Allocentric information is used for memory-guided reaching in depth: A virtual reality study

Mathias Klinghammer [a], Immo Schütz [b], Gunnar Blohm [c], Katja Fiehler [a],*

[a] Justus-Liebig-University, Experimental Psychology, Otto-Behaghel-Str. 10F, 35394 Giessen, Germany
[b] TU Chemnitz, Institut für Physik, Reichenhainer Str. 70, 09126 Chemnitz, Germany
[c] Queen's University, Centre for Neuroscience Studies, 18, Stuart Street, Kingston, Ontario K7L 3N6, Canada

ABSTRACT

Previous research has demonstrated that humans use allocentric information when reaching to remembered visual targets, but most of the studies are limited to 2D space. Here, we study allocentric coding of memorized reach targets in 3D virtual reality. In particular, we investigated the use of allocentric information for memory-guided reaching in depth and the role of binocular and monocular (object size) depth cues for coding object locations in 3D space. To this end, we presented a scene with objects on a table which were located at different distances from the observer and served as reach targets or allocentric cues. After free visual exploration of this scene and a short delay the scene reappeared, but with one object missing (=reach target). In addition, the remaining objects were shifted horizontally or in depth. When objects were shifted in depth, we also independently manipulated object size by either magnifying or reducing their size. After the scene vanished, participants reached to the remembered target location on the blank table. Reaching endpoints deviated systematically in the direction of object shifts, similar to our previous results from 2D presentations. This deviation was stronger for object shifts in depth than in the horizontal plane and independent of observer-target-distance. Reaching endpoints systematically varied with changes in object size. Our results suggest that allocentric information is used for coding targets for memory-guided reaching in depth. Thereby, retinal disparity and vergence as well as object size provide important binocular and monocular depth cues.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The human brain makes use of egocentric (relative to the observer) and allocentric (relative to objects in the environment) reference frames (Battaglia-Mayer, Caminiti, Lacquianiti, & Zago, 2003; Colby, 1998; Klatzky, 1998) to encode object locations for action in the environment. Previous studies demonstrated that egocentric, and in particular gaze-centered, reference frames are predominantly utilized when planning and executing reaching movements toward the remembered location of a visual target (e.g. Cohen & Anderson, 2002; Fiehler, Schütz, & Henriques, 2011; Thompson & Henriques, 2011). However, other studies also revealed evidence for the use of allocentric reference frames for memory-guided reaching (e.g. Diedrichsen, Werner, Schmidt, & Trommershäuser, 2004; Krigolson, Clark, Heath, & Binsted, 2007; Krigolson & Heath, 2004; Obhi & Goodale, 2005) arguing for a combined use of both classes of coding schemes (Byrne & Crawford, 2010; Schütz, Henriques, & Fiehler, 2013, 2015).

Since most of the previous work used rather artificial stimuli like dots and bars, recent work aimed to increase ecological validity of the outcomes by using more naturalistic stimuli (Camors, Jouffrais, Cottereau, & Durand, 2015; Fiehler, Wolf, Klinghammer, & Blohm, 2014; Klinghammer, Blohm, & Fiehler, 2015). For example, in a previous study we presented computer generated images of a breakfast table on a computer screen and asked participants to memorize the locations of six objects on the table (Klinghammer et al., 2015). Then, the whole scene vanished and after a brief delay the scene reappeared for 1000 ms with one of the objects missing and the remaining objects shifted either to the left or to the right. Participants were instructed to reach to the location of the missing object on a grey screen while keeping gaze fixed. Reaching endpoints systematically deviated into the direction of the shifts of the remaining objects suggesting that allocentric information

---

was used to encode the location of the reach target which was then integrated into the reach plan. In the present study, we aim to extend the outcomes of our preceding work from 2D to 3D space by transferring our established paradigm to virtual reality. This allows us to examine the use of allocentric information for memory-guided reaching not only in the horizontal axis but also in depth in real-world-like situations and to determine the role of binocular and monocular (i.e., object size) depth cues for allocentric encoding of memorized object locations when reaching in depth.

So far, we presented 2D stimuli and shifted objects in the left-right (horizontal) plane (Fiehler et al., 2014; Klinghammer et al., 2015). But how do object shifts in depth affect memory-guided reaching movements? It has been demonstrated that delayed pointing to a single target in the dark leads to pointing errors in the horizontal plane that are uncorrelated with pointing errors in the depth plane arguing for two independent subsystems for retaining target locations for action (Chieffi & Allport, 1997). Moreover, research on the Induced Roelofs Effect (IRE) (Coello, Richaud, Magne, & Rossetti, 2003), which describes the misestimation of the position of a target dot placed within a frame into the direction of the closest edge of this frame, shows that the orientation of the surrounding frame influences perception and action differently. While for a horizontally oriented frame the misestimation of the target dot was only found for perceptual judgements, for a frame orientation in depth this misestimation was also observed for memory-guided reaching movements. This suggests that the reach system is especially sensitive to contextual information, when the processing of depth cues is emphasized. By applying a similar IRE paradigm, Neely, Heath, and Binsted (2008) in contrast showed that reaching endpoints were influenced by both orientations of the frame. The authors concluded that one unitary visual system integrates allocentric and egocentric information for both orientation and distance of reaching movements. Thus, it is still unclear whether reaching targets are similarly or differently affected by allocentric information in the distance versus the directional axis. Here, we investigate the use of allocentric information for memory-guided reaching in the horizontal and the depth plane in a more naturalistic environment.

To perceive depth in a visual environment without self-motion, the human brain makes use of monocular (e.g., occlusion, height in the visual field, relative size) and binocular (e.g., binocular disparity, accommodation, vergence) depth cues. Depending on the distance between the observer and object locations in depth, the multiple depth cues are weighted and combined in different ways (Armbrüster, Wolter, Kuhlen, Spijkers, & Fimm, 2008; Cutting, 1997; Knill, 2005; Landy, Maloney, Johnston, & Young, 1995). One strong binocular depth cue for estimating objects' distances in depth is binocular disparity (Bingham, Bradley, Bailey, & Vinner, 2001). A virtual-reality-device such as the Oculus Rift DK2 (Oculus VR, LLC, Menlo Park, CA, USA) makes use of binocular disparity by presenting a slightly shifted perspective of the same scene to the two eyes, mimicking real world perception. In that sense, vergence can also be used providing a reliable depth cue within reaching space (Tresilian, Mon-Williams, & Kelly, 1999; Viguier, Clément, & Trotter, 2001). However, especially actions like prehension of objects need accurate metric depth information which cannot be provided by binocular cues alone (Hibbard & Bradshaw, 2003), but require the use of additional monocular depth cues for accurate depth perception (Bruno & Cutting, 1988; Magne & Coello, 2002).

For example, in a virtual environment study by Naceri, Chellali, and Hoinville (2011), a sphere located in different depths in front of the participants was used as pointing target. In one condition, the absolute size of the sphere was manipulated in a way that irrespective of its actual location in depth, the angular size (i.e., the retinal size) was kept constant. The results demonstrated that the absolute size manipulation influenced depth perception in a subgroup of participants. Regardless of the actual depth position of the sphere, they pointed to the same position as indicated by the constant angular size of the sphere. Based on this finding, the authors concluded that the object size was used as the main depth cue for pointing. The remaining participants were not influenced by the size manipulation and pointed to the correct location of the sphere according to its position in depth. This suggests that in this group of participants vergence was used as the dominant depth cue. In a later study, they again found that around half of the participants relied on object size and misjudged target depth when they verbally estimated target distances in a virtual reality, whereas the other half made use of vergence and correctly reported object distances (Naceri, Moscatelli, & Chellali, 2015). Hence, object size provides one important depth cue which can influence the perceived location of targets for action.

In this study, we aimed to answer two major questions. First, in order to test for potential differences when reaching to objects in virtual reality, we wanted to replicate our previous findings from a 2D paradigm (Klinghammer et al., 2015) in 3D virtual reality. Second, with the possibility of extending space to the third dimension, we wanted to know whether and how allocentric information is utilized for encoding the location of targets in depth for memory-guided reaching and how this is influenced by binocular and monocular (object size) depth cues.

For this purpose, we conducted two experiments. In experiment 1, we transferred our paradigm of Klinghammer et al. (2015) to 3D virtual reality and shifted objects on a breakfast table horizontally before reaching to the remembered location of a visual target. Moreover, we placed objects at three different distances from the observer to test whether 2D effects were consistent across different depth planes. In experiment 2, we used the same paradigm but this time shifted objects in depth and additionally manipulated the depth plane and the size of the objects serving as allocentric cues.

2. Experiment 1

2.1. Introduction

In order to extend the findings from our previous studies (Fiehler et al., 2014; Klinghammer et al., 2015) to a more realistic environment, we aimed to replicate the results from the 2D paradigm in 3D virtual reality. Participants wore a head-mounted display and had to encode the location of several virtual objects on a virtual table before performing a memory-guided reaching movement to the location of a remembered target object. Between scene viewing and reaching, the remaining objects were shifted horizontally. Moreover, object clusters were placed in three different distances to the observer. Based on our previous findings using 2D images (Klinghammer et al., 2015), we expect a similar systematic deviation of reaching endpoints in the direction of lateral object shifts. Since coding of reach targets in the horizontal plane should be independent from coding of reach targets in the sagittal plane (Chieffi & Allport, 1997), we expect lateral deviations of reaching endpoints to be independent of the observer-target distance.

2.2. Methods

2.2.1. Participants

Thirteen volunteers participated in the experiment (6 female), aged 19–31 years (mean 23.7 ± SD 3.9 years). All had normal or corrected-to-normal vision and intact stereo vision as determined by the Graded circle test (part of the Stereo fly test, STEREO OPTI-

CAL CO., INC., Chicago, IL, USA). Individual inter-ocular distances (IODs) were detected using the eye tracker integrated into the head-mounted display (HMD; see below) and entered into the presentation software to adjust stereo rendering (mean IOD 60 ± SD 2 mm). Participants were right-handed as confirmed by the Edinburgh Handedness Inventory (EHI, Oldfield, 1971; *mean handedness quotient* 85.8 ± SD 17.8) and reported no known visuo-motor or neuromuscular deficits. The study was approved by the local ethical committee and followed the statutes of the Declaration of Helsinki (2008). All participants gave written informed consent and received money or course credits for their participation.

2.2.2. Apparatus

Participants were seated at a table, which was equipped with a fixed chin rest and a decimal keyboard at the participant's left side on which one key was used as a button to control the experiment. They were instructed to keep their head stationary throughout the experiment. The chin rest was adjusted so that the participant's eyes were 35 cm above the front table edge. The table surface was otherwise blank to ensure unimpeded reaching. Visual stimuli were generated using Vizard 5.1 (WorldViz, LLC, Santa Barbara, CA, USA) and presented stereoscopically within the Rift DK2 HMD at a resolution of 960 × 1080 pixels per eye and a refresh rate of 75 Hz. Objects in the virtual reality were aligned to and presented at the same position as their counterpart in the real world (table, start position, hand position). Head rotation angles were recorded using the DK2's integrated positional tracker at a rate of 1000 Hz and used to update participants' virtual view point. Eye movements within the virtual reality were recorded using an infrared camera based eye tracker custom made for the HMD (SensoMotoric Instruments GmbH, Teltow, Germany) at 60 Hz. Reach movements were tracked using Optotrak Certus (NDI, Waterloo, ON, Canada) at 250 Hz with one infrared marker attached to the right index finger. All gaze and motion data was resampled to 75 Hz and recorded from the Vizard 3D presentation software.

2.2.3. Materials

Participants stereoscopically viewed a virtual room consisting of a black floor and beige back wall (distance from viewpoint: 1.35 m), as well as a brown cube which served as a table matching the position and dimensions of the real table in front of them (height × width × depth = 71 cm × 80 cm × 80 cm; see also Fig. 1A). A small (2 cm to a side) black cube indicated the location of the start position on the real table. A red sphere (3 cm in diameter) was added to indicate the position of the right index finger tip to the participant. For this purpose the finger location was permanently updated by the data recorded from the Optotrak. On the table surface, we presented six table objects (TO) as possible reach targets (apple, butter tray, espresso cooker, egg, mug and jam jar) in 12 different arrangements. Objects were taken from the open access online 3D-gallery of SketchUp (Trimble Navigation Ltd., Sunnyvale, CA, USA) and converted for the use in Vizard 5.1. Because original object sizes were too big to be able to shift objects left- or rightwards without occluding the fixation cross, we decreased object sizes by 10% in every direction (height, width, depth) from the original object sizes (for object properties see Table 1). The 12 arrangements were used as *encoding scenes*. The TOs were placed in three different depth clusters, balancing target object positions across the clusters. Each of these clusters contained three horizontal depth lines that were 20.0, 27.5, and 35.0 cm (near cluster), or 27.5, 35.0, and 42.5 cm (medium cluster), or 35.0, 42.5, and 50.0 cm (far cluster) away from the front edge of the table with minimal 1 and maximal 3 objects placed in one depth line (see Fig. 1). Moreover, object positions were chosen such that they did not occlude the fixation cross and were never positioned at or beyond the table edge or close to the edge of the visual field, which was restricted by the HMD. Furthermore, each object never occluded more than 20% of another object. Based on the encoding scenes we defined *test scenes* in which one TO was missing (= *reach target*). Every TO served as target equally often and in random order. In two-thirds of these test scenes the remaining TOs were shifted together horizontally by 4 degrees of visual angle (2.15–4.98 cm depending on the cluster and depth line and distance to the table midline; 50% leftward shifts). In the remaining third of test scenes, which served as control condition, no objects were shifted. All in all we defined 72 different encoding scenes leading to 72 test scenes of the control condition and 144 test scenes of the shift conditions.

Moreover, we defined a *mask scene* consisting of 300 grey cubes rendered at an angle of 45° (20 cm side length) that were placed randomly in the participant's field of view (extending 3 m horizontally and vertically and 5 m in depth).

2.2.4. Procedure

Fig. 2 depicts the procedure of an example trial. After participants placed their right index finger on the start position, each trial began with the presentation of one of the 12 encoding scenes. Participants freely explored this scene visually at their own pace and pressed the button with their left index finger to proceed. Then, the mask image to prevent afterimages was presented for 250 ms. The mask was followed by a delay of 1800 ms, during which only the black floor, the table with the start position, the red sphere for the fingertip and the wall were visible. In addition a fixation cross was presented at the center of the middle depth line of the respective depth cluster in order to control for the use of a gaze-centered egocentric reference frame (Thompson & Henriques, 2011). Participants were instructed to fixate the cross and keep their gaze at this location until the end of the trial. After the delay, the test scene appeared for 1000 ms with the target object missing. Then a blank table with the fixation cross, the start position, and the finger position was presented in front of the wall. As soon as the test scene vanished, a short tone was presented cueing participants to perform a reaching movement with their right hand to the remembered target location on the table while fixating on the fixation cross. They were instructed to press the button with their left hand while their right index fingertip remained at the desired target location on the table. After participants pressed the button, the trial ended with a black screen and they returned the right index finger to the start position. After a brief delay the next trial started. All in all, every participant completed 216 trials. Trials within a session were presented in randomized order. The session was repeated once after a short break leading to a total number of 432 trials and an overall experiment duration of approx. 1.30 h per participant.

2.2.5. Data Reduction and Statistical Analysis

Reach and eye movement data was processed using MATLAB R2007b (TheMathWorks, Inc, Natick, MA, USA). Statistical analyses were performed using R 3.2.2 (R Development Core Team). An alpha level of 0.05 was used to evaluate all statistical tests. If correction for multiple testing was necessary, Bonferroni-Holm correction was applied. In case the assumption of sphericity for an analysis of variance (ANOVA) was violated (tested with Mauchly's sphericity test), Greenhouse-Geisser correction was applied.

Before data preprocessing, 20 trials of each participant had to be omitted due to an error in the stimulus material. Thus, the number of trials was reduced from 432 to 412 for each participant. First, we inspected data for incorrect fixation behavior. To this end, we calculated eye velocity in every trial for every participant, from 300 ms after the onset of the delay until the end of the reaching movement (fixation period). The 300 ms delay was chosen to exclude participants' initial saccade toward the fixation cross at the beginning of
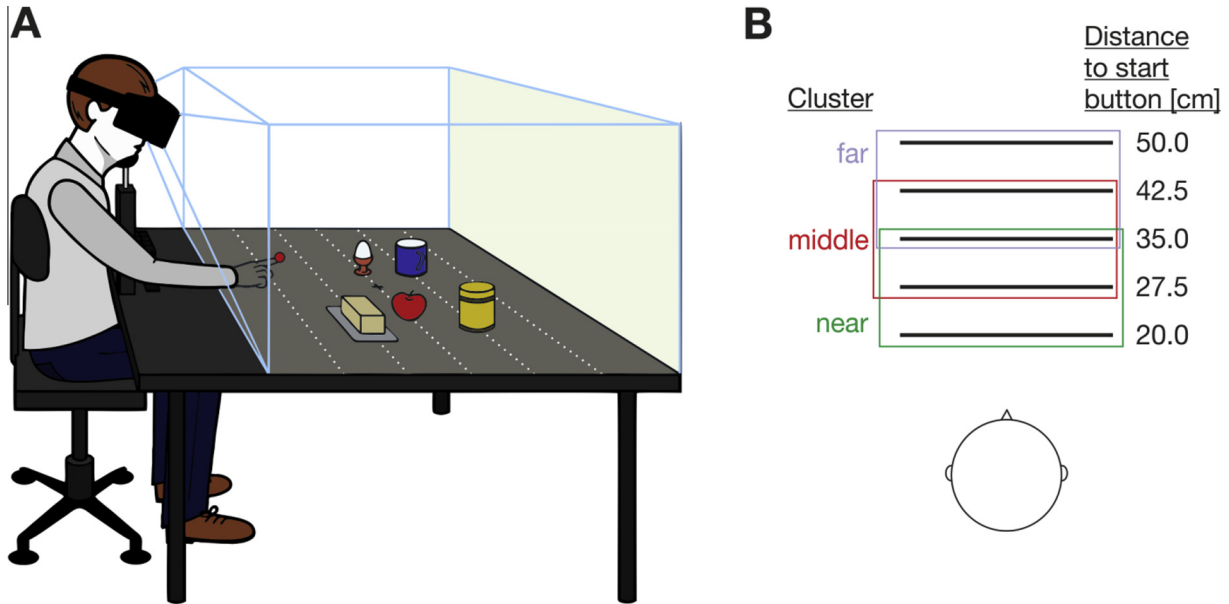
Fig. 1. (A) Schematic view of the VR setup. Participants wore the HMD and sat at a table with their head rested on a chin rest. The presented virtual environment (i.e., table, objects) was aligned with the real world properties of the table. Thus, when participants performed a reaching movement on the table, the red sphere (representing the tip of the right index finger) touched the virtual table at the same time and position as their fingertip touched the physical table. In this example, objects are positioned in the middle depth cluster on three depth lines. For the near and far depth cluster, objects were placed one depth line closer to the participant or one depth line farther away, respectively. (B) Schematic top view on the table representing depth clusters and corresponding depth lines. Distances of the depth lines to the start button are depicted in cm. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1

Height, width and depth of table objects in cm, based on the actual properties in the virtual reality (*increased size by 10%/**original size**/decreased size by 10%*).

Object	Height	Width	Depth
Apple	7.7/**7**/6.7	7.7/**7**/6.7	7.7/**7**/6.7
Butter tray	5.5/**5**/4.5	8.8/**8**/7.2	13.2/**12**/10.8
Egg	6.5/**6**/5.4	4.4/**4**/3.6	4.4/**4**/3.6
Espresso cooker	15.4/**14**/12.6	15.4/**14**/12.6	16.5/**15**/13.5
Vegemite jar	12.1/**11**/9.9	7.7/**7**/6.3	7.7/**7**/6.3
Mug	8.8/**8**/7.2	12.1/**11**/9.9	8.8/**8**/7.2

the delay phase. Within a trial, frames with velocities above 500 deg/s were excluded, as this velocity typically represents eye blinks rather than saccades (Ostendorf, Fischer, Finke, & Ploner,

2007). We then calculated the mean velocity of the fixation period and excluded trials with a mean velocity above 20 deg/s, indicating a saccadic eye movement within this critical time period (226 trials = 4.22%). Subsequently, we analyzed reaching data to detect movement on- and offsets. Movement onsets were defined as the first of four consecutive time frames with a velocity higher than 3 cm/s and acceleration greater than 2 cm/s². Movement offset was defined as the first time frame after movement onset when the movement velocity dropped below 3 cm/s. Offsets with physically impossible coordinates below the table surface indicated measurement errors and were excluded. Moreover, trials with at least 20 consecutive missing data frames within the critical period, i.e. when participants should touch the table, or trials where this critical period could not be determined, were discarded. Trials without a
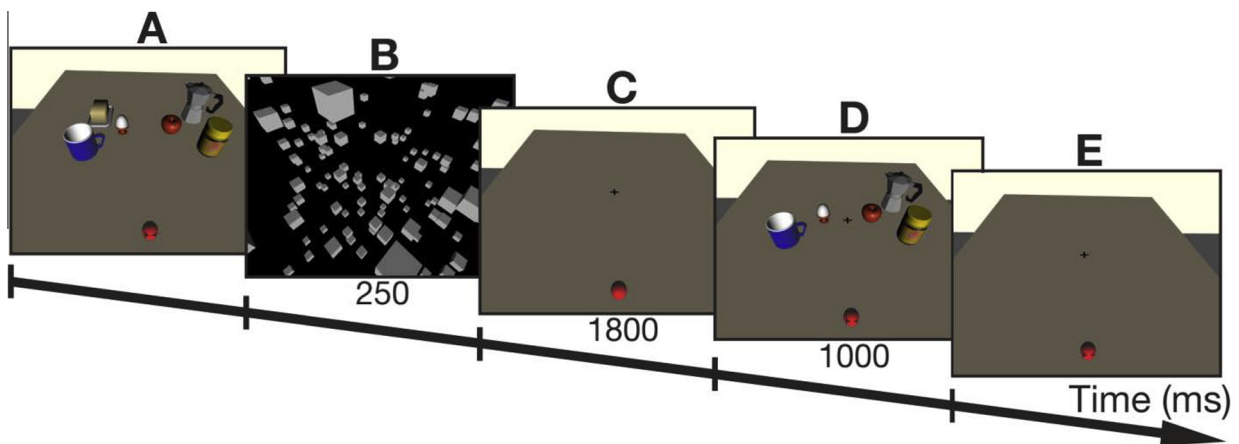
Fig. 2. Time course of an example trial. The presentation of the encoding image (A) was followed by a mask (B) to prevent afterimages. After a delay from which on a fixation cross appeared in the middle of the depth cluster (C) the test scene was shown (D) with one target object missing (in this example the butter tray). Participants had to fixate the fixation cross until the end of the trial. Then the scene vanished and participants performed a reaching movement toward the remembered location of the target object on the blank table with the little red sphere indicating the position of the tip of their right index finger (E). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

movement or with a movement onset before the auditory cue were also excluded (in total 856 trials = 16.86%).

After we extracted reaching endpoints at the offset of the reaching movement, we performed outlier correction for the control condition, excluding trials when the reaching endpoint deviated more than 2.5 SD from the group mean of the corresponding combination of scene arrangement and reach target in the horizontal and depth axis. As no objects were shifted in this condition, we expected participants to reach precisely to the perceived target location. Thus, we used the group means of every arrangement and target combination of the control condition as the actual target positions and calculated horizontal reaching errors and errors in depth by subtracting reaching endpoints in the shift conditions from the corresponding group mean of the control condition. These reaching errors were also outlier corrected by excluding trials with reaching errors deviating more than 2.5 SD from the group mean, separately for each combination of condition and direction of object shifts. Taken together, 174 trials (= 4.08%) were classified as outliers. All in all from originally 5346 trials, 4091 trials entered into further analysis (= 76.38%).

To investigate the influence of allocentric information on reaching endpoints, we calculated allocentric weights for every participant and condition using a linear fit between the actual horizontal reaching errors (i.e., errors in the direction of object shifts) and *maximal expected reaching errors* (MERE). MERE are the expected errors in case a participant completely relies on allocentric information when performing a reaching movement. For example, when objects were shifted by 4 cm to the left, we expect a maximum reaching error of 4 cm to the left. To calculate MERE of different objects in different horizontal and depth locations, we averaged the individual horizontal shift distances of all remaining objects. Thus, the slope of the linear fit between actual and maximal expected reaching errors can be defined as a measure representing to which extent allocentric information was taken into account when reaching to a remembered target, with a slope of 1 for complete reliance on allocentric information and a slope of 0 for no reliance.

We performed two-sided one-sampled *t*-tests to check whether allocentric weights of the different conditions differed significantly from zero. For a more direct comparison of our current results and results from corresponding conditions of our previous study using 2D images (Klinghammer et al., 2015), we performed a two-sided *t*-test for independent samples with allocentric weights of the current experiment averaged over the depth clusters and allocentric weights of the previous study for corresponding conditions with also horizontal shifts of five task-relevant objects. In the previous study, these were the condition with five shifted table objects in experiment 1 (TO-5) and the condition with five shifted background objects in experiment 2 (BO-5). To test for an influence of the distance between observer and target we entered allocentric weights into a one-way repeated measure ANOVA with the factor depth cluster (near, middle, and far). We performed the same ANOVA on standard deviations of the horizontal reaching errors to investigate differences in reaching variability. For both ANOVAs, we conducted two-sided post hoc *t*-tests for paired samples in case of significant main effects. To investigate the influence of the distance between objects and observer on reaction times (time between go cue and reach onset) and movement durations (time between reach onset and offset), we conducted a one-way repeated measure ANOVA for each of these dependent variables.

2.3. Results

In Table 2, the descriptive data for horizontal reaching errors (i.e., errors in the direction of object shifts) and the corresponding means of the MEREs are summarized. Note, that objects in different

depth clusters were shifted by different absolute distances to keep the visual angle of these shifts constant (as also represented by increasing MEREs), thus, biasing the absolute reaching errors. However, when calculating allocentric weights we normalized for these differences by taking different object shift distances as predictors into account.

As depicted in Fig. 3A, averaged reaching endpoint errors for single participants deviated systematically in the direction of horizontal object shifts. Fig. 3B shows the linear fit between MEREs and actual reaching errors of one exemplary participant for horizontal object shifts in the second depth cluster.

We quantified reaching errors by calculating allocentric weights as described above. Averaged weights for horizontal object shifts differed significantly from zero in all depth clusters (see Table 3).

Mean allocentric weights over all clusters of the current study and allocentric weights of corresponding conditions of our previous research are depicted in Fig. 4A. The two-sided *t*-test for independent samples between averaged allocentric weights of the current experiment and those of our previous study (Klinghammer et al., 2015) revealed no differences (current vs. TO-5 in previous Exp. 1: $t(20.385) = 0.423$, $p = 0.656$; current vs. BO-5 in previous Exp. 2: $t(18.898) = 0.325$, $p = 0.749$).

One-way repeated measures ANOVA assessing the influence of distance between target and observer obtained no main effect on allocentric weights ($F(2,24) = 0.997$, $p = 0.359$). The one-way repeated measures ANOVA investigating the influence of distance between target and observer on the variability of reaching endpoints revealed a main effect for the depth clusters ($F(2,24) = 4.578$, $p = 0.021$; see also Fig. 4B). Post-hoc *t*-tests showed only differences between the middle and far depth clusters ($t(12) = -3.493$, $p = 0.013$) with a higher variability for reaching to targets in the far than the middle depth cluster. Other pairwise comparisons did not reach significance (all $p > 0.176$).

The one-way repeated measures ANOVA assessing the influence of depth clusters on reaction times revealed no differences ($F(2,24) = 0.671$, $p = 0.521$; *mean near/middle/far Cluster:* 284 ms/279 ms/281 ms). As expected, due to longer reaches to targets further in depth, the one-way repeated measures ANOVA on movement durations revealed a main effect for the factor depth cluster ($F(2,24) = 136.667$, $p < 0.001$; *mean near/middle/far Cluster:* 553 ms/604 ms/662 ms). Post-hoc *t*-tests revealed longer movement times for the depth clusters middle than near ($t(12) = -12.145$, $p < 0.001$), far than near ($t(-12) = -12.109$, $p < 0.001$), and far than middle ($t(12) = -10.283$, $p < 0.001$).

3. Experiment 2

3.1. Introduction

The aim of experiment 2 was to investigate whether allocentric information is also utilized when encoding the location of memory-guided reach targets in depth. We used the same paradigm as in experiment 1, but this time shifted objects in depth. Based on the findings by Coello et al. (2003) and Neely et al. (2008) suggesting that memory-guided reaching movements are especially sensitive to contextual information, when the processing of depth cues is emphasized, we expect systematic deviations of reaching endpoints in the direction of object shifts. Moreover, allocentric weights should be sensitive to the manipulation of depth cues influencing contextual information.

To examine the use of binocular depth cues on allocentric coding of remembered reach targets in depth, we presented objects in three different depth clusters. In our paradigm, participants could use vergence and retinal disparity during scene encoding in which they freely move their gaze. However, in the test scene participants

Table 2
Summary of horizontal reaching errors for all depth clusters and direction of object shifts in cm. Range, mean and standard deviation of the sample are listed. Additionally, the means of the MEREs for every condition are listed in cm as well. Negative values are assigned to leftward and positive values to rightward object shifts.

Cluster	Shift direction	Range	Mean	SD	Mean MERE
Near	Left	−3.452 to 1.401	−1.526	2.564	−3.355
	Right	0.830 to 4.026	1.707	2.388	3.345
Middle	Left	−3.570 to −0.802	−2.040	2.288	−3.731
	Right	−0.484 to 3.220	1.792	2.426	3.655
Far	Left	−3.535 to −0.629	−2.139	2.831	−4.150
	Right	−0.216 to 3.526	2.143	2.621	4.150
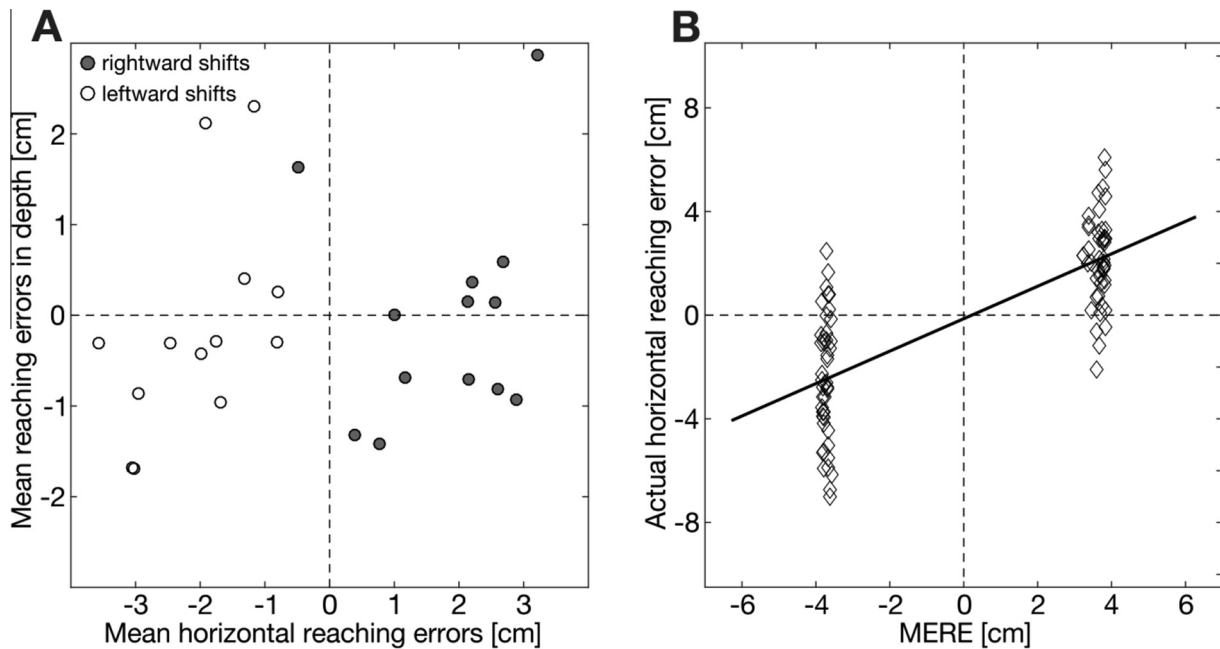
Fig. 3. (A) Mean reaching errors for single participants in cm for horizontal object shifts in the middle depth cluster. (B) Example of a linear fit between MEREs and actual horizontal reaching errors for one participant for horizontal object shifts in the middle depth cluster.

Table 3
Summary of allocentric weights for all depth clusters. Range, mean and standard deviation of the sample are listed together with the results of the two-sided one-sampled t-tests of allocentric weights against 0, Bonferroni-Holm corrected.

Cluster	Range	Mean	SD	t-test results
Near	0.13–0.84	0.48	0.20	$t(11) = 8.627$, $p < 0.001$
Middle	0.20–0.80	0.52	0.18	$t(11) = 10.213$, $p < 0.001$
Far	0.20–0.72	0.52	0.16	$t(11) = 11.605$, $p < 0.001$

had to fixate on the fixation cross, and thus, could only use retinal disparity but not vergence as depth cue. If both vergence and retinal disparity provide reliable depth cues within reaching space, as it has been suggested previously (Bingham et al., 2001; Cutting, 1997; Mon-Williams, 1999; Mon-Williams, Tresilian, & Roberts, 2000; Viguier et al., 2001), we expect allocentric weights to be independent of the observer-target distance.

In order to investigate the influence of monocular depth cues on allocentric coding of reach targets in depth, we manipulated the size of the objects in the test scene. In the *no change condition*, the retinal size of the object varied naturally when shifting objects toward or away from the observer, mimicking a real world situation. Besides this, we did not manipulate the absolute object size. In the *magnification condition*, we magnified the natural change in retinal object size. As the retinal size of objects is used for estimating the object's distance to the observer (Sousa, Brenner, & Smeets, 2011; Sousa, Smeets, & Brenner, 2013), magnifying the object size changes the spatial representation so that the objects appear closer or further away than they physically are. In the *con-*

flict condition, we reversed the size manipulation and hence, created a conflict between the change of the retinal object size and the direction of the object shift. That means, the objects' size was magnified if they were shifted away from the observer and reduced if they were shifted toward the observer. As monocular depth cues provide reliable depth information within reaching space (Bruno & Cutting, 1988; Magne & Coello, 2002; Naceri, Chellali, et al., 2011; Naceri, Moscatelli, et al., 2015), we expect a systematic influence of the manipulation of object size on sagittal deviations of reaching endpoints. In comparison to the no change condition, allocentric weights should be higher in the magnification and smaller in the conflict condition.

3.2. Methods

3.2.1. Participants & apparatus

Fifteen right-handed volunteers with normal or corrected-to-normal vision participated in the experiment. Two participants were excluded due to non-compliance with the fixation instruc-
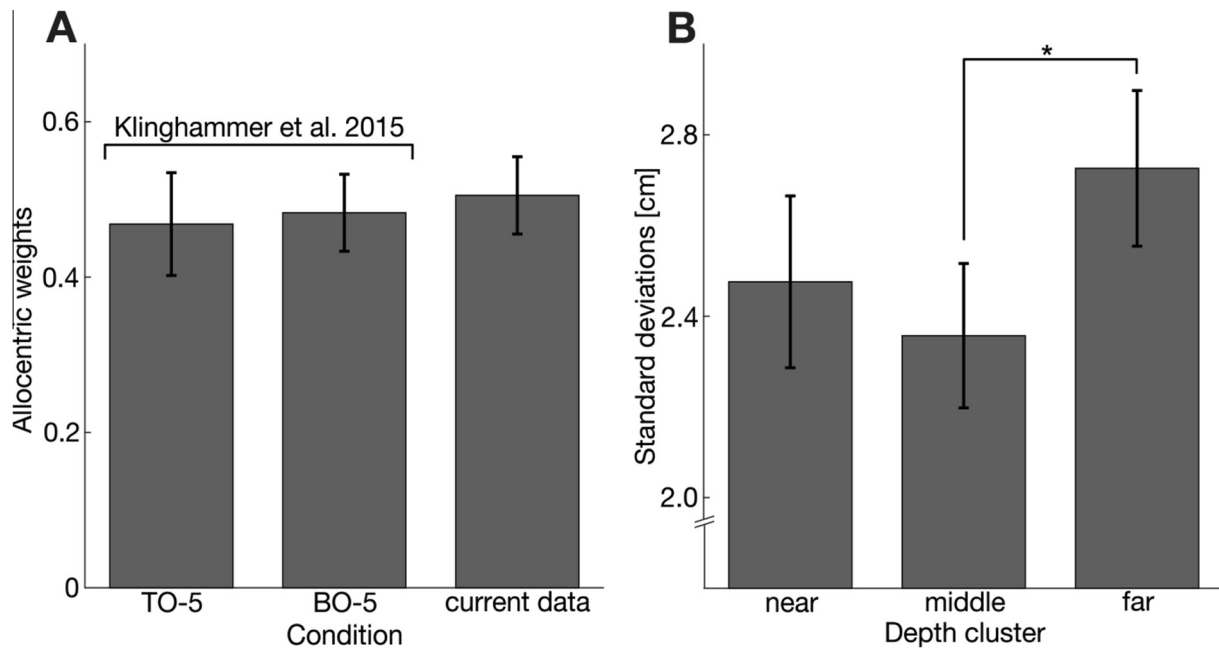
Fig. 4. (A) Mean allocentric weights from our previous study where five table objects (TO-5) or five background objects (BO-5) were shifted and our current experiment which also shifts five table objects. (B) Mean standard deviations of reaching errors for the three depth clusters in cm. Error bars represent 1 SEM and asterisks indicate significant differences (*: $p < 0.05$).

tions (greater than 25% invalid trials). The final sample therefore consisted of 13 participants (6 female), aged 19–29 years (mean 23 ± SD 3.2 years). By using the same measuring techniques as in experiment 1, we ensured all participants had intact stereo vision and measured IOD (mean IOD 60 ± SD 2 mm) and handedness (mean handedness quotient 85.8 ± SD 17.8). The study was approved by the local ethical committee and followed the statutes of the Declaration of Helsinki (2008). All participants gave written informed consent and received money or course credits for their participation.

The experimental set-up was identical to experiment 1.

3.2.2. Materials

We created a new set of encoding scenes as described in experiment 1. Since objects were now shifted in depth instead of horizontally, object shifts could not lead to an occlusion of the fixation cross. We therefore used the objects' original sizes for the encoding scenes (for object properties see Table 1). Based on the encoding scenes we again defined test scenes in which one TO was missing (= reach target) in one of the three different depth clusters. Every TO served as target equally often and in random order. Example images of the different conditions used in experiment 2 can be found in Fig. 5. In 75% of these test scenes the remaining TOs were shifted together by 4 degrees of visual angle (calculated based on the table plane; 3.12–8.27 cm depending on the cluster and depth line position) in depth (50% away from the participant). In one-third of the tests scenes where objects were shifted, we did not manipulate the size of the remaining objects (= no change condition). In another third of these scenes, we increased the object size (depth, width, height) by 10% when objects were shifted toward the participant and decreased them by 10% when objects were shifted away from the participants (= magnification condition). Thus, we magnified the natural change in retinal object size, i.e., objects appeared bigger when they were closer and smaller when they were further away from an observer. In the last third of these scenes, we reversed the magnification condition and thus introduced a conflict between the direction of

object shift and the change in retinal object size. Objects which were shifted towards the participant became smaller and bigger when they were shifted away (= conflict condition). In the remaining 25% of the overall test scenes no objects were shifted. These were used as control condition. All in all, we defined 72 different encoding scenes leading to 72 test scenes in the control condition and 216 test scenes in the other experimental conditions. Masking scenes were created identically to experiment 1.

3.2.3. Procedure

The overall trial procedure was the same as in experiment 1. All in all, every participant completed 288 trials within one session per day, split into two blocks separated by a short break. The trials were the same in each block but were presented within a block in randomized order. The order of blocks was also randomized. Every session was repeated twice on different days leading to a total number of 864 trials per participant. The overall experiment duration for one participant was approx. 3 h.

3.2.4. Data Reduction and Statistical Analysis

Data preprocessing and analysis were performed using the same software and procedures as in experiment 1. We discarded data without correct fixation behavior, which applied to 250 trials (= 2.23%), as well as trials with movement recording or timing errors as described in experiment 1 (1897 trials = 17.27%). After extracting reaching endpoints, we performed an outlier correction for the control condition. Then, reaching errors in the horizontal and depth axis for the other shift conditions were calculated and outlier-corrected. Taken together, 338 trials (= 3.72%) were classified as outliers. All in all from originally 11,232 trials, 8747 trials entered into further analysis (= 77.88%).

To investigate the influence of allocentric information on reaching endpoints, we again calculated allocentric weights as described for experiment 1, this time using the actual object shifts in depth as MERE and the reaching errors in depth for the linear fits.

We performed two-sided one-sampled t-tests to investigate whether group allocentric weights for the different conditions dif-
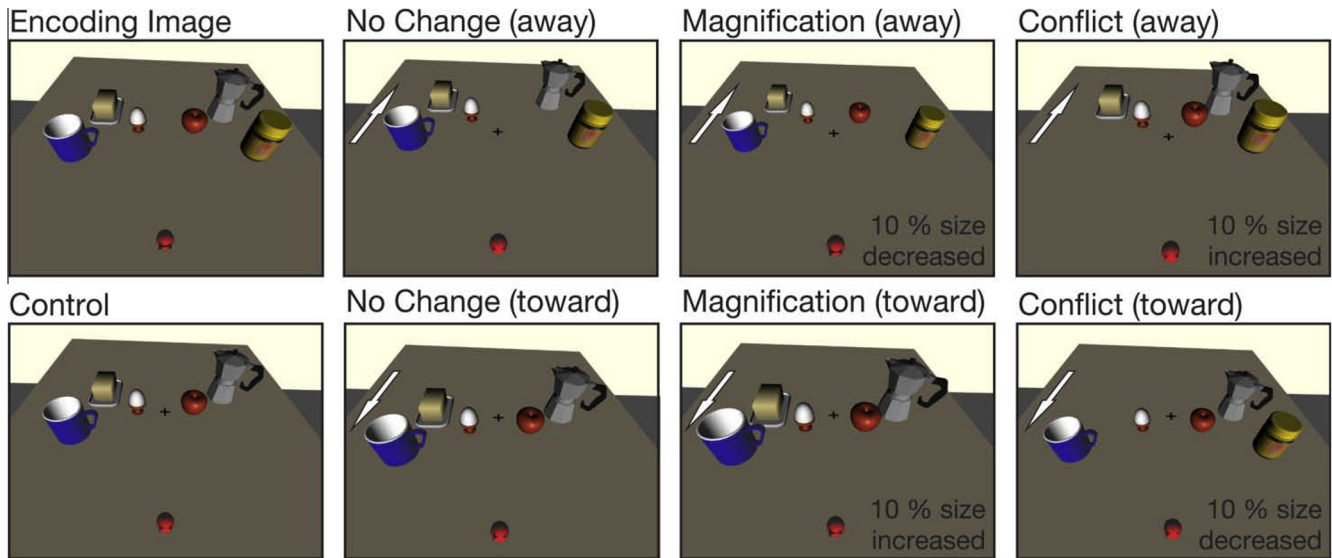
Fig. 5. Example images of an encoding and corresponding test scenes for the different conditions. The arrow indicates the direction of object shifts.

fered significantly from zero. To draw conclusions about the use of allocentric information for coding object locations in the horizontal or depth axis, we performed a two-way ANOVA on the allocentric weights of experiment 1 and the weights of the no change condition of experiment 2 with the factors depth cluster and object shift direction (i.e., horizontal from experiment 1 and in depth from experiment 2). In order to investigate a potential influence of the object size manipulation and the distance between observer and target on allocentric weights, we performed a two-way repeated measures ANOVA with the factors condition (no change, magnification, and conflict) and depth cluster (near, middle, and far). To assess differences in variabilities of reaching endpoints between conditions with manipulations of object size and observer-target distances, we entered standard deviations of reaching endpoints in a two-way repeated measures ANOVA with the factors condition (no change, magnification, and conflict) and depth cluster (near, middle, and far). For both ANOVAs, we conducted two-sided post hoc t-tests for paired samples or one-way repeated measures ANOVAs in case of significant main effects or interactions. To investigate the influence of the object size manipulation and the distance between objects and observer on reaction times (time between go cue and reach onset) and movement durations (time between reach onset and offset), we conducted a two-way repeated measures ANOVA for each of these dependent variables. Again, we conducted two-sided post hoc t-tests for paired samples for both ANOVAs in case of significant main effects.

3.3. Results

In Table 4, the descriptive data for reaching errors in depth (i.e., errors in the direction of object shifts) and the means of the MEREs are summarized. As in experiment 1, objects in different depth clusters were shifted by different absolute distances to keep the visual angle of these shifts constant. Thus, the absolute reaching errors are biased and were transformed to allocentric weights to normalize for this fact.

As indicated in Fig. 6A, averaged reaching endpoint errors for single participants deviated systematically into the direction of object shifts in depth. As an example, in Fig. 6B we depict the linear fit between MEREs and actual reaching errors for one exemplary participant in one condition with object shifts in depth in the second depth cluster without object size manipulation.

We quantified reaching errors by calculating allocentric weights as described in the Methods section. Averaged weights for shifts in depth differed significantly from zero in all conditions and in all depth clusters (see Table 5).

The two-way ANOVA investigating the influence of depth cluster and object shift direction (i.e., horizontal shifts from experiment 1 and shifts in depth from experiment 2) on allocentric weights revealed a main effect of the object shift direction ($F(1,72) = 13.094$, $p < 0.001$). Allocentric weights in experiment 2 were higher than in experiment 1. There was no main effect for depth cluster ($F(2,72) = 0.407$, $p = 0.667$) and no interaction of the two factors ($F(2,72) = 0.041$, $p = 0.960$).

The two-way repeated measures ANOVA assessing the influence of object size manipulation and distance between target and observer on allocentric weights revealed a main effect for the object size manipulation ($F(2,24) = 22.169$, $p < 0.001$, see Fig. 7A). Post-hoc t-tests revealed higher weights in the magnification compared to the no change condition ($t(12) = -5.083$, $p < 0.001$), higher weights in the no change than conflict condition ($t(12) = 2.71$, $p = 0.019$), and higher weights in the magnification than conflict condition ($t(12) = 5.576$, $p < 0.001$). We did not find a main effect for the distance between target and observer ($F(2,24) = 0.331$, $p = 0.721$) but an interaction between the two factors ($F(4,48) = 3.464$, $p = 0.014$). However, one-way repeated measures ANOVAs for each condition with the factor cluster on allocentric weights failed statistical significance after correction for multiple testing ($ps > 0.125$).

The two-way repeated measures ANOVA investigating the influence of object size manipulation and distance between target and observer on the variability (standard deviations) of reaching endpoints revealed a main effect for observer-target distance ($F(2,24) = 32.577$, $p < 0.001$, see Fig. 7B). Post-hoc t-tests revealed a higher variability for far than near depth clusters ($t(12) = -5.906$, $p < 0.001$) for far than middle depth cluster ($t(12) = -6.689$, $p < 0.001$), but not near and middle ($t(12) = 1.089$, $p = 0.298$). We neither found a main effect for the object size manipulation ($F(2,24) = 5.564$, $p = 0.056$) nor an interaction between the two factors ($F(4,48) = 1.141$, $p = 0.349$).

The two-way repeated measures ANOVA assessing the influence of object size manipulation and depth clusters on reaction times revealed only a main effect for depth clusters ($F(2,24) = 4.696$, $p = 0.019$; *mean near/middle/far Cluster:*
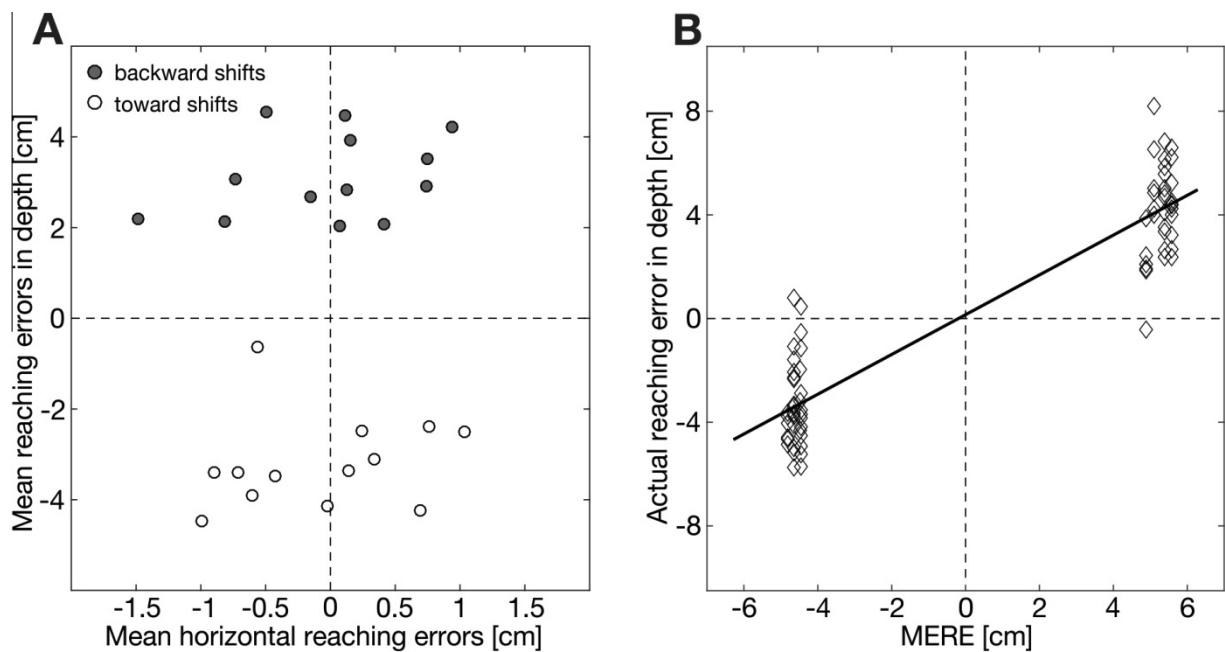
Table 4
Summary of reaching errors in depth for all depth clusters, object size conditions and directions of object shifts (backward = away from the participant; toward = in the direction of the participant) in cm. Range, mean and standard deviation of the sample are listed. Additionally, the means of the MEREs for every condition are listed in cm as well. Negative values are assigned to shifts toward the participant and positive values to shifts away from the participant.

Condition	Shift direction	Range	Mean	SD	Mean MERE
Near cluster					
No change	Toward	−3.895 to 1.729	−2.423	2.010	−3.767
	Backward	0.452 to 4.289	2.513	2.205	4.265
Magnification	Toward	−4.109 to 1.057	−2.459	2.027	−3.764
	Backward	1.185 to 4.674	2.853	2.360	4.280
Conflict	Toward	−3.627 to 0.663	−2.468	2.116	−3.764
	Backward	1.249 to 4.268	2.508	2.231	4.272
Middle cluster					
No change	Toward	−4.463 to −0.635	−3.194	2.036	−4.597
	Backward	2.027 to 4.540	3.123	2.086	5.318
Magnification	Toward	−5.013 to −0.531	−3.423	2.100	−4.596
	Backward	1.665 to 5.101	3.441	2.048	5.319
Conflict	Toward	−4.349 to −0.823	−3.194	2.142	−4.593
	Backward	1.439 to 4.893	3.087	2.126	5.321
Far cluster					
No change	Toward	−6.051 to −1.372	−3.916	2.520	−5.622
	Backward	2.851 to 5.987	4.220	2.574	6.773
Magnification	Toward	−6.284 to −1.179	−4.001	2.466	−5.623
	Backward	3.190 to 7.104	4.471	2.979	6.754
Conflict	Toward	−5.832 to −1.262	−3.532	2.383	−5.610
	Backward	1.866 to 5.952	3.467	3.329	6.761

Fig. 6. (A) Mean reaching errors for all single participants in cm for object shifts in depth (middle depth cluster, no change condition; backward = away from the participant; toward = in the direction of the participant). (B) Example of a linear fit between MEREs and actual reaching errors in depth for one participant for object shifts in depth in the second depth cluster in the no change condition. Negative values are assigned to shifts toward the participant and positive values to shifts away from the participant.

289 ms/284 ms/296 ms). However, post hoc *t*-tests did not reach significance (all *p* > 0.095). The two-way repeated measures ANOVA on movement durations revealed only a main effect for the factor depth cluster ($F(2,24) = 117.407$, $p < 0.001$; *mean near/ middle/far Cluster:* 566 ms/612 ms/664 ms). Post-hoc *t*-tests revealed higher movement durations for the depth clusters middle than near ($t(12) = −8.786$, $p < 0.001$), far than near ($t(−12) = −11.285$, $p < 0.001$), and far than middle ($t(12) = −11.687$, $p < 0.001$).

4. General discussion

Object locations are represented in egocentric (e.g. Cohen & Anderson, 2002; Lacquaniti & Caminiti, 1998; Thompson & Henriques, 2011) and allocentric reference frames (e.g. Diedrichsen et al., 2004; Krigolson & Heath, 2004; Krigolson et al., 2007; Obhi & Goodale, 2005; Schütz et al., 2013, 2015). An increasing number of studies provide evidence that both classes of reference frames are used and integrated when humans perform memory-guided reaching movements (Byrne & Crawford, 2010; Schütz et al., 2013, 2015). There are recent attempts at studying the underlying coding schemes of reaching movements in more naturalistic environments by increasing ecological validity. For example, photographs of rich and complex scenes are presented providing multiple allocentric cues for coding of reach targets in space (Camors et al., 2015; Fiehler et al., 2014). One important limitation of these studies is their restriction to the 2D monitor space. Here, we aimed to overcome this limitation by transferring our

Table 5
Summary of allocentric weights for all conditions and depth clusters. Range, mean and standard deviation of the sample are listed together with the results of the two-sided one-sampled *t*-tests of allocentric weights against 0, Bonferroni-Holm corrected.

Condition	Range	Mean	SD	*t*-test results
Near cluster				
No change	0.32–0.81	0.61	0.15	$t(12) = 14.681, p < 0.001$
Magnification	0.44–0.92	0.66	0.14	$t(12) = 16.937, p < 0.001$
Conflict	0.45–0.81	0.62	0.10	$t(12) = 22.406, p < 0.001$
Middle cluster				
No change	0.35–0.81	0.64	0.11	$t(12) = 17.735, p < 0.001$
Magnification	0.46–0.85	0.69	0.11	$t(12) = 22.240, p < 0.001$
Conflict	0.47–0.80	0.63	0.11	$t(12) = 21.318, p < 0.001$
Far cluster				
No change	0.46–0.85	0.66	0.11	$t(12) = 20.832, p < 0.001$
Magnification	0.41–0.97	0.69	0.14	$t(12) = 17.276, p < 0.001$
Conflict	0.40–0.80	0.57	0.11	$t(12) = 18.994, p < 0.001$
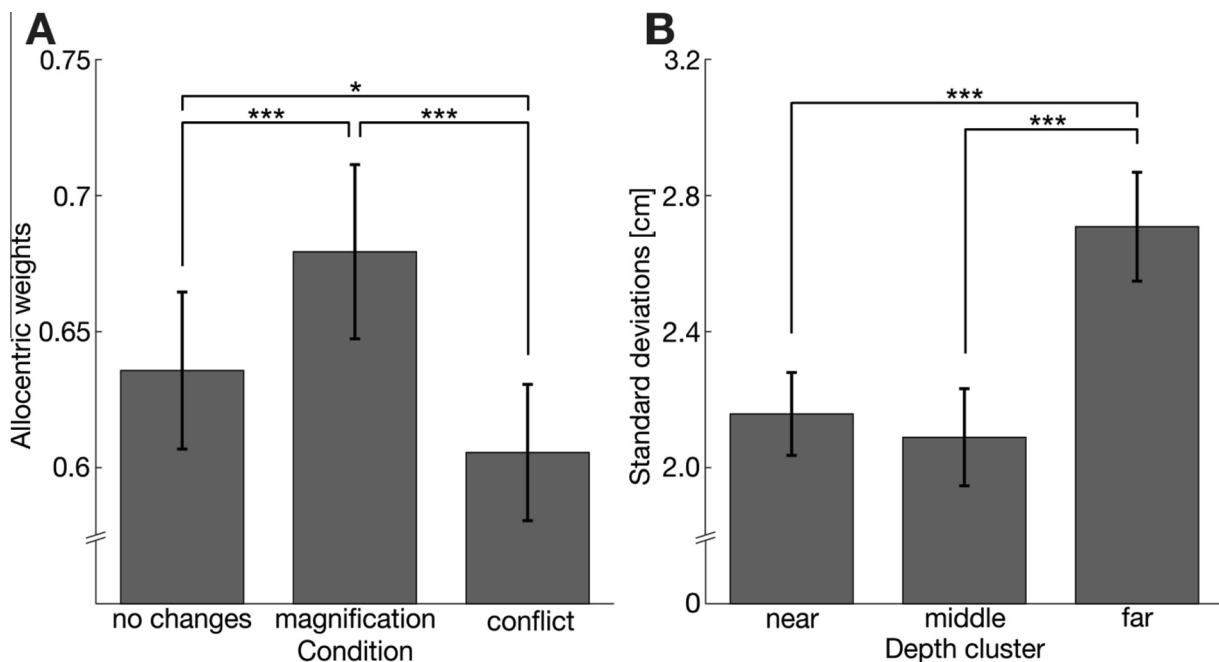
Fig. 7. (A) Mean allocentric weights for the three object size manipulation conditions. (B) Mean standard deviations of reaching errors for the three depth clusters in cm. Error bars represent 1 SEM and asterisks indicate significant differences (*: $p < 0.05$; ***: $p < 0.001$).

previous paradigm (Fiehler et al., 2014; Klinghammer et al., 2015) to 3D virtual reality. This also allowed us to investigate whether and how allocentric information is utilized for encoding the location of reach targets in depth and how this is influenced by binocular and monocular (object size) depth cues.

It is still an unresolved question whether reaching targets are similarly or differently affected by allocentric information when reaching to memorized targets in the horizontal versus the depth plane (for conflicting results see, Coello et al., 2003; Neely et al., 2008). Here, we addressed this point by directly comparing results from horizontal and sagittal object shifts in a more naturalistic environment.

In our first experiment, we aimed to replicate the results from our previous experiments using 2D images (Fiehler et al., 2014; Klinghammer et al., 2015) in a 3D virtual reality setup. Based on the previous results, we predicted systematic reaching errors in the direction of object shifts if participants took information from the surrounding objects serving as allocentric cues into account. We showed allocentric weights similar to those found in our 2D-study (Klinghammer et al., 2015) suggesting that participants make use of allocentric information not only when they reach to remembered objects presented on a 2D monitor but also

when they reach to memorized objects in 3D virtual reality. Thus, we were able to generalize our results from 2D stimuli to 3D virtual-reality and can exclude that methodological differences may cause different results. Our finding that the visuo-motor system makes use of allocentric information from objects in a virtual environment agrees well with previous evidence showing smaller reach errors in rich (i.e., containing multiple objects, linear perspective or texture cues) compared to poor (containing a target object and nothing else) virtual scenes (Naceri, Chellali, et al., 2011). Allocentric cues provided in rich environments are not only used effectively for memory-guided reaching, but also in perceptual tasks such as matching the position of two objects in virtual reality (Murgia & Sharkey, 2009). This indicates the utilization of allocentric information for both perception and action in 3D space. As weights in our first experiment were ranging from 0.48 to 0.52 (see Table 3), we conclude that movement planning and execution were affected to about 50% by the allocentric information of the object shifts. The remaining 50% could be attributed to the influence of an egocentric reference frame. However, as the environment also provided some other, more stable landmarks (i.e., table, fixation cross, edges of the HMD) we cannot exclude an influence of additional allocentric cues.

In our second experiment, we investigated whether and how allocentric information is used for coding locations of targets for memory-guided reaching in depth and how this is influenced by binocular and monocular (object size) depth cues. We used the same paradigm as in experiment 1, but this time shifted objects in depth, i.e., toward or away from the observer, instead of horizontally. We found systematic deviations of reaching endpoints into the direction of object shifts in depth. These weights ranged from 0.61 to 0.66 (see Table 5, no change condition) which indicates that around 64% of movement planning and execution were influenced by the allocentric information of the shifted objects. The remaining 36% can be attributed to the influence of egocentric and/or additional allocentric representations of the object locations. However, participants seem to rely more strongly on allocentric than egocentric representations when coding object locations for memory-guided reaching movements in depth.

In order to investigate whether allocentric information is used differently for horizontal reaching movements than reaching movements in depth, we compared our results from experiment 1 (horizontal object shifts) with the corresponding results of experiment 2 (object shifts in depth in the condition without object size manipulation). As expected, we found that allocentric information was used for memory-guided reaching movements in both experiments regardless of shift direction. This is in line with previous work (Neely et al., 2008) arguing for a similar integration of egocentric and allocentric information for both movement distance and movement direction. In contrast to Coello et al. (2003), our results do not support the claim that memory-guided reaching movements are prone to allocentric information in the horizontal axis. We even observed a stronger weighting of allocentric information when participants had to encode object locations in depth than in the horizontal axis. This suggests different allocentric representations of target distance and direction, as has been previously proposed by Chieffi and Allport (1997) for representing reach targets in an egocentric frame of reference. Hence, our results extend this finding to allocentric coding of targets for memory-guided reaching. Additional evidence for independent mechanisms comes from perceptual experiments showing that by increasing the complexity of a visual scene, participants' underestimation of a pointing target decreased in the distance axis, but not in the directional axis (Coello & Magne, 2000). Thus, it is conceivable that in the study by Neely et al. (2008), results in the condition with a frame oriented in depth revealed smaller allocentric effects compared to our paradigm which used a more complex visual environment.

Our overall finding that allocentric information influences reaching endpoints of memory-guided movements is in line with previous research. For example, movement parameters such as maximum grip aperture (Franz, Hesse, & Kollath, 2009; Westwood, McEachern, & Roy, 2001) or reaching amplitude (Gentilucci, Chieffi, Daprati, Saetti, & Toni, 1996) are more strongly affected by visual illusions of the target in an open-loop than a closed-loop movement task. If visual feedback of the target is not available during movement, the brain seems to rely more strongly on allocentric representations which make use of relational metrics and thus, are prone to visual illusion. In our study, memory-guided reaching may have strengthened the use of allocentric information of the objects provided in the scene. Future studies are needed investigating the contribution of allocentric information for visually-guided actions.

By placing objects in three different depth clusters, we examined whether binocular depth cues such as vergence and retinal disparity can be efficiently used for allocentric coding of reach targets in depth. Neither in experiment 1 nor experiment 2 we found evidence for an influence of the distance between the target and the observer on reaching endpoint accuracy. Our results suggest that binocular depth cues provide important information for coding object locations in depth and thus, lead to a consistent use of the allocentric information across varying observer-target distances. This is in line with previous findings indicating that vergence can be effectively used as absolute depth cue within reaching space (Medendorp & Crawford, 2002; Naceri, Chellali, et al., 2011; Tresilian et al., 1999; Viguier et al., 2001). If observer-target distances exceed 55 cm, reach endpoint accuracy seems to decrease significantly (Naceri, Chellali, et al., 2011). Here, we did not exceed this range even after shifting objects away from the observer. However, we found that variability of reaching endpoints increased for the targets located further away from the observer (but still within reachable space) in both experiments, which is in line with previous findings (Messier & Kalaska, 1997; Schmidt, Zelaznik, Hawkins, Frank, & Quinn, 1979). Moreover, in both experiments we revealed an effect of observer-object distance on movement durations which is likely caused by the longer hand transportation phase to the reach target. Since we did not find an influence of different depth clusters or object size manipulation on reaction times, these factors seem to have no impact on the movement planning phase.

In the second experiment, we also manipulated absolute object size as a monocular depth cue during the presentation of the test scene. In one condition we magnified the natural depth effect so that objects' retinal size became larger when they were shifted toward and smaller when they were shifted away from the observer while the shift distance was always the same. In a second condition we reversed this effect and decreased absolute object size when objects were shifted toward and increased it when they were shifted away from the observer, thus creating a conflict between shift direction and change of the perceived object size. We found increased allocentric weights in the magnification condition whereas allocentric weights were decreased in the conflict condition compared to a condition with no change in object size (natural depth effect). Based on our findings, we conclude that besides binocular depth cues, object size is an important monocular depth cue for allocentric coding of reach targets in 3D space (c.f., Bruno & Cutting, 1988; Magne & Coello, 2002; Naceri, Chellali, et al., 2011; Naceri, Moscatelli, et al., 2015; Sousa, Brenner, et al., 2011; Sousa, Smeets, et al., 2013). As the differences between the size change conditions were relatively small, it is likely that other depth cues (e.g., object occlusion, binocular disparity) have been utilized as well and partially compensated for this manipulation. This assumption is supported by the model of modified weak fusion (Landy et al., 1995) which states that depth perception is specified by different independent depth cues and that these cues are weighted and combined depending on the location of an object and the situation of observation. Thus, it is likely that manipulating only one cue may lead to smaller effects in depth perception. Another possibility for the relatively small effects we found may relate to the observation that the retinal object size of a trial within an experiment is used for estimating the distance to the same object at the same location in a consecutive trial, even though the object sizes were slightly different between these trials (Sousa, Brenner, et al., 2011; Sousa, Smeets, et al., 2013). This can lead to systematic misestimations of the object's distance in the consecutive trial. Hence, in our paradigm the retinal size of objects in the encoding scene might have affected the distance perception of shifted and size manipulated objects in the test scene.

Recently, Naceri, Moscatelli, et al. (2015) compared verbal estimates of object distances between settings in virtual reality and the real world. Results revealed a better performance in the real world condition arguing for additional depth cues used for estimating object's depth location. While vergence and retinal disparity are reliable depth cues in VR settings, real world accommodation of the eye lenses cannot be mimicked as the distance between eyes

and display does not change in a VR setting. This effect leads to a vergence-accommodation conflict between these two depth cues resulting in less precise depth perception in VR (see also, Bingham et al., 2001). Thus, we cannot claim that our results can be entirely transferred to real world situations. Nevertheless, we are convinced that our approach is still an important step from classical laboratory to more realistic settings. Moreover, VR provides a good compromise to approximate real world settings while still offering an easy but powerful possibility to control for various parameters such as the object positions within an experiment.

Overall, our findings demonstrate that allocentric information is utilized when coding target locations for memory-guided reaching in depth. Besides binocular depth cues, object size as monocular depth cue plays an important role, whereas additional depth cues might contribute as well.

Acknowledgments

References

Armbrüster, C., Wolter, M., Kuhlen, T., Spijkers, W., & Fimm, B. (2008). Depth perception in virtual reality: distance estimations in peri- and extrapersonal space. *Cyberpsychology & Behavior, 11*(1), 9–15. http://dx.doi.org/10.1089/cpb.2007.9935.

Battaglia-Mayer, A., Caminiti, R., Lacquianiti, F., & Zago, M. (2003). Multiple levels of representation of reaching in the parieto-frontal network. *Cerebral Cortex, 13* (10), 1009–1022.

Bingham, G. P., Bradley, A., Bailey, M., & Vinner, R. (2001). Accommodation, occlusion, and disparity matching are used to guide reaching: A comparison of actual versus virtual environments. *Journal of Experimental Psychology: Human Perception and Performance, 27*(6), 1314–1334.

Bruno, N., & Cutting, J. E. (1988). Minimodularity and the perception of layout. *Journal of Experimental Psychology: General, 117*(2), 161–170.

Byrne, P. A., & Crawford, J. D. (2010). Cue reliability and a landmark stability heuristic determine relative weighting between egocentric and allocentric visual information in memory-guided reach. *Journal of Neurophysiology, 103*, 3054–3069. http://dx.doi.org/10.1152/jn.01008.2009.

Camors, D., Jouffrais, C., Cottereau, B. R., & Durand, J. B. (2015). Allocentric coding: Spatial range and combination rules. *Vision Research, 109*, 87–98. http://dx.doi.org/10.1016/j.visres.2015.02.018.

Chieffi, S., & Allport, D. A. (1997). Independent coding of target distance and direction in visuo-spatial working memory. *Psychological Research Psychologische Forschung, 60*(4), 244–250.

Coello, Y., & Magne, P. (2000). Determination of target distance in a structured environment: Selection of visual information for action. *European Journal of Cognitive Psychology, 12*(4), 489–519.

Coello, Y., Richaud, S., Magne, P., & Rossetti, Y. (2003). Vision for spatial perception and vision for action: a dissociation between the left-right and the near-far dimensions. *Neuropsychologia, 41*(5), 622–633.

Cohen, Y. P., & Anderson, R. A. (2002). A common reference frame for movement plans in the posterior parietal cortex. *Nature Reviews Neuroscience, 3*(7), 553–562. http://dx.doi.org/10.1038/nrn873.

Colby, C. L. (1998). Action-oriented spatial reference frames in cortex. *Neuron, 20*(1), 15–24.

Cutting, J. E. (1997). How the eye measures reality and virtual reality. *Behavior Research Methods, Instruments & Computer, 29*(1), 27–36.

Diedrichsen, J., Werner, S., Schmidt, T., & Trommershäuser, J. (2004). Immediate spatial distortions of pointing movements induced by visual landmarks. *Perception & Psychophysics, 66*, 89–103. http://dx.doi.org/10.3758/bf03194864.

Fiehler, K., Schütz, I., & Henriques, D. Y. (2011). Gaze-centered spatial updating of reach targets across different memory delays. *Vision Research, 51*(8), 890–897. http://dx.doi.org/10.1016/j.visres.2010.12.015.

Fiehler, K., Wolf, C., Klinghammer, M., & Blohm, G. (2014). Integration of egocentric and allocentric information during memory-guided reaching to images of a natural environment. *Frontiers in Human Neuroscience, 8*, 636. http://dx.doi.org/10.3389/fnhum.2014.00636.

Franz, V. H., Hesse, C., & Kollath, S. (2009). Visual illusions, delayed grasping, and memory: no shift from dorsal to ventral control. *Neuropsychologia, 47*(6), 1518–1531.

Gentilucci, M., Chieffi, S., Daprati, E., Saetti, M. C., & Toni, I. (1996). Visual illusion and action. *Neuropsychologia, 34*(5), 369–376.

Hibbard, P. B., & Bradshaw, M. F. (2003). Reaching for virtual objects: binocular disparity and the control of prehension. *Experimental Brain Research, 148*(2), 196–201. http://dx.doi.org/10.1007/s00221-002-1295-2.

Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In *Spatial cognition* (pp. 1–17). Berlin Heidelberg: Springer.

Klinghammer, M., Blohm, G., & Fiehler, K. (2015). Contextual factors determine the use of allocentric information for reaching in a naturalistic scene. *Journal of Vision, 15*(13), 24. http://dx.doi.org/10.1167/15.13.24. 1–13.

Knill, D. C. (2005). Reaching for visual cues to depth: the brain combines depth cues differently for motor control and perception. *Journal of Vision, 5*(2), 103–115.

Krigolson, O., Clark, N., Heath, M., & Binsted, G. (2007). The proximity of visual landmarks impacts reaching performance. *Spatial Vision, 20*, 317–336. http://dx.doi.org/10.1163/156856807780919028.

Krigolson, O., & Heath, M. (2004). Background visual cues and memory-guided reaching. *Human Movement Science, 23*(6), 861–877.

Lacquaniti, F., & Caminiti, R. (1998). Visuo-motor transformations for arm reaching. *European Journal of Neuroscience, 10*, 195–203.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research, 35*(3), 389–412.

Magne, P., & Coello, Y. (2002). Retinal and extra-retinal contribution to position coding. *Behavioral Brain Research, 136*(1), 277–287.

Medendorp, W. P., & Crawford, J. D. (2002). Visuospatial updating of reaching targets in near and far space. *NeuroReport, 13*(5), 633–636.

Messier, J., & Kalaska, J. F. (1997). Differential effect of task condition on errors of direction and extent of reaching movements. *Experimental Brain Research, 115* (3), 469–478.

Mon-Williams, M. (1999). Some recent studies on the extraretinal contribution of distance perception. *Perception, 28*(2), 167–181. http://dx.doi.org/10.1068/p2737.

Mon-Williams, M., Tresilian, J. R., & Roberts, A. (2000). Vergence provides depth perception from horizontal retinal image disparities. *Experimental Brain Research, 133*(3), 407–413. http://dx.doi.org/10.1007/s002210000410.

Murgia, A., & Sharkey, P. M. (2009). Estimation of distances in virtual environments using size constancy. *The Journal of Virtual Reality, 8*(1), 67–74.

Naceri, A., Chellali, R., & Hoinville, T. (2011). Depth perception within peripersonal space using head-mounted display. *Presence, 20*(3), 254–272. http://dx.doi.org/10.1162/PRES_a_00048.

Naceri, A., Moscatelli, A., & Chellali, R. (2015). Depth discrimination of constant angular size stimuli in action space: Role of accommodation and convergence cues. *Frontiers in Human Neuroscience, 9*, 511. http://dx.doi.org/10.3389/fnhum.2015.00511.

Neely, K. A., Heath, M., & Binsted, G. (2008). Egocentric and allocentric visual cues influence the specification of movement distance and direction. *Journal of Motor Behavior, 40*(3), 203–213. http://dx.doi.org/10.3200/JMBR.40.3.203-213.

Obhi, S. S., & Goodale, M. A. (2005). The effects of landmarks on the performance of delayed and real-time pointing movements. *Experimental Brain Research, 167*, 335–344. http://dx.doi.org/10.1007/s00221-005-0055-5.

Ostendorf, F., Fischer, C., Finke, C., & Ploner, C. J. (2007). Perisaccadic compression correlates with saccadic peak velocity: Differential association of eye movement dynamics with perceptual mislocalization patterns. *Journal of Neuroscience, 27*(28), 7559–7563.

Schmidt, R. A., Zelaznik, H., Hawkins, B., Frank, J. S., & Quinn, J. T. Jr., (1979). Motor-output variability: A theory for the accuracy of rapid motor acts. *Psychological Review, 47*(5), 415–451.

Schütz, I., Henriques, D. Y., & Fiehler, K. (2013). Gaze-centered spatial updating in delayed reaching even in the presence of landmarks. *Vision Research, 87*, 46–52. http://dx.doi.org/10.1016/j.visres.2013.06.001.

Schütz, I., Henriques, D. Y., & Fiehler, K. (2015). No effect of delay on the spatial representation of serial reach targets. *Experimental Brain Research, 233*(4), 1225–1235. http://dx.doi.org/10.1007/s00221-015-4197-9.

Sousa, R., Brenner, E., & Smeets, J. B. (2011). Judging an unfamiliar object's distance from its retinal image size. *Journal of Vision, 11*(9), 10. http://dx.doi.org/10.1167/11.9.10. 1–6.

Sousa, R., Smeets, J. B., & Brenner, E. (2013). The influence of previously seen objects' size in distance judgments. *Journal of Vision, 13*(2), 2. http://dx.doi.org/10.1167/13.2.2. 1–8.

Thompson, A. A., & Henriques, D. Y. (2011). The coding and updating of visuospatial memory for goal-directed reaching and pointing. *Vision Research, 51*(8), 819–826. http://dx.doi.org/10.1016/j.visres.2011.01.006.

Tresilian, J. R., Mon-Williams, M., & Kelly, B. M. (1999). Increasing confidence in vergence as a cue to distance. *Proceedings of the Royal Society B, 266*, 39–44.

Viguier, A., Clément, G., & Trotter, Y. (2001). Distance perception within near visual space. *Perception, 30*(1), 115–124.

Westwood, D. A., McEachern, T., & Roy, E. A. (2001). Delayed grasping of a Müller-Lyer figure. *Experimental Brain Research, 141*, 166–173.