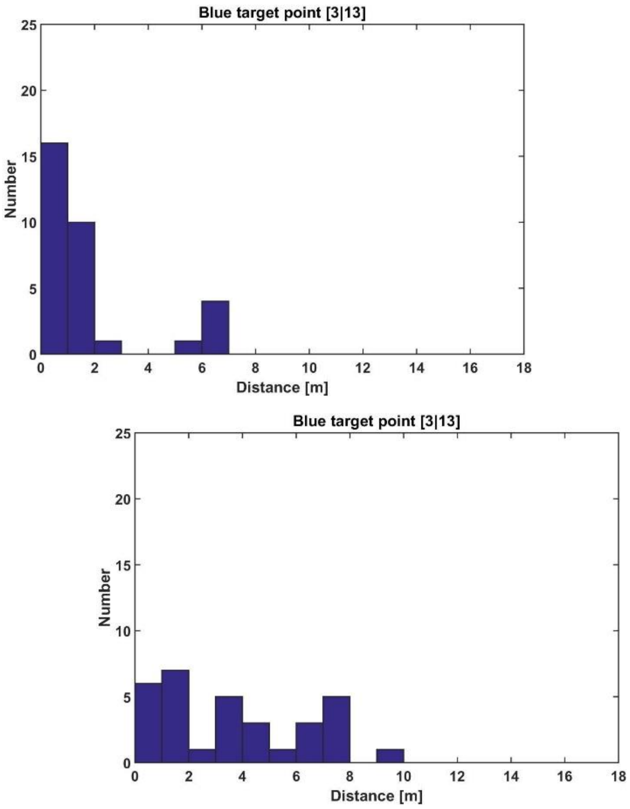


Figure 13: Histogram of the distances to the actual target point finding the blue target; (a) during the texture condition; (b) during the random dot condition

(a) **(b)**

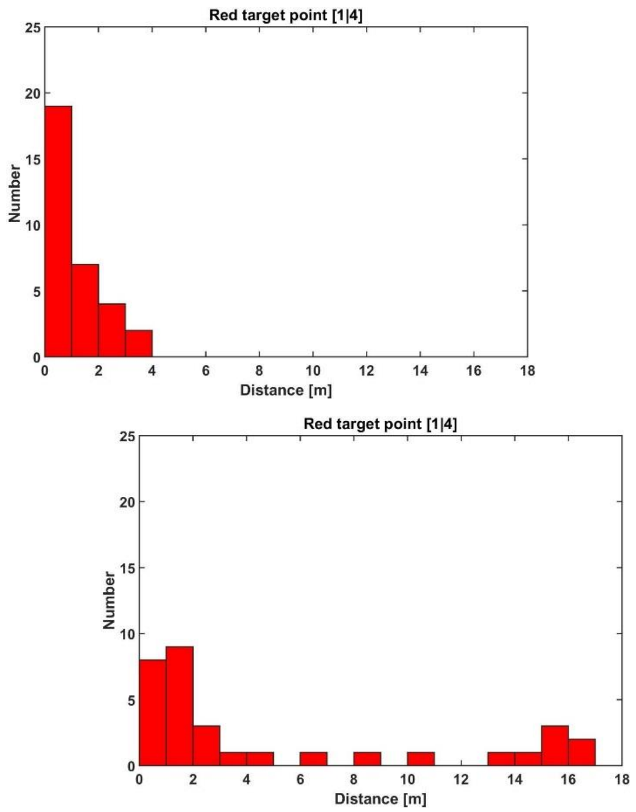


Figure 14: Histogram of the distances to the actual target point finding the red target; (a) during the texture condition; (b) during the random dot condition

(a) **(b)**

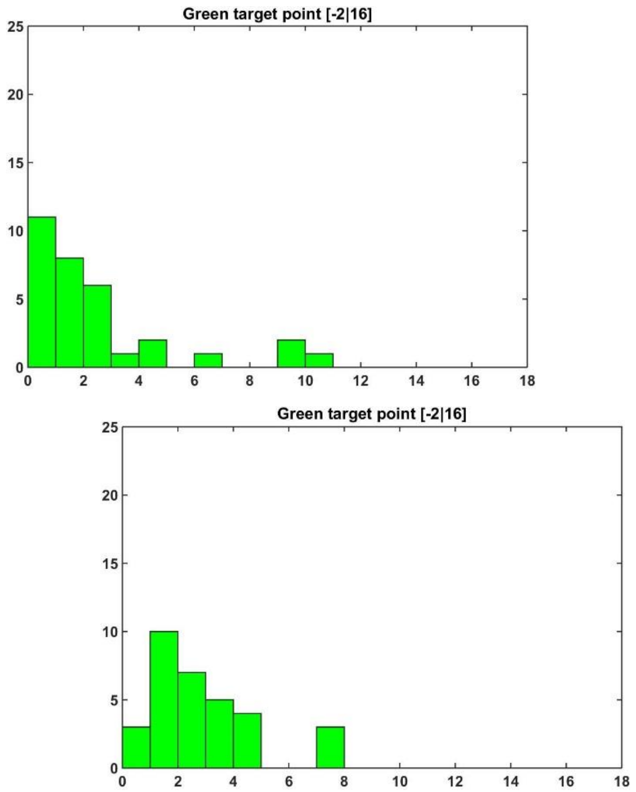


Figure 6: Histogram of the distances to the actual target point finding the green target; (a) during the texture

condition; (b) during the random dot condition

Experiment 2

In experiment 2 in both texture and random dot conditions there is a peak with a distance between 0 and 1 meter finding each target point.

Furthermore in the texture condition finding the blue target point the 8 participants managed to stay within 2 meters of the actual target point in 29 of 32 trials. During 2 trials they stayed within 3 meters and only 1 trial was between 6 and 7 meters of distance to the target (cf. figure 16 (a)). Finding the red target 30 trials ended up being 2 meters at maximum away from the actual target position. There were 2 outliers between 11 and 15 meters as seen in figure 17 (a). Finding the green target point resulted in the most distanced answers compared to the other two. 25 trials landed at a distance between 0 and 2 meters, 3 trials between 2 and 4 meters, also 3 between 4 and 7 meters and 1 outlier between 11 and 12 meters (cf. figure 18 (a)).

In the random dot condition the peaks at 0-2 meters were lower than in the texture condition but still distinctive. Finding the blue target there were still values until 11 meters even though these were only single ones evenly scattered (cf. figure 16 (b)). Finding the red target there were no answers given between 2 and 7 meters and then single ones until 16 meters. (cf. figure 17 (b)). Additionally while trying to find the green target the number of answers given decreased evenly from 0 to 8 meters and 2 outliers between 11 and 12 meters and between 13 and 14 meters (cf. figure 18 (b)).

In summary the histograms show that the participants were closer to the actual target points in the texture condition than in the random dot condition and farther in the first experiment with the monocular vision compared to the second experiment with the "rounded corners". In addition we can see that the participants seemed to be closer to the searched target point while trying to find the red target during both experiments.

(a) **(b)**

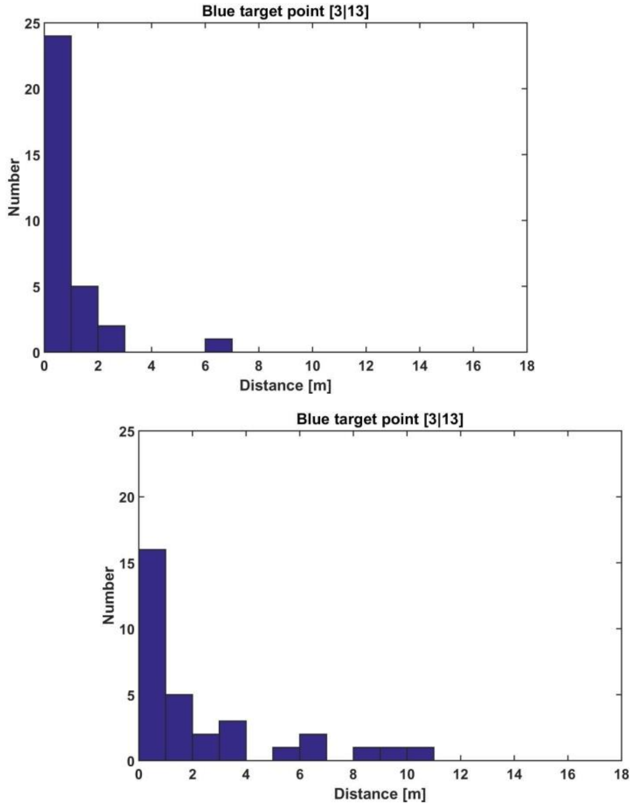


Figure 16: Histogram of the distances to the actual target point finding the blue target;(a) during the texture condition; (b) during the random dot condition

(a) **(b)**

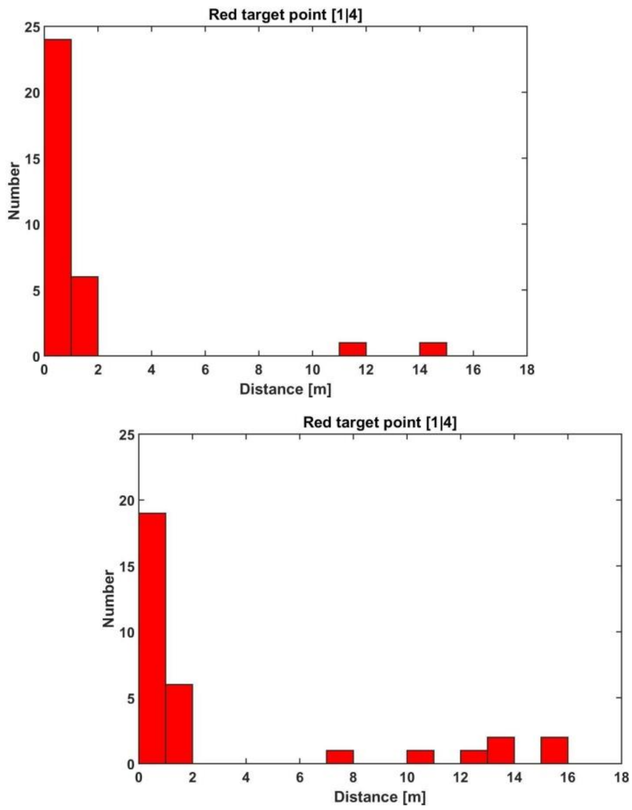


Figure 17: Histogram of the distances to the actual target point finding the red target; (a) during the texture

condition; (b) during the random dot condition

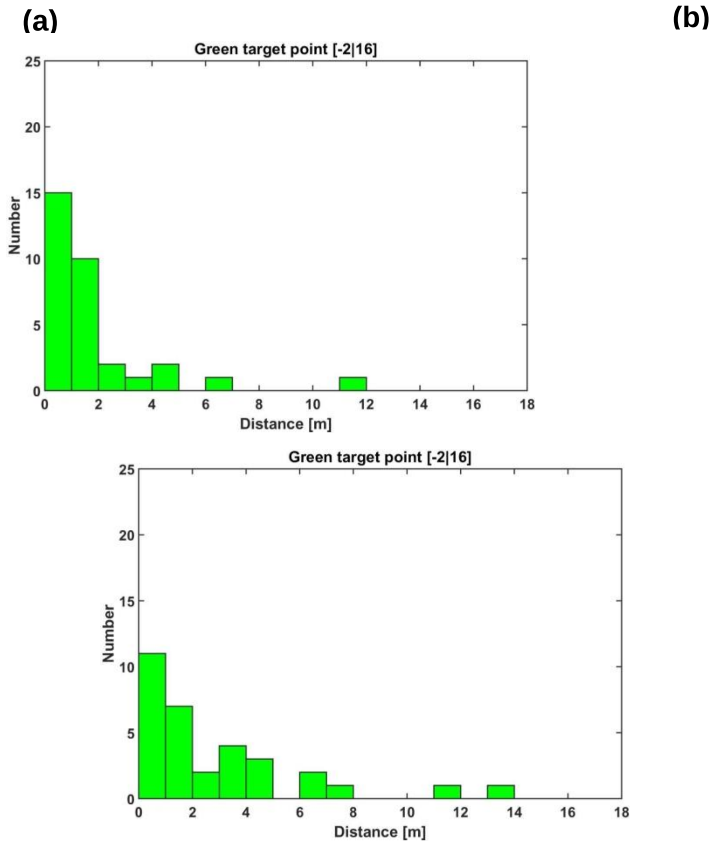


Figure 18: Histogram of the distances to the actual target point finding the green target; (a) during the texture condition; (b) during the random dot condition

Comparison of error rates

Experiment 1

When looking at the percentage of errors in the first experiment in figure 19 we can see that only during finding the red target the relative error was significantly lower in the texture condition than in the random dot condition (Wilcoxon ranksum test, p=0.026). Whereas looking at the bars of the blue target and the overall results there seems to be a trend which is not significant between the texture and random dot condition (blue target: p=0.159; overall: p=0.080). Meanwhile the participants actually had a higher rate of relative error in the texture condition than in the random dot condition while trying to find the green target (p=0.730).

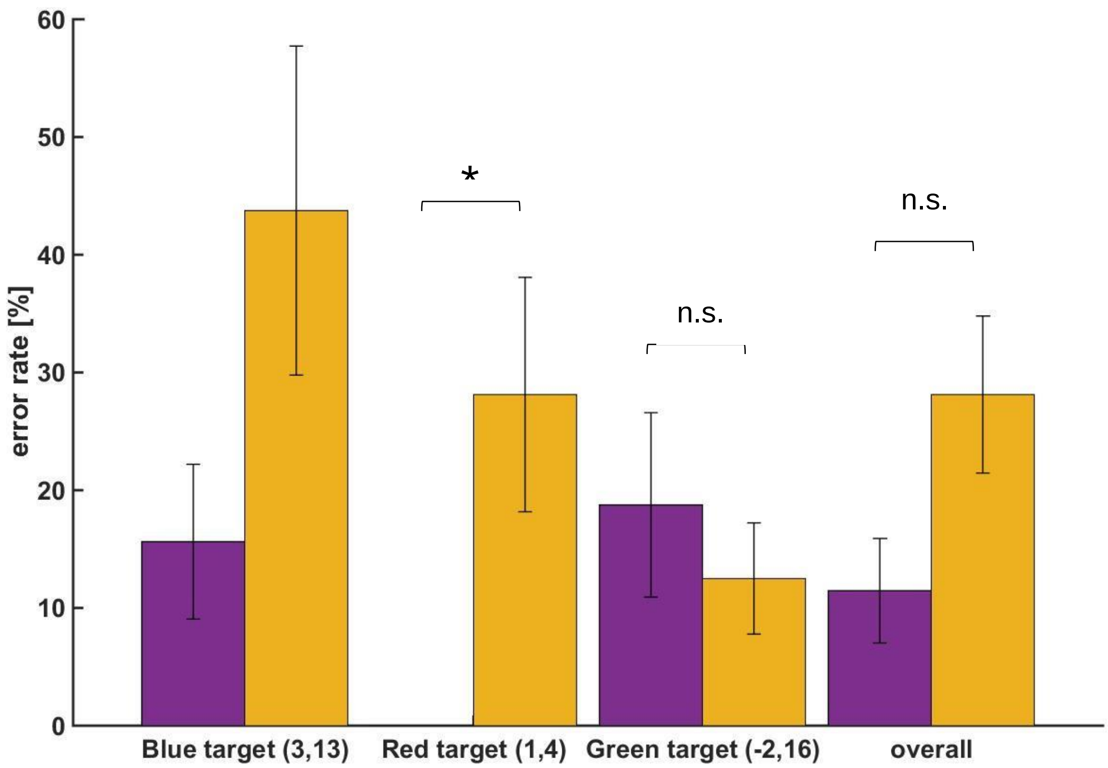


Figure 19: error rate in % of each target point and the overall error rate during experiment 1 (monocular view)

Experiment 2

In the second experiment the relative errors were also lower in the texture condition than in the random dot condition as seen in figure 20 although none of the results differed significantly (Wilcoxon ranksum test, blue target: p=0.323; red target: p=0.467; green target: p=0.641; overall results p=1).

Finally we compared the overall results of all 3 experiments which again showed no significant difference between the random dot conditions (Wilcoxon ranksum test, monocular versus rounded corners: p=0.255; monocular vs. previous results: p=0.135; rounded corners vs. previous results: p=0.348) shown in figure 21.

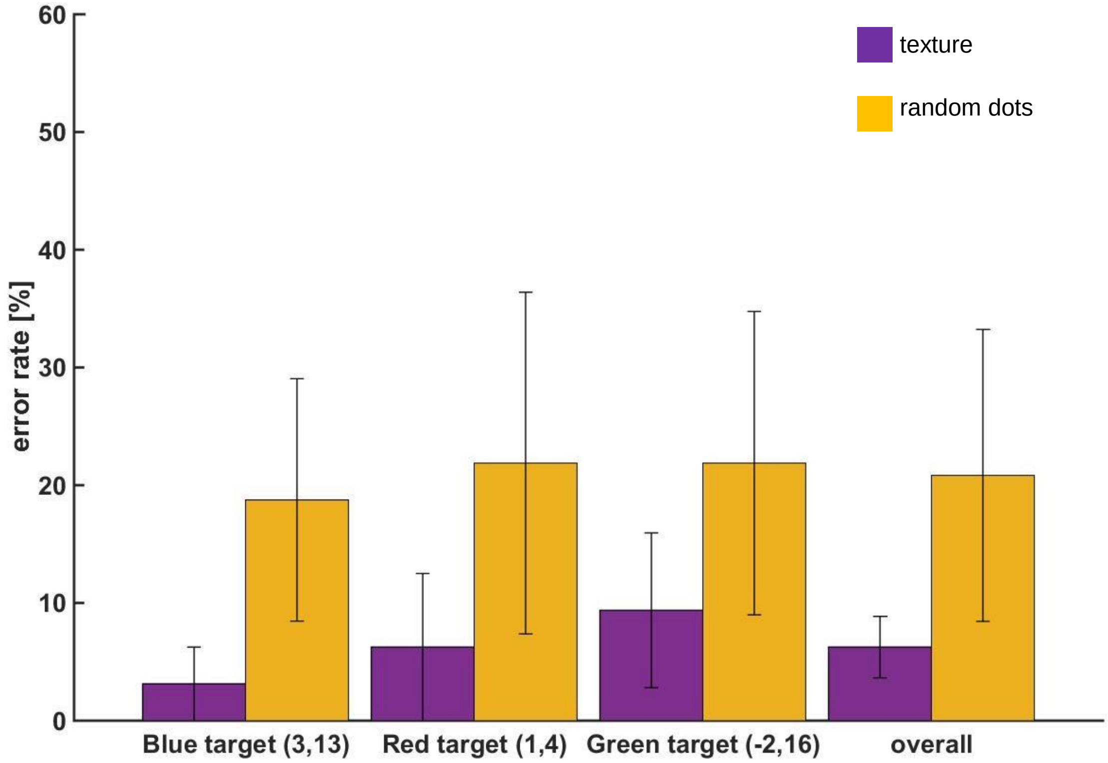


Figure 20: error rate in % of each target point and the overall error rate during experiment 2 (rounded corners)

Comparison of experiment 1, experiment 2 and the previous results

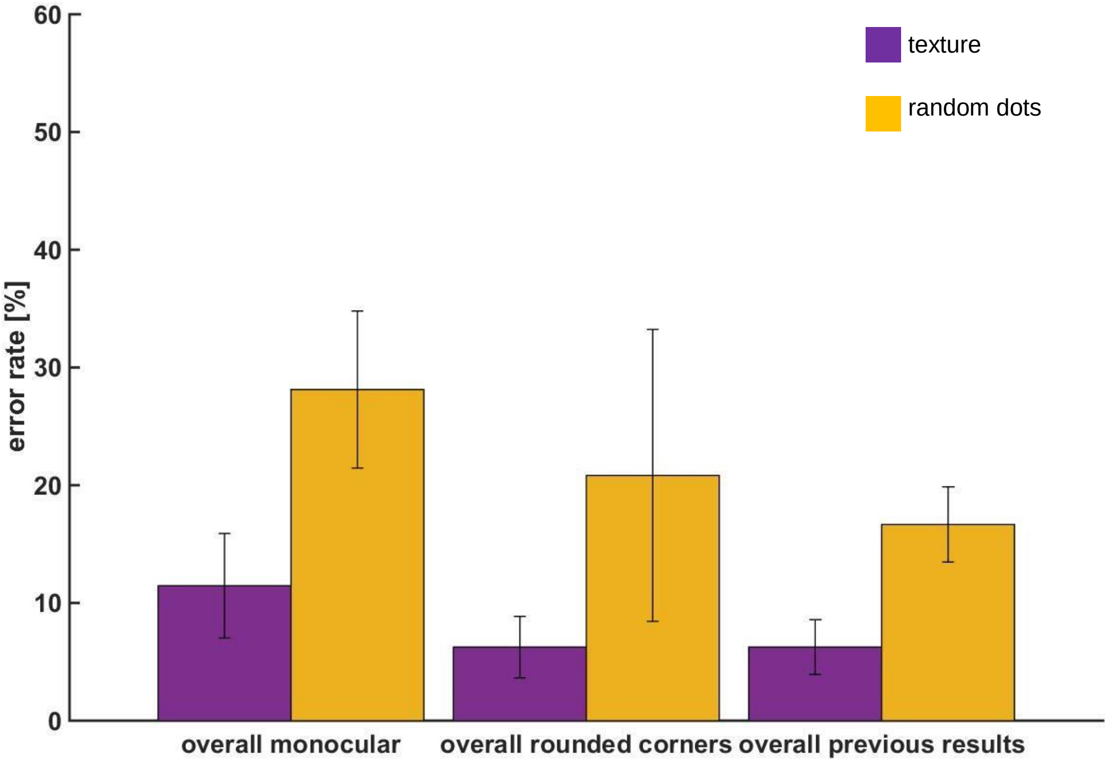


Figure 21: comparison of all overall error rates of experiment 1 & 2 and of the previous results by Prozmann

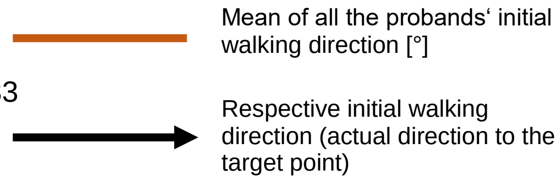
Histograms of initial walking directions

Furthermore we examined the directions in which the participants started walking as they were looking for the target point. We display these results in a histogram showing the different angles in relation to the angle of the actual target point (to simplify the histogram we always placed this angle on the 0° mark for every target point).

Experiment 1

Examining the histograms of the first experiment we can see a clear difference between the two conditions. During the texture condition the mean of the participants' initial walking direction is very close to the actual direction to the target point as seen in figure 22. Additionally the length of the line showing the mean of the direction is close to the length of the arrow showing into the direction of the actual target point, which means a lot of the participants were walking into the direction of the calculated mean. The values aren't scattered over the circle but directed (v-test, p=0, v=76.655) and primarily 20°± around the 0° mark.

In contrast to this the starting angles in the random dot condition in figure 23 are scattered all over the circle and therefore non-directional (p=0.075, v=9.971). Though the mean does actually lay on the 0° mark its length is very short and therefore only a few participants actually went into this direction.



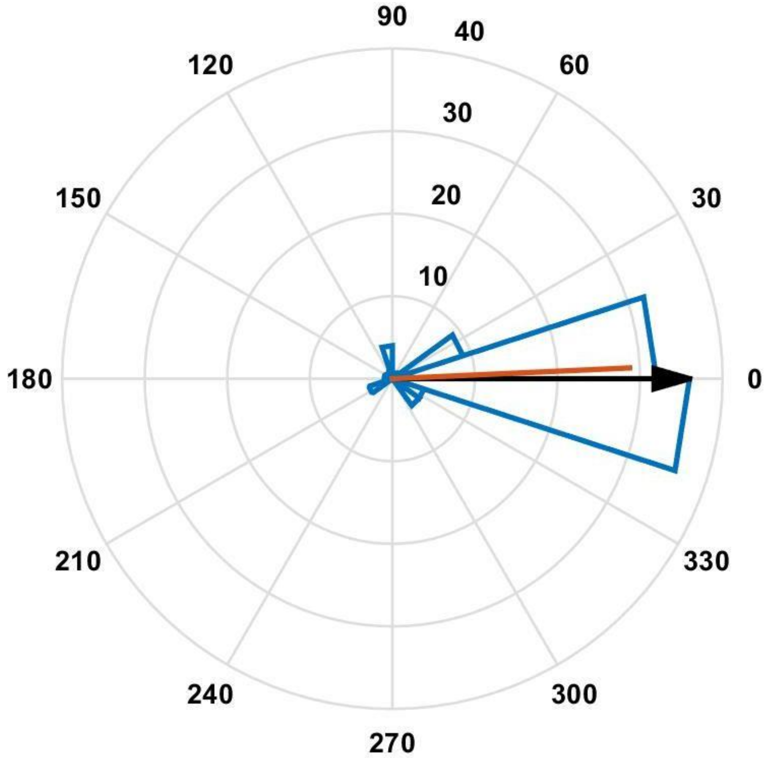


Figure 22: mean of the initial walking direction in experiment 1 during the texture condition; the mean being a result of the walking directions to all 3 target points, each respective initial walking direction laid on the 0° mark

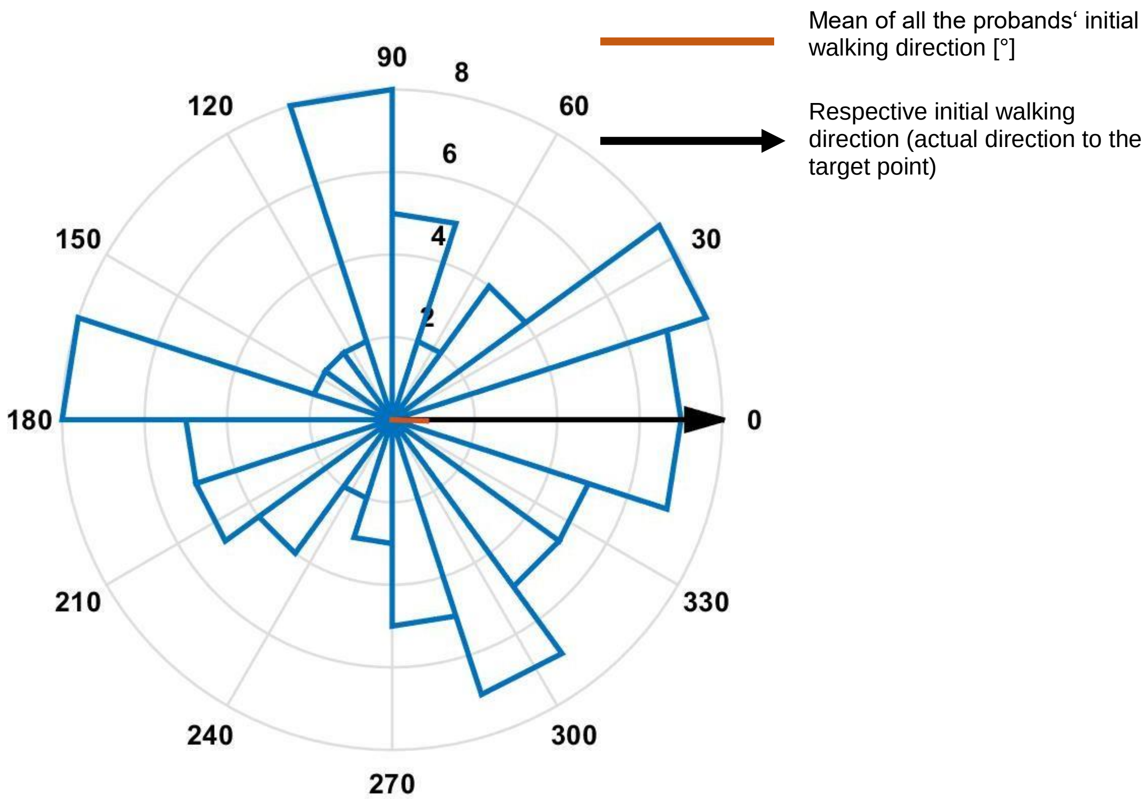


Figure 23: mean of the initial walking direction in experiment 1 during the random dot condition; the mean being a result of the walking directions to all 3 target points, each respective initial walking direction laid on the 0° mark

Experiment 2

On the other hand in experiment 2 the difference between the texture condition and the random dot condition isn't as distinct as in experiment 1. During the texture condition the means of the starting directions were very close to the actual direction as seen in figure 24 and are according to this directed (v-test, $p<0.001$, $v=78.934$). The length of the line showing the mean is also close to 1 and thus in the majority of trials participants' initial walking direction was close to the mean. The highest peak is again ± 20° around the 0° mark.

Looking at the histogram of the random dot condition the walking directions are slightly more scattered around the circle and the mean is further away from the 0° mark as seen in figure 25. The length of the line showing the mean is also shorter than in the texture condition showing a slightly less directed histogram ($p<0.001$, $v=44.676$).

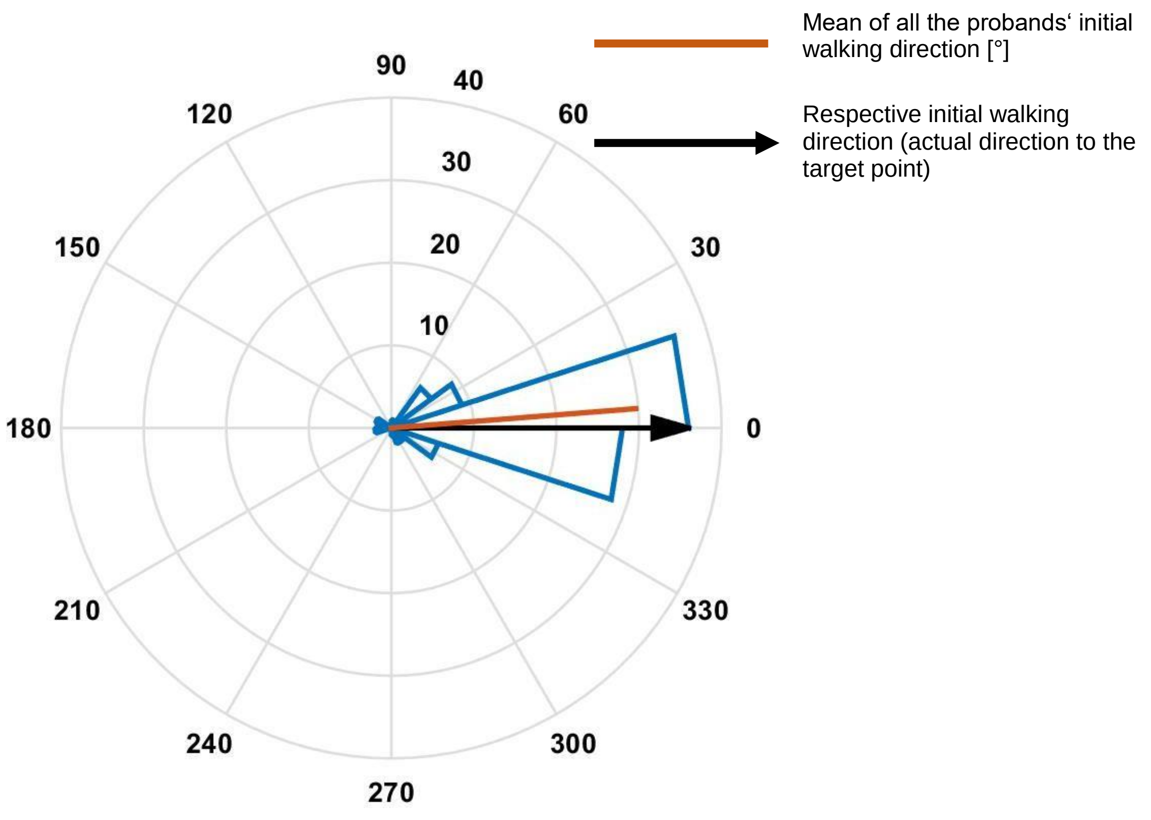
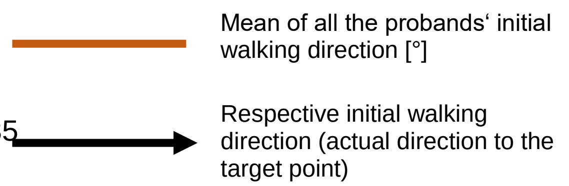


Figure 24: mean of the initial walking direction in experiment 2 during the texture condition; the mean being a result of the walking directions to all 3 target points, each respective initial walking direction laid on the 0° mark



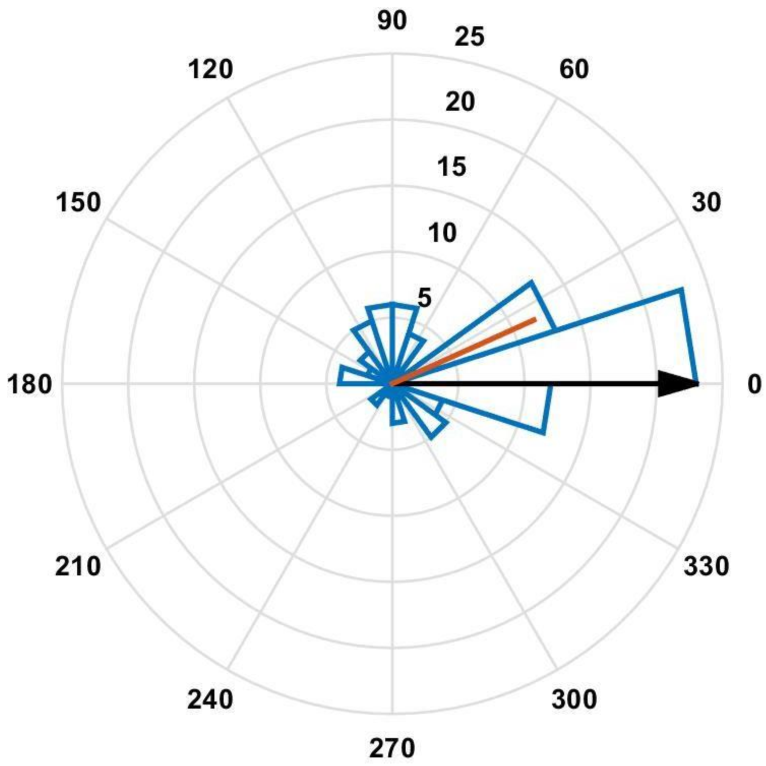


Figure 25: mean of the initial walking direction in experiment 2 during the random dot condition; the mean being a result of the walking directions to all 3 target points, each respective initial walking direction laid on the 0° mark

Heatmaps

At last we looked at the heatmaps of each target point during each condition in both experiments. The distinction between the two different conditions is obvious in both experiments. While in experiment 1 the participants seemed to have stayed in almost every part of the room in the random dot condition, they seem to have stayed in a more confined area during the texture condition as seen in figures 26 and 27.

This also applied to the heat maps of experiment 2 whereas the participants were more confined during both the texture condition as well as the random dot condition as seen in figures 28 and 29. Furthermore in both experiments and conditions they seemed the most directed during the task of finding the red target point.

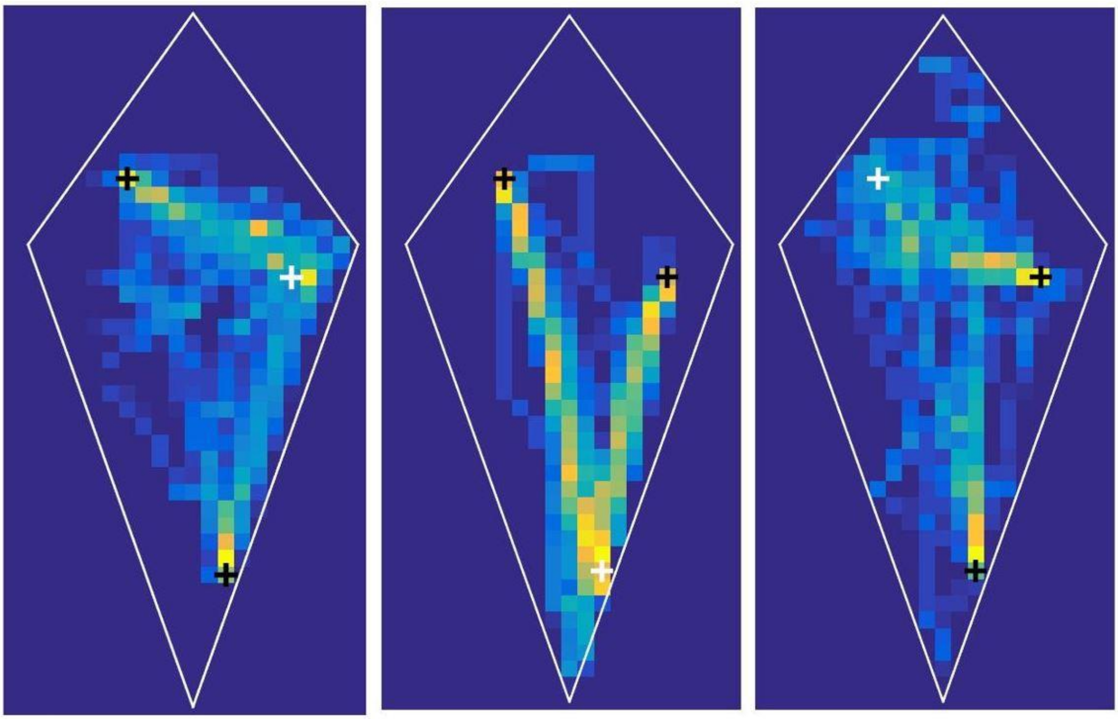


Figure 26: Heat maps of the positions remained finding the 3 different target points (blue, red and green) during the texture condition in experiment 1

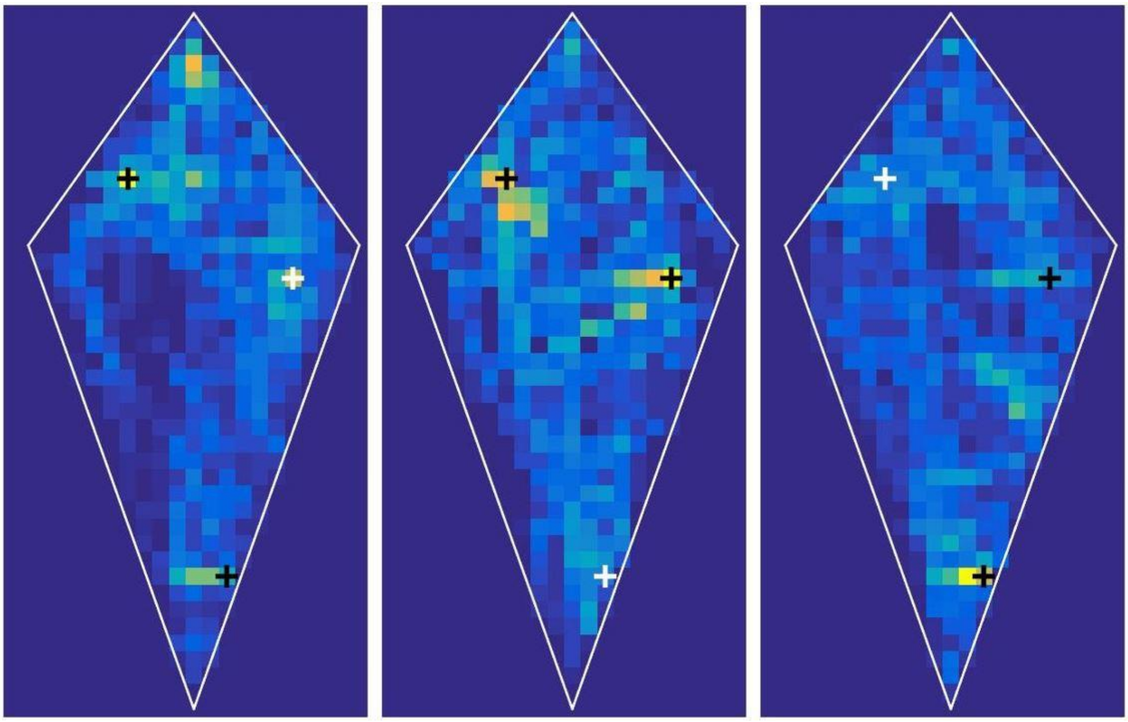


Figure 27: Heat maps of the positions remained finding the 3 different target points (blue, red and green) during the random dot condition in experiment 1

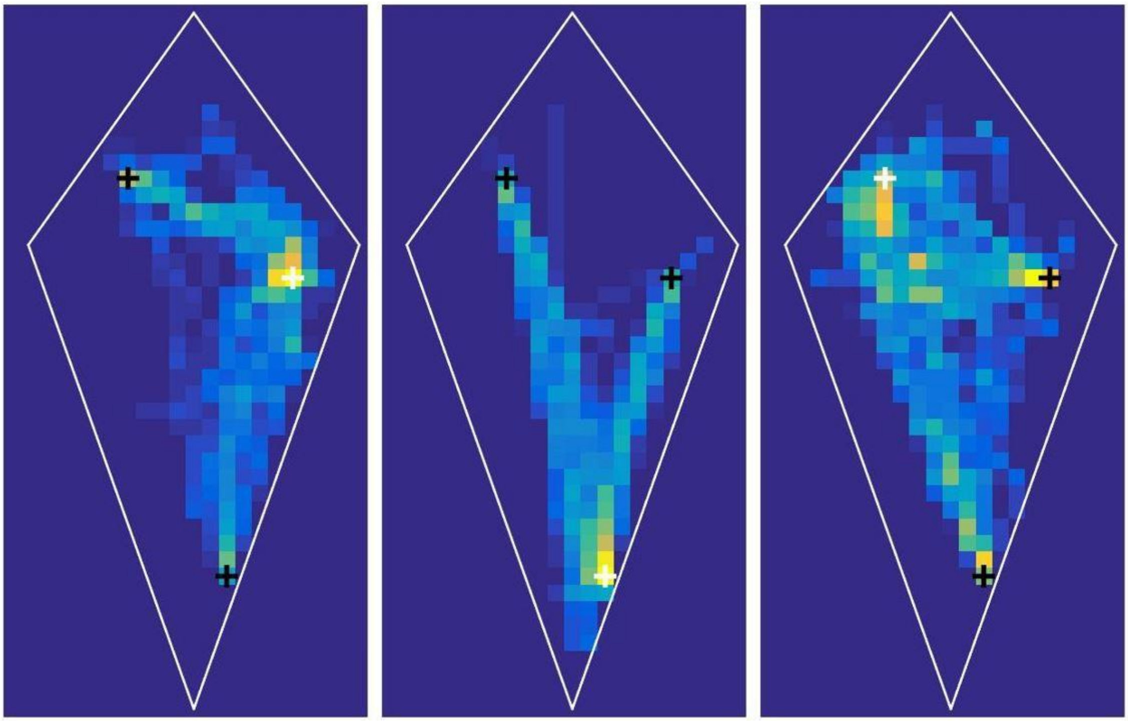


Figure 28: Heat maps of the positions remained finding the 3 different target points (blue, red and green) during the texture condition in experiment 2

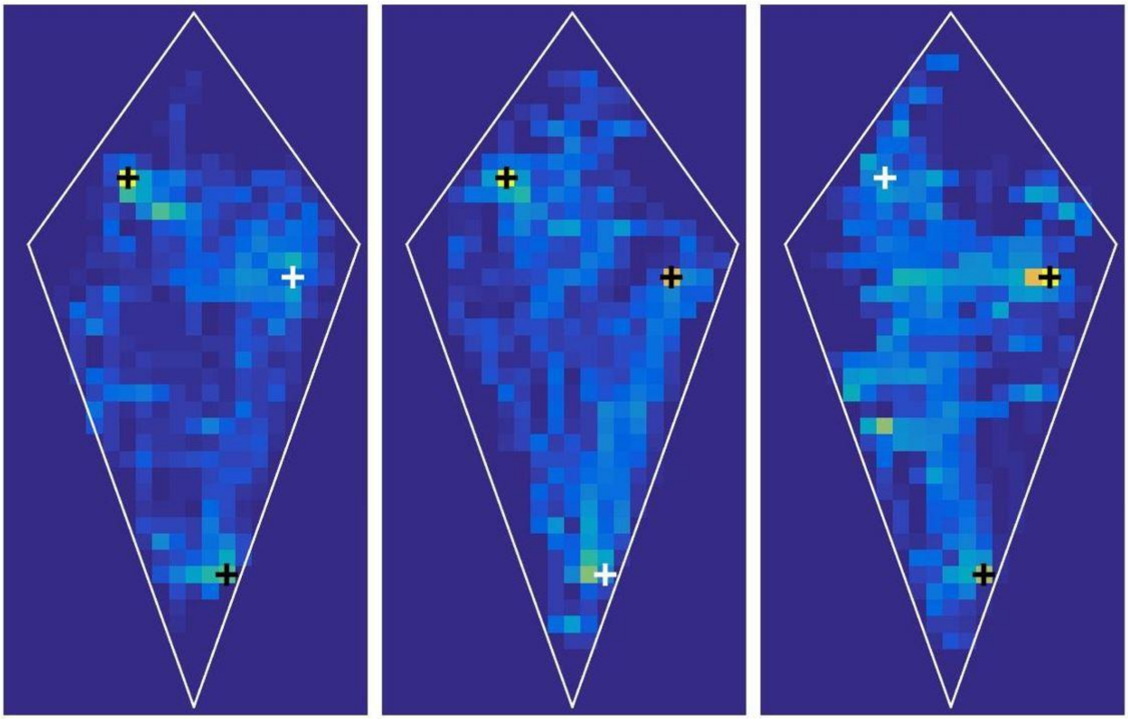


Figure 29: Heat maps of the positions remained finding the 3 different target points (blue, red and green) during the random dot condition in experiment 2

Discussion

In general the results are as expected as stated above. The results indicate that although some participants may have used the sharper corners in the previous study as landmarks they are also able to complete the task without this help by only using the depth perception. Also the results suggest that having only motion parallax as a depth cue limits one's ability to fulfill the homing task.

Looking at the results of experiment 1 we can see that the participants completed the tasks with less errors in the texture condition than in the random dot condition as well as having smaller error ellipses. This applies to every target point except for the green target point. While the area of the error ellipses were also smaller in the texture condition, the number of errors was bigger (6 errors in the texture condition and 4 errors in the random dot condition). This may be caused by rotational errors, although the area of the error ellipse show a trend that the participants had a clearer idea of where the target point was supposed to be as the error ellipse shows the distribution of the "correct" answers. In general the results were as expected since the texture condition gave additional cues such as intensity information and not only motion parallax could be used. The errors also seemed to be rotational errors during the texture condition while some of the errors during the random dot condition weren't rotational errors but other errors as they were not in the opposite corner from the correct position. Also the statistical errors were more scattered in the random dot condition opposed to the texture condition as seen by looking at the area of the error ellipses.

In experiment 2 the participants had more errors during the random dot condition than in the texture condition as well as having larger error ellipses in the random dot condition. Comparing these results with the results of experiment 1 we can see that altogether the participants completed the homing tasks better in experiment 2 than in experiment 1 where they had a bigger number of errors and larger error ellipses. These first results are implying that even without any landmark information present homing tasks remain manageable. However taking away all binocular cues limits one's ability to perceive depth strongly.

Furthermore we compared the actual distances to the 3 different target points, here all points (statistical errors as well as rotational and other errors) are taken into consideration.

In experiment 1 the participants' confirmed positions were always closer to the actual target point during the texture condition than in the random dot condition. This was as we had expected. Also the results imply that the red target point was easiest to be found probably caused by it being in the sharpest corner. The results in the random dot condition, which show some distances up to 17 meters, may again be explained by the participants making rotational errors as they must have been in the opposite corner of the kite-shaped room (because of the maximum length of the room being 21 meters). The participants seemed to find the blue target easier than the green target during the texture condition while it was harder for them during the random dot condition.

In experiment 2 the participants' confirmed positions were mostly within 2 meters of distance to the actual target point as the conditions seemed to have been easier than the conditions in experiment 1. Again the results imply that the red target point was easiest to be found also being supported by the questionnaire in which most participants said this was the easiest target point to be found.

In conclusion these histograms indicate that the participants used the red target point for orientation probably due to the fact that it was still in the sharpest corner as its angle was the narrowest. On the other hand our thesis that the participants need more depth information than they can gain from motion parallax is strengthened by these histograms.

Additionally we examined the distribution of statistical errors, i.e. the mean distances to the respective goal position without the rotational errors or other errors. These results show that during the random dot condition the participants were further away from the actual goal position than during the texture condition with a highly significant difference in the monocular experiment as well as during the previous results. This is caused by the additional information during the texture condition such as intensity information and clear landmark information during the monocular experiment as well as the previous tests. (Gillner, Weiß, & Mallot, 2008)

Also the results indicate that during the monocular experiment the participants seemed to have less of an idea where the goal position actually was than during the binocular experiment with the rounded corners. Looking at these results it is evident that taking away any landmark information doesn't impair participant's ability to fulfil the task. This is again indicated when comparing the results of the experiment with the binocular view and the rounded corners with the previous results as the participants were actually even significantly closer to the goal position in the experiment with the rounded corners than in the previous results. This can be explained by the fact that the comparison is not within participant but in between participant besides the fact that during the previous results data of more participants were included into the results.

When looking at the percentage of errors in experiment 1 we can see that the red target point was easiest to be found during the texture condition but harder to be found during the random dot condition than the green target. This may be explained by the fact that during the texture condition the corner in which the red target point lays may still be used as a landmark while it couldn't be used as a landmark during the random dot condition.
While these single results aren't as expected we have to mention that none of the results (except for the comparison of the texture and random dot condition finding the red target) were significant and only show a slight trend.
Looking at the error rates in experiment 2 the texture condition always resulted in a lower error rate than the random dot condition. Here the participants had the lowest error rate while trying to find the blue target during the texture condition.
Overall the results again show that more errors were made during experiment 1 than during experiment 2 and slightly more errors during experiment 2 than in the previous studies. The results show that trying to fulfil the task with only monocular cues i.e. only motion parallax as a depth indicator is not possible. These results also indicate that using rounded corners doesn't change the results i.e. landmark information is not needed.

These conclusions are also underpinned when looking at the questionnaire, which was filled out by the participants after the experiment. The participants scored experiment 1 (monocular view) as being more difficult and said they were less motivated during this experiment than the participants in the second experiment.

Furthermore we examined the angles of the first walking directions in contrast to the direction of the actual target point to find out whether the participants immediately knew the rough direction to the target point.
The results indicate that the participants immediately had an idea of which direction they had to walk to during the texture condition of experiment 1 as they are directed. The orange line showing the mean of the initial walking direction of all participants during experiment 1 is very close to the actual direction (only differing by about 2 or 3 degrees) and also pretty long indicating how many of the participants actually started in this direction.
In contrast to the results of the texture condition the random dot condition results are non-directional. Although the orange mean line may be directed to the actual target point, its length shows that only a few of the participants actually started towards that direction. The histogram is scattered around the angles having lots of participants actually starting in the opposite direction (180°) or 90° off the actual goal direction.

During experiment 2 the results were similar during the texture condition to the results during this condition in experiment 1 and also directed. But the results of the random dot condition weren't as scattered as in experiment 1. While the mean was a little further away from the actual direction (nearly 30° away from it) the results still show a directed histogram.
These results again confirm our thesis that place recognition is indeed manageable without any landmark information but rather difficult without any stereopsis possible.

At last we compared the heat maps of the positions the participants remained the most during the different experiments as well as the different conditions while trying to find each of the 3 goal positions. These confirm our thesis but the question whether the sharp corner close to the red target point might have still been used as a landmark emerges again. Additionally the heat maps give us some information on the searching behavior of the participants. The heat maps show that the participants moved around a lot and went to every part of the room during the random dot condition in experiment 1. On the other hand during the random dot condition in experiment 2 the participants seemed to move a little less and didn't go to every corner of the room but seemed a little more directed in their search.

Also the results from the preliminary test did not differ from the previous results. As in the studies from Prozmann we did not find a correlation between participant's performance in the pre-test and the main experiment.

Comparing our results with former studies you can find very conflicting results: our results seem to contradict other studies of which some show that motion parallax is supposed to be as strong of a depth indicator as stereopsis like the study by Rogers & Graham (1982). Although in their experiment the task differed from ours as there was no homing task but only needed the participants to judge or infer the shape of a three dimensional surface. Additionally the participants were able to move their head while the image they had to observe was stationary (Rogers & Graham, 1982).

Meanwhile other studies come to similar results like in this study by saying that the threshold for depth perception is higher when only motion parallax is available (Bradshaw & Rogers, 1996). Although we should consider that during the second experiment with the rounded corners and the binocular view not only disparity was a cue for depth perception. Motion parallax was another helpful indicator for depth perception as the participants could still move their view of the kite-shaped room. The interaction between these two cues was tested by Bradshaw & Graham in 1996, resulting in the assumption that binocular disparity and motion parallax are not two independent mechanisms but actually increase their effect when occurring together in a non-linearly way (Bradshaw & Rogers, 1996). This idea is supported by other studies in which cells were found that are active during both mechanisms (Poggio & Talbot, 1981) (Maunsell & van Essen, 1983).

Looking into the results in detail the assumption of the so called rotational errors is supported by several other studies. Rats for example were found to confuse diagonally opposite positions when being tested in a rectangular room. This led scientists to believe there was a so-called geometric module in the rat's brain which is responsible for dealing with geometric information (Cheng, 1986). However (Graham, Good, McGregor, & Pearce, 2006) found that rats showed poor performances in studies which tested the use of geometric information. Nevertheless the phenomenon of mistaking opposite positions in a room was also found in human children as well as in adults during specific circumstances (Learmonth, Nadel, & Newcombe, 2002) (Hermer-Vazquez, Spelke, & Katsnelson, 1999). These studies were also supported later on by (Stürzl, Cheung, Cheng, & Zeil, 2008). Animals and human children seem to make these rotational errors even with other visual features available. In our study there were no visual objects that the participants could make use of except for the depth estimations, which could be the reason why even adults made these rotational errors. According to (Stürzl, Cheung, Cheng, & Zeil, 2008) snapshot memory is the reason for systematic rotational error as the snapshots of opposite corners had the least image differences.

The reason for the fewer numbers of rotational errors during the random dot condition than the texture condition could be explained with the theory of (Stürzl, Cheung, Cheng, & Zeil, 2008). Snapshots rely on fixed image features which aren't available during the random dot condition.

In conclusion our results indicate that motion parallax alone does not give enough depth information to fulfil a homing task and therefore a binocular view must be available to the participants. Also our results show that landmark information is not needed to fulfil the homing task as already assumed by Prozmann.

Bibliography

Bradshaw, M. F., & Rogers, B. J. (1996). The Interaction of Binocular Disparity and Motion Parallax in the Computation of Depth. *Vision Res., Vol.36, No.21*, 3457-3468.

Cartwright, B. A., & Collett, T. S. (1983). Landmark Learning in Bees. Experiments and Models. *Journal of Comparative Physiology 151*, 521-543.

Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition, 23*, 149-178.

Durier, V., Graham, P., & Collett, T. S. (2003). Snapshot Memories and Landmark Guidance in Wood Ants. *Current Biology*, 1614-1618.

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

Graham, M., Good, M. A., McGregor, A., & Pearce, J. M. (2006). Spatial learning based on the shape of the environment is influenced by properties of the objects forming the shape. *Journal of Experimental Psychology: Animal Behavior Processes, Vol. 32, No. 1*, 44-59.

Graham, P., & Collett, T. S. (2002). View-based navigation in insects: how wood ants (Formica rufa L.) look at and are guided by extended Landmarks. *The Journal of Experimental Biology*, 2499-2509.

Hermer, L., & Spelke, E. S. (1994). A geometric process for spatial reorientation in young children. *Nature*, 57-59.

Hermer-Vazquez, L., Spelke, E. S., & Katsnelson, A. S. (1999). Sources of flexibility in human cognition: Dual-task studies of space and language. *Cognitive Psychology, Vol. 39, No. 1*, 3-36.

Kollin, J., & Hollander, A. (2007). Re-engineering the Wheatstone stereoscope. *The International Society for Optical Engineering*.

Learmonth, A. E., Nadel, L., & Newcombe, N. S. (2002). Children's use of landmarks: Implications for modularity theory. *Psychological Science, 13*, 337-341.

Lynch, K. (1960). *The image of the city.* Cambridge: MIT Press.

Maunsell, J. H., & van Essen, D. C. (1983). Functional properties of neurons in the middle temporal visual are of the macaque monkey, II: Binocular interactions and sensitivity to binocular disparity. *Journal of Neuroscience, 49*, 1148-1167.

Poggio, G. F., & Talbot, W. H. (1981). Mechanisms of static and dynamic stereopsis in foveal cortex of the rhesus monkey. *Journal of Physiology, 315*, 469-492.

Prozmann, V. (2015). The Role of Stereoscopic Depth Cues in Place Recognition.

Purves, D., Brannon, E. M., Cabeza, R., Huettel, S. A., LaBar, K. S., Platt, M. L., & Woldorff, M. G. (2008). Principles of Cognitive Neuroscience. Sunderland, Massachusetts USA: Sinauer Associates, INc.

Rogers, B., & Graham, M. (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision Research, 22*, 261-270.

Rossel, S. (1983). Binocular stereopsis in an insect. *Nature*, 821-822.

Siegel, A. W., & White, S. (1975). The development of spatial representations of large scale environments. *Advances in child development and behavior: Academic Press*, 10-55.

Stürzl, W., Cheung, A., Cheng, K., & Zeil, J. (2008). The Information Content of Panoramic Images I: The Rotational Errors and the Similarity of Views in Rectangular Experimental Arenas. *Journal of Experimental Psychology: Animal Behaviour Processes, Vol. 34, No. 1*, 1-14.

Appendix

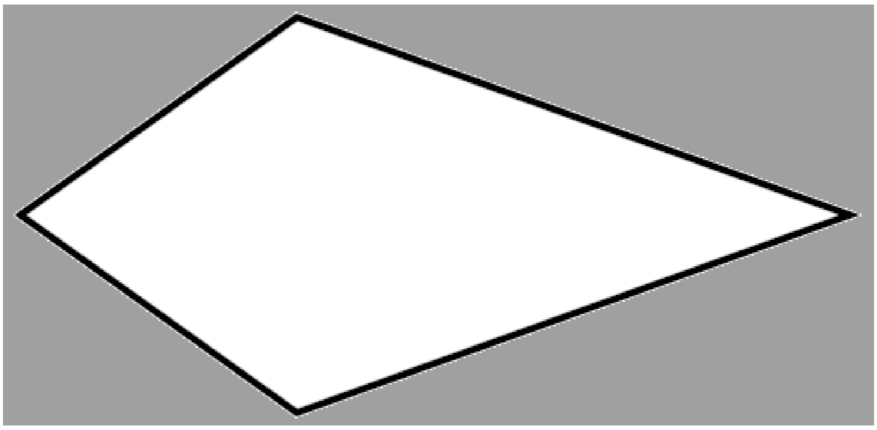
Experimental Description

Liebe(r) Proband(in),

die beiden heutigen Experimente bestehen aus einem graphischen Programm welches vor Dir auf den beiden Bildschirmen präsentiert wird. Dabei solltest Du während der Experimente nur auf den Bildschirm vor Dir blicken. Der Spiegel vor Dir wird dabei so positioniert, dass Du nur mit einem Auge auf den vorderen Bildschirm schaust und mit dem anderen Auge auf den Spiegel.

Im ersten Experiment wirst Du zwei unterschiedlich weit entfernte Quadrate zu sehen bekommen, welche durch flimmernde Punkte auf schwarzen Grund dargestellt werden. Dabei sollst Du jeweils entscheiden, welches Quadrat Dir ***näher*** erscheint und dies durch Drücken der jeweiligen Maustasten (linke Taste = linkes Quadrat, rechte Taste = rechtes Quadrat) bestätigen.

Im zweiten Experiment wirst Dich in einem drachenförmigen Raum wiederfinden, dessen Struktur, ebenso wie die Quadrate aus dem ersten Experiment, durch flimmernde, weiße Punkte auf schwarzem Grund festgelegt ist.



Zu Beginn jedes Versuchsdurchganges kannst Du den Durchgang durch Drücken der Leertaste starten. Du befindest Dich nun an einem Platz (Startpunkt) innerhalb des Raumes. Hier kannst Du Dich durch Bewegen der Maus im Raum umdrehen und Dich in einem kleinen Bereich vorwärts ("W"), rückwärts ("S") und seitwärts ("A", "D") bewegen. Dabei sollst Du Dir die Position genau anschauen und merken. Bist Du Dir sicher, dass Du Dir die Position gemerkt hast, kannst Du den zweiten Teil durch erneutes Drücken der Leertaste beginnen. Jetzt befindest Du Dich wieder erneut an einem Ort im Raum. Ziel ist es nun, zum zuvor gezeigten Startpunkt zurückzulaufen. Dabei kannst Du Dich mithilfe der Maus drehen und Dich vorwärts ("W"), rückwärts ("S") und seitwärts ("A", "D") bewegen. Dabei solltest Du Dich zu Beginn wieder umschauen und dann möglichst zügig und ohne Umwege zum Startpunkt zurücklaufen. Bist Du Dir sicher den Startpunkt gefunden zu haben, bestätige dies durch Drücken der Leertaste. Damit ist ein Durchgang abgeschlossen. Der nächste Durchgang beginnt mit einem schwarzen Bildschirm und den Worten „Durchgang starten durch Drücken der Leertaste“. Möchtest Du eine Pause einlegen, ist dies die Gelegenheit. Die anderen Teile des Experimentes führe bitte zügig aber genau aus.

Der Versuchsleiter wird Dir die einzelnen Teile nochmal genau erklären.

Viel Spaß! ☺

Questionnaire

Fragebogen

Geschlecht: m

Alter: 22

Studiengang: Biologie (Master Neurobiologie)

Kreise die Zahl die Deiner Antwort am ehesten entspricht:

- Hat Dir das Experiment Spaß gemacht?

  sehr wenig 1 – 2 – 3 – 4 – 5 – 6 – 7 sehr viel

- Warst du motiviert?

  sehr wenig 1 – 2 – 3 – 4 – 5 – 6 – 7 sehr viel

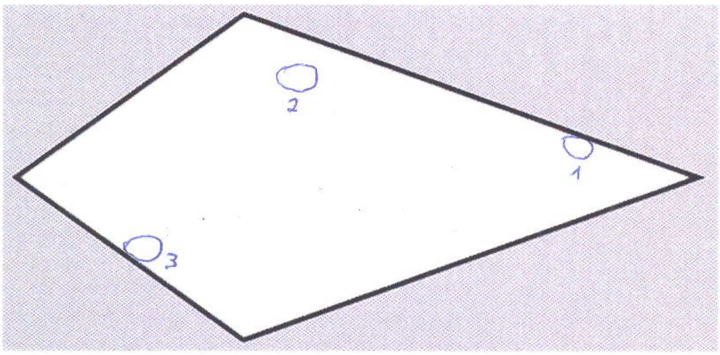
- Wie schwierig fandest du das Experiment?

  sehr leicht 1 – 2 – 3 – 4 – 5 – 6 – 7 sehr schwer

- Gab es Positionen bei denen Du es einfacher fandest, sie wieder zu finden?

  Position 1

- Kannst Du Deine Zielpunkte in den dargestellten Raum eintragen:



Vielen Dank für Deine Teilnahme!

Fragebogen

Geschlecht: weiblich

Alter: 23

Studiengang: Biologie Master

Kreise die Zahl die Deiner Antwort am ehesten entspricht:

- Hat Dir das Experiment Spaß gemacht?

  sehr wenig 1 – 2 – 3 – 4 – 5 – 6 – 7 sehr viel

- Warst du motiviert?

  sehr wenig 1 – 2 – 3 – 4 – 5 – 6 – 7 sehr viel

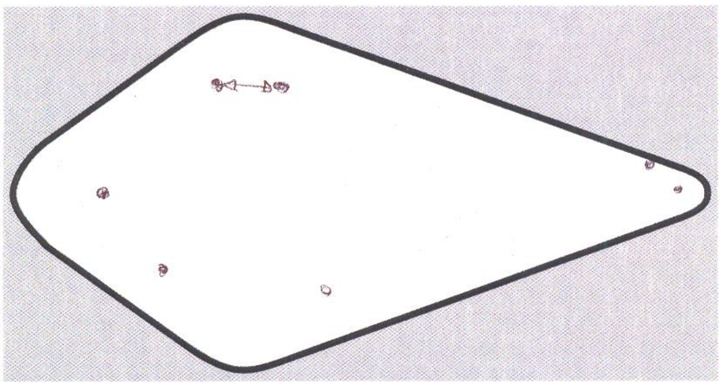
- Wie schwierig fandest du das Experiment?

  sehr leicht 1 – 2 – 3 – 4 – 5 – 6 – 7 sehr schwer

- Gab es Positionen bei denen Du es einfacher fandest, sie wieder zu finden?

  Positionen in der langen Ecke

- Kannst Du Deine Zielpunkte in den dargestellten Raum eintragen:



Vielen Dank für Deine Teilnahme!

Fragebogen

Geschlecht: weiblich

Alter: 21

Studiengang: Bio Bsc.

Kreise die Zahl die Deiner Antwort am ehesten entspricht:

- Hat Dir das Experiment Spaß gemacht? Nein

  sehr wenig 1 – 2 – 3 – 4 – 5 – 6 – 7 sehr viel

- Warst du motiviert?

  sehr wenig 1 – 2 – 3 – 4 – 5 – 6 – 7 sehr viel

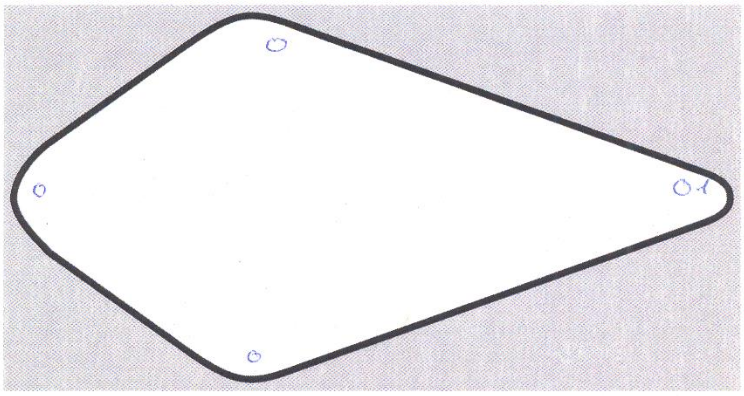
- Wie schwierig fandest du das Experiment?

  sehr leicht 1 – 2 – 3 – 4 – 5 – 6 – 7 sehr schwer

- Gab es Positionen bei denen Du es einfacher fandest, sie wieder zu finden?

  o' am leichtesten

- Kannst Du Deine Zielpunkte in den dargestellten Raum eintragen:



Vielen Dank für Deine Teilnahme!