Investigating spatial reference frames for reaching in a naturalistic scenario

Justus-Liebig-Universität Gießen

Fachbereich 06

Psychologie und Sportwissenschaften

Otto-Behaghel-Straße 10F

35394 Gießen

September 2016

Inauguraldissertation zur Erlangung des

Doktorgrades der Naturwissenschaften

Vorgelegt von

Mathias Klinghammer

geboren am 10.09.1986

in Arnstadt

Gutachter:

Prof. Dr. Katja Fiehler (Gießen, Psychologie)

Prof. Karl Gegenfurtner, Ph.D. (Gießen, Psychologie)

Mitglieder der Prüfungskommission:

Prof. Dr. Reinhard Lakes-Harlan (Gießen, Biologie)

Prof. Dr. Jörn Munzert (Gießen, Sportwissenschaft)

Acknowledgements

Working for a doctoral degree is a long and sometimes exhausting process that I would not have survived without the help of many people.

Especially, I would like to thank…

… Katja Fiehler for patiently dealing with my lack of knowledge and teaching me the scientific right from wrong. Especially I would like to thank her for giving me the opportunity to accomplish this huge step in my career.
… Gunnar Blohm for being an extremely relaxed and informative Co-Supervisor, and for hosting me in the most welcoming atmosphere during my stay in Kingston.
… Stefanie Müller and Immo Schütz for mastering the most difficult task: sharing an office with me. Thank you for answering almost every stupid question, for giving me the best support in every situation, for providing an awesome working atmosphere, and for standing my anti-jokes.
… Hanna Gertz for being a great colleague and especially for proofreading this manuscript and giving valuable suggestions.
… Tobias Möhler, Dimitris Voudouris, and the whole Department of Experimental Psychology for giving advice, supporting me, and spending their free time with me.
… all the doctoral candidates and PIs working in the IRTG for coming up with research ideas, spending countless hours at talks and workshops with me, and for some really enjoyable get-togethers.
… the members of Gunnar Blohm's lab for helping me with my Bayesian struggles and making my stay in Kingston an outstanding experience.
… many student assistants, Bachelor students, and Master students for collecting most of the data and doing the jobs that would have really bored me.
… Prof. Karl Gegenfurtner for assessing this thesis and Prof. Jörn Munzert and Prof. Reinhard Lakes-Harlan for their examination of this thesis.
… my family.

Zusammenfassung

In unserem alltäglichen Leben führen wir eine Vielzahl unterschiedlicher Handlungen mit Objekten in unserer Umgebung aus, wie beispielsweise der Griff nach der Kaffeetasse am Frühstückstisch. Wiewohl diese Handlungen ohne größeren kognitiven Aufwand erfolgen, erweisen sich die zugrunde liegenden Mechanismen als durchaus komplex. Einer dieser Mechanismen ist die Kodierung der Position jenes Objekts, welches Ziel unserer Handlung sein soll. Diese Kodierung erfolgt in sogenannten *Referenzsystemen*, welche Koordinatensysteme sowie -ursprünge darstellen, zu welchen die Objektpositionen in Bezug gesetzt werden können. Man unterscheidet hierbei zwischen dem *egozentrischen* Referenzsystem, bei welchem jeder Teil des Körpers als Koordinatenursprung fungieren kann, und dem *allozentrischen* Referenzsystem, in welchem jedes Objekt in der Umgebung diese Funktion übernehmen kann. Die Nutzung dieser Referenzsysteme war Gegenstand vieler Untersuchungen in der Vergangenheit. Jedoch wurde in diesen Studien meist auf abstrakte Stimuli wie Punkte oder Balken, die auf einem ansonsten schwarzen Bildschirm präsentiert wurden, zurückgegriffen, was an der Übertragbarkeit der Ergebnisse ins tägliche Leben zweifeln lässt. Darüber hinaus wurden Probanden in einigen Studien dahingehend beeinflusst, dass sie das eine oder andere Referenzsystem bevorzugt anwandten.

Ziel der vorliegenden Arbeit war es diese Mängel durch die Nutzung natürlicher und komplexer Stimuli zu überwinden ohne dabei die Probanden zum Gebrauch eines bestimmten Referenzsystems zu verleiten. Hierfür nutzte ich ein experimentelles Paradigma in welchem die Probanden gedächtnisbasierte Zeigebewegungen auf ein Zielobjekt innerhalb einer komplexen Frühstücksszene ausführen sollten, welche auf einem Computerbildschirm oder in virtueller Realität dargeboten wurde. Auf Basis von vier Studien war ich in der Lage neue Einblicke in die Integration der beiden Referenzsysteme zum Kodieren von Objektpositionen in einer komplexen Umgebung zu geben. Erstens bestätigte ich die Nutzung beider Referenzsysteme zur Kodierung der Position eines Zielobjekts für gedächtnisbasierte Zeigebewegungen. Zum Zweiten, konnte ich einige Faktoren aufdecken, welche die Nutzung von Objekten als allozentrische Hinweisreize bedingen, wie die Aufgabenrelevanz von Objekten, die Distanz zwischen einem Zielobjekt und aufgabenrelevanten Objekten sowie die Reliabilität der allozentrischen Information. Zum Dritten konnte ich zeigen dass die allozentrische

Information zu einem stärkeren Grad zum Kodieren von Zielposition in die Tiefe verglichen mit Zielpositionen in der horizontalen Ebene genutzt wird. Und zum Vierten habe ich herausgefunden, dass die Objektgröße ein wichtiger Tiefenreiz für die allozentrische Kodierung von Objektpositionen in die Tiefe darstellt.

Abstract

Humans perform numerous interactions with objects in the environment, like reaching for a mug on the breakfast table, in everyday live. Even though these interactions can be done without bigger cognitive effort, the underlying mechanisms are fairly complex. One crucial mechanism is to code the location of the object that we want to act on (*target*), which is done by using so called *reference frames*. These reference frames provide coordinate systems and origins to which targets can be related to. In the *egocentric* reference frames, any body part can serve as the origin of the coordinate system whereas in the *allocentric* reference frame, any object in the environment could serve as the origin. The use of these reference frames for coding a target location to perform an action was research question of several studies in the past. However, these studies manly used rather artificial stimuli like dots and bars on black screens, lacking ecological validity. Moreover, in some studies participants were biased to prefer one reference frame over the other.

In the present thesis, I aimed to overcome these limitations by using naturalistic and complex stimuli which provide multiple potential allocentric cues for coding a target location without biasing participants coding behavior. To this end, I applied a paradigm in which participants had to perform a memory-guided reaching movement to a target location that was represented by an object in a breakfast scene that was presented either on a computer screen or in virtual reality. By conducting four studies, I was able to provide novel insights into the integration of the two reference frames in complex environments. First, I confirmed that both reference frames are used for coding a target location for memory-guided reaching. Second, I revealed several factors that determine the use of objects as allocentric cues in a complex environment, such as the task-relevance of objects, the distance between a target and task-relevant objects and the reliability of the allocentric information. Third, I have shown that the use of the allocentric information is higher for coding target locations in depth compared to coding a target location in the horizontal plane. And fourth, I found that object size is an important depth cue for allocentric coding of object locations in depth.

Contents

1 Introduction

In our everyday life, we frequently interact with objects in the environment. Reaching for a mug, grasping the key or reaching for the light switch are just very few examples. We do many of these seemingly simple interactions without bigger effort but in fact the underlying mechanisms are fairly complex. One very basic mechanism is encoding the location of an object that we want to act on. In the past, several studies investigated this mechanism using rather abstract stimuli like dots and bars on black screens in their experiments. But these studies do not account for the fact that we usually reach for mugs, keys, or other objects in a more complex environment and hence, lack ecological validity. Therefore, I aim to overcome these limitations by using more naturalistic and complex stimuli for my research and thus, will give novel insights into the use of *spatial reference frames* as a core feature of spatial encoding mechanisms underlying reaching movements. In this context, I will also provide answers to related questions like how object locations are coded in situations when multiple objects are provided in the environment and which contextual factors influence this coding mechanism.

1.1 Two main classes of reference frames

When reaching for a desired object (*target*) in a real-world setting the human brain not only needs to encode the position of the target object, but also to relate the target to other objects in the surroundings. This is important as these objects may stand very closely to the target or between the reaching hand and the target and thus, we also need to take their locations into account in order to avoid knocking them over. Suddenly we are dealing with an entity consisting of the target and multiple other objects that need to be brought into relation. To do so, the human brain makes use of different *frames of reference*, which are classified by their coordinate systems and the underlying origins (Soechting & Flanders, 1992). Reference frames can be split into two broad classes, namely *egocentric* and *allocentric* reference frames (Battaglia-Mayer, Caminiti, Lacquianiti, & Zago, 2003; Colby, 1998; for an example see figure 1.1). In an egocentric frame of reference any body part (e.g., eyes, head, or trunk) of the observer who is about to perform an action, serves as the origin of the coordinate system to which all objects in the environment are related.

For example, the location of the mug we want to reach for can be described as a vector with the direction and amplitude from the observer's eyes, head, or trunk to the mug (e.g., red dashed-dotted arrow in figure 1.1). In contrast, for the allocentric reference frame any object or landmark in the environment could serve as origin of the coordinate system (yellow dotted arrows in figure 1.1). Also a structured background could be used as allocentric cue. Thus, it is independent of the observer's position but depends on cues in the surrounding world and therefore, is also called *object-centered* or *world-centered* reference frame (Battaglia-Mayer, Caminiti, Lacquianiti, & Zago, 2003; Colby, 1998; Klatzky, 1998).

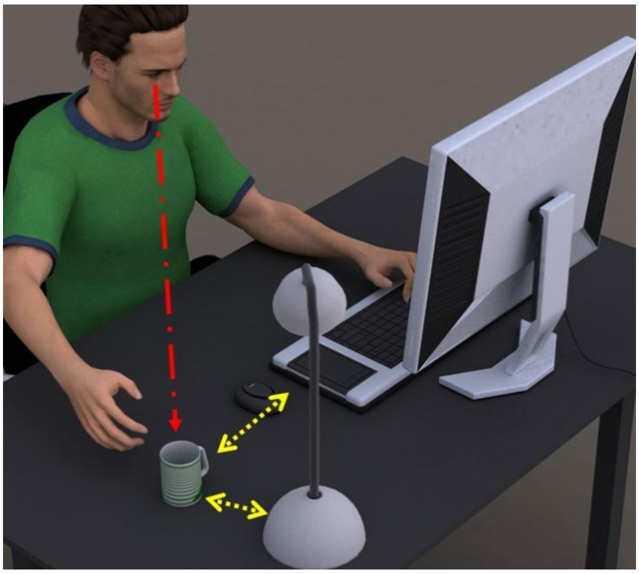

Figure 1.1. Illustration of different reference frames for coding a target location (mug) for reaching. In the egocentric frames of reference, targets are coded relative to body parts of the observer (here: relative to the eye; red, dashed-dotted arrow). In the allocentric reference frame, a target is coded relative to other objects in the environment (here: relative to the keyboard and the lamp; yellow, dotted arrows).

The use of reference frames for reaching varies between situations when one has to reach for a target that is still visible (i.e., *visually-guided reaching*, *closed-loop reaching*) and when one has to reach for a target that is not visible (i.e., *memory-guided reaching*, *open-loop reaching*) when performing the reaching movement. Usually the influence of allocentric information on reaching behavior is bigger in open-loop than closed-loop reaching (Krigolson & Heath, 2004; Sheth & Shimojo, 2004; Westwood & Goodale, 2003). In the experiments I conducted, visual feedback about the target and the allocentric information was always occluded at the moment when a movement-cue initiated participants' reaching onset (open-loop reaches). Therefore, empirical conclusions about visually-guided reaching cannot be provided.

1.2 Egocentric reference frames for reaching

There is a vast number of studies in the literature showing that egocentric reference frames are used for planning and executing reaching movements toward the remembered location of a target. Of particular interest for my work were visual targets, which are assumed to be coded in a *gaze-dependent* or *gaze-centered* reference frame (Crawford, Henriques, & Medendorp, 2011; Thompson & Henriques, 2011; see also figure 1.1). This is somewhat intuitive, as the visual input of the reach target is encoded first in the reference frame that is related to the sensory organ that also receives the input. A seminal study which demonstrates the use of a gaze-dependent reference frame for goal-directed movements in humans was conducted by Henriques, Klier, Smith, Lowy, & Crawford (1998). Participants had to reach to the location of a remembered, visually presented target-LED in an otherwise dark environment (i.e., without any allocentric information or visual feedback of the hand). In different conditions, participants had to fixate their gaze on various locations relative to the target when it was presented or had to move their eyes to a certain location between target presentation and movement onset. Results revealed reaching errors (reaching endpoints that deviated from the actual target location) that varied systematically with fixation location relative to the target. Based on the systematic influence of gaze location on reaching errors, the authors concluded that the reach targets were encoded in a gaze-dependent frame of references. In the past, this finding was supported by many other studies, also arguing for a gaze-dependent reference frame for memory-guided reaching (Ambrosini, Ciavarro, Pelle, Perrucci, Galati, Fattori, Galletti, & Committeri, 2012; Beurze, Van Pelt, & Medendorp, 2006; Fiehler, Schütz, & Henriques, 2011; Khan, Pisella, Rossetti, Vighetto, & Crawford, 2005; Medendorp & Crawford, 2002; Medendorp, Beurze, Van Pelt, & Van Der Werf, 2008; Selen & Medendorp, 2011; Thompson & Henriques, 2008). Even though the gaze-dependent reference frame seems to have a major role, there are also other egocentric reference frames for coding target locations. For instance, also *body-centered* (Bernier & Crafton, 2010) or *head-centered* (Avillac, Denève, Olivier, Pouget, & Duhamel, 2005; Mullette-Gillman, Cohen, & Groh, 2009) egocentric reference frames have been identified for coding target locations egocentrically, mostly mixed with a gaze-centered representation. Moreover, for the final movement execution the spatial information about the reach target location needs to be integrated and transferred to an egocentric representation that also includes proprioceptive information about the hand position for generating a motor command (Sober & Sabes, 2003; Sober & Sabes, 2005). The precise nature of this transformation and the action execution itself is beyond the scope of this thesis but can be found elsewhere (Crawford et al., 2011).

However, besides the availability of egocentric information, in real world situations a target is usually surrounded by many other objects or background structures which are providing additional allocentric information. How this allocentric information is also used for coding the location of a target is described in the next chapter.

1.3 Allocentric reference frame for reaching

How allocentric information is used for memory-guided reaching movements has been addressed in previous research. In an exemplary experiment on this topic (Krigolson & Heath, 2004), participants had to encode the location of a briefly flashed target-LED. This flash was either accompanied by an illumination of four background-LEDs that surrounded the target in the shape of a square or was presented alone. After the LEDs were turned off, a tone instructed participants to perform a memory-guided reaching movement to the location of the target-LED. Results revealed higher accuracy and precision in the condition when the background LEDs were presented. The authors concluded that the allocentric information given by the background-LEDs was taken into account when planning the reaching movement. The finding that landmarks increase accuracy and/or precision of memory-guided reaching movements to a memorized target was confirmed by other studies (Krigolson, Clark, Heath, & Binsted, 2007; Lemay, Bertram, & Stelmach, 2004; Obhi & Goodale, 2005).

But not only accuracy and precision of reaching endpoints demonstrate the use of allocentric information for memory-guided reaching movements. The location of reaching endpoints itself can also be manipulated by displacing allocentric cues. For instance, in a study by Byrne and Crawford (2010) a target dot was briefly presented within a square of four landmark dots. After the stimuli disappeared, the four landmark dots reappeared but were coherently shifted away from their former location. After the landmark dots disappeared, participants performed a reaching movement to the remembered location of the target dot. The rationale behind the landmark shift was that if participants take the allocentric information by the landmarks into account, reaching endpoints should also deviate into the direction of the landmark shifts. In contrast, if participants purely use an egocentric representation of the target location, then reaching endpoints should not have been affected by the landmark shift. The authors found reaching endpoints between these two extrema, and argued that both egocentric and allocentric information were used for performing the memory-guided reach. This is remarkable because participants were instructed to ignore the surrounding landmarks. Thus, the allocentric information could not be actively ignored, but was integrated into the movement plan. In a different study with a similar paradigm the allocentric information was available while participants performed the reaching movement (Sheth & Shimojo, 2004). Furthermore, they were instructed to perform the reach to the target location relative to the landmark and hence, to prefer the use of the allocentric information. Again, the reaching endpoints were highly affected by the shift of the allocentric information, confirming findings by Byrne and Crawford (2010).

Current research suggests that if allocentric information is provided, multiple factors influence the extent to which it is used for performing memory-guided reaching movements. One of these factors is the distance between the reach target and potential allocentric cues. With an increasing distance, the influence of the allocentric information decreases (Camors, Jouffrais, Cottereau, & Durand, 2015; Krigolson et al., 2007). Furthermore, when multiple potential allocentric cues in different distances are provided, the landmark closest to the target is used for coding a target location allocentrically (Diedrichsen, Werner, Schmidt, & Trommershäuser, 2004; Spetch, 1995). Moreover, with decreasing landmark stability (e.g., by introducing a vibration of landmarks) the use of allocentric information decreases (Byrne & Crawford, 2010). Finally, the anticipation of availability of visual feedback leads participants to rely more on the egocentric information. In contrast, when they expected no target visibility or were unsure about this, they preferred an allocentric coding of the target (Neely, Tessmer, Binsted, & Heath, 2008).

Taken together, besides research showing the use of a gaze-dependent reference frame for coding target locations, there is an increasing number of studies providing evidence that also the allocentric information is taken into account, with benefits for reaching endpoint accuracy and precision. Moreover, multiple factors that determine the use of allocentric information have been identified.

1.4 Reference frames in real world situations

It is noteworthy that almost all of these studies chose rather artificial and arbitrary stimuli like LEDs (Carrozzo, Stratta, McIntyre, & Lacquaniti, 2002; Krigolson & Heath, 2004; see figure 1.2A for an example), cylindrical objects (Lemay et al., 2004; Obhi & Goodale, 2005), or geometric patterns on a screen (Byrne & Crawford, 2010; Diedrichsen et al., 2004; Krigolson et al., 2007; Neely et al., 2008; Spetch, 1995; Schütz et al., 2013; Seth & Shimojo, 2004). The question arises to which extent results from these studies can be transferred to a real world scenario where multiple objects could serve as allocentric cues in a complex environment. Previous research from other fields has shown that results from laboratory experiments do not always transfer to real-world settings. For example, for decades research on visual attention mainly used sets of much reduced and artificial stimuli like letters or arrows to draw conclusions about the nature of the human visual attention system (e.g. Treisman & Gelade, 1980; Posner, 1980). However, more recent studies have shown that older theories (e.g. Feature-Integration Theory of attention or findings about covert visual attention) do not fully account for findings about visual attention and eye-movements in naturalistic scenes (e.g., Henderson, Brockmole, Castelhano, & Mack, 2007; Wolfe, 1994; Wolfe, Võ, Evans, & Greene, 2010). This led to the demand "that research needs to be grounded in the real world and not in experimental paradigms" (Kingstone, Smilek, Ristic, Friesen, & Eastwood, 2003). Similar to research on visual attention, it is conceivable that the use of more naturalistic stimuli for the investigation of the use of reference frames for memory-guided reaching leads to an extension of previous results of studies with more abstract stimuli.

Another shortcoming of previous research is that reach targets and allocentric cues had no functional relation to each other (e.g. Byrne & Crawford, 2010; Diedrichsen et al., 2004; Krigolson et al., 2007; Krigolson & Heath, 2004; Neely et al., 2008; Spetch, 1995; Schütz et al., 2013; Seth & Shimojo, 2004; see figure 1.2A for an example). For example, when reaching in the real world for a mug on an office desk while looking on the monitor, one has to take the locations of other objects like a keyboard, a glass, or office supplies into account to prevent collisions while reaching for the mug. Thus, objects in this environment are functionally linked to the target. In contrast, when presenting dots, bars, or geometric patterns, they are not functionally related to each other. Although previous studies used these functionally unrelated stimuli, they found a robust influence of allocentric information on reaching movements. However, the use of allocentric information in a more complex environment with objects that are functionally linked together could reveal different results compared to a functionally unrelated set of stimuli. Finally, in some studies (Schenk, 2006; Thaler & Goodale, 2011a,b; Zaehle, Jordan, Wüstenberg, Baudewig, Dechent, & Mast, 2007) participants were explicitly instructed to use the egocentric or allocentric reference frame or were given a task that biased participants' behavior toward one of the two reference frames for coding a target location. It is unlikely that participants showed a naturalistic target location encoding and reaching behavior in these tasks. In real world scenarios, no explicit instruction is given and the task comes from the situation itself, namely to perform a memory-guided reach for an object in a complex environment with multiple objects.

Taken together, the described differences between experiments performed in the laboratory and real world situations could lead to findings that do not entirely represent the nature of human behavior. To overcome these limitations and to do a first step into the direction of a more realistic approach for investigating the use of allocentric information for memory-guided reaching, I conducted a series of experiments which I will summarize chapter 1.6. In these experiments I used naturalistic, complex scenes with multiple objects that were functionally related to each other (for an example see figure 1.2B) without biasing participants' behavior by giving a certain instruction. However, most of my studies were carried out by presenting stimuli in 2D space on a computer screen. For an even more realistic real world scenario, one has to consider 3D space as well, which is the topic of the next chapter.

Figure 1.2. (**A**) Example of stimuli from a top-view perspective adapted from Krigolson and Heath (2004). A target-LED could appear at different distances from the participant, who performed reaching movement from the home position to the target location. Four background-LEDs served as allocentric cues and were lit together with the target-LED when participants encoded its location. Note that there is no functional relation between target and allocentric information. (**B**) Example of a stimulus image taken from my second study. Participants encoded the location of objects in a more naturalistic scene. As different objects could serve as potential reach target and objects are belonging together with regard to their content (breakfast scene), target and allocentric cues are related.

1.5 Coding of target locations in depth

A special case in the laboratory, but an everyday situation in real world is that coding of object locations is not restricted to the 2D-plane (i.e. computer screen) but rather happens in 3D space. To perceive the location of an object in depth, the human brain makes use of several *depth cues*, which can be classified as monocular (e.g., occlusion, height in the visual field, relative size) and binocular depth cues (e.g., binocular disparity, accommodation, vergence). All depth cues are combined and weighted, depending on distance between the observer and object locations in depth (Armbrüster, Wolter, Kuhlen, Spijkers, & Fimm, 2008; Cutting, 1997; Knill, 2005; Landy, Maloney, Johnston, & Young, 1995). Especially, if one wants to act on an object, very precise depth information about the object location is necessary, which cannot be provided by binocular depth cues alone (Hibbard & Bradshaw, 2003) but also depends highly on the use of monocular depth cues (Bruno & Cutting, 1988; Magne & Coello, 2002), especially the object's size (Naceri, Chellali, & Hoinville 2011).

Besides different depth cues, how are egocentric and allocentric information used for coding a target locations in depth? Even though there is evidence that shows gaze-centered coding (Van Pelt & Medendorp, 2008), other studies revealed an influence of allocentric information on coding target locations in depth for reaching (Coello, Richaud, Magne, & Rossetti, 2003; Neely, Heath, & Binsted, 2008). A paradigm that was utilized to investigate the use of allocentric information for reaching to a target in depth employs the *Induced Roelofs Effect* (IRO). Here, participants had to encode the location of a target (dot) within a rectangular frame, whose long axis was off-centered relative to the target. After target and frame vanished, participants performed a memory-guided reaching movement to the target location. By using this paradigm, Coello and colleagues (2003) found reaching errors that deviated systematically in the direction of the closest edge of the frame, but only in conditions when the frame was oriented in depth but not when it was oriented horizontally in front of the participants. It can be concluded that the allocentric information given by the frame was taken into account when participants planed their reaching movements. However, it seems that this influence was only present for coding in depth. In contrast, Neely and colleagues (2008) performed an experiment using the same paradigm. Here, the authors did not find a difference between different frame orientations, arguing for a similar use of allocentric information irrespective of horizontal target coding or coding in depth.

Taken together, coding of target locations for actions in depth depends on the availability of different depth cues. Especially object size seems to play an important role. Nevertheless, if available, allocentric information also contributes to the coding of reach targets. However, it is rather unclear whether this influence varies between horizontal coding and coding in depth. Moreover, also the IRO does not represent a real world scenario as stimuli are rather abstract. To deepen our knowledge about the use of allocentric information for reaching in depth, I conducted one study in which I used naturalistic and complex stimuli as in my 2D studies, but presented them in 3D virtual reality and compared allocentric location coding of reaching targets in the horizontal plane to their coding in the depth plane. Moreover, I also investigated the influence of object size as an important monocular depth cue on the encoding of allocentric reach targets in depth.

1.6 Outline

The paradigm of all studies reported in this thesis was held constant for better comparison between them (see figure 1.3) and was based on the approach by Byrne and Crawford (2010) and Sheth and Shimojo (2004). In their experiments, landmarks were coherently shifted between encoding of a target location and a memory-guided reach toward it. They found that reaching endpoints deviated systematically in the direction of the shifted landmarks, indicating the use of the allocentric information in coding reach target locations for memory-guided reaching.

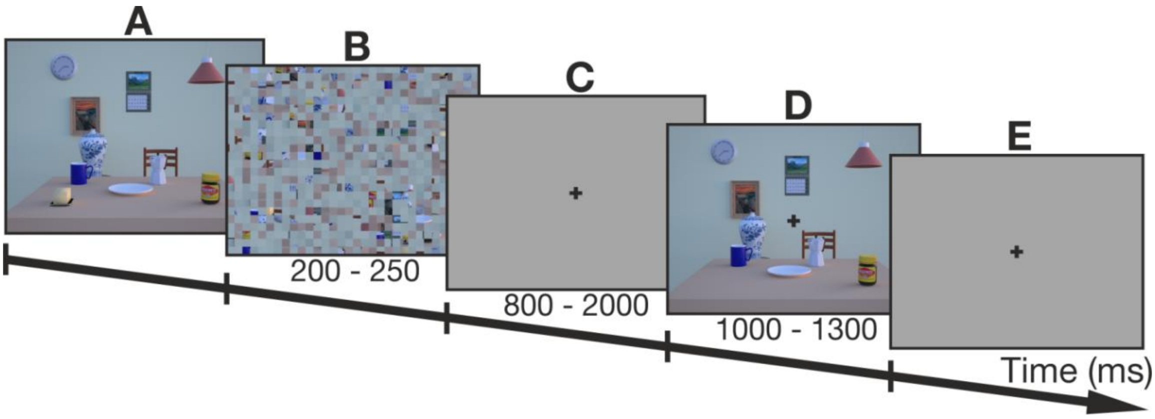

Figure 1.3. Example trial scheme of the paradigm that I used in my studies. Different durations are due to slight adjustments of timings between them. Stimulus images are taken from the third study. (**A**) First, the encoding image was presented and participants terminated the exploration of the image by a button press. (**B**) Then, a scrambled version of the encoding image was presented (except for the first study), followed by (**C**) a delay. (**D**) Thereafter, the test image with one of the task-relevant objects missing (here: butter dish) was presented before (**E**) a tone prompted participants to reach to the target onto a gray screen while fixating the cross at the center of the screen. Note that in the fourth study, stimuli were not presented on a computer screen but in virtual reality. Therefore, no objects were placed in the background, participants' viewpoint was closer to the table and the viewing angle was steeper onto the table. Moreover, the delay and reaching image did not consist of a gray screen but of an empty table with the fixation cross on it.

In the first study (chapter 2) I was aiming to gain basic insights into the use of egocentric and allocentric reference frames for memory-guided reaching by using more naturalistic and complex stimuli. Participants encoded the locations of objects in a scene that consisted of a breakfast table with some objects on the table and some in the background, which was presented a photograph on a computer screen. After this encoding scene vanished, a short delay was inserted (gray screen with a fixation cross to keep the gaze-centered reference frame constant). Next, the breakfast scene reappeared as a test scene, but with one object missing, which served as the reach target location. After the test scene vanished, a gray screen was presented and participants were asked to perform a memory-guided reaching movement to the location of the missing object on the screen. They were not instructed to reach to the target location either relative to their egocentric position or relative to the locations of the other objects and thus, were able to show realistic real world behavior. Besides removing one object in the test scene, I also shifted some of the remaining objects either to the left or to the right. As in the experiments by Byrne and Crawford (2010) and Sheth and Shimojo (2004), I expected that by shifting landmarks (i.e. the objects), reaching endpoints should deviate in the direction of object shifts if the allocentric information was used to code the target location. In contrast, if participants completely rely on their unchanged egocentric reference frame, reaching endpoints should be unaffected by the object shifts. Moreover, I varied the number of shifted objects to investigate whether there is an influence of the magnitude of changes in the scene on the use of allocentric information. Results revealed a systematic deviation of reaching endpoints in the direction of object shifts of about 40 to 50 % of shift magnitude. Furthermore, I found that the reaching errors scaled with the number of shifted objects. Moreover, I observed a difference in the use of allocentric information depending on whether the shifted objects were located either on the breakfast table or in the background. A possible explanation for this difference could have been the fact that only one background object was shifted at a time whereas one, three, or five objects on the table could have been displaced. Thus, the amount of changes of background objects could have been too small to influence participants' reaching behavior. Another possibility was that only objects on the table served as potential reach targets and thus were highly relevant to perform the task. In contrast, background objects never served as potential reach target and hence were rather irrelevant to perform the reaching task. To test for these two options, I conducted two experiments in my second study.

In the second study (chapter 3), I applied the same experimental paradigm as in the first study, but I used 3D-rendered scenes presented on a computer screen instead of photographs, as 3D-rendered images are easier to create. In the first experiment, I increased the number of shifted background objects up to five compared to the first study, but still only objects on the table could serve as potential reach targets. In the second experiment, I shifted the same number objects in the background or on the table as in experiment 1 but this time, only objects in the background were task-relevant in that they could serve as potential reach targets. In both experiments, I found an influence of object shifts on reaching endpoints only in conditions in which objects were shifted that could serve as potential reach targets, irrespective of whether the objects were located on the table or in the background. Similar as in the first study, reaching errors scaled with the number of shifted objects. It can be concluded that task-relevance of objects is an important factor that determines whether an object is used as allocentric cue or not.

The aim of third study (chapter 4) was to examine further factors that determine the use of objects as allocentric cues. Again, the experimental paradigm was similar as in the two previous studies. However, in these studies the task-relevant objects were grouped as a spatially distinct object cluster either on the table or in the background, which could have affected the way participants used the task-relevant or task-irrelevant allocentric information. To overcome this limitation, in the third study I positioned task-relevant and task-irrelevant objects not grouped but distributed over the whole scene in the background and on the table. Task-relevant and task-irrelevant objects appeared mixed instead of spatially distinct. To enable participants to perform the task, they were trained on task-relevance of the objects before the experiment. Another factor that could affect the use of allocentric information for coding a target location is the reliability of the allocentric cues. If one would perceive the allocentric information to be unstable, one may rather rely on a more stable representation of the target location, in this case the egocentric reference frame (Byrne & Crawford, 2010). If so, reaching endpoints should be influenced by object shifts to a lesser extent. To manipulate the reliability of the allocentric information, I introduced conditions in which objects were not shifted coherently in one direction but incoherently in opposite directions. Thus, participants would perceive object locations in the test scene as very different from the locations of the encoding scene. Consequently, the reliability of the object locations and hence, the reliability of the allocentric information given by these objects should be very low. First, I found an overall reduction in the use of the allocentric information, also in conditions when task-relevant and task-irrelevant objects were shifted coherently in the same direction. Second, incoherent object shifts (i.e. shifts of objects in opposite directions) led to a dramatically reduced use of the allocentric information.

In my fourth study (chapter 5) I continued working on the main goal of this series of experiments to extend laboratory findings to more naturalistic settings. Even though, I used naturalistic and complex stimuli in the previous studies, they were still presented in 2D on a computer screen. The next logical step toward a real world situation was to present stimuli in 3D space in a virtual reality. To this end, I transferred the paradigm of my previous studies to a virtual reality setup. In a first experiment I wanted to validate the new technique. Therefore, task-relevant objects were only placed on a table and were shifted only coherently in the horizontal direction. Thus, I was able to compare results of this virtual reality setup to corresponding conditions of my second study. Moreover, I wanted to investigate whether the use of allocentric information depends on the distance between objects and observer. To this end, I placed objects in three depth clusters with increasing distances between the observer and the objects. Results revealed reaching errors that did not differ from those I found in the second study. No influence of different distances between observer and objects was found. In a second experiment within this study, I followed the controversy that the direction and distance of a reaching target seems to be encoded in different subsystems (Chieffi, & Allport, 1997) but is affected similarly by allocentric information oriented in these two dimensions (Neely, et al., 2008). Therefore, I used the same paradigm as in the first experiment but shifted objects in depth instead of in the horizontal direction and compared the results to those of the first experiment. Moreover, it is unclear to which extent different depth cues (e.g. binocular disparity, vergence, object size) in the perception of depth contribute to the allocentric coding of object locations in 3D space. As a first step, I explored the role of object size and introduced conditions in which the retinal object size due to object shifts was manipulated. In one condition the object size changed naturally when objects were shifted toward or away from a participant (i.e. retinal object size increases or decreases). In a second condition, this effect was magnified by increasing the object size additionally when objects were shifted toward the participant and decreased additionally when shifted away. In a third condition, I reversed this effect and thus created a conflict of object shift direction and change in retinal object size. Results revealed a higher use of the allocentric information compared to the first experiment when objects have been shifted horizontally. Regarding the object size manipulation, I found an increased use of allocentric information when changes in retinal object size were magnified and a reduction when a conflict between object shift direction and changes in retinal object size was created.

2 Integration of egocentric and allocentric information during memory-guided reaching to images of a natural environment.

A similar version of this manuscript has been published as:

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, doi: 10.3389/fnhum.2014.00636.

When interacting with our environment we generally make use of egocentric and allocentric object information by coding object positions relative to the observer or relative to the environment, respectively. Bayesian theories suggest that the brain integrates both sources of information optimally for perception and action. However, experimental evidence for egocentric and allocentric integration is sparse and has only been studied using abstract stimuli lacking ecological relevance. Here, we investigated the use of egocentric and allocentric information during memory-guided reaching to images of naturalistic scenes. Participants encoded a breakfast scene containing six objects on a table (local objects) and three objects in the environment (global objects). After a 2 s delay, a visual test scene reappeared for 1s in which one local object was missing (= target) and of the remaining, one, three or five local objects or one of the global objects were shifted to the left or to the right. The offset of the test scene prompted participants to reach to the target as precisely as possible. Only local objects served as potential reach targets and thus were task-relevant. When shifting objects we predicted accurate reaching if participants only used egocentric coding of object position and systematic shifts of reach endpoints if allocentric information were used for movement planning. We found that reaching movements were largely affected by allocentric shifts showing an increase in endpoint errors in the direction of object shifts with the number of local objects shifted. No effect occurred when one local or one global object was shifted. Our findings suggest that allocentric cues are indeed used by the brain for memory-guided reaching towards targets in naturalistic visual scenes. Moreover, the integration of egocentric and allocentric object information seems to depend on the extent of changes in the scene.

2.1 Introduction

When reaching to a visual target in a naturalistic environment, the brain can make use of absolute or relative spatial information for reach planning. This can be formalized in terms of two broad classes of reference frames: an egocentric reference frame that represents the absolute position of an object with respect to the observer and an allocentric reference frame coding the position of an object relative to other objects in the environment (Colby, 1998). While egocentric reference frames depend on eye, head, body, etc. position and orientation, allocentric reference frames are relatively observer-invariant. It is well known that for goal-directed reaching movements, a gaze-dependent, egocentric reference frame is used preferentially as demonstrated by electrophysiological studies in monkeys (Batista et al., 1999; Buneo et al., 2002) and behavioral (Fiehler et al., 2011; Henriques et al., 1998; Medendorp & Crawford, 2002) and brain imaging studies (Bernier & Grafton, 2010; Medendorp et al., 2003) in humans.

Despite the dominance of gaze-dependent representations for reach planning, allocentric information also contributes to the encoding of reach target location. For example, visual landmarks provided during target presentation lead to an increase in accuracy and precision of reaching movements (Krigolson & Heath, 2004; Krigolson et al., 2007; Obhi et al., 2005). The effect of reduced reach endpoint variability was even more pronounced when the landmarks were placed close to the reach target (Krigolson et al., 2007). If landmarks are present while participants reach to remembered targets updated in their visual periphery, the influence of gaze-dependent spatial coding has been found to decrease suggesting a combined use of egocentric and allocentric information (Schütz et al., 2013). Such combination of egocentric and allocentric reference frames is supposed to occur after the intervening saccade at the time of action (Byrne et al., 2010) and depends on heuristics for external cue stability as well as the reliability of egocentric and allocentric cues which determines the weighting in memory-guided reaching (Byrne & Crawford, 2010; McGuire & Sabes, 2009). In addition, the proximity of the landmarks and the target seems to affect reach endpoints showing systematic distortions toward the nearest landmark (Diedrichsen et al., 2004). However, this effect only occurred when landmarks were available during target encoding but not during reaching. Moreover, structured visual background placed close to the target led to more precise reaching movements than distal visual background presumably linked to the proximity of veridical target location (Krigolson et al., 2007). The use of allocentric cues in addition to egocentric representations has even been demonstrated for imagined landmarks which were not physically present during target encoding or reaching but represented a virtual straight line (Carrozzo et al., 2002). The authors argued for the use of concurrent and independent coexisting egocentric and allocentric target representations used for memory-guided reaching.

Here we set out to address a series of controversies and gaps in the literature. (1) So far, isolated visual targets together with abstract, task-irrelevant landmarks on an otherwise blank screen have been used to investigate the underlying reference frames for reaching movements. However, it is not a given that findings from such abstract studies will hold in natural situations, where we are surrounded by a vast number of visual features creating a complex visual scene. (2) Moreover, previous studies (e.g. Thaler & Goodale, 2011a, 2011b; Schenk, 2006; Zaehle et al., 2007) explicitly asked participants to use a predefined egocentric or allocentric reference to perform the task probably covering individual spatial coding strategies. Therefore, one aim of our study was to examine the contribution of egocentric and allocentric information to reaching to images of a natural scene without biasing participants' behavior to use either one or the other reference frame. (3) It has been suggested that object proximity is an important factor biasing reach endpoint (Diedrichsen et al., 2004); we will challenge this view here. (4) We will further test whether allocentric information influences reach trajectory planning (Burns & Blohm, 2010) versus feedback-based control processes (Krigolson et al., 2007). Participants reached to a remembered location of an object on a breakfast table while we varied the location of the surrounding objects by applying a leftward or a rightward shift (allocentric cue). Spatial shifts were either applied to surrounding objects on the table which could be potential targets and were thus task-relevant (local objects) or to objects in the environment which never served as a target (global objects). Since the position of gaze, head and body were kept constant, we expected no systematic reach errors if participants relied on an egocentric target representation alone. If participants represented the target with respect to other objects on the table and/or in the environment, i.e. they used an allocentric representation, we predicted reach errors which vary as a function of object shifts. We show that memory-guided reaches to images of naturalistic environments are planned using both egocentric and local allocentric information, but not global allocentric cues.

2.2 Methods

Participants

Data were recorded from 14 participants with normal or corrected to normal vision. One participant was excluded from further analysis because of poor fixation behavior (< 1 % valid trials), another participant because of frequent movement onsets while the test scene was still displayed (29.2 %). The final sample consisted of 12 participants (3 female; 3 left-handed, self-report) ranging in age from 20 to 37 years (*mean* 24 ± 4 years). All procedures were conducted in agreement with the ethical guidelines of the local ethics committee of the University of Giessen and were approved by the Queen's University Ethics Committee in compliance with the Declaration of Helsinki.

Materials

Participants viewed photographic color images showing a breakfast scene with 6 *local objects* (coffee mug, plate, espresso cooker, marmalade jar, butter dish, and egg cup) on a table that was placed in front of a white wall and 3 *global objects* (table [T], table cloth [C], and painting on the wall [P]) in the scene (see figure 2.1A). The object properties are summarized in table 2.1.

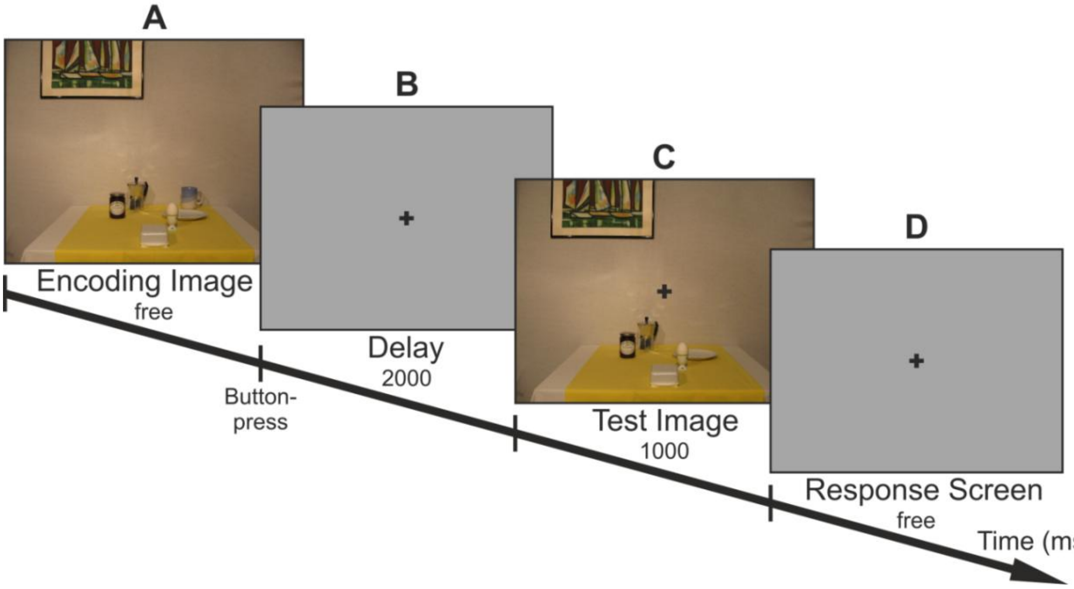

Figure 2.1. Trial-Procedure. Participants first viewed 1 of 18 encoding images (**A**) without time limit and free gaze. After a 2 s blank gray screen (**B**) the test image appeared for 1 s (**C**) in which one of the objects on the table was missing (= target; here: cup). They were instructed to reach to the remembered target as soon as the response screen was presented (**D**). After the encoding period, fixation had to be maintained at the fixation cross until the end of the reach.

Table 2.1. Maximum height and width of objects in the scene in cm.

object	height	width
plate	2.1	19
butter dish	4.9	8.5
marmalade jar	10.6	6.5
coffee mug	10.3	8
egg cup	10.1	4.1
espresso cooker	15	15
painting	41	51
table	75.4	78
table cloth	/	60

The 6 local objects were arranged in 18 different configurations on the table to minimize memory effects (*encoding image*). To this end, the objects were assigned to one of four possible locations in depth (8 cm, 28 cm, 48 cm, or 68 cm from the front table edge) and to a randomized horizontal position. Configurations were pseudo-randomized and fulfilled the following criteria: (i) at least one object was placed at every depth position, (ii) objects were placed with a minimum horizontal distance of 8cm away from the edges of the table cloth in order to enable horizontal displacement on the table cloth and (iii) < 50 % of each object was occluded. In addition to the encoding images, we created *test images* lacking one of the 6 local objects (= reach target). In 2/3 of the test images, local or global objects were physically displaced in the horizontal direction on the table by 8 cm either to the left or to the right (50 % leftward displacement) prior to taking photographs. Due to the finite camera distance, this corresponds to different shifts on the image (and thus also on the screen), depending on the depth position of the object, i.e. whether it was located in the proximal, first medial, second medial or distal depth plane. Thus resulting visual shifts on the screen images could be 4.24 deg, 3.80 deg, 3.36 deg and 2.92 deg for proximal, first medial, second medial or distal object depth respectively. In the remaining 1/3 of the test images, the remaining objects in the scene were not shifted. In order to ensure precise and reproducible object placement in the images, a grid was projected from above on the table before the photographic image was taken with a resolution of 2048 x 1536 pixels.

In total, 342 photographic images were taken including 18 encoding images and 324 test images with 108 images without object displacement, 108 images with local object displacement and 108 images with global object displacement. Separate photographic images were taken for each target (6) in each configuration (18) and experimental condition (3; control, local and global).

Apparatus

Stimuli were presented on a 19“ (40.64 cm x 30.48 cm) CRT monitor with a resolution of 1920 x 1200 pixels and a refresh rate of 60 Hz using the Psychtoolbox (Brainard, 1997) in Matlab (The Mathworks, Inc., Natick, MA, USA). Monitor/image edges were visible. Participants sat at a table with their head stabilized on a chin rest guaranteeing an eye-monitor distance of 47 cm. They performed the task in complete darkness but the use of a computer screen resulted in some limited illumination of the hand. Participants executed right arm reaches from an elevated start position placed 27 cm in front of the screen at the level of the lower screen edge. Reaches were recorded with an Optotrak Certus (NDI, Waterloo, Ontario, Canada) infrared marker-based motion tracking system with a sampling rate of 250 Hz using one marker attached to the fingertip. In order to control for correct fixation behavior, we also recorded eye movements using an EyeLink 1000 tracking system (SR Research, Osgoode, ON, Canada) with a sampling rate of 1000 Hz. Participants initiated the trials by a left-hand button press on a game controller located on the table in front of their left shoulder.

Procedure

The trial procedure is illustrated in figure 2.1. Participants started each trial by a button press with their left hand. An encoding image containing all local and global objects was displayed on the screen until participants continued the trial with a button press on the controller. They were instructed to encode the location of the local objects in the scenes while freely moving the eyes. Participants had as much time as desired and were instructed to press the game controller with the left hand in order to pursue the trial. The encoding phase was followed by a central fixation cross that appeared on a uniform gray background for 2 s prompting participants to maintain fixation at this location until the end of the reach. Then, the test image without one of the 6 local objects was presented for 1 s, superimposed with a fixation cross. After the test image disappeared, the fixation cross was displayed on a uniform gray background and participants were asked to reach with their right hand to the remembered location of the missing object (= target) on the screen. Thus, reaches were performed while fixating at the center of the screen and without any visual information about the scene. Whenever participants were unsure about the location of the target, they were instructed not to reach but to continue with the next trial.

Participants performed three experimental conditions (figure 2.2). In the *allo-local* condition, we manipulated the number of local objects shifted in the scene of the test image before reaching. In particular, 1, 3 or all 5 remaining local objects were horizontally misplaced by 8 cm (in physical space) to the left or to the right (loc1, loc3, loc5) without affecting the position of the global objects. Within one trial, objects were always shifted in the same direction. In the *allo-global* condition, one global object was shifted by 8 cm (in physical space) leftwards or rightwards by leaving the location of the local objects unchanged (gloT, gloC, gloP). In the control condition, no object shifts occurred.

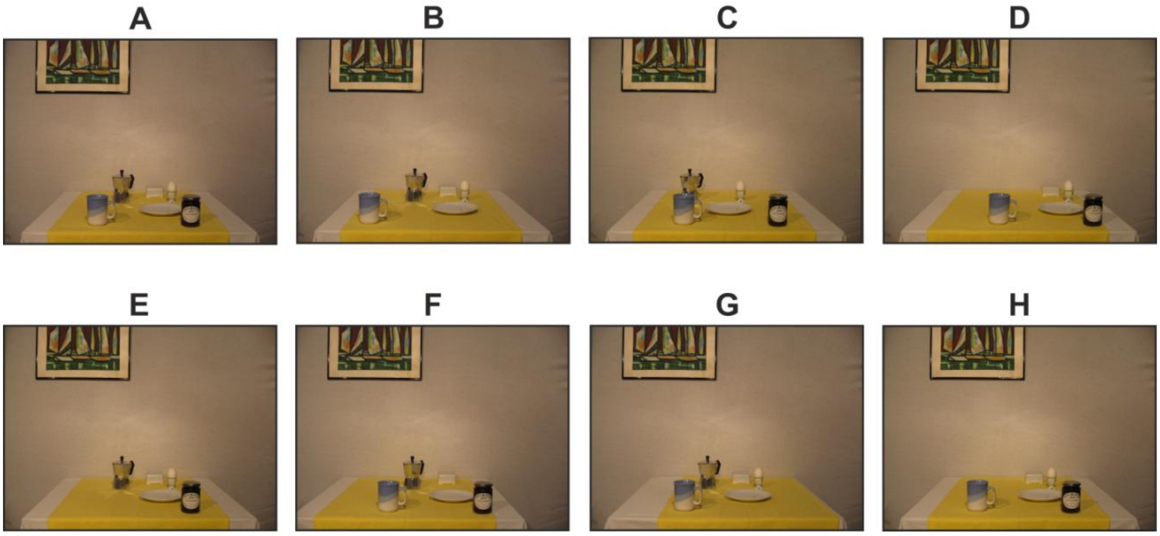

Figure 2.2. Example images of one encoding image and seven corresponding test images. (**A**) Encoding image with all six objects. (**B**) Test image of the local 1 condition (loc1) with the marmalade jar missing and the cup shifted to the left. (**C**) Local 3 condition (loc3) with missing butter dish and espresso, egg and plate shifted to the left. (**D**) Example image from the local 5 condition (loc5): The espresso cooker is missing, all other objects are shifted to the right. (**E**) Control condition with missing cup. (**F**) Global-Table (gloT) condition with the egg missing and the table shifted to the left. (**G**) Global-Table cloth (gloC) condition with the marmalade jar missing and the table cloth shifted to the right. (**H**) Global-Painting (gloP) condition with the espresso cooker missing and the painting shifted to the right.

Each participant completed 648 trials split up in 18 blocks consisting of 36 trials each. Before the start of the experiment, each participant completed a training block of 18 control trials. Data of each participant were recorded in three 1h sessions on different days consisting of 6 blocks each.

Data Reduction and Statistical Analysis

Data preprocessing and analyses were done using MATLAB and final inferential statistics were computed in SPSS (Version 21.0). An alpha level of .05 was used for evaluating all effects.

First, we analyzed eye tracking data in order to control for correct fixation. Trials were classified as invalid and excluded from further analyses if gaze deviated more than ± 2.5° from the fixation location. This applied to 564 trials (7.25 %). Second, reach endpoints were determined as the position where reach velocity and screen distance were minimal. Reaching endpoints in screen coordinates were then computed from camera coordinates using quaternion transformation (Leclercq et al., 2013).We excluded trials in which reach endpoints deviated more than 2.5 SD from the average reach endpoint per test image (figure 2.3B). This resulted in removing 638 trials of the remaining trials (8.2 %). In 187 trials (2.4 %), participants responded before the test image disappeared. To test memory-guided reaching without visual scene information, these trials were also removed for analysis. In total, 6387 out of 7776 trials remained for analysis.

To investigate eye movement behavior during the scene encoding phase, we computed the relative frequency of fixations (figure 2.3A). To do so, we averaged fixation positions (excluding saccades) across all encoding phase time frames and convolved the result with a Gaussian filter of 1.5 deg width. The result was plotted as a heatmap and overlaid onto an example encoding image.

In order to investigate the influence of allocentric information in the scene on reach endpoints, we computed allocentric weights using linear regressions. In a first step, we calculated the group-mean reaching endpoint for every combination of object configuration and target identity in the control condition. These values served as subjective target location in the scene. In a second step, for every single reaching response in the allocentric conditions, its horizontal deviation from the subjective target location of the corresponding control image (same target and arrangement) was computed and compared to the expected allocentric deviation. Expected allocentric deviations were calculated for every test image as the average value by which the reference objects were shifted in the scene. For example, a visual leftward shift of three reference objects by 4.24 deg, 3.80 deg and 3.36 deg (loc3 condition; objects placed at different locations in depth) would result in an expected allocentric deviation of 3.80 cm (average of the three individual object shifts) if the target were solely represented in an allocentric reference frame, i.e. relative to other objects in the scene. In general, leftward deviations were coded as negative values and rightward deviations as positive values. In a third step, the observed horizontal deviation from the subjective target location for a leftward and for a rightward shift of the same target in the same arrangement were plotted against the expected allocentric deviations for each individual and each allocentric condition. Finally, a regression line was fitted to the data and the slope of the regression line determined the allocentric weight.

We applied one-sampled t-tests to examine whether individual local and global allocentric weights significantly differed from zero. Since allocentric weights are computed on the basis of the results of the control condition, a test against zero corresponds to a statistical comparison to the control condition. To compare individual allocentric weights across conditions, we then computed one-way repeated measures ANOVAs with 3 levels for the local condition (loc1, loc3, loc5) and the global condition (gloT, gloC, gloP), separately. Significant results were followed-up with post-hoc t-tests. Based on our hypotheses, t-tests were calculated one-sided and corrected for multiple comparisons using Bonferroni-Holm correction.

To test for differences in movement initiation and duration depending on the experimental conditions, we examined response latencies and movement times respectively. Response latencies were determined as the time from the disappearance of the test image until the start of the reaching movement which was defined as the point in time when the right index finger exceeded a velocity of 50 mm/s for 200 ms. Movement time was determined as the time from the start of the movement until the end defined as the time point when the velocity of the index finger fell below 50 mm/s for 100ms and distance to the screen was minimal. Individual median response latencies and movement times were compared between the experimental conditions by computing separate one-way repeated measures ANOVAs with 4 levels for the local condition (loc1, loc3, loc5, control) and for the global condition (gloT, gloC, gloP, control). Two-sided post-hoc t-tests were calculated and corrected for multiple comparisons using Bonferroni-Holm correction.

We instructed participants to perform no reach movement if they were uncertain about the location and/or identity of the target. Frequency of trials in which participants did not respond was computed per condition and tested against the assumption that those trials are equally distributed across all conditions by using a Friedman's test.

To determine whether allocentric influences were part of the overall movement plan or whether they emerged only during online corrections (cf. Burns & Blohm, 2010;

Krigolson et al., 2007), we analyzed reaching trajectories using functional data analysis (FDA; Ramsay & Silverman, 2005). Some trials were excluded from the analysis due to the following reasons: (a) less than 50 data frames were collected per reaching movement due to Optotrak marker visibility problems; (b) moving velocities exceeded 600 cm/s during one reaching movement; and (c) trials lacked more than 20 consecutive data frames. Following these criteria only three trials (less than 0.1 %) were discarded.

First, we shifted the movement onset (i.e. the first data frame) of each trajectory to the coordinate point 0/0/0 (x-, y-, and z-direction in 3D Cartesian space) and aligned the subsequent data frames. Second, we spatially normalized the trajectories by fitting order 6 splines to each of the three dimensions (x, y, and z) with a spline at every data frame. Third, we smoothed the data using a roughness penalty on the fourth derivative and $\lambda = 1^{-10}$ (within 0.008 of the generalized cross-validation estimate). Out of this mathematical definition we evaluated for each trajectory 1200 equally spaced data points. Then, 120 out of 1200 points were extracted resulting in spatially normalized trajectories. This procedure had also the advantage that missing data frames within one reaching movement were interpolated (for further details see also, Chapman & Goodale, 2010). As reaching endpoints differed between different stimulus' images (due to different target locations on the screen) within one condition, trajectories had to be rotated to one single reaching endpoint per condition to be able to average reach trajectories. Therefore each trajectory was transformed to the polar coordinate system. For every possible combination of object arrangements and targets, we calculated the mean angle of the last data point of the control conditions for every participant. This value was then subtracted from every angle value of the control condition and any other condition of the corresponding arrangement-target combination, resulting in a rotation of the trajectories of the control condition to the center of the display and a respective rotation of the trajectories from the other conditions. Consequently, the distances and proportions between control trajectories and the trajectories from other conditions remained unaffected. Afterwards the rotated trajectories were converted back to the Cartesian coordinate system. Finally we averaged trajectories over every condition for every participant.

For statistical analysis the preprocessed, normalized and averaged trajectories were entered into four functional-ANOVAs (Ramsey & Silverman, 2005), two for global and two for local conditions including one for right- and one for leftward object shifts. The functional-ANOVA models were single factor designs with 4 levels (control, loc1, loc2, loc3 and control, gloT, gloC, gloP). Functional pairwise comparisons (equivalence to a paired t-test) between the control condition (no object shift) and every experimental condition (with object shift) were conducted post-hoc (one comparison for each shift direction).

2.3 Results

In the present study we investigated whether or not allocentric coding schemes are used when people reach to remembered targets in a natural scene. We manipulated the location of the reference objects by shifting the objects to the left or to the right before reaching. Reference objects were either potential reach targets (local condition) or other objects in the scenes (global condition). First, we sought to quantify eye movement behavior during the encoding phase. Figure 2.3A illustrates the relative frequency of fixations overlaid on an example encoding image (see Methods for details). Clearly, participants visually explored relevant portions of the image, i.e. local object regions where potential reach targets were located. The screen center, the position of the future fixation cross, naturally resulted in the most frequent fixation location (red). Figure 2.3B depicts a typical example of individual reach endpoints for one participant and the applied exclusion criteria towards one target (egg). Clearly, only real outliers were removed.

Figure 2.4 represents the reach endpoints for all participants observed in the local and global conditions. As the overall pattern shows, reach endpoints were influenced by left- and rightward shifts of 3 or 5 reference objects in the local conditions (figure 2.4A) but were hardly affected in the local and global conditions when only 1 reference object was shifted (figure 2.4B). In particular, reach errors were distributed along the horizontal axis and increased with the number of local objects shifted in the scene.

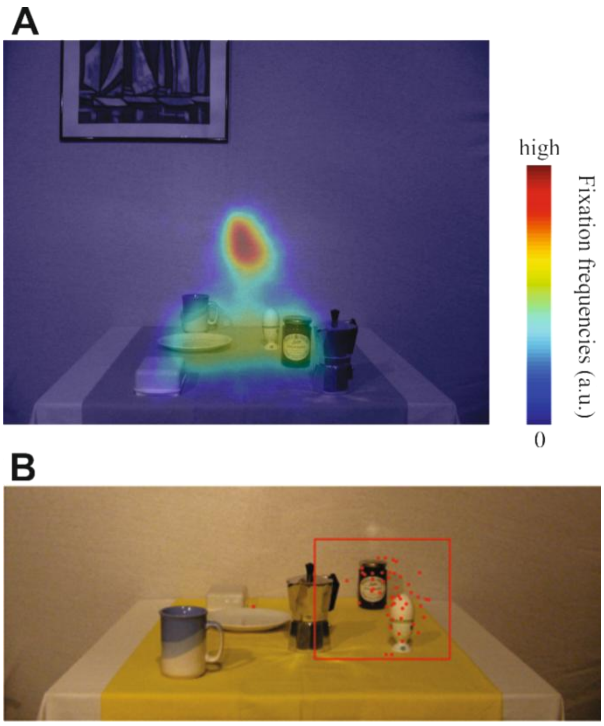

Figure 2.3. (**A**) Heatmap of relative fixation frequency during the encoding phase plotted against an example encoding image. Blue colors denote few or no fixations, whereas a red colors denote many fixations in that region. (**B**) Typical example plot of reaching endpoint towards the egg in one of the 18 configurations. Red dots are reaching endpoints from individual trials (local + global). The red square represents the outlier criterion of 2.5 SD relative to the mean reach endpoint in the control condition. All data within the red square have been considered for data analysis, data points outside have been treated as outliers.

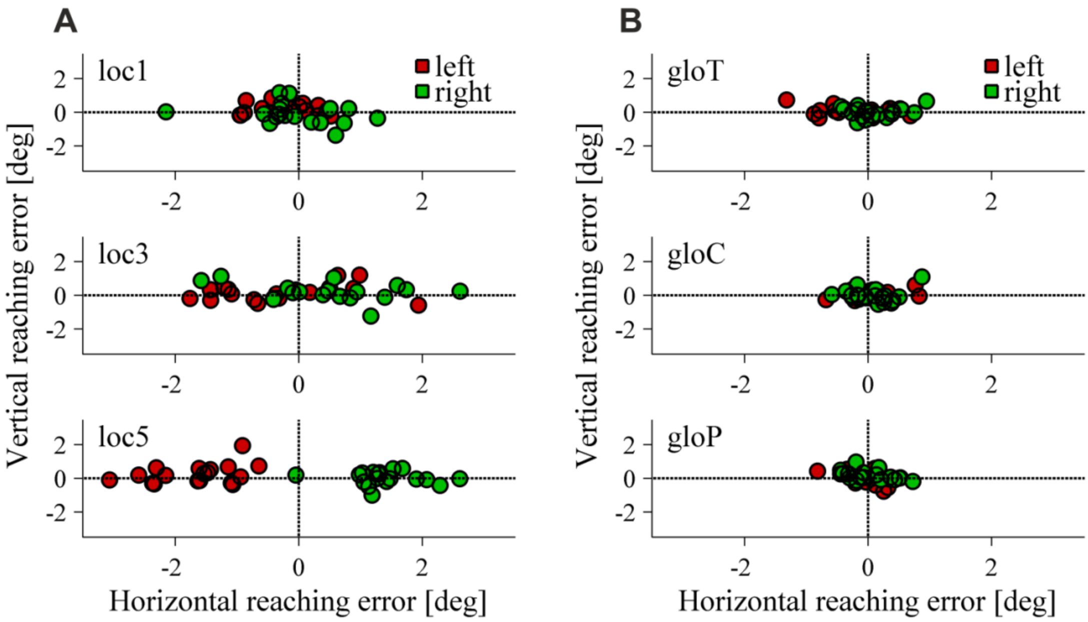

Figure 2.4. Horizontal and vertical reaching error in the local (**A**) and the global conditions (**B**). Each data point represents the average reaching endpoint for one test image across participants. Red colors indicate a leftward, green colors a rightward target shift.

Figure 2.5 displays the observed horizontal reach errors as a function of the predicted allocentric reach errors for each test image. Reach errors varied within the expected direction of the shift of the reference objects in the loc5 and loc3 conditions where 5 or 3 local objects were shifted before the reach. The allocentric weights ranged between 1 % to 43 % in the local conditions and 1 % to 4 % in the global conditions. Table 2.2 summarizes the *mean* (*SD*) reach errors for each individual participant and for loc3 and loc5 conditions separately. A leftward shift of the reference objects resulted in reach endpoints left of the target location and vice versa. This was confirmed by the allocentric weights (= slope of the regression line) which significantly differed from zero in the loc5 ($t(11) = 9.90$, $p < .001$) and the loc3 ($t(11) = 2.43$, $p = .017$) conditions. We found a smaller but non-significant effect for the gloT condition where the table was shifted in the scene ($t(11) = 2.36$, $p = .019$; critical p-value = .0166).

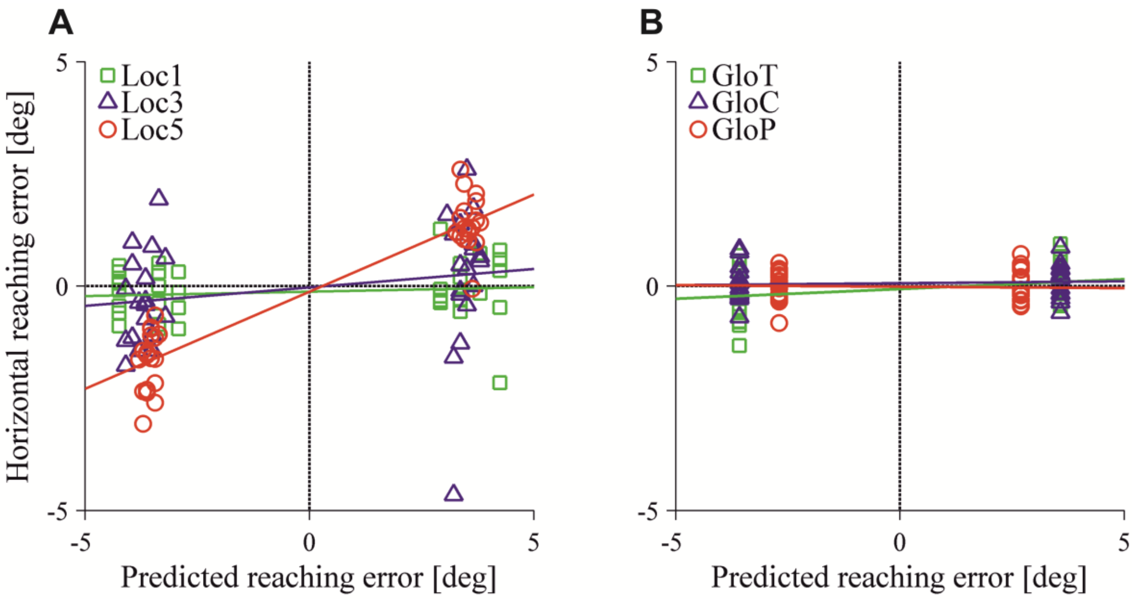

Figure 2.5. Horizontal reaching errors as a function of predicted allocentric reaching errors for the local (**A**) and the global conditions (**B**). Each symbol specifies mean reach endpoints for one test image. Colored lines represent regression fits for each allocentric condition.

Table 2.2: *Mean* (*SD*) reaching endpoints relative to control condition for every participant in the loc3 and loc5 condition, split up by the direction of the allocentric shift. Each data point is based on 36 trials (minus disregarded trials). Negative values indicate a leftward shift relative to control condition. All values are reported in degree visual angle.

	Loc3				**Loc5**			
Partici-pant	**left**		**right**		**left**		**right**	
1	1.16	(4.65)	0.11	(4.91)	-1.78	(4.14)	0.22	(4.56)
2	-1.73	(3.03)	-0.41	(3.38)	-3.09	(1.95)	0.92	(2.46)
3	-1.12	(2.64)	-0.37	(2.19)	-2.89	(1.51)	1.21	(1.37)
4	-1.16	(3.21)	-1.29	(2.78)	-1.94	(1.25)	0.60	(1.55)
5	-2.39	(1.54)	0.29	(2.00)	-2.22	(2.67)	1.32	(1.40)
6	0.09	(2.79)	-0.18	(1.66)	-1.44	(1.11)	0.66	(1.50)
7	-0.75	(1.31)	0.82	(2.04)	-1.27	(0.99)	2.82	(1.17)
8	-0.45	(2.44)	-0.04	(1.22)	-1.09	(1.13)	0.39	(1.69)
9	0.45	(1.38)	0.76	(2.91)	0.02	(1.01)	2.06	(1.47)
10	0.17	(1.52)	1.86	(1.57)	-1.18	(1.10)	2.83	(0.98)
11	-0.59	(1.66)	0.56	(1.67)	-0.60	(0.93)	1.58	(1.33)
12	0.31	(2.41)	0.13	(3.15)	-0.72	(1.80)	0.60	(2.90)

To compare the individual allocentric weights within the allo-local and the allo-global conditions, we computed one-way repeated measures ANOVAs which revealed a significant main effect of condition for allo-local ($F(2,22) = 59.35$, $p < .001$) but no effect for allo-global ($F(2,22) = 2.438$, $p = .111$). Post-hoc t-tests indicated that allocentric weights in the loc5 condition were significantly higher than in the loc3 ($t(11) = 8.935$, $p < .001$) and the loc1 ($t(11) = 9.448$, $p < .001$) conditions. In addition, allocentric weights in the loc3 condition were higher than in the loc1 condition ($t(11) = 2.348$, $p = .019$). Thus, allocentric weights increase with an increasing number of local reference objects shifted in the horizontal plane (figure 2.6).

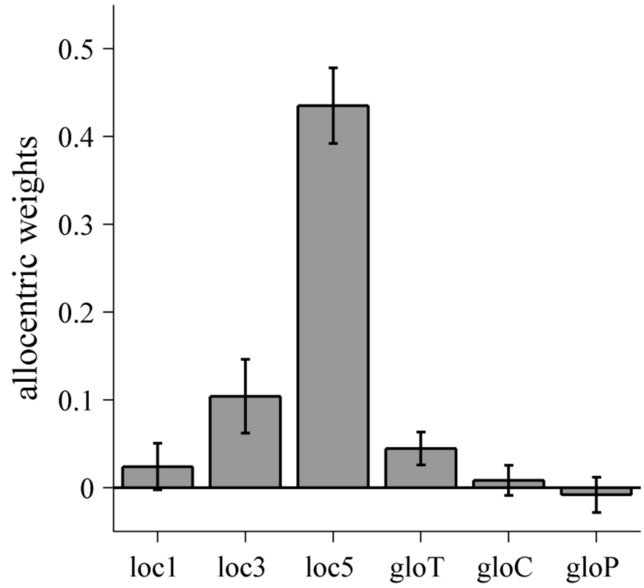

Figure 2.6. Allocentric weights for the allo-local and allo-global conditions. Data are averaged over individual allocentric weights with error bars denoting one standard error of variability between observers. Individual allocentric weights range from -0.15 to 0.18 (loc1), -0.14 to 0.39 (loc3), 0.21 to 0.61 (loc5), -0.06 to 0.14 (gloT), -0.09 to 0.10 (gloC) and -0.16 to 0.08 (gloP).

It has previously been shown that landmarks can influence reach trajectories and that this effect is distance dependent (Diedrichsen et al., 2004). Therefore, we also tested for the effect of proximity in the loc1 and loc3 conditions by correlating the observed reaching error with the mean distance of the shifted object/s with respect to the target. However, we could neither find a correlation for the loc1 (r = -0.09, p = .615) nor for the loc 3 (r = -0.01, p = .962) conditions.

Response latencies of reaches for the allo-local and the allo-global conditions are illustrated in figure 2.7A. Response latencies did not significantly differ between the allo-global conditions (F(3,33) = 0.372, p = .774) but significantly differed between the allo-local conditions (F(3,33) = 14.54, p < .001). In comparison to the control condition, reaches were slower in the loc1 (t(11) = 5.643, p < .001) and the loc3 (t(11) = 6.64, p < .001) conditions. Moreover, reaches in the loc3 condition were also initiated more slowly than in the loc5 condition (t(11) = 3.616, p = .004). Movement times did neither vary between allo-local conditions (F(3,33) = 0.560, p = .645) nor between allo-global conditions (F(3,33) = 0.44, p = .726).

To assess task difficulty, we tested whether the frequency of trials in which participants did not respond differed across all conditions and thus violates the assumption of equal trial distribution across conditions. The results of the Friedman test rejected the assumption that those trials are equally distributed across all conditions (χ^2 = 46.6, p < .001). As depicted in figure 2.7B, participants showed more frequent no reaching responses in the local compared to the global conditions with the highest frequency in the condition where 3 local objects were shifted (loc3).

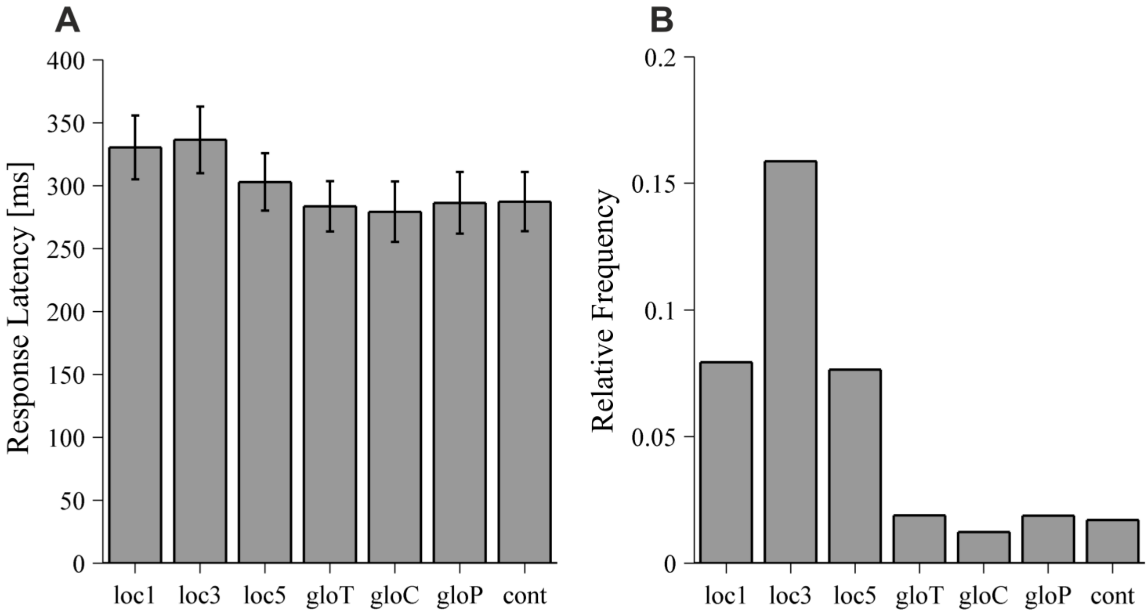

Figure 2.7. (**A**) Response latencies of reaching movements in ms for the local, global and control (cont) conditions. Values are averaged across median response latencies of individual observers. Error bars denote one standard error of variability between observers. (**B**) Relative frequency of trials where participants with no reach response for the local, global and control (cont) conditions. For each condition, the relative frequency is computed as the amount of trials without a reaching response divided by the total amount of trials in that condition.

To examine whether reaching errors due to allocentric object shifts emerged early during the reaching movement (due to different motor plans) or late during the reaching movement (due to error correction mechanisms), we used four functional ANOVAs (one for each experimental condition and shift direction) and functional pairwise comparisons to compare reaching trajectories of different allocentric conditions and the control condition. The functional ANOVAs revealed that trajectories of local object shifts differed in the horizontal plane (x-axis, parallel to the screen). Trajectories for both leftward and rightward shifts started to differ roughly at half-distance (≈ 48.75 % = 11.7 cm) of the reach trajectory (figure 2.8A, significant regions indicated by the grey vertical bars). Functional ANOVAs for global object shifts showed significant differences for leftward shifts starting from roughly the last third (68.3 % = 16.4 cm) up to the end and for rightward shifts just for a small area right after half-distance (from 57.5 % = 13.8 up to 68.3 % = 16.4) of the reaching movement. Subsequent functional pairwise comparisons between every local condition and the control condition for the two shift directions showed that only trajectories in the loc5 condition differed significantly from the control condition. Loc5 trajectories for leftward and for rightward shifts started to differ slightly earlier than half-distance of the reaching movement (leftward: 43.3 % = 10.4 cm; rightward: 48.3 % = 11.6 cm). Differences increased until the end of the movement (figure 2.8A, indicated by the red vertical significance bars). Functional pairwise comparisons for global conditions revealed only a significant difference between gloT and the control condition for leftward object shifts for roughly the last third of the reaching movement (starting from 70.8 % = 17 cm till the end; figure 2.8B, indicated by the blue vertical significance bar).

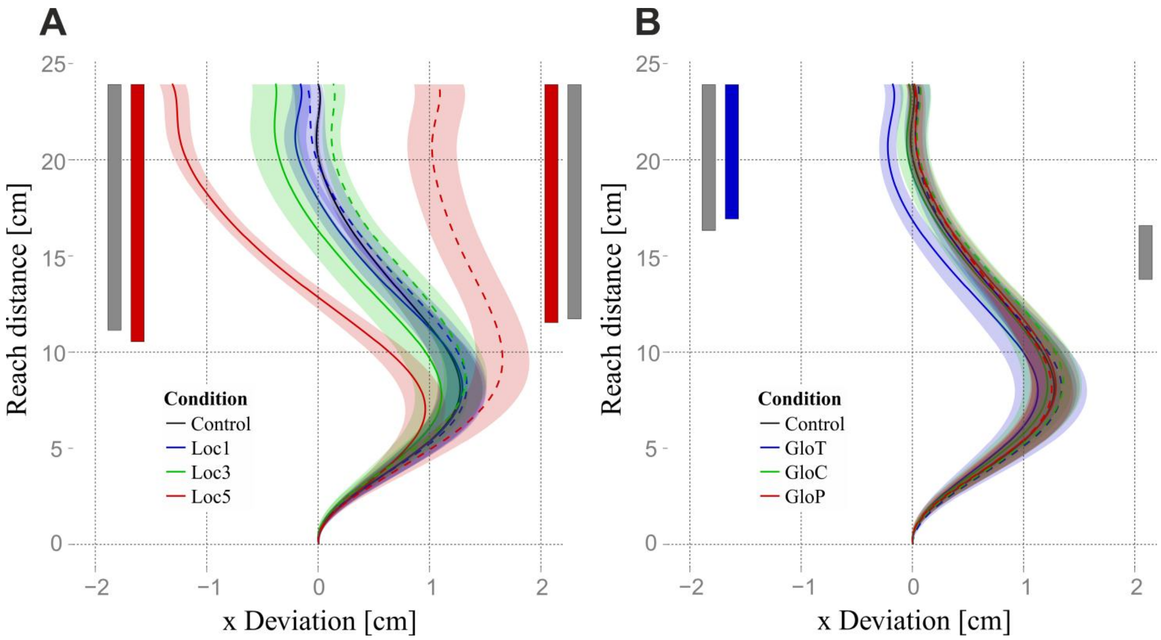

Figure 2.8. Trajectories and results of functional analysis. Mean trajectories of all participants are plotted as the deviation on the x-axis (parallel to the screen) against the reaching distance (axis orthogonal to the screen). Trajectories for leftward object shifts are plotted with solid lines and rightward object shifts with dashed lines. Grey bars indicate the area where trajectories of left- or rightward shifts for local or global conditions showed a significant main effect. (**A**) Mean trajectories for all local conditions and the control condition are displayed. Red bars indicate the area where leftward and rightward shifts of the loc5 condition significantly differed from the control condition. (**B**) Mean trajectories for all global conditions and the control condition. The blue bar indicates the area where leftward shifts in the gloT condition significantly differed from the control condition. Shaded areas in (A) and (B) express one standard error of the mean of the corresponding mean trajectory.

2.4 Discussion

In this study, we investigated the use of egocentric and allocentric information during memory-guided goal-directed reaching using a naturalistic visual scene. Allocentric information was varied by shifting objects on the table (local objects) or objects in the environment (global objects) leftwards or rightwards after scene encoding and before reaching. Memory-guided reaching movements were performed without visual information about the scene while gaze and body position remained fixed. We predicted accurate reaching movements if participants relied only on egocentric object coding, i.e. representing the target relative to gaze or body position, and systematic shifts of reach endpoints if they used allocentric cues (local or global) for goal-directed reaching. Our results demonstrated that reach endpoints varied as a function of objects shifted in the scene. The more local objects were horizontally misplaced the larger were the reach errors in the direction of the objects shifted. The present findings suggest that allocentric cues are indeed used during goal-directed reaching, but only if a substantial change of allocentric information is present in complex visual scenes.

Previous studies consistently reported that reach targets are represented relative to gaze direction, i.e. in an egocentric frame of reference (e.g., Henriques et al., 1998; Medendorp & Crawford, 2002). Beyond egocentric coding, allocentric cues also contribute to reaching movements as has been demonstrated in studies using visual landmarks (Byrne & Crawford, 2010; Byrne et al., 2010; Obhi & Goodale, 2005), imagined landmarks (Carrozzo et al., 2002) or structured visual backgrounds (Krigolson & Heath, 2004; Krigolson et al., 2007). While these studies examined reaching movements in rather unnatural tasks using isolated visual targets presented together with abstract, task-irrelevant landmarks, here we studied reaching behavior with more naturalistic stimuli by using photographic images of a breakfast scene. Despite the stable and reliable egocentric information of body and gaze position, we found large effects of allocentric cues on reach endpoints in line with the previous findings based on less ecologically valid experimental tasks (e.g., Byrne & Crawford, 2010). Since the target was defined as the missing local object in the shifted target scene, object shifts seem to be incorporated into the memory representation of the target established during scene encoding resulting in a combined representation which is used for calculating the reach plan. This is supported by the reaching trajectories in the object shift condition (loc5) which started to deviate from the no-shift condition early after reach onset. In sum, our results suggest that allocentric cues are even effective if they are provided after target encoding.

The present results demonstrated that the number of local objects shifted in the scene systematically affected reaching movements. We found larger distortions of reach endpoints with an increasing number of local objects shifted in the scene. Reach errors were most pronounced when all remaining local objects (loc5) were shifted, intermediate when 3 local objects (loc3) were shifted and absent for shifts of one local object (loc1). This result implies that substantial changes of allocentric cues in complex visual scenes are required to influence reaching movements. It is important to note that after object shifts the spatial relations between the objects in the loc5 condition remained constant while they completely changed in the loc3 condition. This resulted in a higher number of no-response trials and slower response latencies in the loc3 condition indicating higher task difficulty. Nevertheless, allocentric coding was still present in the loc3 condition, but the effect was diminished compared to the loc5 condition. Based on the present data, we cannot disentangle whether the reduced effect of allocentric coding is caused by larger task difficulty or fewer changes in the scene image. Previous findings on the Roelofs effect argue for the latter factor showing that the amount of a perceived target displacement when the whole frame around the target was shifted equaled the sum of a perceived target shift when only parts of the frame were shifted (Walter & Dassonville, 2006). Accordingly, we observed that allocentric weights were highest in conditions, when five local objects were shifted and lowest, when only one object was moved with the weights of three shifted local objects in between. We exclude a potential effect of proximity of target and allocentric cues on reach endpoints (c.f., Diedrichsen et al., 2004) because local object shifts appeared in the immediate vicinity of the target. Thus, we suggest that in a realistic visual environment it is the number of changed allocentric cues rather than distance that determines integration weight.

Local objects might also function as potential obstacles in real world situations which are especially important for movement programming. Obstacles constitute spatial constraints on movement execution and thus are not considered as distractors but rather as task-relevant non-target information (Tresilian, 1998) which is represented together with the target information in the attention system (Baldauf & Deubel, 2010; Tipper et al., 1997). As a consequence, the presence of obstacles requires additional anticipatory processing of movements leading to slower movement initiation (Biegstraaten et al., 2003; Sailing et al., 1998). Accordingly, we observed longer response latencies when local objects (loc1 and loc3) were shifted, but not for global object shifts.

The absence of an influence of global allocentric cues on reach endpoints can be explained by multiple factors. First, the changes of global objects in the scene were undersized due to only one global object being shifted (instead of multiple as in the local conditions). Therefore, global conditions might be more similar to the loc1 condition. One can speculate that an increase in the number of shifted global objects might lead to similar results as we observed for the local object shifts. Second, it is also possible that it is the object displacement relative to object size that plays a role, in which case smaller objects should have larger influences on allocentric coding. Third, we cannot entirely rule out that the visibility of the frame of the presentation screen throughout the experiment has acted as a strong global allocentric cue. Since the screen never moved but the frame of the screen was a very salient visual feature (i.e. high contrast), it might have overridden more subtle global allocentric cues within the images. Fourth, local and global objects differ in task relevance, in the way that local objects represented potential reach targets in contrast to global objects which never served this function. This information was given by task instruction and thus may have influenced strategic behavior. Task relevance has been shown to affect overt attention in naturalistic tasks resulting in more fixations on task-relevant than task-irrelevant objects (Ballard & Hayhoe, 2009; Land & Hayhoe, 2001). These findings are consistent with the fixation behavior we observed during the encoding phase which was spatially restricted to locations of the local objects. Fixations also frequently occurred at the table/table cloth placed right underneath the local objects; however, these global objects did not affect reaching behavior. In support of this finding, previous work demonstrated that object features which are task-irrelevant are not attended even if the respective object is fixated (Triesch et al., 2003). Together with the fact that working memory capacity for spatial information is limited to up to 4 items (Luck & Vogel, 1997) and retention of task-relevant objects is prioritized (Maxcey-Richard & Hollingworth, 2013), it is conceivable that participants encoded the location of local objects, i.e. task-relevant information, which were then incorporated into the reach plan while ignoring the location of the global objects, i.e. task-irrelevant information in the environment. Whether or not task relevance of allocentric information is a central factor in reach planning should be examined in future studies. Finally, the global allocentric cues lacked of a causal relationship to the reach target as discussed in the next paragraph.

We believe that our findings can be explained in the framework of causal Bayesian integration (Körding & Tenenbaum, 2007; Körding et al., 2007). The gradual increase of allocentric cue effects with the number of shifted local objects is consistent with more reliable allocentric cue information when more local objects are shifted. In that sense, the more local objects are shifted, the smaller the variance associated with allocentric information and thus the higher the allocentric weight in the integration of egocentric (probably body and gaze) and allocentric position. But how does this explain the absence of global allocentric cue effects? We believe that the concept of causality in Bayesian integration might be a key in understanding this. First, one can argue that there is no real causal link between the global objects and the local objects, as the picture frame is totally task-irrelevant and the exact position of the table and table cloth are not important, unless local objects had been positioned at the edge (and could thus fall off), which was not the case. Second, the spatial extend of the table and table cloth might have simply resulted in less precise positional information due to their large spatial extent. Third, and maybe more importantly, when the table cloth or table moved, local objects stayed fixed in space (i.e. did not move with the table and table cloth). Thus, the causal link between table/cloth and local objects on the table was broken, since normally objects would move with the table/cloth. In that case, causal Bayesian integration discounts any global allocentric cue effects due to a lack of a causal relationship between table/cloth movement and target location.

Our observations that movement endpoints are systematically shifted by local allocentric cues could result from two different sources: reach trajectory planning (Burns & Blohm, 2010) or feedback-based control processes (Krigolson et al., 2007). Indeed, allocentric information could be included in the reach plan right from the start as is the case in visual-proprioceptive integration (Burns & Blohm, 2010; Sober & Sabes, 2003, 2005), in which case one would expect manifestations of allocentric influences on the reach plan early on in the reach trajectory. Alternatively, allocentric information could only be incorporated during feedback corrective processes (i.e. later on in the movement), which would be consistent with observations of allocentric visual background influences on reaches (Krigolson et al., 2007). Our data on reaching trajectories is consistent with the former hypothesis and shows that local allocentric information might influence reach planning differently than allocentric background information.

In the present study we examined egocentric-allocentric cue integration for memory-guided (not visually-guided) reaches. Reaches were performed immediately after the presentation of the test scene; a condition which is usually defined as immediate reaching (cf. Bridgeman, et al., 2000; Hay & Redon, 2006). However, here we asked participants to reach to the missing object in the test scene which required to build up representations of potential reach targets during the encoding scene which were then updated on the basis of the test scene after a 2s delay. Delay is believed to have an important influence on spatial coding. For example, Hay and Redon (2006) found that delayed reaching accuracy declined in darkness but remained constant when a structured visual background was available. They explain their findings with a decaying egocentric representation and a more permanent allocentric representation of target location. This is also consistent with observations that visual landmarks increase space constancy (Deubel et al., 2010) and decrease egocentric, gaze-dependent coding of reach targets (Schütz et al., 2013). Furthermore, allocentric information has a stronger impact on delayed than immediate reaches showing increased reach errors in the direction of a shifted landmark with longer delays between stimulus offset and motor response (Bridgeman, et al., 1997, 2000). An interesting prediction from these findings is that shorter (resp. longer) delays should lead to lower (resp. higher) allocentric weights because egocentric information is initially more accurate but decays faster than allocentric information.

Overall, we have shown that allocentric information is used by the brain to plan memory-guided reaches towards targets in naturalistic visual images. Our data is generally consistent with Bayesian causality principles and demonstrates that egocentric-allocentric cue integration is highly flexible and task-dependent. It would be interesting to further examine the role of causality in egocentric-allocentric cue integration, in particular with respect to the causal relationship between visual landmarks.

3 Contextual factors determine the use of allocentric information for reaching in a naturalistic scene

A similar version of this manuscript has been published as:
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, 1–13, doi: 10.1167/15.13.24.

Numerous studies have demonstrated that humans incorporate allocentric information when reaching towards visual targets. So far, it is unclear how this information is integrated into the movement plan when multiple allocentric cues are available. In this study we investigated whether and how the extent of spatial changes and the task-relevance of allocentric cues influence reach behavior. To this end, we conducted 2 experiments where we presented participants 3D-rendered images of a naturalistic breakfast scene on a computer screen. The breakfast scene included multiple objects (= allocentric cues) with a subset of objects functioning as potential reach targets; i.e., they were task-relevant. Participants freely viewed the scene and after a short delay, the scene reappeared with 1 object missing (= target) and other objects being shifted left- or rightwards. Afterwards, participants were asked to reach towards the target position on a gray screen while fixating the screen center. We found systematic deviations of reach endpoints in the direction of object shifts which varied with the number of objects shifted, but only if these objects served as potential reach targets. Our results suggest that the integration of allocentric information into the reach plan is determined by contextual factors, in particular by the extent of spatial cue changes and the task-relevance of allocentric cues.

3.1 General Introduction

When performing reaching movements to a remembered visual target, the brain uses two classes of reference frames, the ego- and the allocentric reference frame (Colby, 1998; Battaglia-Mayer, Caminiti, Lacquianiti, & Zago, 2003). While a vast number of studies have shown that egocentric reference frames are used to plan and execute goal-directed reaching movements (e.g. Cohen & Anderson, 2002; Lacquaniti & Caminiti, 1998; Thompson & Henriques, 2011), an increasing number of studies also demonstrated an additional influence of allocentric information on reaching behavior arguing for the integration of multiple reference frames (e.g. Byrne & Crawford, 2010; Diedrichsen, Werner, Schmidt, & Trommershäuser, 2004; Krigolson & Heath, 2004; Krigolson, Clark, Heath, & Binsted, 2007; Obhi & Goodale, 2005; Schütz, Henriques, & Fiehler, 2013, 2015), even in conditions when allocentric cues are covert and fall along a virtual line (Carrozzo, Stratta, McIntyre, & Lacquaniti, 2002). Byrne and Crawford (2010) further examined how egocentric and allocentric information are combined for reaching. They asked participants to reach for a remembered, visual target (light dot) which was either presented in isolation or was surrounded by four landmarks. After this display disappeared, only the landmarks reappeared but were shifted in space. Even though participants were instructed to ignore the landmarks (using solely egocentric information), averaged reaching endpoints deviated systematically in the direction of the landmark shift. This result shows that allocentric information was integrated into the reach plan and influenced reaching endpoints.

Overall, the scope of the previous studies is rather limited. First, these studies mainly investigated if allocentric information (in addition to egocentric information) contributes at all to reaching behavior but hardly examined how allocentric information is used if multiple cues are available in the environment. Second, previous studies presented rather isolated and abstract stimuli on a blank screen which had little relevance for the task. Third, task instructions usually prompted participants to apply either an egocentric or an allocentric reference frame, biasing the participants’ behavior in one direction. In this study, we aim to investigate whether and how different allocentric cues contribute to reaching movements depending on (i) the change and (ii) the task-relevance of allocentric cues in a naturalistic everyday scene. Therefore, we utilized a memory-guided reaching task which did not bias participants to use either an egocentric or an allocentric reference frame.

A recent study (Fiehler, Wolf, Klinghammer, & Blohm, 2014) provided first answers to the question which factors influence the integration of allocentric information for planning and executing reaching movements. In the experiment, participants freely explored an image of a breakfast scene presented on a computer screen, which contained multiple objects on a table (*table objects*) and in the background (*background objects*).

After scene encoding and a short delay, a test scene reappeared for 1 sec with 1 table object missing indicating the reach target. Moreover, 1, 3, or 5 of the remaining table objects or 1 background object was shifted either to the left or to the right. The test scene was followed by a blank screen asking participants to perform a memory-guided reaching movement toward the target position (i.e. the position of the missing object). The main finding was that reach endpoints deviated in the direction of the object shifts and scaled with the number of objects shifted, but only if table objects were shifted, whereas a shift of background objects led to no deviations of reach endpoints. The authors concluded that allocentric information contributes to reaching movements to targets in naturalistic scenes, but this contribution depends on the extent of changes in the scene.

However, the study by Fiehler et al. (2014) cannot explain the lack of an effect when background objects were shifted. Finding an explanation for this observation could reveal factors which are important to understand how allocentric information is integrated into the movement plan. The goal of the present study was to test two possible explanations. First, we examined if a change of allocentric information in a scene must exceed a certain extent in order to influence reaching behavior. Given the fact that a shift of at least 3 table objects was necessary to produce a significant deviation of reach endpoints in the previous study (Fiehler et al., 2014), it is conceivable that 1 shifted background object was not sufficient to cause any behavioral effect. Second, we investigated if only allocentric information relevant for the task contributes to reaching movements. In the previous study (Fiehler et al., 2014), table objects served as potential reach targets and thus were highly relevant to successfully perform the task whereas background objects never served as reach targets. To this end, we performed two experiments in which we manipulated the number of shifted objects (Experiment 1) and the task-relevance of shifted objects (Experiment 2) in the naturalistic scene.

3.2 Experiment 1

3.2.1 Introduction

In order to test whether a certain amount of allocentric cue change in a scene is necessary to influence reach behavior, we applied a similar task as in the study by Fiehler et al. (2014), but this time we systematically varied the number of shifted background objects (instead of table objects). To this end, we added background objects and shifted either 1, 3 or 5 background objects left- or rightwards, comparable to the table object shifts applied in the previous study. Again, table objects served as potential reach targets. If a minimum number of shifted objects in the scene is needed to cause a substantial effect on reach endpoints, we expect an increase in the deviation of reach endpoints in the direction of the background object shifts, as was found for table object shifts before (Fiehler et al., 2014). Moreover, the amount of reach endpoint displacements should be similar for shifts of background and table objects.

3.2.2 Methods

Participants

We recorded data from 15 participants. Two participants were rejected due to fixation behavior (< 50 % trials with correct fixation) and one participant was rejected because 63 % of trials were classified as unsure response or no response was given. Thus, the final sample consisted of 12 participants (5 female) with normal or corrected to normal vision. Participants were right-handed as assessed by the Edinburgh handedness inventory (EHI, Oldfield, 1971; *mean* 95.73 ± 12.4) and ranged in age from 20 to 31 years (*mean* 24.4 ± 2.9 years). They received course credit or were paid for their participation. The experiment was conducted in agreement with the ethical guidelines of the local ethics committee of the University of Giessen in compliance with the Declaration of Helsinki (2008).

Materials

Stimuli consisted of 3D-rendered images of a breakfast scene. Images were created in SketchUp Make 2013 (Trimble Navigation Ltd.) and afterwards rendered with Indigo Renderer 3.6.26 (Glare Technologies Ltd.) with a resolution of 3562 x 2671 pixels. The breakfast scene contained 6 *table objects* (TO) consisting of a coffee mug, a plate, an espresso cooker, a Vegemite jar, a butter dish, and an egg cup on a brown table that was placed 90 cm in front of a gray wall. Furthermore, 5 *background objects* (BO), consisting of a table, a plant, a chair, a floor lamp, and a painting on the wall, were located in the surrounding area. Objects were taken from the open access online 3D-gallery of SketchUp. Object properties are summarized in table 3.1. We set the 6 TO in 40 different arrangements whereas the BO always appeared at the same position (*encoding image*).

The TO in any arrangement were placed at one of three possible horizontal depth lines that were equally spaced (19.5 cm starting from the front table edge) on the table following three criteria: (i) minimal 1 and maximal 3 objects were positioned at every depth line, (ii) objects were placed with a distance to the edges of the table so that in case of object displacement it never stood right at the table edge or in the air next to the table, and (iii) < 20 % of an object was occluded by another object. Based on the encoding images, we created *test images*, in which 1 of the TO was missing (= *reach target*). In 7/9 of the test images, other objects (TO and/or BO) in the scene were shifted horizontally between 3.08 deg and 4.08 deg (*mean* 3.61 deg ± 0.37 deg) either to the left or to the right (50 % leftward displacement). Variations in the horizontal displacement arose from the fact that objects were placed at different depth lines relative to the virtual camera position. Hence, similar physical shifts of objects at different depth lines in 3D-space would result in different displacements in the 2D-image. In the remaining 2/9 of the test images, no objects were shifted. These images served as control condition.

Table 3.1. Maximum height, width and distance to camera of objects in the scene in cm, based on the actual properties in SketchUp. Table objects had no fixed distance to the camera as they were randomly placed on one of three different depth lines. However, the reported size relates to their absolute values in SketchUp. Some background objects were not fully visible due to an overlap with other background objects or partial cutting by the image borders. In that case, the absolute size of the actually visible object part is reported here.

Object	**Height (visible)**	**Width**	**Distance to camera**
Plate	1.96	19.27	variable
Butter dish	4.91	8.40	variable
Egg	7.45	4.92	variable
Espresso cooker	15.10	8.47	variable
Vegemite jar	11.44	6.72	variable
Mug	9.62	7.90	variable
Table	8.48	78.00	154.00
Plant	51.28	37.52	212.50
Painting	25.63	42.75	232.52
Chair	15.40	30.48	193.50
Lamp	54.40	24.53	212.50

In total, 1120 images were created, including 40 encoding images, 840 test images (120 with TO shifts, 720 with BO or BO and TO shifts) and 240 control images. Furthermore, from each of the 40 encoding images, a scrambled version of the image made up of 768 randomly arranged squares was created and used for masking of the encoding image.

Apparatus

Stimuli were presented on a 19” (40.5 cm x 30 cm) CRT monitor (Ilyama MA203DT) with a resolution of 1280 x 960 pixels and a refresh rate of 85 Hz. To reduce the influence of a high-contrast frame around the scene, a black cardboard (70 cm x 50 cm) frame was attached to the monitor. Participants sat at a table with their head stabilized on a chin rest with a distance of roughly 47 cm from the eyes to the center of the screen. A decimal-keyboard was placed in front of the participants with the start button aligned to the chin rest and the center of the screen with a distance of 24 cm from the screen. Reaches were performed with the right index finger and recorded with an Optotrak Certus (NDI, Waterloo, Ontario, Canada) tracking system with a sampling rate of 150 Hz using one infrared marker attached to the fingertip of the right index finger. To control for correct fixation behavior, eye movements were recorded using an EyeLink II system (SR Research, Osgoode, Ontario, Canada) with a sampling rate of 500 Hz. To run the experiment and to control the devices we used Presentation 16.5 (Neurobehavioral Systems, Inc.).

Procedure

The procedure of an example trial is depicted in figure 3.1. At the onset of every trial, a fixation cross on a gray screen appeared prompting participants to fixate and press a button in order to perform a drift correction for the EyeLink II. Thereafter, the encoding image of the breakfast scene was presented. participants freely explored the scene without any time constraints and terminated the encoding phase by pressing the start button. Then, a scrambled version of the encoding scene appeared for 200 ms to avoid afterimages followed by a delay phase of 1800 ms with a gray screen and a central fixation cross. Participants were instructed to fixate the cross until the end of the reaching movement in order to provide a stable egocentric reference and to reduce the inter-subject variability of reach endpoints due to gaze shifts during the delay period (cf., Byrne & Crawford, 2010). After the delay, the test image which lacked 1 TO was presented for 1000 ms. The trial continued with a short tone which signaled the participants to perform the reaching movement towards the remembered location of the target object onto a gray screen. Thus, reaches were performed with gaze kept on the fixation cross and without any visual information of the encoding or test images. In this way we ensured that allocentric information could not be used for subsequent online corrections during the reaching movement, which would have led to an allocentric bias.

Participants were instructed to reach to the location of the missing object as accurately as possible; we did not instruct them to reach to a specific part of an object. However, reach endpoints in the control condition were very consistent irrespective of the target object width (standard deviation of reach endpoints, x-axis: 1.38-1.64 cm, y-axis: 0.29-0.30 cm), suggesting a uniform reaching behavior across all objects. Whenever they were unsure about the target location or identity, they had to reach to a marked location at the lower right edge of the monitor. These invalid trials were repeated at the end of the experiment. If participants released the button before the go-signal, they received feedback and these invalid trials were also repeated at the end of the experiment.

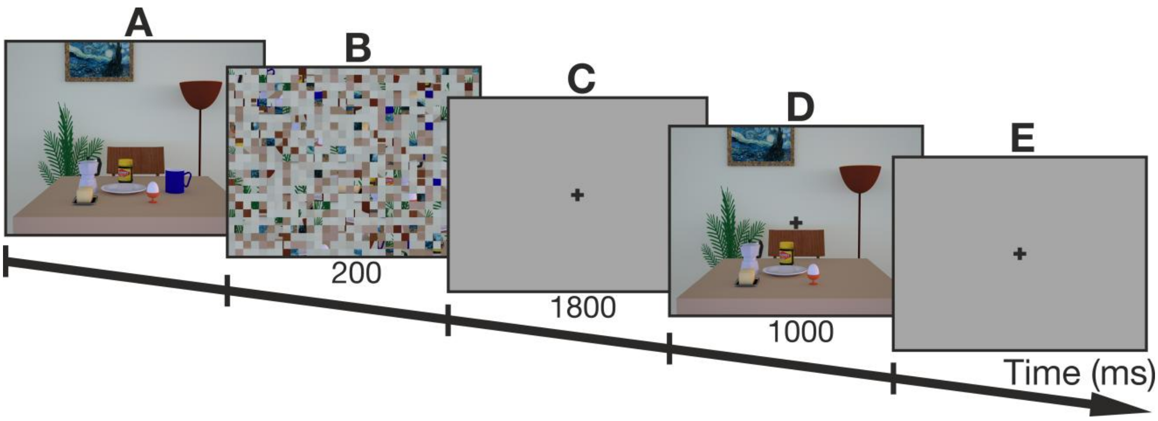

Figure 3.1. Trial scheme of one example trial (control condition). (**A**) First, the encoding image was presented and participants terminated the exploration of the image by button press. (**B**) Then, a scrambled version of the encoding image was presented for 200 ms, followed by (**C**) a delay which lasted for 1800 ms. (**D**) Thereafter, the test image was presented for 1000 ms before (**E**) a tone prompted participants to reach to the target onto a gray screen while fixating the cross at the center of the screen.

Participants performed five experimental conditions (for examples, see figure 3.2). In all experimental conditions, 1 of 6 TO was always removed from the test image, which served as the reach target. In the *TO-5* condition, the remaining 5 TO were shifted either to the left or to the right. In the *BO* conditions, 1, 3 or all 5 BO were shifted left- or rightward (*BO-1*, *BO-3*, *BO-5*). In the *control* condition, all objects remained stationary. In all conditions, left- and rightward object shifts were balanced with 50 % of trials in each direction. There were two additional conditions, in which the BO and the TO were shifted in the same or in the opposite direction. These conditions, however, are not relevant for answering the research question of this article and will be not presented here.

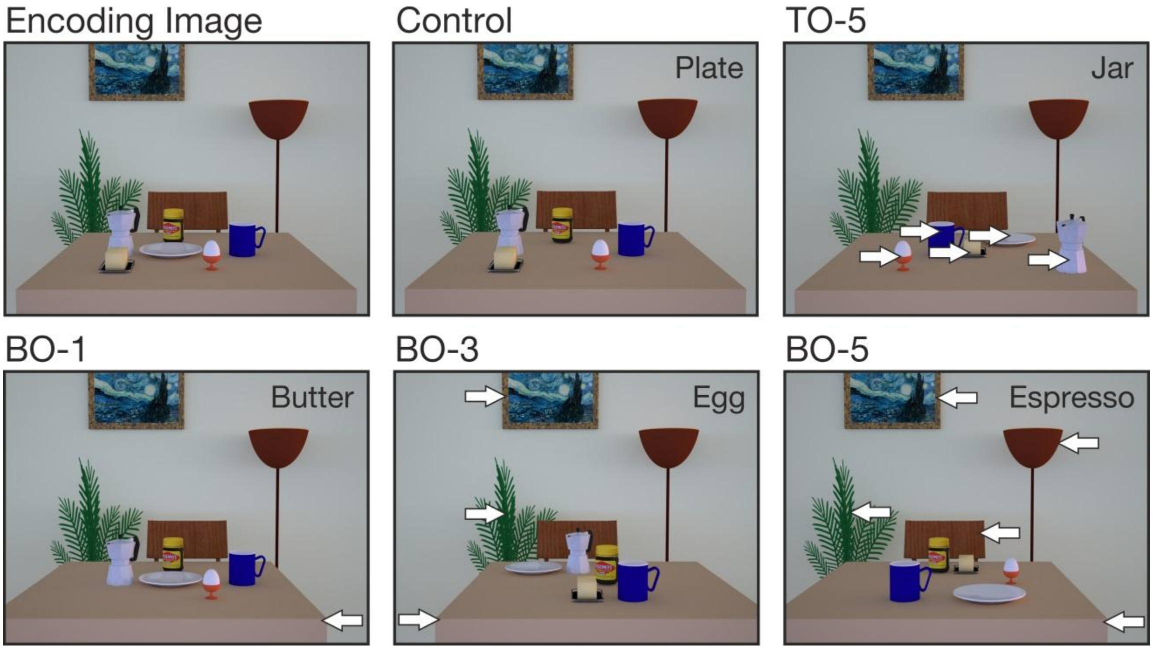

Figure 3.2. Examples of encoding and test images for different conditions. Object names in the box indicate the reach target (= missing object on the table).

Each participant completed a minimum of 1320 trials. Because some trials were repeated (criteria see above), the actual number of performed trials varied from 1322 to 1435 trials. Trials were separated in five sessions with one session per day which lasted about one hour with one break in between. Trials were presented in pseudo-randomized order with a random sequence of conditions and encoding images within a session but fixed trial combinations between sessions. A trial was never followed by a trial containing the same encoding image.

Every test image was presented once (except the repeated trials due to unsure or too early responses, see above), apart from test images of the control condition, which were presented twice because of their importance for calculating reaching errors (see data reduction).

Data Reduction and Statistical Analysis

Data preprocessing was performed with MATLAB R2007b (The MathWorks Inc.) and inferential statistics with R 3.1 (The R Foundation of Statistical Computing). All statistical tests were computed with an alpha-level of .05. If correction for multiple testing was necessary, Bonferroni-Holm correction was applied. In case the assumption of sphericity for an ANOVA was violated (tested with Mauchly's sphericity test), Greenhouse-Geisser correction was applied.

First, we inspected the eye tracking data and discarded trials from further data analysis in which participants' gaze deviated more than 2.5° from the center of the fixation cross during a period beginning from delay onset till the end of the reaching movement. All in all 597 trials (3.76 %) were rejected due to bad fixation. Second, reaching onsets and offsets were defined for every trial. The moment participants released the response button determined the reaching onset. Reach offsets were calculated from Optotrak data and defined as the first time point during the movement when velocity dropped below 20 mm/s if the index finger reached a maximum distance of about 3 cm from the screen. Reach endpoints were extracted at the time of reach offset. Some trials were excluded because reaching offsets or endpoints could not be extracted due to rarely occurring interferences of the infrared markers of the Optotrak with the IREDs of the EyeLink II (198 trials = 1.3 %). Third, we excluded trials in which reaching endpoints deviated more than ± 2.5 *SD* in horizontal or vertical direction from the group mean in each condition for each object shift direction (291 trials = 1.93 %). Taken together, from originally 15840 trials of all participants, 14755 valid trials (93.14 %) remained. From these trials, 10828 trials (= 93.99 % of 11520 trials) belonged to the conditions of interest (control, TO-5, BO-1, BO-3, BO-5) and were entered into further analysis. Reported results refer only to these conditions.

To investigate the influence of object shifts (i.e., allocentric information) on reaching endpoints, we calculated allocentric weights for every participant and every condition by linear regression fit. First, we determined reaching errors as the horizontal distance of the reach endpoint and the actual target position of the encoding image. Therefore, we averaged reach endpoints of the control condition of all participants for every combination of object arrangements and target objects separately. Since none of the objects were displaced in the control condition, no systematic reaching errors should have occurred. These averaged reach endpoints were used as actual target positions. Then, we calculated the differences between reaching endpoints of the other experimental conditions from the corresponding target position in the horizontal plane. This resulted in positive values for misestimations to the right and negative values for misestimations to the left. Moreover, we determined *maximal expected reaching errors* (MERE) for every image after an object shift by assuming that participants completely relied on the allocentric information of shifted objects and thus produced reaching endpoints equal to the amount of the objects' displacement. To this end, we averaged the amount of displacement of the shifted objects for every image. If only 1 object was shifted, only this was taken into account. In case that more than 1 object were shifted, the different displacements were averaged. Please note that the perceived object shift depended on the object's location in depth. For the regression fit, the MERE was used as a predictor and the actual reaching error as a dependent variable for the two shift directions within one condition for every participant. The resulting slope of the regression line indicated the extent to which a participant relied on the allocentric information of object displacements and thus was used as allocentric weight for further analysis.

We performed two-sided one-sampled t-tests to investigate whether group allocentric weights for the different conditions differed significantly from zero. Since reaching errors and thus allocentric weights were computed on the basis of the results of the control condition, a test of weights against zero corresponds to a statistical comparison to the control condition. In order to assess the impact of allocentric information of BO on reaching endpoints, we conducted a one-way repeated measures ANOVA for conditions with BO shifts with the factor number of shifted BO and three levels (BO-1, BO-3, BO-5). In case of significant main effects, we conducted two-sided post-hoc t-tests for paired samples.

In order to investigate a potential influence of selective overt visual attention for task-relevant objects on the observed pattern of reach endpoints, we computed a heatmap of fixation densities of the encoding phase. To this end, we calculated the mean fixation point for every fixation starting from the second fixation until the end of the encoding phase for each participant in all conditions. Then, we collapsed the mean fixation points in one heatmap to generate an overview about participants' average fixation behavior. For a better comparison of the fixation behavior in Experiment 1 and 2, we present the heatmaps of both experiments together in the Results section of Experiment 2 (see figure 3.8).

3.2.3 Results

As illustrated in figure 3.3, reaching errors in condition TO-5 deviated in the direction of table object shifts, whereas no systematic reaching errors occurred when 1 (BO-1), 3 (BO-3) or 5 (BO-5) background objects were shifted. Figure 3.4A depicts the actual reaching errors and the corresponding MERE of one prototypical participant for the conditions TO-5 and BO-5 for leftward (negative values) and rightward (positive values) object displacements. The slope of the regression line defined the allocentric weight of the respective condition. Allocentric weights of TO shifts significantly differed from zero ($t(11) = 7.075$, $p < .001$) and were substantially higher compared to BO shifts irrespective of the number of shifted objects (figure 3.4B; table 3.2).

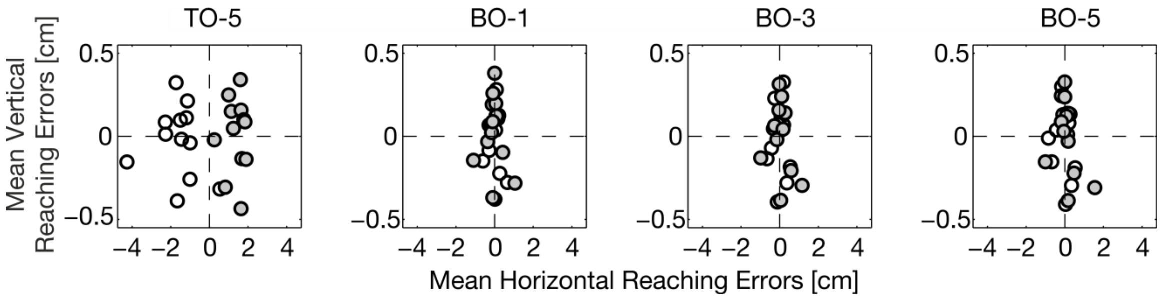

Figure 3.3. Mean horizontal and vertical reaching errors (in cm) of every participant for each condition. Leftward object shifts are depicted in white, rightward objects shifts in gray.

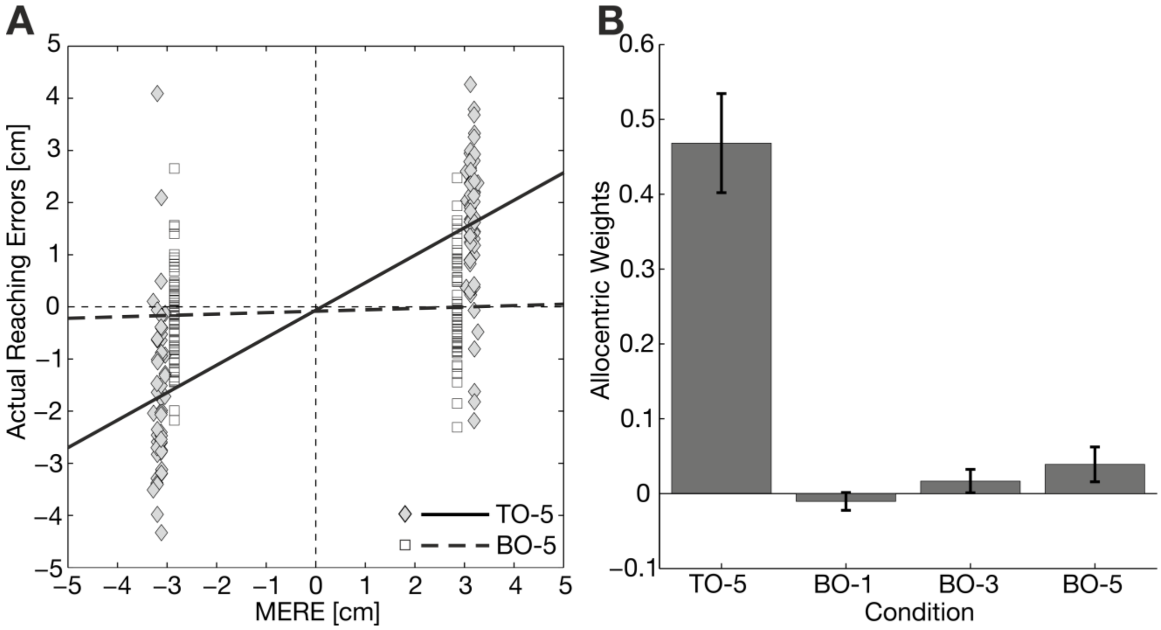

Figure 3.4. (**A**) Example reaching errors of one participant for the condition with shifts of 5 table objects (TO-5) and 5 background objects (BO-5). Actual reaching errors are plotted against the maximal expected reaching error (MERE). Moreover, the linear fit (slope = allocentric weight) is depicted. (**B**) Allocentric weights for every condition averaged over all participants. Error bars indicate 1 *SEM*.

To assess the effect of an increasing number of BO shifted on reaching endpoints, we conducted a one-way repeated measure ANOVA. We found a main effect for the number of shifted BO ($F(2,22) = 6.121$, $p = .008$). Post-hoc t-tests indicated that allocentric weights of BO-1 were smaller than of BO-3 ($t(11) = -2.913$, $p = .042$) and of BO-5 ($t(11) = -2.895$, $p = .042$), whereas allocentric weights of BO-3 and BO-5 did not differ ($t(11) = -1.495$, $p = 0.163$). However, the effect size was very small ($\eta^2 = 0.11$) indicating a rather weak influence of BO shifts on reaching endpoints.

Table 3.2. Summary of allocentric weights for all conditions. Range, mean and standard deviation of the sample are listed. Results of two-sided one-sampled t-tests are Bonferroni-Holm corrected.

Condition	Range	Mean	SD	t-test results
TO-5	0.04 – 0.94	0.47	0.23	$t(11) = 7.075$, $p < .001$*
BO-1	-0.08 – 0.07	-0.01	0.04	$t(11) = -0.879$, $p = .616$
BO-3	-0.06 – 0.14	0.02	0.05	$t(11) = 1.069$, $p = .616$
BO-5	-0.06 – 0.21	0.04	0.08	$t(11) = 1.679$, $p = .364$

3.2.4 Discussion

In this Experiment, we tested whether a decreased number of background objects shifted in the scene (1 background objects vs. 3 or 5 table objects) could account for the lacking influence of background object shifts on reaching endpoints in the study by Fiehler et al. (2014). Therefore, we manipulated the allocentric information by shifting 1, 3 or 5 background objects left- or rightwards after scene encoding and before participants performed a reaching movement to the remembered target position. No visual information about the scene was given during the reaching movement and gaze and body position remained fixed. If a minimum extent of spatial change of allocentric cues is necessary to influence reaching movements, we predicted a similar increase of mean allocentric weights with the number of shifted background objects in Experiment 1 as was observed previously in conditions with increasing table object shifts (Fiehler et al., 2014).

We found no influence of background object shifts on reaching endpoints irrespective of the number of shifted background objects. In all experimental conditions including background object shifts, averaged allocentric weights did not differ from zero. Nevertheless, allocentric weights of the condition with 1 shifted background object were smaller than conditions with 3 or 5 shifted background objects. Thus, there was a small increase in displacements of reaching endpoints with an increasing number of shifted background objects. However, as the effect size of this result was very small and allocentric weights of these conditions did not differ from zero, this influence is negligible. This becomes even more evident if we compare the allocentric weights observed here for 3 (0.02) and 5 (0.04) background object shifts with the allocentric weights of 3 (0.1) and 5 (0.44) table object shifts of the previous study (Fiehler et al., 2014) or 5 table object shifts (0.47) assessed in this experiment (compare to figure 3.7B below). In sum, the lack of an effect of allocentric information of background object stimuli on reaching endpoints can't be explained by a decreased number of shifted background objects.

The results of Experiment 1 further demonstrated that shifts of 5 table objects led to a displacement of reaching endpoints into the direction of the object shift. This supports the results of the previous study (Fiehler et al., 2014) and confirms that the slight adjustments of the procedure and the use of 3D-rendered scenes instead of photographic images did not significantly influence the present results.

3.3 Experiment 2

3.3.1 Introduction

A second possible explanation for the absence of an effect in conditions with background object shifts in the study by Fiehler et al. (2014) is the fact that background objects never served as potential reach targets and hence, were not relevant to perform the task. Thus, participants may have paid no attention to these objects and in turn did not integrate this information into the reach plan. To address this hypothesis, we conducted a second experiment that was very similar to the paradigm that we used in our first experiment. The crucial difference was that background objects instead of table objects served as reach targets. We applied the same procedure but slightly adjusted the stimuli to match the task requirements (see methods for further details). If the task-relevance of an object is essential for its use as allocentric cue for reaching, we expect systematic reaching errors only in conditions with background object shifts. Furthermore, these errors should increase with the number of shifted background objects and should reveal a pattern similar to the results for table object shifts found by Fiehler et al. (2014). In contrast, we expect no systematic reaching errors in conditions with table object shifts in Experiment 2 as these objects are not relevant to perform the task.

3.3.2 Methods

Participants

We recorded data from eleven participants. Two participants were rejected due to bad fixation behavior (< 50 % trials with correct fixation). Thus, data of nine participants (six female) with normal or corrected to normal vision was further analysed. Participants were right-handed as assessed by the Edinburgh handedness inventory (EHI, Oldfield, 1971; *mean* 74.67 ± 26.91) and ranged in age from 19 to 27 years (*mean* 22.11 ± 3.14 years). They received course credit or were paid for their participation. The experiment was conducted in agreement with the ethical guidelines of the local ethics committee of the University of Giessen in compliance with the Declaration of Helsinki (2008).

Materials

Stimuli were created as described for Experiment 1. The breakfast scenes of Experiment 2 differed from the scenes used in Experiment 1 in the following ways. First, instead of 6, only 5 TO were placed on the table (the egg was omitted) in order to create a comparable non-target stimulus set to Experiment 1 (5 non-target objects). Accordingly, we also increased the number of BO from 5 to 6 to create a comparable target-stimulus set to Experiment 1 (6 target objects). In addition, we replaced the BO by 6 other objects (chair, vase, painting, calendar, clock, and ceiling lamp) that were more equal in size and generally smaller in width and thus should reduce the variance of reaching endpoints across the target objects. The object properties of the BO target objects used in Experiment 2 are summarized in table 3.3.

Table 3.3. Maximum height, width and distance to camera of new background objects in the scene in cm, based on the actual properties in SketchUp. Painting, calendar and clock had no fixed distance to the camera because their position altered on three different height levels on the wall. Some background objects were sometimes not fully visible due to an overlap with other objects. Therefore visible heights may vary from the actual height depending on the used object arrangement.

Object	**Height (visible)**	**Width**	**Distance to camera**
Chair	18.00	22.06	229.75
Vase	33.45	19.87	249.01
Painting	29.28	22.59	variable
Calendar	31.21	19.28	variable
Clock	20.45	20.45	variable
Ceiling lamp	12.41	20.18	182.13

We created 14 different object arrangements. Placements of TO followed the same criterion as in Experiment 1. The BO were positioned (i) with a distance to the edges of the image so that in case of object displacements they would be still completely visible, and (ii) that they were never occluded > 20 % by another object even after object displacement. The painting, calendar and clock were placed at three different heights with (i) minimum 1 object placed at every height level, and (ii) the calendar never placed on the highest level in order to minimize unrealistic object arrangements in the scene. The distance of the low height level from the ground was 107.55 cm, of the middle level 125.99 cm and of the high level 144.43 cm. Distances from the height levels to the camera were 278.97 cm, 279.51 cm, and 281.13 cm for the low, middle, and high level, respectively. The positions of the other BO were fixed on one horizontal line for each object in different distances to the camera (see table 3.3). Again, we created test images, but this time with 1 BO (instead of 1 TO) missing which served as reach target. In 2/3 of the test images, all 5 TO or 1, 3 or 5 BO were displaced either to the left or to the right (50 % leftward displacements). Amounts of horizontal displacement varied between 3.56 deg and 4.47 deg (*mean* 3.86 deg ± 0.33 deg). In the remaining 1/3 of the test images, no objects were shifted (= control condition).

In total, 266 images were rendered, including 14 encoding images, 168 test images (42 with TO shifts, 126 with BO shifts) and 84 control images. Moreover, from each of the 14 encoding images, a scrambled version made up of 768 randomly arranged squares of the image was created and used to mask the encoding image.

Apparatus & Procedure

The same set-up and trial procedure were used as in Experiment 1. Participants performed five experimental conditions (for examples see figure 3.5). In contrast to Experiment 1, 1 of the 6 BO was always removed from the test image and served as reach target. Each participant completed a minimum of 504 trials. Trials in which participants responded too early, did not respond or were unsure about the target position were repeated at the end of the experiment. Thus, the number of actually performed trials varied from 504 to 524 trials. Trials were presented in pseudo-randomized order, but a trial was never followed by a trial containing the same encoding image. The experiment was conducted in two sessions with all trials in the first and a repetition of every trial in the second session. Sessions were performed on two different days with one break in between.

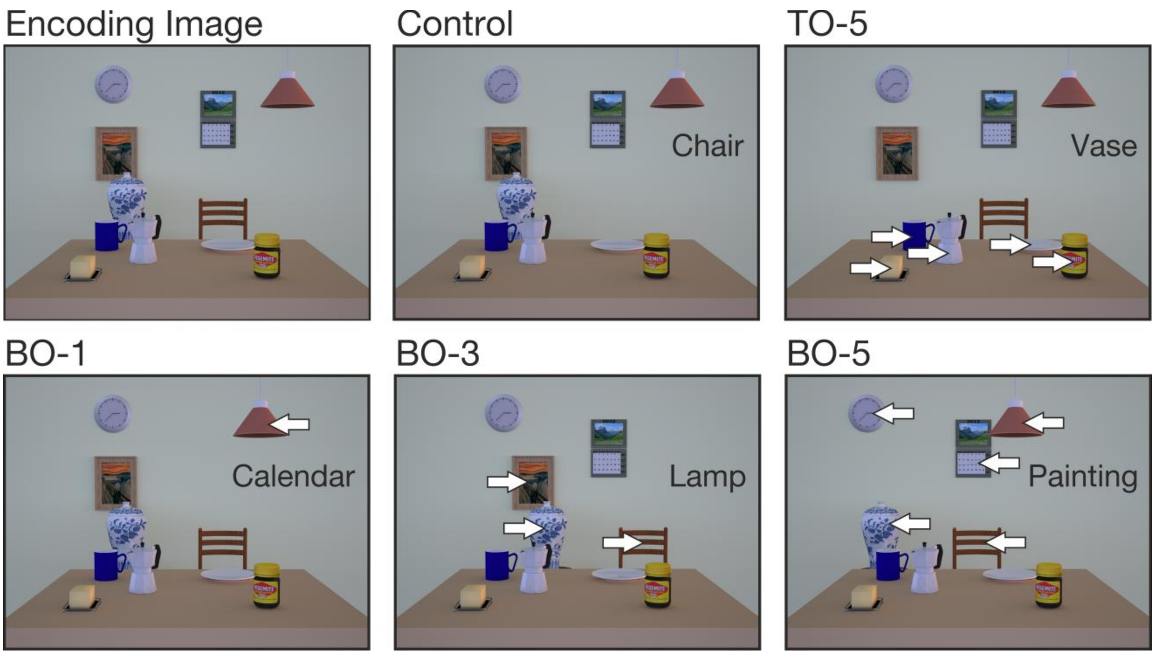

Figure 3.5. Examples of encoding and test images for different conditions. Object names in the box indicate the reach target (= missing object in the background).

Data Reduction and Statistical Analysis

Data preprocessing was performed with MATLAB R2007b (The MathWorks Inc.) and inferential statistics with R 3.1 (The R Foundation of Statistical Computing). All statistical tests were computed with an alpha-level of .05. If correction for multiple testing was necessary, correction following Bonferroni-Holm was applied. In case the assumption of sphericity for an ANOVA was violated (tested with Mauchly's sphericity test), Greenhouse-Geisser correction was applied.

As in Experiment 1, first, eye tracking data was inspected in order to detect incorrect fixations (deviations > 2.5° from fixation cross) and correspondent trials were discarded from further analysis. This applied to 479 trials (10.56 %). Next, reaching onsets and offsets were defined following the same criteria as in Experiment 1. Two trials (0.05 %) had to be rejected from further analysis because reaching endpoints could not be determined. Finally, we excluded trials in which reaching endpoints deviated more than ± 2.5 *SD* in horizontal or vertical direction from the group mean in each condition for every object shift direction (106 trials = 2.61 %). From originally 4536 trials, 3948 valid trials (87.04 %) were entered into statistical analysis.

To investigate the influence of object shifts on reaching endpoints, we calculated allocentric weights as described for Experiment 1. We performed two-sided one-sampled t-tests to investigate whether group allocentric weights for every condition differed significantly from zero. In order to assess the impact of allocentric information by BO on reaching endpoints, we conducted a one-way repeated measures ANOVA for conditions with BO shifts with the factor number of shifted BO and three levels (BO-1, BO-3, BO-5).

Again, we computed a heatmap of fixation densities for the encoding phase as in Experiment 1. While in Experiment 1 target objects were placed in the lower part of the scene, in Experiment 2 they were located in the upper part of the scene around the future location of the fixation cross. Thus, participants tended to fixate this region more often as in Experiment 1, even though objects were never placed in this area. To take this into account, we excluded fixations that fell into the area of the fixation cross from computing the heatmap.

3.3.3 Results

Figure 3.6 depicts horizontal and vertical reaching errors for every condition. Right- and leftward BO shifts led to systematic reach errors most pronounced for 5 object shifts (BO-5), smaller for 3 shifted objects (BO-3) and absent for 1 object shift (BO-1). In contrast to Experiment 1, we found no systematic reach errors if 5 table objects (TO-5) were shifted. Accordingly, allocentric weights differed significantly from zero in BO-5 and BO-3 but not in BO-1 and TO-5 (table 3.4).

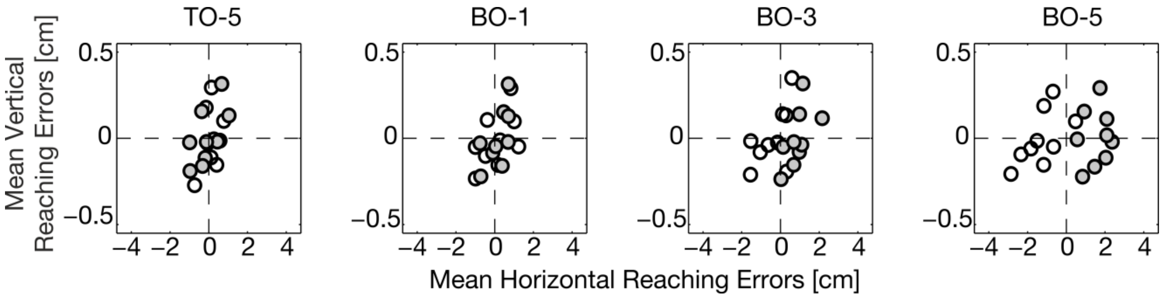

Figure 3.6. Mean horizontal and vertical reaching errors (in cm) of every participant for each condition. Leftward object shifts are depicted in white, rightward objects shifts in gray.

Table 3.4. Summary of allocentric weights for all conditions. Range, mean and standard deviation of the participant group are listed. Results of two-sided one-sampled t-tests are Bonferroni-Holm corrected.

Condition	Range	Mean	SD	t-test results
TO-5	-0.13 – 0.08	-0.03	0.06	$t(8) = -1.599$, $p = .297$
BO-1	-0.48 – 0.69	0.06	0.35	$t(8) = 0.546$, $p = .600$
BO-3	0.10 – 0.55	0.35	0.17	$t(8) = 6.256$, $p < .001$*
BO-5	0.28 – 0.74	0.48	0.15	$t(8) = 9.737$, $p < .001$*

We performed a one-way repeated-measure ANOVA for the BO conditions on allocentric weights to evaluate the influence of an increasing number of shifted BO on reaching endpoints. We found a main effect for the number of shifted BO ($F(2,16) = 7.272$, $p < .001$, $\eta^2 = 0.378$). Post-hoc t-tests indicated that allocentric weights of BO-1 were significantly smaller than weights of BO-5 ($t(8) = -3.379$, $p = .029$). Allocentric weights of BO-3 did not differ from weights of BO-1 ($t(8) = -2.124$, $p = 0.133$) and weights of BO-5 ($t(8) = -2.032$, $p = 0.133$). Figure 3.7A shows the mean allocentric weights for every condition. For comparison, we also depicted the mean allocentric weights published by Fiehler et al. (2014) in figure 3.7B.

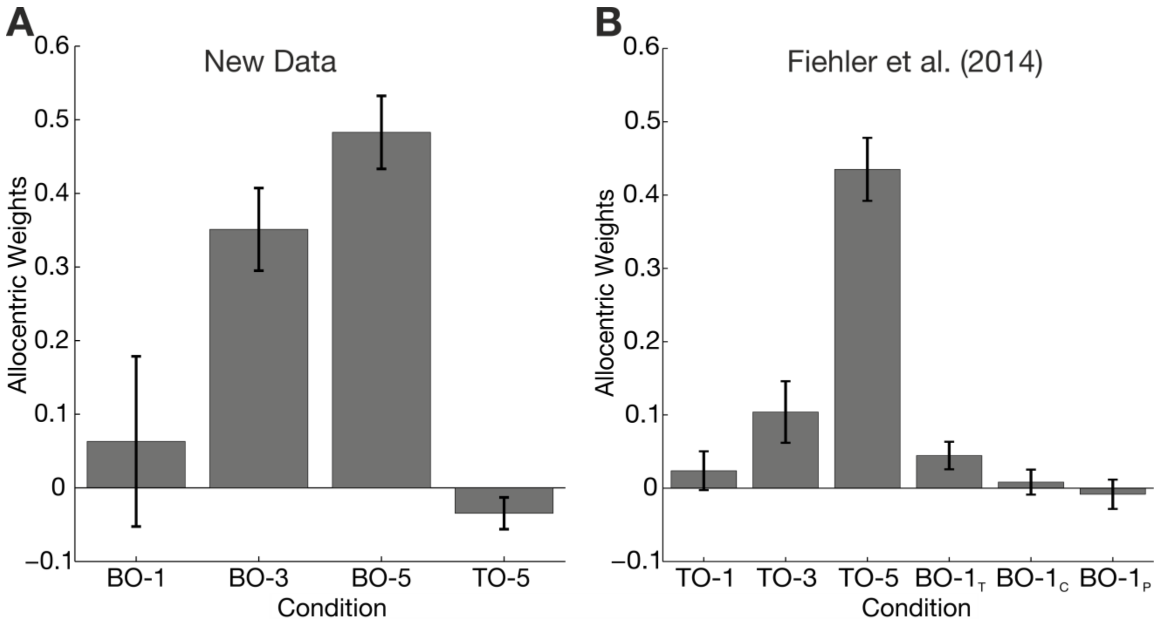

Figure 3.7. (**A**) Mean allocentric weights for the table object condition and all background object conditions. Error bars represent 1 *SEM*. (**B**) Mean allocentric weights reported by Fiehler et al. (2014). Instead of background objects, table objects served as reach targets. In table object conditions, 1, 3, or 5 table objects were shifted either to the right or to the left. In background object conditions (BO-1_T, BO-1_C, BO-1_P), only 1 of the 3 background objects was displaced.

For Experiments 1 and 2, we computed heatmaps of fixation densities over all participants and conditions for the encoding phases. As illustrated in figure 3.8, participants mainly fixated the objects on the table and hardly fixated the objects in the background in Experiment 1 (figure 3.8A) whereas in Experiment 2 the reversed fixation pattern occurred (figure 3.8B). Here, most of the fixations fell on objects in the background while nearly no fixation was kept on objects on the table. This result pattern clearly shows that participants fixate on objects which serve as potential reach targets and thus are relevant for the task, i.e. table objects in Experiment 1 and background objects in Experiment 2.

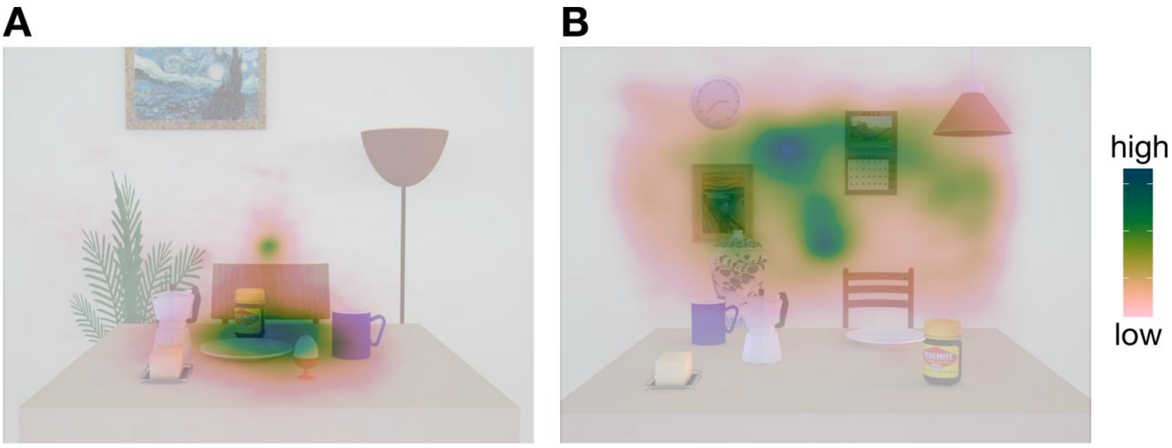

Figure 3.8. Heatmaps of fixation densities for the encoding phase averaged across all participants and conditions for **(A)** Experiment 1 and **(B)** Experiment 2. The high density of fixations at the center of the image corresponds to the location of the upcoming fixation cross after image encoding (cf. trial schedule in figure 1).

3.3.4 Discussion

In Experiment 2, we tested the hypothesis whether task-relevant allocentric cues only are integrated into the reach plan and thus influence reaching behavior. In contrast to Experiment 1, background objects (instead of table objects) served as reach targets and thus became relevant to perform the task. We predicted a similar pattern of mean allocentric weights for background object shifts as was observed for table object shifts in the previous study (Fiehler et al., 2014). In line with our prediction, we observed systematic endpoint errors in the direction of background object shifts. Moreover, these endpoint error deviations increased with the number of shifted background objects. This was confirmed by averaged allocentric weights which revealed a similar increase compared to the allocentric weights in the study of Fiehler et al. (2014). In contrast to Experiment 1, we found no effect on reach endpoints for conditions with table object shifts despite the fact that all 5 table objects were shifted. Therefore, we conclude that the task-relevance of an object determines whether an object is used as an allocentric cue or not.

3.4 General Discussion

There is converging evidence that reach targets are represented in both egocentric (e.g. Cohen & Anderson, 2002; Lacquaniti & Caminiti, 1998; Thompson & Henriques, 2011) and allocentric reference frames (e.g. Byrne & Crawford, 2010; Diedrichsen et al., 2004; Krigolson et al., 2007; Obhi & Goodale, 2005; Schütz et al., 2013, 2015). So far, little is known about the factors contributing to how allocentric information is used for reaching when multiple environmental cues are available in more complex situations, such as in naturalistic images. Here, we followed the approach of Fiehler et al. (2014) and tested two hypotheses derived from their study. First, we examined whether the spatial change of allocentric cues in a scene must exceed a certain extent to influence reach behavior. Second, we tested if task-relevance of allocentric cues is an important factor for determining whether or not they are used for reaching. In our first experiment, we shifted 1, 3, or 5 background objects or 5 table objects while table objects served as potential reach targets. Here, we replicated the findings of Fiehler et al. (2014) and found a systematic shift of reach endpoints in the direction of the table object shifts. However, an increasing number of shifted background objects did not influence reach endpoints, even if all available background objects were shifted. In our second experiment, again 1, 3, or 5 background objects or 5 table objects were shifted; but this time background objects served as potential reach targets. Here, we found the reversed effect of Experiment 1, i.e. reaching errors varied systematically with the number and direction of shifted background objects while being unaffected by shifts of table objects. Based on these findings, we conclude that task-relevance is a crucial factor for the use of allocentric cues for reaching.

If task-relevance is given for a group of objects, we observed that the number of shifted objects determines their influence on reaching endpoints. We found allocentric weights increasing from 0 % for 1 to 48 % for 5 shifted objects, irrespective of whether task-relevant objects were placed in the background or on the table. This result is comparable to the study by Fiehler et al. (2014) who also found a systematic increase of allocentric weights with an increasing number of shifted task-relevant table objects varying from 1 % for 1 to 43 % for 5 shifted objects. Accordingly, previous work on the Roelofs effect showed that the mislocalization of a target within a frame increases when the full frame is shifted compared to conditions in which only parts of the frame are being displaced (Walter & Dassonville, 2006). This indicates that an increasing amount of changing allocentric information has a cumulative impact on the perceived target localization, especially if it comprises all available allocentric cues. However, as we revealed in Experiments 1 and 2 this finding crucially depends on the condition that the group of misplaced objects is task-relevant. We suggest that both egocentric and allocentric information are integrated into the reach plan depending on bottom-up and top-down processes, namely the amount of changes in the scene and the task-relevance of allocentric cues, respectively.

Our findings also support results on visual attention and task-relevance of objects in real world and natural scenes. It has been demonstrated that overt visual attention is mainly distributed to objects that are relevant to perform a task reflected in more fixations on task-relevant than -irrelevant objects (Ballard & Hayhoe, 2009; DeAngelus & Pelz, 2009; Land & Hayhoe, 2001) and longer fixation durations (Mills, Hollingworth, Van der Stigchel, Hoffman, & Dodd, 2011). In order to test whether participants in our experiments also showed an attentional preference for task-relevant objects during image encoding, we computed heatmaps illustrating fixation densities for the encoding images of all participants (figure 3.8). Figure 3.8 shows that participants fixated the area around the table objects more often than the area around the background objects in Experiment 1 while the reversed fixation pattern is visible in Experiment 2, i.e. more fixations on background than on table objects. Consistent with previous findings on top-down control of eye movements in real world and natural scenes (Ballard & Hayhoe, 2009; DeAngelus & Pelz, 2009; Land & Hayhoe, 2001; Mills et al., 2011), fixations (i.e., overt visual attention) were predominantly shifted to objects that were relevant to perform the task. Moreover, the fixation behavior found here is in line with the study by Fiehler et al. (2014) where participants mainly fixated the area around the task-relevant table objects. We conclude that overt visual attention is mainly distributed to areas with task-relevant allocentric cues which are consequently integrated into the reach plan.

Previous research suggests that task-relevance of objects in a scene can improve the detection of changes of object properties (Triesch, Ballard, Hayhoe, & Sullivan, 2003) and prioritize their retention in visual working memory (Maxcey-Richard & Hollingworth, 2013). In our task, a short delay was implemented between the encoding and the test image which required maintaining object positions in visual working memory. This information was subsequently used to detect the target location and to perform the reaching movement. Participants' fixation behavior (figure 3.8) suggests that they prioritized the position of task-relevant objects during scene encoding which may have facilitated target detection and the selection of task-relevant allocentric cues for reaching.

Besides task-relevance, the distance between target and allocentric cues seems to influence how and to which extent these cues contribute to reaching movements (Camors, Jouffrais, Cottereau, & Durand, 2015; Diedrichsen et al., 2004; Krigolson et al., 2007). In our experiments task-relevant objects, which served as reach targets, were always placed in the direct vicinity of each other, excluding a potential impact of target-landmark-distance. However, some (but not all) task-irrelevant objects were presented further away from the target which might have led to an attenuated contribution of task-irrelevant objects on reaching. Nevertheless, we believe that a larger distance between task-irrelevant objects and target cannot fully account for the lack of an effect of irrelevant objects. It is rather likely that participants initially ignored task-irrelevant objects during image encoding as evidenced by the heatmaps in figure 3.8 and thus did not integrate this allocentric information into the reach plan.

Furthermore, a temporal delay between stimulus presentation and action performance can influence the weighting of egocentric and allocentric information with stronger weighting of allocentric information for delayed than immediate movements (Bridgeman, Peery, & Anand, 1997; Chen, Byrne, & Crawford, 2011; Hay & Redon, 2006; Obhi & Goodale, 2005). Importantly, allocentric information are also incorporated in immediate reaches as recently demonstrated by a study on the Roelofs effect (Taghizadeh & Gail, 2014) and as we demonstrated in our experiments. Since allocentric coding is supposed to be stronger for delayed than immediate reaches (Bridgeman et al., 1997; Chen et al., 2011; Hay & Redon, 2006; Obhi & Goodale, 2005), the allocentric weights we observed here could have even been higher with a temporal delay before the reach.

The extent to which results from studies using photographs or 3D-rendered scenes can be transferred to real world situations poses an interesting question for future research. It has been demonstrated that real-world objects were better recalled and recognized than photographs or line drawings of these objects (Snow, Skiba, Coleman, & Berryhill, 2014). Based on this finding, one could argue that the present results cannot be generalized to naturalistic behavior. However, we believe that our approach is an important intermediate step in order to transfer outcomes from laboratory settings to the real world.

Overall, our findings extend the current scientific body showing that task-relevance of allocentric cues determines their contribution to reaching movements if multiple cues are available in a more complex and naturalistic environment. Moreover, this influence on reaching behavior further depends on the extent of changes of task-relevant allocentric information.

4 Scene configuration and object reliability affect the use of allocentric information for memor-guided reaching

A similar version of this manuscript is submitted as:

Klinghammer, M., Blohm, G., & Fiehler K. (submitted). Scene configuration and object reliability affect the use of allocentric information for memory-guided reaching.

Previous research has shown that egocentric and allocentric information is used for coding target locations for memory-guided reaching movements. Especially, task-relevance determines the use of objects as allocentric cues. Here, we investigated the influence of scene configuration and reliability as a function of task-relevance on allocentric coding for memory-guided reaching. For that purpose, we presented participants a breakfast scene with five objects on a table and six objects in the background. Six of these objects served as potential reach-targets (= task-relevant objects). Participants explored the scene and after a short delay, a test scene appeared with one of the task-relevant objects missing, indicating the location of the reach target. After the test scene vanished, participants performed a memory-guided reaching movement toward the target location. Besides removing one object from the test scene, we also shifted the remaining task-relevant and/or task-irrelevant objects left- or rightwards either coherently in the same direction or incoherently in opposite directions. By varying object coherence, we manipulated the reliability of task-relevant and task-irrelevant objects in the scene. In order to examine the influence of scene configuration (distributed vs. grouped arrangement of task-relevant objects) on allocentric coding, we compared the present data with our previously published data set (Klinghammer et al., 2015). We found that reaching errors systematically deviated in the direction of object shifts, but only when the objects were task-relevant and their reliability was high. However, this effect was substantially reduced when task-relevant objects were distributed across the scene leading to a larger target-cue distance compared to a grouped configuration. No deviations of reach endpoints were observed in conditions with shifts of only task-irrelevant objects or with low object reliability irrespective of task-relevancy. Moreover, when solely task-relevant objects were shifted incoherently, the variability of reaching endpoints increased compared to coherent shifts of task-relevant objects. Our

results suggest that the use of allocentric information for coding targets for memory-guided reaching depends on the scene configuration, in particular the average distance of the reach target to task-relevant objects, and the reliability of task-relevant allocentric information.

4.1 Introduction

It is well established that 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) for coding object locations in the environment. For memory-guided reaching movements, egocentric (Cohen & Anderson, 2002; Fiehler, Schütz, & Henriques, 2011; Lacquaniti & Caminiti, 1998; Thompson & Henriques, 2011) and allocentric information (e.g. Byrne & Crawford, 2010; Diedrichsen, Werner, Schmidt, & Trommershäuser, 2004; Krigolson & Heath, 2004; Krigolson, Clark, Heath, & Binsted, 2007; Obhi & Goodale, 2005) is integrated into a movement plan. Recent work from our lab confirmed the contribution of allocentric information for memory-guided reaching by using naturalistic 2D images of complex scenes (Fiehler, Wolf, Klinghammer, & Blohm, 2014) or 3D virtual reality (Klinghammer, Schütz, Blohm, & Fiehler, under review). In our previous studies (Fiehler et al., 2014; Klinghammer et al., 2015), we presented 2D naturalistic images of a breakfast scene on a computer screen containing objects on a table and in the background. We instructed participants beforehand that either table or background objects function as potential reach targets and thus, are relevant for the task. Participants first encoded the scene with free gaze and after a short delay they briefly viewed a test scene with one of the task-relevant objects missing indicating the reach target location. After the test scene vanished, participants performed a reach to the remembered location of the missing object on a gray screen while gaze was fixed. Besides removing one object from the test scene, we also shifted some of the remaining objects either to the left or to the right. We found that reaching endpoints systematically deviated into the direction of object shifts, but only when we shifted task-relevant objects. When we shifted task-irrelevant objects that never became a reach target, reaching endpoints remained unchanged compared to a control condition with no object shifts. Based on these findings, we concluded that first, allocentric information is utilized for memory guided reaching and second, objects' task-relevance (i.e., do they serve as potential reach target or not) is an important factor determining whether they are used as allocentric cues or not. Here, we followed a similar approach and first aimed to investigate how the scene configuration influences allocentric coding for memory-guided reaching. Second, we wanted to know whether and how the reliability of task-relevant and task-irrelevant objects, i.e., the reliability of allocentric information, affects the use of objects as allocentric cues.

In our experiments so far (Fiehler et al., 2014; Klinghammer et al., 2015), task-relevant objects were located either on the table or in the background forming a spatial cluster that was separated from the cluster containing task-irrelevant objects (see figure 4.1). This spatial arrangement influenced participants' encoding behavior in a way that their overt spatial attention was mainly directed to the relevant objects' cluster while ignoring the area containing the irrelevant objects (see fixation density maps, figure 4.1). The question arises whether spatial grouping of task-relevant information facilitates allocentric coding and thus, would be impeded if task-relevant objects are distributed across the whole scene. Moreover, spatial grouping of objects in task-relevant table or background objects also led to a smaller mean distance between the reach target and the task-relevant than task-irrelevant objects. There is evidence that with an increasing distance between target and landmark (i.e., allocentric cue), the influence of the landmark becomes less effective (Camors, Jouffrais, Cottereau, & Durand, 2015; Krigolson et al., 2007). Furthermore, endpoints of pointing movements are most affected by the closest landmark if multiple landmarks are available (Diedrichsen et al., 2004; Spetch, 1995). In this study, we aimed to examine the influence of the scene configuration by randomly placing task-relevant objects on the table *and* in the background within the same scene. By doing so, we not only increased the mean distance from the target to the task-relevant objects but at the same time also reduced the mean distance from the target to the task-irrelevant objects which also occurred in close proximity to the target comparable to the task-relevant ones. Based on the findings reported above, we predict a decrease in allocentric coding compared to our previous study (Klinghammer et al., 2015), in which task-relevant objects were spatially grouped and therefore, placed in the closer vicinity to the target than task-irrelevant objects.

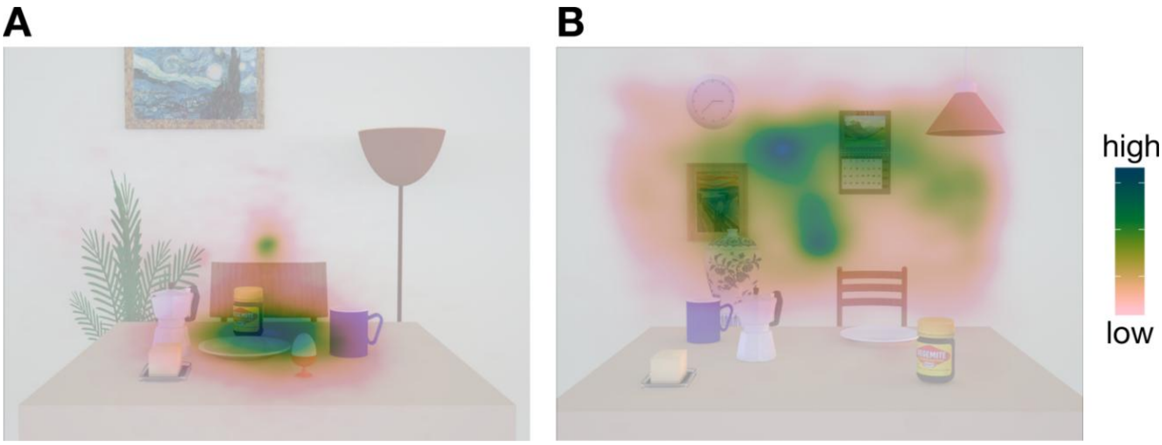

Figure 4.1. Fixation density maps from Klinghammer et al. (2015) containing examples images of the stimuli. In (**A**), only objects on the table served as reach targets (task-relevant objects) and formed a spatial object cluster which is spatially distinct from the irrelevant objects in the background. As a result, participants mainly fixated this area while ignoring the background. In (**B**), only objects in the background served as reach targets and formed an object cluster. Consequently, we observed the reversed fixation pattern.

Beyond the scene configuration, the reliability of allocentric cues might influence their use for coding reach targets in space. Byrne and Crawford (2010) investigated the influence of the reliability and stability of landmarks on allocentric coding of target locations for memory-guided reaching. Landmarks consisted of four dots that were arranged in an imagined square around a target dot. They varied their stability and reliability by manipulating the amplitude of the dots' vibration. The authors hypothesized that with increasing the vibration amplitude the landmark stability and reliability should decrease and thus, the weighting of the allocentric information, should also decrease. As expected, they found a larger influence of low vibrating landmarks on absolute reaching endpoints compared to landmarks with high vibration amplitude; however, there was no effect on the variability of reaching endpoints. The authors concluded that landmark stability influences the weighting of the allocentric information. This is in line with previous studies on memory-guided reaching showing that humans use stable and reliable landmarks leading to increased reaching endpoint accuracy (Krigolson & Heath, 2004; Obhi & Goodale, 2005). Similar results have been reported for changes in spatial object configurations influencing the reliability of allocentric cues. For example, when asking participants to detect a shift of one of multiple objects from an encoding to a probe display they were more accurate in conditions in which the objects were shifted coherently in one direction (minimal change) than they were arranged in a new, random fashion (maximal change) (Jiang, Olson, & Chun, 2000). This suggests that the global configuration of objects in the display was taken into account and used for representing the single objects’ location. Transferred to our paradigm, we would expect that breaking up the coherence of the spatial object configuration, due to object shifts in opposite directions in the test scene, the reliability of these objects should decrease and thus, their contribution to allocentric coding of memorized reach targets.

Taken together, in the current study we investigated whether the scene configuration and the reliability of task-relevant and task-irrelevant objects serving as potential allocentric cues influence the use of allocentric information for memory-guided reaching in naturalistic scenes. To this end, we used a similar paradigm as published by Klinghammer et al. (2015), but this time we distributed task-relevant and task-irrelevant objects across the whole scene preventing spatially distinct clusters. Based on our previous findings (Klinghammer et al., 2015), we expect systematic deviations of reaching endpoints in the direction of task-relevant but not of task-irrelevant object shifts. However, reach endpoint deviations resulting from task-relevant object shifts should be smaller than the ones observed by Klinghammer et al. (2015) due to the increased distance between the target and the task-relevant objects and/or the placement of task-irrelevant objects in closer proximity to the potential reach targets. In order to manipulate the reliability of task-relevant and task-irrelevant objects, we introduced conditions where we either keep the coherence of the whole object arrangement intact (same object shift direction) or break it up (opposite object shift direction). We expect that systematic deviations of reaching endpoints in the shift-direction are larger in the condition with high (intact coherence) than low (broken coherence) object reliability. Assuming that task-irrelevant objects are widely ignored and thus, unused for allocentric coding of memorized reach targets (Klinghammer et al., 2015), reach endpoint deviations in the direction of object shifts should be strongly influenced by the reliability of task-relevant objects, but hardly affected by the reliability of task-irrelevant objects. In particular, we expect an effect of object shifts on reaching endpoints in conditions in which we keep the scene coherence within the group of task-relevant objects intact, which should strongly decrease when we shift task-relevant objects incoherently. In contrast, we expect a substantially reduced influence of object shifts in conditions when task-irrelevant compared to task-relevant objects are shifted alone regardless of the coherence manipulation.

4.2 Methods

Participants

We recorded data from 22 participants. Four of them had more than 30 % of trials without correct fixation and thus, were excluded from further analysis. For one additional participant, we failed to measure reach endpoints or correct fixation behavior in more than 60 % of trials and therefore discarded the data from further analysis. The final sample consisted of 17 participants (8 female) with normal or corrected to normal vision ranging in age from 19 to 30 years (*mean* 25 ± *SD* 3.2 years). They were right-handed as assessed by Edinburgh handedness inventory (EHI, Oldfield, 1971; *mean handedness quotient* 78 ± *SD* 18.8). They received course credit or were paid for their participation. The experiment was conducted in agreement with the ethical guidelines of the local ethics committee of the University of Giessen in compliance with the Declaration of Helsinki (2008).

Apparatus

Stimuli were presented on a 19” (40.5 cm x 30 cm) CRT monitor (Ilyama Vision Master Pro 510) with a resolution of 1280 x 960 pixels and a refresh rate of 85 Hz. To reduce the influence of a high-contrast frame around the scene, a black cardboard (70 cm x 50 cm) frame was attached to the monitor. Participants sat at a table with their head stabilized on a chin rest with a distance of roughly 47 cm from the eyes to the center of the screen. A decimal-keyboard was placed in front of the participants with the start button 24 cm away from the screen and aligned to the chin rest and the center of the screen. Reaches were performed with the right index finger and recorded with an Optotrak Certus (NDI, Waterloo, Ontario, Canada) tracking system with a sampling rate of 150 Hz using one infrared marker attached to the fingertip of the right index finger. To control for correct fixation behavior, eye movements were recorded using an EyeLink II system (SR Research, Osgoode, Ontario, Canada) with a sampling rate of 500 Hz. To run the experiment and to control the devices we used Presentation 16.5 (Neurobehavioral Systems, Inc., Berkeley, CA).

Materials

Stimuli consisted of 3D-rendered images of a breakfast scene. Images were created in SketchUp Make 2015 (Trimble Navigation Ltd., Sunnyvale, CA) and afterwards rendered with Indigo Renderer 3.8.21 (Glare Technologies Ltd.) with a resolution of 3562 x 2671 pixels. The breakfast scene contained 5 objects consisting of a coffee mug, a plate, an espresso cooker, a Vegemite jar, and a butter dish placed on a brownish table that stood 86 cm in front of a gray wall. Furthermore, 6 objects, consisting of a chair, vase, painting, calendar, clock, and ceiling lamp were located behind the table in the background. Objects were taken from the open access online 3D-gallery of SketchUp. Object properties are summarized in table 4.1.

We set all objects in 18 different arrangements (*encoding image*). They were placed so that < 20 % of an object was occluded by another object and with a distance to the edges of the table or the image so that in case of object displacement no object stood in the air next to the table or outside of the image. In any arrangement, objects on the table were placed at one of three possible horizontal depth lines that were equally spaced (19.5 cm starting from the front table edge) on the table with minimal 1 and maximal 2 objects positioned at every depth line. The painting, calendar and clock were placed at three different heights at the wall with 1 object placed at every height level, and the calendar never placed on the highest level in order to minimize unrealistic object arrangements in the scene. The distance of the low height from the ground was 107.55 cm, of the middle height 126.38 cm and of the high height 145.20 cm. Distances from the height levels to the camera were 278.97 cm, 279.51 cm, and 281.30 cm for the low, middle, and high height, respectively. The positions of the vase, chair, and ceiling lamp were fixed on one horizontal line for each object in different distances to the camera (see table 4.1). Based on the encoding images, we created *test images*, in which 1 of 6 pre-defined objects (3 table objects and 3 background objects) was missing (= *reach target*). These 6 pre-defined objects served as potential reach targets and thus, were highly relevant to perform the task (= *relevant objects* (RO)). The remaining 5 objects never served as reach targets and thus, were task-irrelevant (= *irrelevant objects* (IO)). In 2/3 of the test images, objects (RO and/or IO) were shifted horizontally between 3.56 deg and 4.47 deg (*mean* 3.86 deg ± *SD* 0.33 deg) either to the left or to the right (50 % leftward displacement) in the same (= coherent object shift) or in opposite directions (= incoherent object shift). Variations in the horizontal object displacement arose from the fact that objects were placed at different depth lines relative to the virtual camera position. Hence, similar physical shifts of objects at different depth lines in 3D-space would result in different displacements in the 2D-image. In the remaining 1/3 of the test images, no objects were shifted. These images served as control condition.

In total, 360 images were rendered, including 18 encoding images, 228 test images (76 with only RO shifts, 76 with only IO shifts, 76 with RO and IO shifts) and 114 control images. Moreover, from each of the 18 encoding images, a scrambled version made up of 768 randomly arranged squares of the image was created and used to mask the encoding image.

Table 4.1. Maximum height, width and distance to camera of all objects in the scene in cm, based on the actual properties in SketchUp. Objects on the table, painting, calendar and clock had no fixed distance to the camera because they were randomly placed on one of three different depth lines on the table or their position altered on three different height levels at the wall respectively. Some background objects were sometimes not fully visible due to object overlap. Therefore, visible heights may vary from the actual height depending on the object arrangement.

Object	Height (visible)	Width	Distance to camera
Plate	1.97	19.32	variable
Butter dish	4.89	8.36	variable
Espresso cooker	15.11	8.62	variable
Vegemite jar	11.44	6.86	variable
Mug	9.80	7.80	variable
Chair	18.00	18.00	229.32
Vase	31.98	19.69	241.52
Painting	29.11	22.59	variable
Calendar	31.21	19.28	variable
Clock	20.45	20.45	variable
Ceiling lamp	12.48	20.18	182.13

Procedure

Participants first read a written instruction about the experimental procedure informing about the RO and their function. Afterwards, they performed a learning block in which the 6 RO were presented together on the computer screen and participants were requested to memorize these objects without time restriction. Next, a picture of only one RO or IO was presented and participants were asked to indicate by button press whether this object was a potential reach target or not. After feedback about the correctness of the response was given, the next picture with a different object appeared on the screen. This was repeated until every object was presented once. The learning block ended if participants correctly classified the presented objects as potential reach target for at least three times in a row. Then, the experiment started after some training trials.

The procedure of an example trial is depicted in figure 4.2. Before every trial, a fixation cross on a gray screen appeared prompting participants to fixate and press a button in order to perform a drift correction for the EyeLink II. Thereafter, the trial started with the presentation of the encoding image of the breakfast scene. Participants freely explored the scene without any time constraints and terminated the encoding phase by pressing the start button. Then, a scrambled version of the encoding scene appeared for 200 ms to avoid afterimages followed by a delay phase of 800 ms with a gray screen and a central fixation cross. Participants were instructed to fixate the cross until the end of the reaching movement in order to control for the use of a gaze-centered egocentric reference frame (Thompson & Henriques, 2011). After the delay, the test image was presented for 1300 ms which lacked 1 RO defining the reach target. The trial continued with a short tone after the test image vanished which signaled the participants to perform the reaching movement towards the remembered location of the target object onto a gray screen. Thus, reaches were performed with gaze kept on the fixation cross and without any visual information of the encoding or test images. In this way we ensured that allocentric information could not be used for subsequent online corrections during the reaching movement, which would have led to an allocentric bias.

Participants were instructed to reach to the location of the missing object as accurately and natural as possible. Whenever they were unsure about the target location or identity, they had to reach to a marked location at the lower right edge of the monitor. These invalid trials were repeated at the end of the experiment. If participants released the button before the go-signal, they received feedback and these invalid trials were also repeated at the end of the experiment.

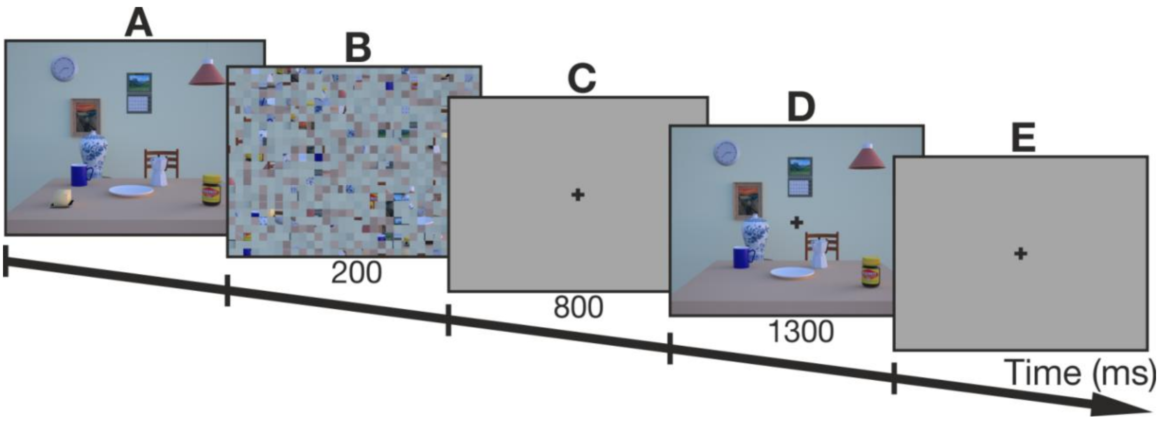

Figure 4.2. Trial scheme of one example trial (control condition). (**A**) First, the encoding image was presented and participants terminated the exploration of the image by button press. (**B**) Then, a scrambled version of the encoding image was presented for 200 ms, followed by (**C**) a delay which lasted for 800 ms. (**D**) Thereafter, the test image with one of the task-relevant objects missing (butter dish) was presented for 1300 ms before (**E**) a tone prompted participants to reach to the target onto a gray screen while fixating the cross at the center of the screen.

Participants performed six experimental conditions (for examples, see figure 4.3). In all experimental conditions, 1 of 6 RO was always removed from the test image, which served as the reach target. In the *RO same* condition, the remaining 5 RO were shifted either to the left or to the right. In the *IO same* condition, all 5 IO were shifted left- or rightward. In the *RO diff* or *IO diff* condition, the 5 relevant or the 5 irrelevant objects were shifted in different directions with 3 objects displaced in one and the remaining 2 objects in the opposite direction, i.e. 3 objects shifted rightward and 2 leftward or vice versa. The direction in which 3 objects were shifted is regarded as the main shift direction. In the *RO+IO same* condition, all relevant and irrelevant objects were shifted in the same direction, whereas in the *RO+IO diff* condition all relevant objects were shifted in the opposite direction of all irrelevant objects. How these different conditions influence the overall scene coherence and the coherence within the group of task-relevant objects is summarized in table 4.2. In all conditions, left- and rightward object shifts were balanced with 50 % of trials in each direction; the same accounts for the direction of the main object shift in the conditions RO diff and IO diff. In the *control* condition, all objects remained stationary.

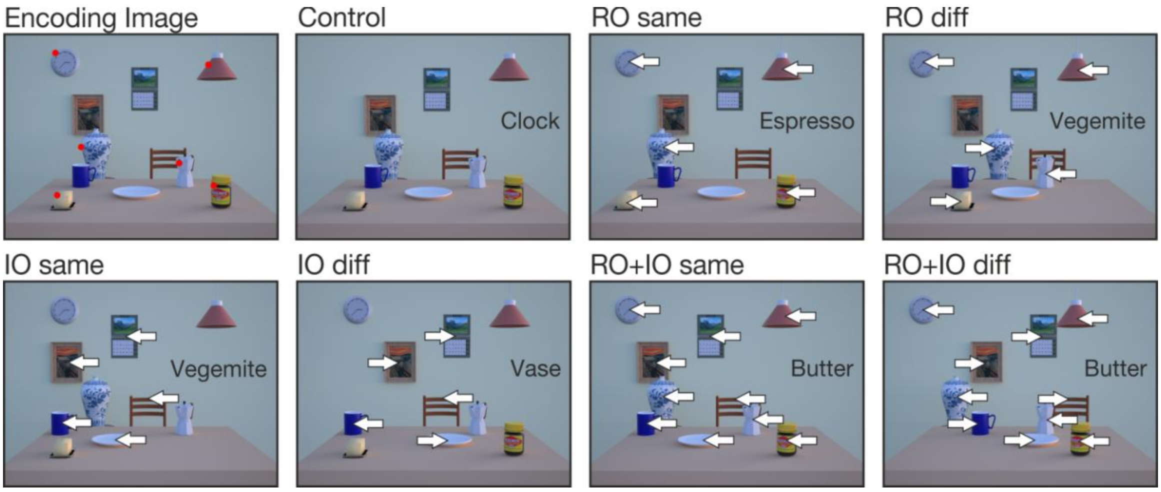

Figure 4.3. Examples of encoding and test images for the 7 different conditions. Object names in the box indicate the reach target (= missing object on the table). Arrows indicate the direction in which objects were shifted. Red dots in the encoding image mark the task-relevant objects (RO), but were absent in the experiment.

Table 4.2. Expected influences of the different object shifts in the 7 conditions on the global scene coherence and the coherence of object locations within the group of RO and IO.

	Influence on…		
Condition	**…global coherence**	**…coherence within RO**	**…coherence within IO**
control	intact	intact	intact
RO same	partially broken	intact	intact
IO same	partially broken	intact	intact
RO diff	partially broken	broken	intact
IO diff	partially broken	intact	broken
RO+IO same	intact	intact	intact
RO+IO diff	broken	intact	intact

Each participant completed a minimum of 648 trials. Because some trials were repeated (criteria see above), the actual number of performed trials varied from 651 to 736 trials. Trials were separated in three sessions with one session per day which lasted about one hour with one break in between. Trials were presented in pseudo-randomized order with a random sequence of conditions and encoding images within a session but fixed trial combinations between sessions. A trial was never followed by a trial containing the same encoding image. Every trial was repeated one time.

Data Reduction and Statistical Analysis

Data preprocessing was performed with MATLAB R2012a (The MathWorks Inc., Natick, MA) and inferential statistics with R 3.1 (R Development Core Team, www.r-project.org). All statistical tests were computed with an alpha-level of .05. If correction for multiple testing was necessary, Bonferroni-Holm correction was applied. In case the assumption of sphericity for an ANOVA was violated (tested with Mauchly's sphericity test), Greenhouse-Geisser correction was applied.

First, we inspected the eye tracking data and discarded trials from further data analysis in which participants' gaze deviated more than 2.5° from the center of the fixation cross during a period beginning from delay onset till the end of the reaching movement. All in all 724 trials (6.57 %) were rejected due to bad fixation. Second, reaching onsets and offsets were defined for every trial. The moment participants released the response button determined the reaching onset. Reach offsets were calculated from Optotrak data and defined as the first time point during the movement when velocity dropped below 20 mm/s if the index finger reached a maximum distance of about 3 cm from the screen. Reach endpoints were extracted at the time of reach offset. Some trials were excluded because reaching offsets or endpoints could not be extracted due to rarely occurring interferences of the infrared markers of the Optotrak with the IREDs of the EyeLink II (134 trials = 1.3 %). Third, we excluded trials in which reaching endpoints deviated more than ± 2.5 *SD* in horizontal or vertical direction from the group mean in each condition for each object shift direction (534 trials = 5.26 %). Taken together, from originally 11.016 trials of all participants, 9.624 valid trials (87.36 %) remained.

To investigate the influence of object shifts (i.e., allocentric information) on reaching endpoints, we calculated allocentric weights for every participant and every condition by linear regression fit. First, we determined reaching errors as the horizontal distance of the reach endpoint and the actual target position of the encoding image. Therefore, we averaged reach endpoints of the control condition of all participants for every combination of object arrangements and target objects separately. Since none of the remaining objects were displaced in the control condition, we assume no systematic reaching errors. These averaged reach endpoints were used to define the target positions. Then, we calculated the differences of the reaching endpoints of the other experimental conditions from the corresponding target position in the horizontal plane. This resulted in positive values for misestimations to the right and negative values for misestimations to the left. In the next step, we determined *maximal expected reaching errors* (MERE) for every image after an object shift by assuming that participants completely relied on the allocentric information of the shifted objects and thus produced reaching endpoints equal to the amount of the objects' displacement. To this end, we averaged the amount of displacement of the shifted objects for every image. If objects were shifted in different directions, we either averaged over the shift distances of only relevant objects (RO+IO diff) or averaged over the main shift direction (RO diff, IO diff). For the regression fit, the MERE was used as a predictor and the actual reaching error as a dependent variable for the two shift directions within one condition for every participant. The resulting slope of the regression line indicated the extent to which a participant relied on the allocentric information of object displacements and thus was defined as allocentric weight varying from 0 (no use of the allocentric information of the shifted objects) to 1 (full use of the allocentric information of the shifted objects) for further analysis. To investigate the influence of object shifts on the variability of reaching endpoints, we calculated standard deviations of reaching errors for every participant in every condition. To account for the fact that reaching errors to the left had negative and reaching errors to the right had positive values, we calculated standard deviations for the two shift directions separately and averaged the data afterwards.

To investigate whether participants used a different encoding behavior with respect to their focus of overt visual attention compared to our previous experiments (Klinghammer et al., 2015), we created fixation density maps of participants' fixation behavior during the encoding phase. To this end, we calculated a mean fixation point for every fixation starting from the second fixation until the end of the encoding phase. To examine whether participants fixated relevant objects more often than irrelevant objects, we collapsed fixations of the different object arrangement scenes resulting in 18 different heatmaps. We then visually inspected the heatmaps and descriptively compared them to the ones of our previous study (Klinghammer et al., 2015).

Next, we performed two-sided one-sampled t-tests to investigate whether the group allocentric weights of the different conditions differed significantly from zero. To investigate whether the scene configuration influenced the use of the allocentric information for memory guided reaching, we performed two-sided t-tests for independent samples on allocentric weights from the RO same condition of the current study and the corresponding conditions (conditions when five task-relevant objects were shifted in the same direction) of our previous study published by Klinghammer et al. (2015). In order to assess the impact of scene coherence and thus, the reliability of allocentric cues on reaching endpoints and their variability, we first conducted two-sided one-sampled t-tests for the allocentric weights and standard deviations of reaching errors for the conditions RO+IO same and RO+IO diff. Second, we performed a two-way repeated measures ANOVA for conditions RO same, RO diff, IO same, and IO diff with the two factors shift coherence (same or different shift direction) and object relevance (shifted objects are potential reach targets or not). In case of significant main effects or interactions, we conducted two-sided post-hoc t-tests for paired samples.

4.3 Results

To investigate the participants‘ encoding behavior, we created fixation density maps of the encoding scene for every object arrangement. As an example, we depict one representative heatmap for one exemplary object arrangement in figure 4.4. Fixation density is highest at the locations of task-relevant objects (butter dish, ceiling lamp, clock, espresso cooker, vase, Vegemite jar) whereas it is lower at locations of task-irrelevant objects. The heatmaps of the other object arrangements showed a very similar pattern.

Figure 4.4. Fixation density map with averaged fixations of all participants during the encoding phase for one example object arrangement. On the right side, the task-relevant objects of the experiment are depicted (butter dish, ceiling lamp, clock, espresso cooker, vase, Vegemite jar). Participants show higher fixation frequencies for these task-relevant objects than for the remaining task-irrelevant objects.

In figure 4.5, we illustrate reaching errors of one exemplary participant in the different conditions averaged over the 18 object arrangements. Reaching endpoints in conditions RO same and RO+IO same deviated systematically in the direction of object shifts. RO diff and RO+IO diff also showed horizontal reaching errors, but these errors were independent of the direction of the object shifts. The conditions IO same and IO diff hardly demonstrated deviations of reaching endpoints indicating that in these conditions object shifts had a negligibly influence on reaching behavior.

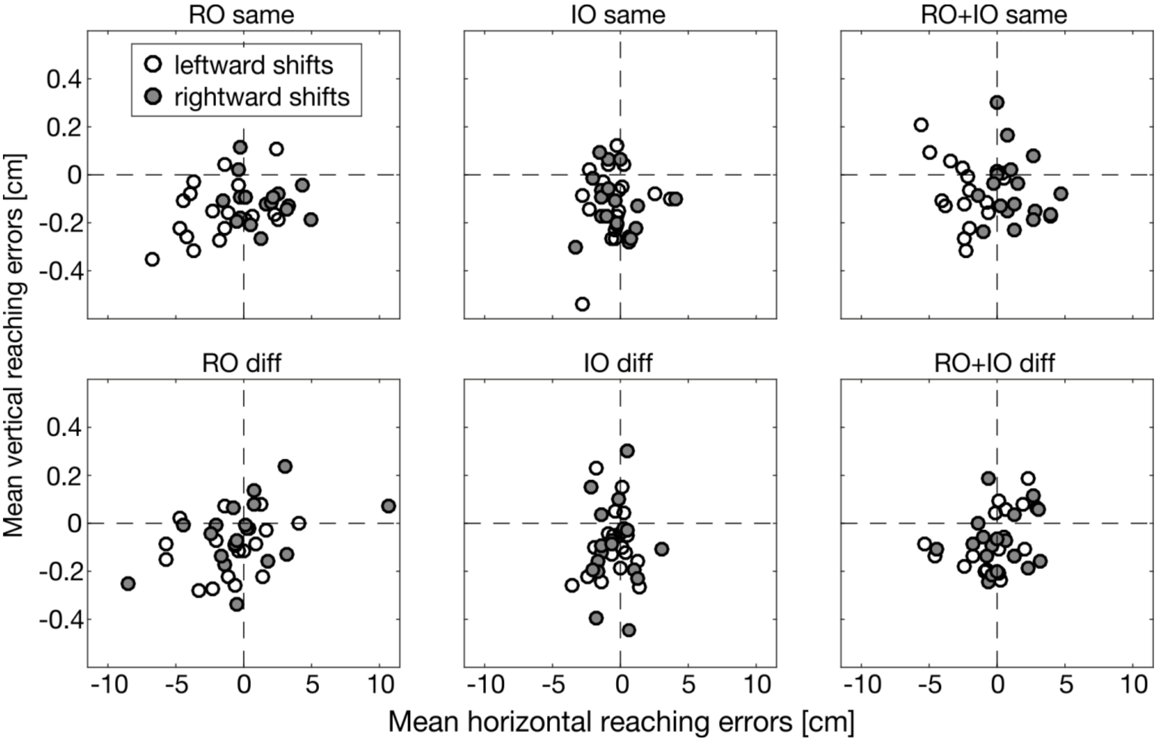

Figure 4.5. Horizontal and vertical reaching errors (in cm) of one exemplary participant averaged across the different object arrangements. Leftward object shifts are represented by white, rightward objects shifts by dark gray dots. Please note the small values on the ordinate.

Figure 4.6A depicts the actual reaching errors and the corresponding MERE of one prototypical participant for the condition RO same for leftward (negative values) and rightward (positive values) object displacements. The slope of the regression line defined the allocentric weight of the respective condition. On the group level, allocentric weights of RO same and RO+IO same significantly differed from zero (RO same: $t(16) = 3.885$, $p = .007$; RO+IO same: $t(6.021)$, $p < .001$), whereas allocentric weights of the other conditions did not (table 4.3).

Table 4.3. Summary of allocentric weights per condition. Range, mean and standard deviation of the sample are listed. Results of two-sided one-sampled t-tests are Bonferroni-Holm corrected. Significant results are indicated by asterisks (**: $p < .01$; ***: $p < .001$).

Condition	Range	Mean	SD	t-Test results
RO same	-0.08 – 0.46	0.12	0.13	$t(16) = 3.885$, $p = .007$**
RO diff	-0.14 – 0.13	0.01	0.10	$t(16) = 0.594$, $p = .561$
IO same	-0.15 – 0.29	0.03	0.10	$t(16) = 1.388$, $p = .369$
IO diff	-0.10 – 0.18	0.05	0.09	$t(16) = 2.061$, $p = .224$
RO+IO same	-0.00 – 0.56	0.20	0.14	$t(16) = 6.021$, $p < .001$***
RO+IO diff	-0.17 – 0.18	0.04	0.09	$t(16) = 1.852$, $p = .248$

The pairwise comparisons of the allocentric weights of the condition RO same with the corresponding conditions of our previous experiments (Klinghammer et al., 2015), in which we also shifted five task-relevant objects, revealed smaller allocentric weights for the current experiment (current vs. previous - shift of 5 table objects (TO-5): $t(16.015) = -4.710$, $p < .001$; current vs. previous – shift of 5 background objects (BO-5): $t(14.612) = -6.123$; $p < .001$; figure 4.6B).

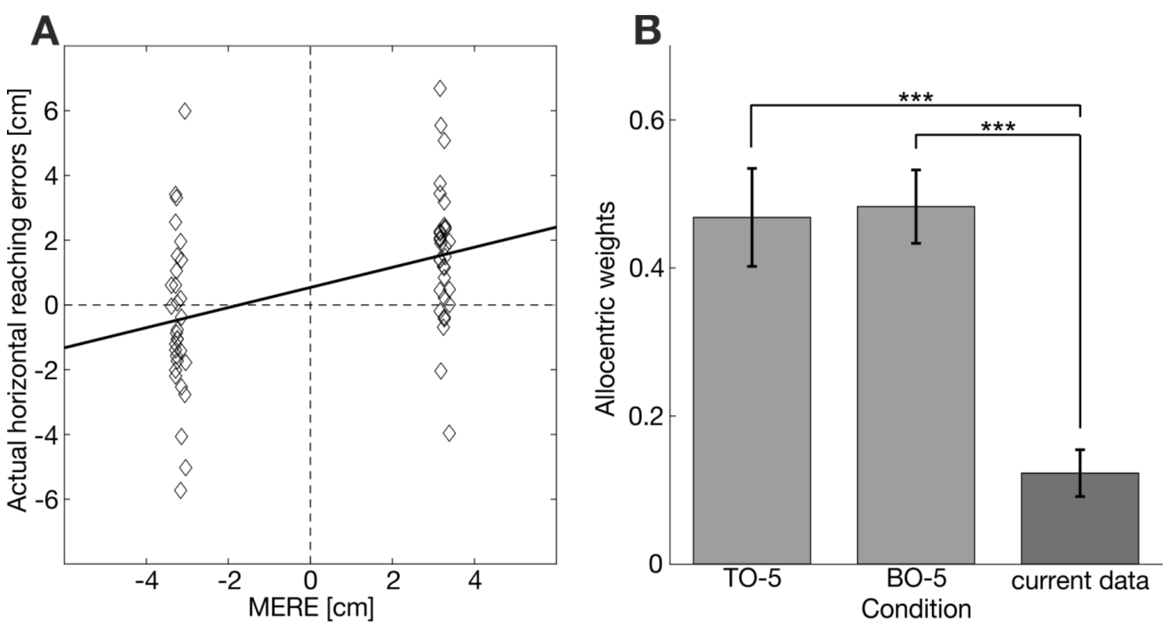

Figure 4.6. (**A**) Example of a linear fit between MERE and actual horizontal reaching errors for one example participant for the condition RO same. The slope of the fit defines the allocentric weight for this participant in this condition. (**B**) Mean allocentric weights from our previous study (Klinghammer et al., 2015; lighter gray bars) where five table objects (TO-5) or five background objects (BO-5) were shifted in the same direction and our current experiment in which also 5 objects distributed across the table and the background were shifted. Asterisks indicate significant differences (***: $p < .001$).

To assess the influence of the reliability of task-relevant and task-irrelevant objects on their use for allocentric coding, we compared conditions with coherent and incoherent shifts of task-relevant and/or task-irrelevant objects. The results of the allocentric weights and the standard deviations of the reaching endpoints are illustrated in figure 4.7A and B. First, we performed a two-sided one-sampled t-test for the allocentric weights and standard deviations of the conditions RO+IO same and RO+IO diff in which we shifted both task-relevant and task-irrelevant objects. We found that allocentric weights differed between the conditions ($t(16) = 4.010$, $p = .001$) showing higher allocentric weights when relevant and irrelevant objects were shifted coherently in the same direction than shifted incoherently in opposite directions. The standard deviations did not differ between RO+IO same and RO+IO diff ($t(16) = 0.256$, $p = .801$). Second, we conducted a two-way repeated measures ANOVA for the conditions RO same, RO diff, IO same, and IO diff. We found a main effect of shift coherence ($F(1,16) = 7.165$, $p = .017$). Coherent object shifts led to higher allocentric weights (*mean* = .078) than incoherent shifts (*mean* = .030). We also observed an interaction of shift coherence and object relevance ($F(1,16) = 5.485$, $p = .032$). Post-hoc t-tests indicated that allocentric weights of RO same were higher than of RO diff ($t(16) = 2.872$, $p = .022$), but did not differ between IO same and IO diff ($t(16) = 0.465$, $p = .648$). Two-way repeated measures ANOVA for standard deviations of reaching endpoints revealed main effects of shift coherence ($F(1,16) = 1.086$, $p = .019$, *mean* coherent shifts = 1.982, *mean* incoherent shifts = 2.147) and object relevance ($F(1,16) = 86.097$, $p < .001$, *mean* relevant object shifts = 2.430, *mean* irrelevant object shifts = 1.699). These main effects were further restricted by an interaction of shift coherence and object relevance ($F(1,16) = 6.977$, $p = .018$). Post-hoc t-tests indicated that standard deviations of reach endpoints in the condition RO same were smaller than in the condition RO diff ($t(16) = -2.880$, $p = .022$), while they were comparable for the conditions IO same and IO diff ($t(16) = -0.243$, $p = .811$). Table 4.4 summarizes the corresponding descriptive statistics.

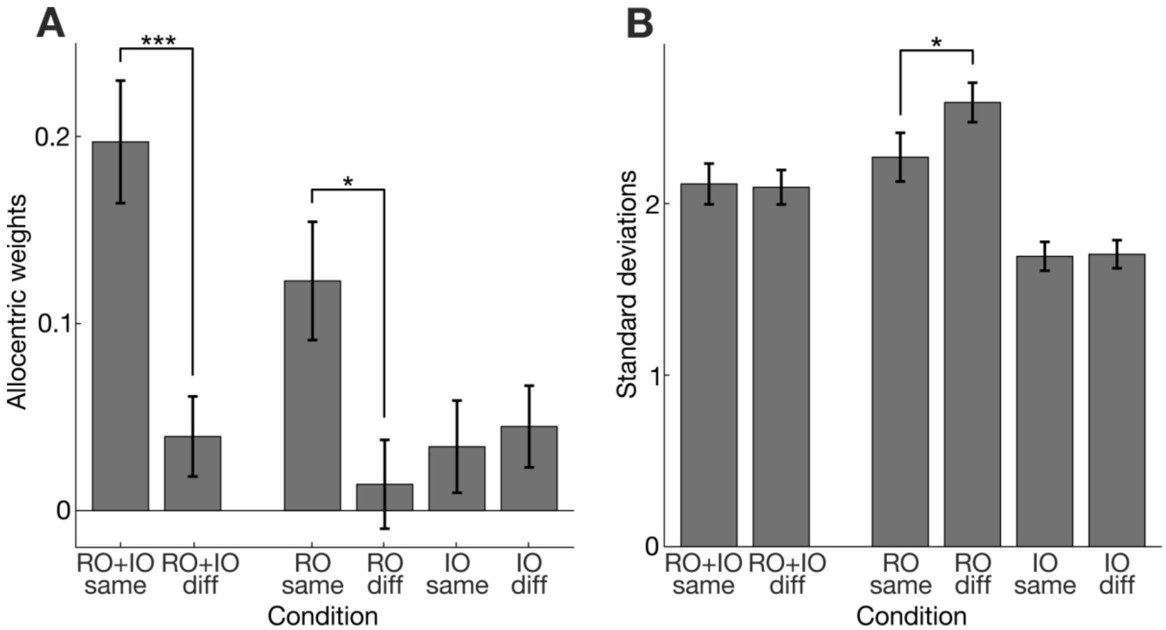

Figure 4.7. (**A**) Allocentric weights averaged over participants. (**B**) Standard deviations of reaching errors in cm averaged over participants. Error bars indicate 1 *SEM*. Asterisks indicate significant differences (*: $p < .05$; ***: $p < .001$).

Table 4.4. Summary of standard deviations in cm. Range, mean and standard deviation of the sample are listed.

Condition	Range	Mean	SD
RO same	1.18 – 3.42	2.27	0.59
RO diff	1.22 – 2.63	1.69	0.35
IO same	1.43 – 3.54	2.12	0.49
IO diff	1.46 – 3.11	2.10	0.41
RO+IO same	1.85 – 3.54	2.59	0.47
RO+IO diff	1.26 – 2.74	1.71	0.34

4.4 Discussion

In the current study, we aimed to investigate how the scene configuration influences allocentric coding for memory-guided reaching. Second, we wanted to know whether and how the reliability of allocentric information, in particular the reliability of task-relevant and task-irrelevant objects, affects their use as allocentric cues. We found that reaching errors systematically deviated in the direction of object shifts, but only when the objects were task-relevant and their reliability was high. However, this effect was substantially reduced when task-relevant objects were distributed across the scene. No deviations of reach endpoints were observed in conditions with shifts of only task-irrelevant objects or with low object reliability irrespective of task-relevancy. Moreover, when solely task-relevant objects were shifted incoherently, the variability of reaching endpoints increased compared to coherent shifts of task-relevant objects.

In order to check for differences in participants' encoding strategies compared to our previous experiments, we created fixation density maps with mean fixation densities for every object arrangement during the encoding phase. We observed that task-relevant objects were fixated more frequently than task-irrelevant objects. This supports our previous findings (Fiehler et al., 2014; Klinghammer et al., 2015) showing that overt visual attention is mainly distributed across areas containing task-relevant information. Thus, participants used the same encoding strategy as in our previous studies even though we arranged task-relevant objects in a distributed instead of a clustered fashion. Our results are in line with work on overt visual attention showing that participants' fixation behavior is highly task-dependent with more fixations on task-relevant objects (Ballard & Hayhoe, 2009; DeAngelus & Pelz, 2009; Land & Hayhoe, 2001).

Overall, we found allocentric weights different from zero in the RO same and RO+IO same condition suggesting that allocentric information of task-relevant objects was used for memory-guided reaching. Weights in these conditions ranged from 0.12 to 0.22 (see table 4.3) which indicates that reaching endpoints were affected by up to 22 % by the shifted task-relevant objects. The remaining percentage could be attributed to the influence of egocentric or additional allocentric representations as the environment also provided other, more stable landmarks, e.g., edge of the screen, edge of the black cardboard, real table. In comparison to our previous experiments (Fiehler et al., 2014; Klinghammer et al., 2015), allocentric weights were relatively small, even in the condition where we kept the object arrangement maximally coherent (RO+IO same). This might be explained by the randomized presentation of the experimental conditions probably leading to a reduced perceived reliability of allocentric information. That means, trials with incoherent object shifts (i.e., RO+IO diff, RO diff) might have also reduced the reliability in trials with coherent object shifts (i.e., RO+IO same, RO same) leading to an overall decrease in the use of allocentric cues. Future studies should consider using a blocked instead of a fully randomized design to avoid potential carry-over effects of low reliability on high reliability trials.

The lower allocentric weights in the present study were confirmed by the comparison of the RO same condition to the corresponding conditions of our previous study (Klinghammer et al., 2015) in which we also coherently shifted five relevant objects. A main difference between these experiments is the spatial configuration of task-relevant and task-irrelevant objects in the scene (distributed vs. grouped). On the first view, the lower allocentric weights for the distributed object configuration seem to confirm other findings showing that participants preferably rely on landmarks in the direct vicinity of the target (Diedrichsen et al., 2004; Spetch, 1995). By placing task-irrelevant objects as close to the target as task-relevant ones, the former may have served as additional but misleading allocentric cues. However, we found that task-relevant but not task-irrelevant objects differed from zero suggesting that task-relevancy but not landmark vicinity determined to which extent objects were used as allocentric cues. This is also supported by the lower reaching endpoint variability in IO same than RO same ($t(16) = 6.000$; $p < .001$). Moreover, the fixation behavior during the encoding phase indicates that participants actually ignored task-irrelevant objects and mainly focused on task-relevant ones. In contrast to clustering the objects on the table or in the background, the distributed object configuration we applied here also increased the distance between the target and the task-relevant objects that may have decreased their use as allocentric cue. This is in line with previous findings showing that the influence of allocentric information decreases with an increasing distance between target and landmarks (Camors et al., 2015; Krigolson et al., 2007). However, we cannot exclude an impact of an overall lower reliability of the allocentric information in the current experiment that arises from the trial randomization as discussed earlier.

By introducing conditions in which we shifted objects coherently in the same direction or incoherently in opposite directions, we were aiming to gain insights on how the reliability of allocentric information affects its use for coding target locations for memory-guided reaching. As expected, we found the highest allocentric weights in the condition in which the coherence of the object relations was maintained between the encoding and the test scene (RO+IO same). Surprisingly, by shifting task-irrelevant objects in the opposite direction of task-relevant ones (RO+IO diff), the allocentric weights were dramatically reduced and did not differ from zero. Even though these objects were task-irrelevant and thus, could be ignored, they had a clear impact on the use of task-relevant objects as allocentric cues. It is likely that by breaking up the coherence of the whole object arrangement, the reliability of the allocentric information was reduced and as a consequence, participants rather relied on egocentric or other, more stable allocentric target representations. We found a similar result pattern when we shifted only task-relevant objects in the same or in different directions (RO same, RO diff). As predicted, allocentric weights were higher when the coherence within the group of task-relevant objects was kept intact compared to the condition in which it was broken with allocentric weights being not different from zero. Again, this supports a stronger use of allocentric information when its reliability is high. In contrast to the findings by Byrne & Crawford (2010), we found a higher variability of reaching endpoints for low compared to highly coherent object configurations. This further strengthens an important role of the reliability of task-relevant allocentric information for memory-guided reaching movements. Interestingly, we observed increased allocentric weights for RO+IO same compared to RO same ($t(16) = -3.112$, $p = .007$). Even though, a coherent shift of irrelevant objects alone had no influence on reaching endpoints (i.e., no difference of IO same from zero), they may have increased the overall object reliability when they were shifted coherently with task-relevant objects. All in all, it can be concluded that if the reliability of only task-relevant objects is decreased, their use as allocentric cues is substantially reduced. Even though, task-irrelevant objects are not directly used as allocentric cues, they seem to increase the overall contribution of allocentric information by increasing the allocentric reliability when shifted coherently with task-relevant objects. Our results could be explained within the framework of causal Bayesian integration (Körding, Beierholm, Ma, Quartz, Tenenbaum, & Shams, 2007; Sato, Toyoizumi, & Ahira, 2007) in which two different target modalities are combined in a Bayes-optimal way. This framework further assumes that the weightings of the target modalities are modulated by the probability of both sharing one source or not. Transferred to our paradigm it is reasonable that by shifting objects in an incoherent manner, the causal link between the target location and positions of the other objects (in terms of their incoherent spatial relation) is broken. In that case, causal Bayesian integration discounts the allocentric information by the remaining non-target objects leading to a low weighting of the allocentric information. Moreover, in incoherent shift conditions the variability associated with the allocentric information increases, which in turn further decreases its weighting. In contrast, if objects are shifted in a coherent way, the causal link between the location of the target and the other objects is maintained leading to a lower variability and thus, higher weighting of the allocentric information.

Overall, our findings demonstrate that the use of allocentric information for memory-guided reaching depends on task-relevancy and the scene configuration, in particular the average distance of the reach target to task-relevant objects. If task-relevancy is given, the reliability of allocentric information determines to which extent allocentric cues are used to code the location of memorized reach targets. Less reliable allocentric cues contribute to a lesser extent and lead to an increased variability of memory-guided reaching movements. However, the reliability of task-relevant allocentric information seems to be further increased by task-irrelevant allocentric cues if they act coherently.

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

A similar version of this manuscript is currently under revision as:
Klinghammer, M., Schütz, I., Blohm, G., & Fiehler K. (under review). Allocentric information is used for memory-guided reaching in depth: a virtual reality study.

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.

5.1 General 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 & Heath, 2004; Krigolson, Clark, Heath, & Binsted, 2007; Obhi & Goodale, 2005) argueing 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.

5.2 Experiment 1

5.2.1 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.

5.2.2 Methods

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 OPTICAL 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.

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 x 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.

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 x width x depth = 71 cm x 80 cm x 80 cm; see also figure 5.1A). 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 5.1). 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 figure 5.1). 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).

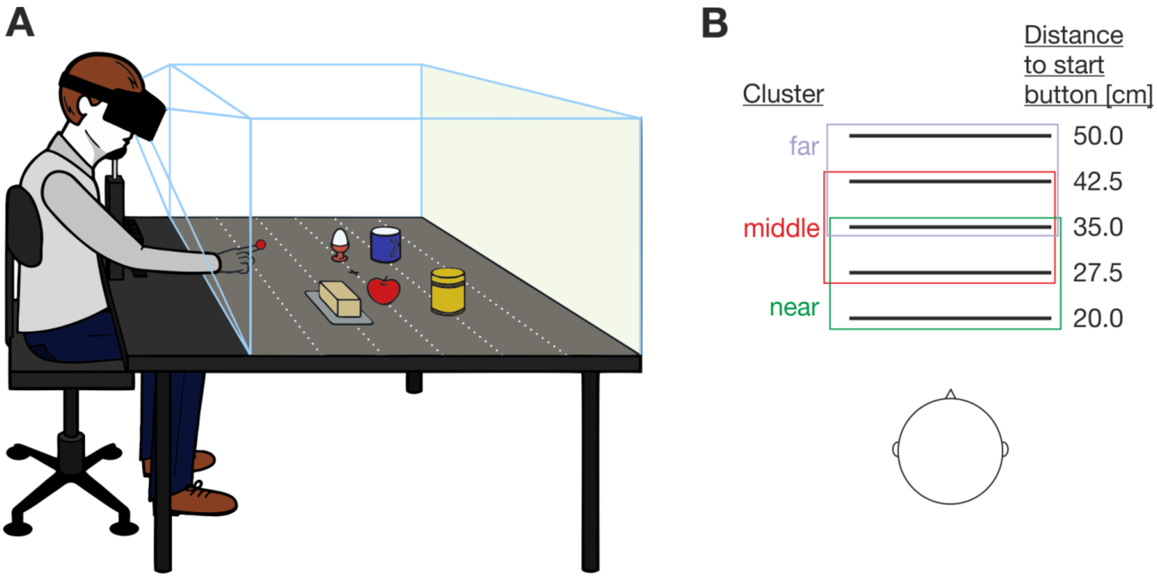

Figure 5.1. (**A**) Schematic view of the VR setup. Participants wore the HMD and sat at a table with their head rested on a chin rest. The presented virtual environment (i.e., table, objects) was aligned with the real world properties of the table. Thus, when participants performed a reaching movement 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.

Table 5.1. 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

Procedure

Figure 5.2 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.

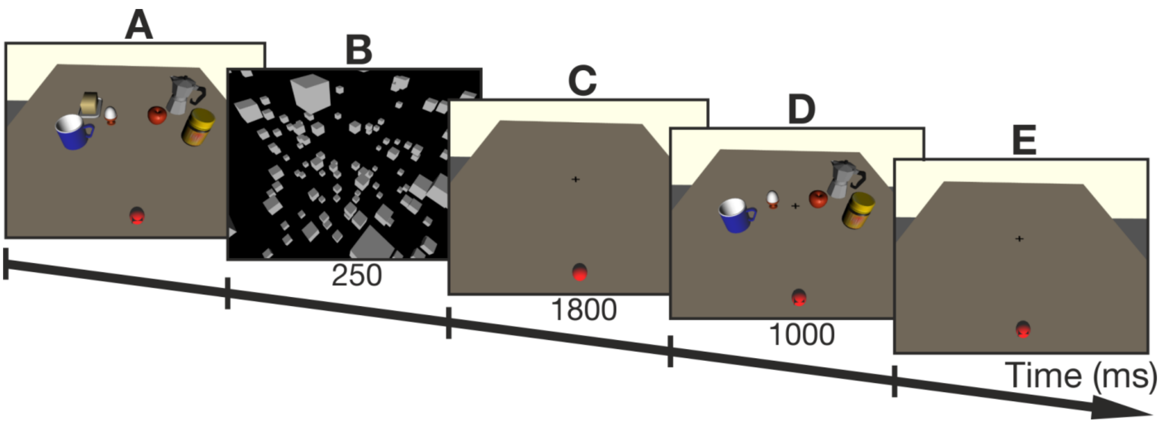

Figure 5.2. Time course of an example trial. The presentation of the encoding image (**A**) was followed by a mask (**B**) to prevent afterimages. After a delay from which on a fixation cross appeared in the middle of the depth cluster (**C**) the test scene was shown (**D**) with one target object missing (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**).

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 .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 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/s^2. 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 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.

5.2.3 Results

In table 5.2, 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.

Table 5.2. 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 – 1.401	- 1.526	2.564	- 3.355
	right	0.830 – 4.026	1.707	2.388	3.345
middle	left	- 3.570 – - 0.802	- 2.040	2.288	- 3.731
	right	- 0.484 – 3.220	1.792	2.426	3.655
far	left	- 3.535 – - 0.629	- 2.139	2.831	- 4.150
	right	- 0.216 – 3.526	2.143	2.621	4.150

As depicted in figure 5.3A, averaged reaching endpoint errors for single participants deviated systematically in the direction of horizontal object shifts. Figure 5.3B shows the linear fit between MEREs and actual reaching errors of one exemplary participant for horizontal object shifts in the second depth cluster.

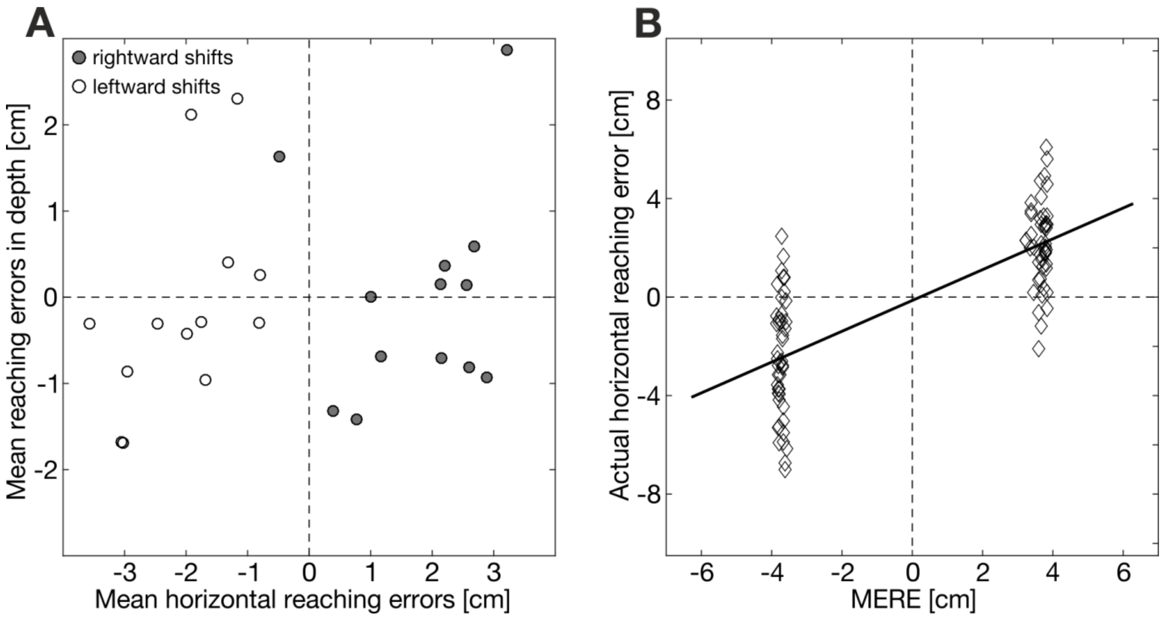

Figure 5.3. (**A**) Mean reaching errors for single participants in cm for horizontal object shifts in depth cluster 2. (**B**) Example of a linear fit between MEREs and actual horizontal reaching errors for one participant for horizontal object shifts in the second depth cluster.

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 5.3).

Table 5.3. 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	$t(11) = 8.627$, $p < .001$
middle	0.20 – 0.80	0.52	0.18	$t(11) = 10.213$, $p < .001$
far	0.20 – 0.72	0.52	0.16	$t(11) = 11.605$, $p < .001$

Mean allocentric weights over all clusters of the current study and allocentric weights of corresponding conditions of our previous research are depicted in figure 5.4A. 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: $t(20.385) = 0.423$, $p = .656$; current vs. BO-5 in previous Exp. 2: $t(18.898) = 0.325$, $p = .749$).

One-way repeated measures ANOVA assessing the influence of distance between target and observer obtained no main effect on allocentric weights ($F(2,24) = 0.997$, $p = .359$). 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* = .021; see also figure 5.4B). Post-hoc t-tests showed only differences between the middle and far depth clusters (*t*(12) = -3.493, *p* = .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* > .176).

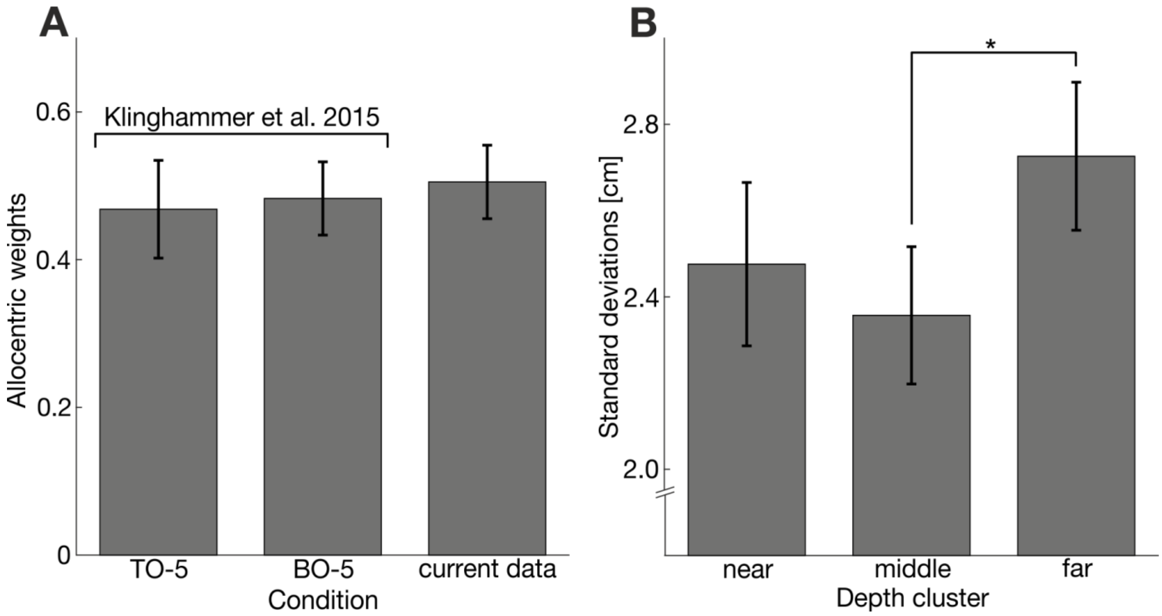

Figure 5.4. (**A**) Mean allocentric weights from our previous study where five table objects (TO-5) or five background objects (BO-5) were shifted and our current experiment which also shifts five table objects. (**B**) Mean standard deviations of reaching errors for the three depth clusters in cm. Error bars represent 1 *SEM* and asterisks indicate significant differences (*: *p* < .05).

The one-way repeated measures ANOVA assessing the influence of depth clusters on reaction times revealed no differences (*F*(2,24) = 0.671, *p* = .521; *mean near / middle / far Cluster:* 284.254 ms / 278.567 ms / 281.031 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* < .001; *mean near / middle / far Cluster:* 552.559 / 603.601 / 662.414). Post-hoc t-tests revealed longer movement times for the depth clusters middle than near (*t*(12) = -12.145, *p* < .001), far than near (*t*(-12) = -12.109, *p* < .001), and far than middle (*t*(12) = -10.283, *p* < .001).

5.3 Experiment 2

5.3.1 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 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 *conflict 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 et al., 2011, 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.

5.3.2 Methods

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 instructions (> 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.

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 5.1). 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 figure 5.5. 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.

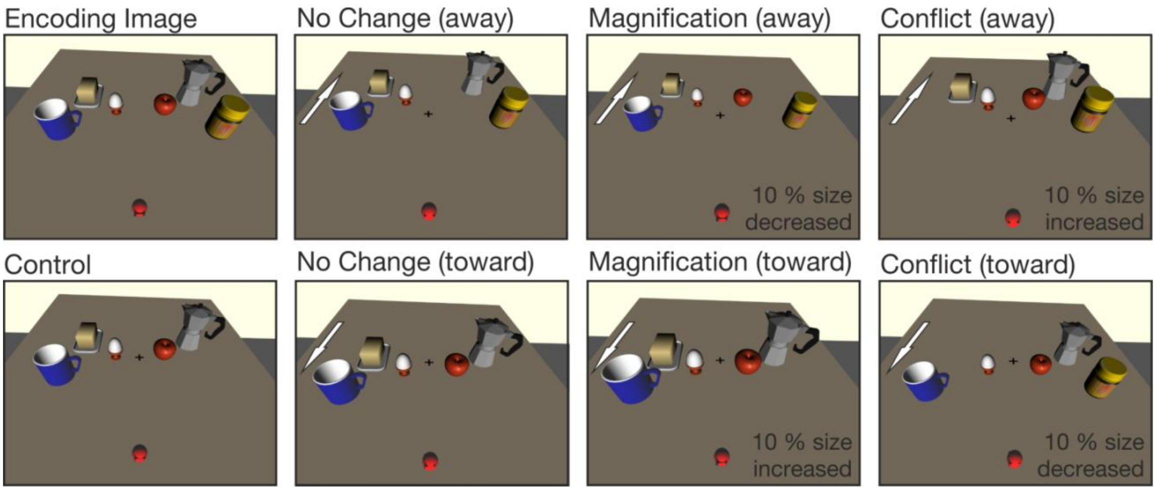

Figure 5.5. Example images of an encoding and corresponding test scenes for the different conditions. The arrow indicates the direction of object shifts.

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.

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 11232 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 differed 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.

5.3.3 Results

In table 5.4, 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.

Table 5.4. 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 – 1.729	- 2.423	2.010	- 3.767
	backward	0.452 – 4.289	2.513	2.205	4.265
magnification	toward	- 4.109 – 1.057	- 2.459	2.027	- 3.764
	backward	1.185 – 4.674	2.853	2.360	4.280
conflict	toward	- 3.627 – 0.663	- 2.468	2.116	- 3.764
	backward	1.249 – 4.268	2.508	2.231	4.272
middle cluster					
no change	toward	- 4.463 – - 0.635	- 3.194	2.036	- 4.597
	backward	2.027 – 4.540	3.123	2.086	5.318
magnification	toward	- 5.013 – - 0.531	- 3.423	2.100	- 4.596
	backward	1.665 – 5.101	3.441	2.048	5.319
conflict	toward	- 4.349 – - 0.823	- 3.194	2.142	- 4.593
	backward	1.439 – 4.893	3.087	2.126	5.321
far cluster					
no change	toward	- 6.051 – - 1.372	- 3.916	2.520	- 5.622
	backward	2.851 – 5.987	4.220	2.574	6.773
magnification	toward	- 6.284 – - 1.179	- 4.001	2.466	- 5.623
	backward	3.190 – 7.104	4.471	2.979	6.754
conflict	toward	- 5.832 – - 1.262	- 3.532	2.383	- 5.610
	backward	1.866 – 5.952	3.467	3.329	6.761

As indicated in figure 5.6A, averaged reaching endpoint errors for single participants deviated systematically into the direction of object shifts in depth. As an example, in figure 5.6B 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.

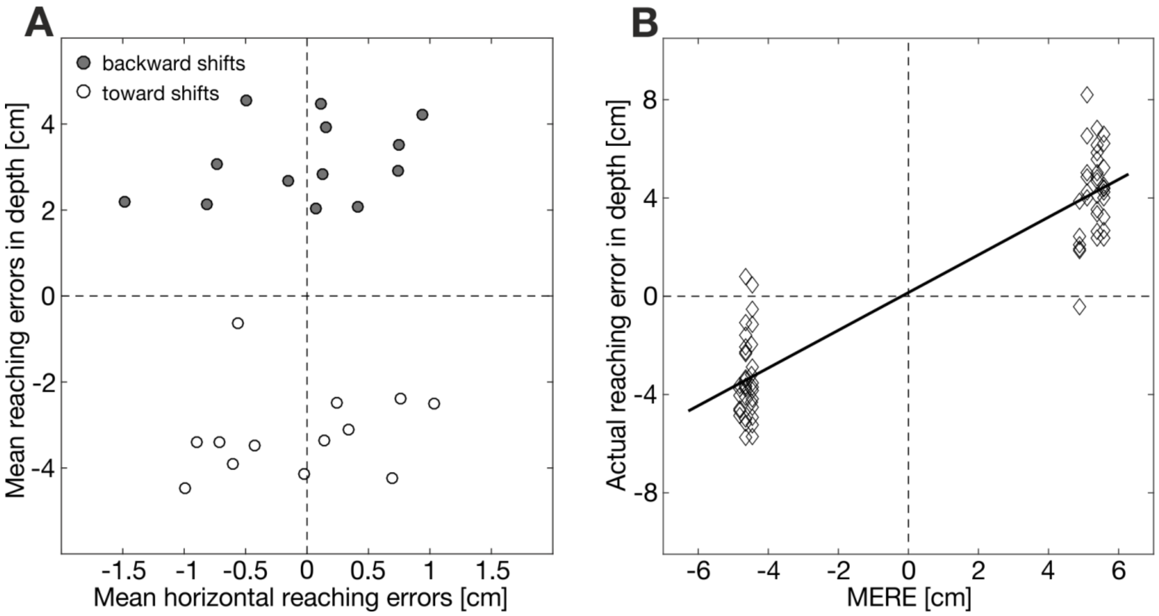

Figure 5.6. (**A**) Mean reaching errors for all single participants in cm for object shifts in depth (middle depth cluster, no change condition; backward = away from the participant; toward = in the direction of the participant). (**B**) Example of a linear fit between MEREs and actual reaching errors 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.

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 5.5).

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 ($F(1,72) = 13.094$, $p < .001$). Allocentric weights in experiment 2 were higher than in experiment 1. There was no main effect for depth cluster ($F(2,72) = 0.407$, $p = .667$) and no interaction of the two factors ($F(2,72) = 0.041$, $p = .960$).

Table 5.5. 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	$t(12) = 14.681$, $p < .001$
magnification	0.44 – 0.92	0.66	0.14	$t(12) = 16.937$, $p < .001$
conflict	0.45 – 0.81	0.62	0.10	$t(12) = 22.406$, $p < .001$
middle cluster				
no change	0.35 – 0.81	0.64	0.11	$t(12) = 17.735$, $p < .001$
magnification	0.46 – 0.85	0.69	0.11	$t(12) = 22.240$, $p < .001$
conflict	0.47 – 0.80	0.63	0.11	$t(12) = 21.318$, $p < .001$
far cluster				
no change	0.46 – 0.85	0.66	0.11	$t(12) = 20.832$, $p < .001$
magnification	0.41 – 0.97	0.69	0.14	$t(12) = 17.276$, $p < .001$
conflict	0.40 – 0.80	0.57	0.11	$t(12) = 18.994$, $p < .001$

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 ($F(2,24) = 22.169$, $p < .001$, see figure 7A). Post-hoc t-tests revealed higher weights in the magnification compared to the no change condition ($t(12) = -5.083$, $p < .001$), higher weights in the no change than conflict condition ($t(12) = 2.71$, $p = .019$), and higher weights in the magnification than conflict condition ($t(12) = 5.576$, $p < .001$). We did not find a main effect for the distance between target and observer ($F(2,24) = 0.331$, $p = .721$) but an interaction between the two factors ($F(4,48) = 3.464$, $p = .014$). However, one-way repeated measures ANOVAs for each condition with the factor cluster on allocentric weights failed statistical significance after correction for multiple testing (*p*s > .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 ($F(2,24) = 32.577$, $p < .001$, see figure 5.7B). Post-hoc t-tests revealed a higher variability for far than near depth clusters ($t(12) = -5.906$, $p < .001$) for far than middle depth cluster ($t(12) = -6.689$, $p < .001$), but not near and middle ($t(12) = 1.089$, $p = .298$). We neither found a main effect for the object size manipulation ($F(2,24) = 5.564$, $p = .056$) nor an interaction between the two factors ($F(4,48) = 1.141$, $p = .349$).

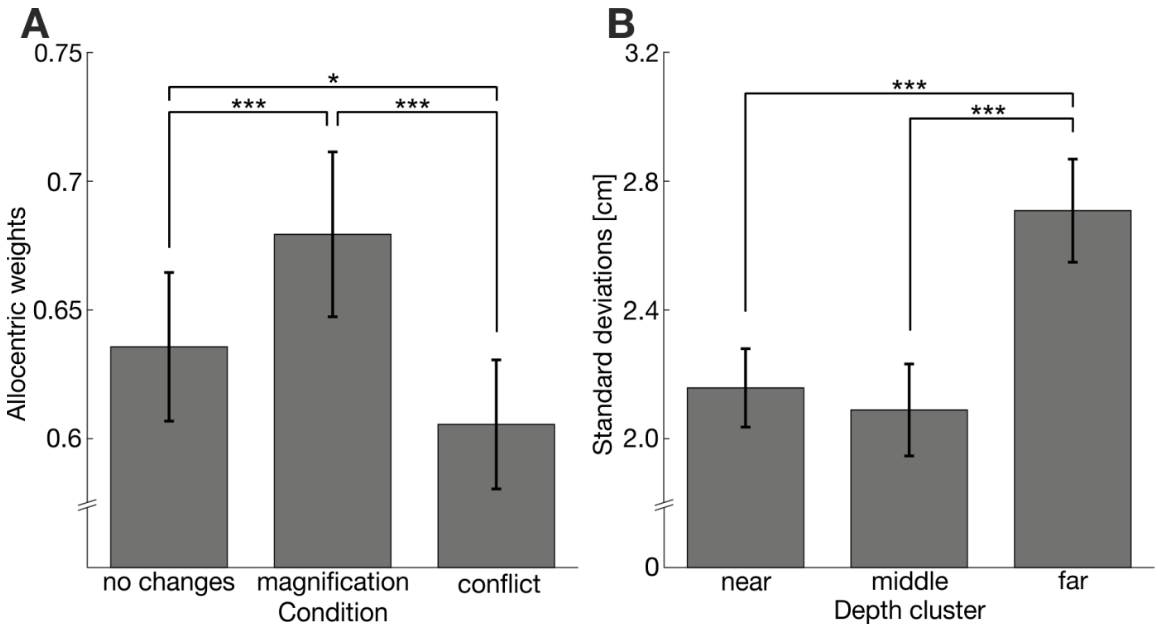

Figure 5.7. (**A**) Mean allocentric weights for the three object size manipulation conditions. (**B**) Mean standard deviations of reaching errors for the three depth clusters in cm. Error bars represent 1 *SEM* and asterisks indicate significant differences (*: p < .05; ***: p < .001).

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 ($F(2,24) = 4.696$, $p = .019$; *mean near / middle / far Cluster:* 289.157 ms / 283.785 ms / 295.781 ms). However, post-hoc t-tests did not reach significance (all $p >$ .095). The two-way repeated measures ANOVA on movement durations revealed only a main effect for the factor depth cluster ($F(2,24) = 117.407$, $p < .001$; *mean near / middle / far Cluster:* 565.847 / 612.491 / 664.274). Post-hoc t-tests revealed higher movement durations for the depth clusters middle than near ($t(12) = -8.786$, $p < .001$), far than near ($t(-12) = -11.285$, $p < .001$), and far than middle ($t(12) = -11.687$, $p < .001$).

5.4 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, Henriques, & Fiehler, 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 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 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 5.3), 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 egocentric representations. 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 5.5, 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 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 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, 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 et al., 2011, 2015; Sousa et al., 2011, 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 et al., 2011, 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 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.

6 Discussion

6.1 Summary

The aim of the presented research was to investigate the use of spatial reference frames for memory-guided reaching movements in a naturalistic and complex real-world-like scenario. To this end, I conducted four studies, which provided new insights in this topic. Here I will briefly summarize the main results and points of discussion, which were discussed in more detail in the corresponding study chapters.

In the first study, I wanted to answer the basic question whether egocentric and allocentric information is integrated for memory-guided reaching in a complex environment. I found reaching errors that varied systematically in the direction of object shifts, but only when objects on the table were shifted. In this case, this effect scaled with the number of objects being shifted. In contrast, if background objects were shifted, no influence on reaching endpoints was observed. By calculating allocentric weights, it can be concluded that the allocentric information affected the movement plan by up to 43 % if all five objects on the table were shifted. The remaining 57 % could be attributed to the influence of the egocentric reference frame or other, more stable allocentric cues in the environment, such as the table in the scene but also real world cues like the frame of the computer screen. It is reasonable that allocentric and egocentric information is integrated for coding a target location for memory-guided reaching, which confirms results from previous studies (Byrne & Crawford, 2010; Sheth & Shimojo, 2004). The finding that only object shifts on the table but not in the background (*global objects* in this experiment) affected participants' behavior raised two possible explanations. First, the changes of background objects were undersized, as only one background object was shifted at a time. Second, participants never attended background objects because they never served as potential reach targets and thus, were irrelevant to perform the task. The goal of the second study was to test for these two possibilities.

In the first experiment of my second study only objects on the table served as potential reach targets. When table objects were shifted, I found reaching errors with a similar magnitude as in the first study. However, shifts of background objects had no influence on reaching behavior, even though up to five objects were shifted. In the second experiment, the same numbers of objects were shifted but instead of table objects, background objects served as reach targets. I found a reversed pattern with reaching endpoints being affected by up to 48 % by the allocentric information of the background object shifts which also scaled with the number of shifted objects in a very similar pattern compared to the first study. In contrast, this time I did not observe an influence of table object shifts. This suggests that the task-relevance of objects (i.e. whether they are potential reach targets or not) is an important factor that determines their use as allocentric cues. If task-relevance is given, the magnitude of the influence of object shifts scales with the number of shifted objects. It seems that participants mainly attend to the location of task-relevant objects which is supported by heatmaps that represent averaged fixation densities during the phase when participants encoded object locations freely with their gaze. Here, I observed that participants mainly fixated the area which contained task-relevant objects, namely the table in experiment 1 and the background in experiment 2. Thus, task-irrelevant objects are rather ignored and not used as allocentric cues. The importance of task-relevance related with a more frequent fixation of task-relevant objects is supported by research on overt visual attention in real world and natural scenes in which participants also mainly fixated objects that were relevant to perform a given task while ignoring other objects (Ballard & Hayhoy, 2009; DeAngelus & Pelz, 2009; Land & Hayhoy, 2001; Mills, Hollingworth, Van der Stigchel, Hoffman, & Dodd, 2011).

In my third study, I aimed to discriminate further factors that determine the use of objects as allocentric cues for memory-guided reaching. Compared to my second study I placed task-relevant and task-irrelevant distributed over the whole scene instead of clustered on the table or in the background. In order to manipulate the reliability of the allocentric information, I shifted objects coherently in one direction (high reliability) but also incoherently in opposite directions (low reliability). I found an overall reduction of reaching errors due to object shifts. This was also true for conditions in which all objects were shifted coherently in one direction and thus, provided maximal reliability. Here, reaching endpoints were only affected about 22 % by the allocentric information. This finding can be attributed to the increased average distance between reach target and surrounding task-relevant objects, which were used as allocentric cues. Previous studies already found such an effect of distance between target and allocentric cues (Camors et al., 2015; Krigolson et al., 2007), supporting my interpretation. In addition, I found a much higher influence of allocentric information on reaching behavior in conditions in which objects were shifted coherently compared to conditions with incoherent object shifts. As incoherent object shifts reduced the reliability of the allocentric information, participants put less weight on this unreliable cue and therefore, rather used the egocentric reference frame (see also chapter 6.2) or maybe also other, more stable allocentric cues (i.e. edge of the computer screen, the table participants were sitting at). In summary, the average distance between the reach target and the task-relevant allocentric cues and the overall reliability of the allocentric information were identified as further factors that determine the use of objects as allocentric cues for memory-guided reaching in a complex environment.

The goal of my fourth study was to transfer this paradigm into a more naturalistic 3D virtual reality setting and thereby investigating the use of allocentric information for coding reach target locations in depth compared to the horizontal plane (Coello et al., 2003). Moreover, I wanted to figure out to which extent object size, as an important monocular depth cue (Naceri et al., 2011), contributes to this coding. First, I found reaching errors in the direction of horizontal object shifts that did not differ from reaching errors of the corresponding conditions of my second study. Thus, I validated virtual reality as a new and promising technique for investigating the use of egocentric and allocentric reference frames in a more realistic, three-dimensional environment. Next, I found that reaching movements were more affected by object shifts in depth than object shifts in the horizontal plane. It can be concluded that the allocentric information is used for coding reach target locations in both the depth and the horizontal plane (Neely et al., 2008). However, it seems that two different subsystems are responsible for coding target locations in these planes leading to different results between these two axes, which is in line with previous research (Chieffi & Allport, 1997; Coello & Magne, 2000). Regarding object size as an important depth cue, I found an increased influence of the allocentric information on reaching behavior in the condition with magnified object size changes and a decreased influence in the condition with a conflict between changes in object size and shift direction compared to the condition in which the retinal size of objects changed in a natural way due to object shifts in depth. It can be concluded that object size is an important depth cue for allocentric coding of reach targets in depth. However, the differences between these conditions were rather small. It is thus very likely that other depth cues like object occlusion and binocular disparity are also integrated (Knill, 2007; Landy et al., 1995) and partially compensate for the object size manipulation.

6.2 Optimal integration of egocentric and allocentric information

One striking result that can be found over all my studies is that both egocentric and allocentric information is used for coding target locations. If task-relevance of allocentric cues is given and if these cues are reliable, then participants used the allocentric information from 22 % (third study) up to 66 % (fourth study) for planning and executing the reaching movement. The finding that not one single reference frame is used but both are integrated is supported by previous research (Byrne & Crawford, 2010; Camors et al., 2015; Sheth & Shimojo, 2004). However, the question arises which mechanism regulates the integration process and the weighting of the two reference frames.

One popular framework that has been shown to explain the integration process of two sources of information from different modalities is optimal Bayesian integration (also known as maximum-likelihood estimation; Battaglia, Jacobs, & Aslin, 2003; Ernst & Banks, 2002; Knill, 2007; Knill & Saunders, 2004; Vaziri, Diedrichsen, & Shadmehr, 2006). In this concept, two sources of information are combined by taking the variability of each source into account. In that sense, the weighting of a source with a high variability is decreased whereas the weighting of the source with lower variability is increased, leading to a combined estimate that is closer to the source with the lower variability. However, both sources of information are used to some extent, yielding a lower variability of the combined estimate compared to the variabilities of the two sources. With respect to the integration of different reference frames, the two sources for coding the target location are the egocentric information (i.e. by the gaze vectors from the fixations in the encoding scene) and the allocentric information (i.e. the relative location of objects to each other in the encoding scene). Both of them have a certain variability which depends on several factors like variability of the visual input or the reliability of the allocentric information. Depending on these variabilities, both sources of information are integrated, which occurs likely at the time of action (Byrne, Cappadocia, & Crawford, 2010). As objects were shifted into a certain direction in the test image, the allocentric information needs to be updated. The egocentric information, however, remains unaffected as participants were not allowed to move and had to fixate their gaze on a fixation cross. Thus, the location of the reach target is a weighted combination of the egocentric and the updated allocentric location, which is used to generate a movement plan toward the target location. This can explain my findings that reaching endpoints lay somewhere between the absolute egocentric location (no reaching error) and the absolute allocentric location (reaching errors with the same magnitude as the shifts of task-relevant objects). Moreover, this approach could also account for findings from my third study that a reduction of the reliability of the allocentric information by incoherent object shifts leads to a reduced influence of the allocentric information on reaching endpoints. As a low reliability of a source can be described in terms of a high variability, in my third study an incoherent object shift (low allocentric reliability) would increase the variability of the allocentric information. By integrating egocentric and allocentric information optimally, the high allocentric variability would lead to a lower weighting of this cue, which is actually reflected by the lower allocentric influence on reaching behavior in conditions with incoherent object shifts.

In fact, other studies already applied a Bayesian approach to investigate the integration of egocentric and allocentric information (Byrne & Crawford, 2010; Camors et al., 2015). Byrne and Crawford (2010) aimed to influence the reliability of the allocentric information by adding a vibration of different amplitudes to the landmarks. However, they did not find an influence of the vibration on the variability of the allocentric information, but still the vibration reduced the weighting of the allocentric information when optimally combined with the egocentric information. The authors solved this inconsistency by applying a stability heuristic to their Bayesian model that modulates the integration of egocentric and allocentric information. This heuristic could have been learned from previous experiences that unstable landmarks provide a less useful allocentric cue than stable ones and therefore should be down weighted when integrated with the egocentric information. Also Camors and colleagues (2015) applied a Bayesian approach, but additionally aimed to include the distance between a reach target and the allocentric information in their model. This has been achieved by adding a factor (i.e. prior) to the model that modulates the coupling of the two sources of information. Consequently, they were able to model a decreased allocentric influence with an increased distance between target and allocentric cues.

A further promising extension of the optimal Bayesian integration was suggested by Körding and colleagues (2007; see also Sato, Toyoizumi, & Ahira, 2007). In their proposed framework of causal Bayesian integration, two sources are not only weighted by taking their individual variabilities into account. This approach further modulates these weightings by considering the probability that both sources are sharing a common cause or not. In my third study I found a reduced allocentric influence when objects were shifted in opposite directions. It is possible that because of these shifts the causal link between the target and the object locations (in terms of their incoherent spatial relation) is broken. Thus, following to the causal Bayesian integration, the weighting of the allocentric information would be decreased, which is supported by the low influence of allocentric information in conditions with incoherent object shifts in my third study.

Taken together, optimal Bayesian integration provides a promising framework that could explain the nature of the integration of egocentric and allocentric reference frames for memory-guided reaching. Previous work has shown that by adding different factors to this framework, various aspects such as landmarks' stability or the distance between the target and allocentric cues can be captured (Byrne & Crawford, 2010; Camors et al., 2015). By adding further factors such as task-relevance of objects or reliability of the allocentric information, which determine the use of allocentric cues to the model, one might be able to provide a more accurate description of egocentric and allocentric integration. One of these factors could be the causal link between the target location and the positions of allocentric cues by using causal Bayesian integration (Körding et al., 2007; Sato et al., 2007).

6.3 Memory-guided versus visually-guided reaching movements and the influence of delay

I conducted only experiments in which participants had to perform open-loop (memory-guided) reaching movements. Thus, I cannot provide conclusions about the influence of allocentric information on closed-loop (visually-guided) reaching behavior. Moreover, participants' reaching onsets followed right after the test scene vanished. Thus, I am also not able to draw empirical conclusions about the effect of a delay between stimulus offset and reaching onset. However, some studies that investigated these aspects have been published. Here, I will summarize some findings of these studies on which I will base some predictions about possible implications for my research.

In the study by Krigolson and Heath (2004), participants had to reach to the location of a target-LED which was either presented alone or surrounded by four background-LEDs. The reach was performed either while the LEDs were still illuminated (visually-guided) or after the LEDs were turned off (memory-guided). Furthermore there were two conditions with different delays (0 and 1 s) between switching off of the LEDs and the go-signal to start the reach. Reaching endpoints were more accurate for visually-guided reaches compared to delayed reaches, and more precise for the visually-guided condition compared to the memory-guided and delay conditions. Both results were found regardless of whether allocentric cues (background-LEDs) were present or not. Thus, memory-guided reaches generally seem to be less accurate and precise compared to visually-guided reaching movements. However, Hay and Redon (2006) reported that a delay between target offset and movement onset diminished accuracy of reaching endpoints only when the target was presented alone (egocentric). When additional landmarks were provided during the encoding phase, no decrease in endpoint accuracy was observed. Furthermore, the presence of landmarks during target encoding can also improve the precision of delayed reaching movements as compared to delayed reaching without landmarks (Obhi & Goodale, 2005; Seth & Shimojo, 2004). Thus, if additional landmarks are presented simultaneously with the target, these allocentric cues are used to encode the location of the target. This type of allocentric representation seems to be more stable over time and in turn leads to higher accuracy and precision of reaching movement after a delay compared to an egocentric representation. However, additional landmarks do not necessarily improve accuracy and precision of reaching endpoints with longer delays (8 and 12 s) between target offset and movement onset (Schütz, Henriques, & Fiehler, 2013).

It seems that participants mainly rely on the allocentric information in open-loop and delayed reaching tasks (see also Franz, Hesse, & Kollath, 2009). Thus, if one would implement a delay of several seconds between stimulus offset (offset of the test scene) and reaching onset in the present paradigm, I would assume that there should be at least no drop of the influence of the allocentric information. However, this effect might be limited up to a certain delay as Schütz and colleagues (2013) could not find an influence of allocentric information on reaching behavior after delays longer than 8 s. It seems unlikely that after delays longer than that, participants switch back to a purely egocentric coding strategy. It is more likely that after longer delays an overall decrease in reaching accuracy and precision would be observed.

Seth and Shimojo (2004) conducted a reaching experiment in which participants encoded a target location which was presented next to a landmark. After a short mask, only the landmark was presented but being slightly shifted. Then, participants had to reach to the target location while the landmark was still visible. They found reaching endpoints being deviated in the direction of the landmark shift by about 68 % of the shift magnitude. This is comparable to the results of my fourth study in which open-loop reaches revealed an allocentric influence of 66 %. Thus, in a closed-loop version of my paradigm one should observe a comparable influence of allocentric information on reaching behavior. Moreover, it is conceivable that decreased endpoint variability should be observed as it has been shown that closed-loop reaches are more precise than open-loop reaching movements (Krigolson & Heath, 2004).

6.4 Conclusion

In my thesis I provide insights in the use of spatial reference frames for memory-guided reaching. In a naturalistic environment with multiple objects that could serve as potential allocentric cues participants integrate both egocentric and allocentric information into the movement plan. Specifically, the integration of allocentric information depends on different factors, such as the task-relevance of potential allocentric cues, the distance of the reaching target to task-relevant allocentric cues, and the reliability of the allocentric information. These findings could be described and formalized in terms of optimal Bayesian integration or causal inference of two sources. Moreover, I demonstrate that allocentric information is also used for coding target locations in depth, even to a higher extent compared to coding target locations in the horizontal plane. Thus, it seems that two different subsystems are responsible for coding target locations in these two spatial planes. Furthermore, I show that object size is an important monocular depth cue for allocentric coding of reach targets in depth, even though the information from this cue is combined with other depth cues (e.g., retinal disparity, vergence).

6.5 Research outlook

The findings that I described in this thesis aim for a deeper understanding of the use of egocentric and allocentric information for performing reaching movements to a remembered target location in a naturalistic, complex environment. I do not only provide new insights into this topic but also established a new paradigm that could be used in future research and therefore, provide the possibility to directly compare upcoming outcomes with results that I have obtained.

So far I only manipulated the allocentric information by shifting objects. However, for a more precise investigation of the integration of egocentric and allocentric information, one should also introduce conditions in which the egocentric information is manipulated. In that sense, it would be possible to vary the variance of this cue which should lead to a decreased weighting of the egocentric information if combined optimally with the allocentric information. This could provide further support for the idea that both reference frames are combined in a Bayes-optimal way and could uncover further factors that determine the use of the egocentric information. One way to manipulate the egocentric reference frame would be to manipulate participants gaze behavior (Byrne & Crawford, 2010). However, even though the gaze-dependent reference frame has a dominant role within the egocentric reference frames, also whole body movements need to be considered as this would lead to maximal changes in the egocentric frame of reference. This could be achieved by using the virtual reality setup that I used in my fourth study which also provides a real-time position tracking of the participant. Thus, one would be able to manipulate the egocentric frame of reference by moving the participant and simultaneously adjust his/her perspective and/or allocentric reference frame in the virtual reality.

In the present paradigm, participants did not know which of the objects will actually serve as reach target when encoding the breakfast scene. Thus, they needed to encode the location of all potential target objects which in turn might lead to an overrating of the allocentric reference frame. In contrast, if participants would know which object will serve as the reach target beforehand, they could focus only on this object while ignoring the other objects. Hence, it seems unlikely that the reach target would be coded relative to the surrounding objects. Currently, the lab of Katja Fiehler is conducting an experiment investigating this question and preliminary results indicate an influence of the allocentric information on reaching endpoints but no difference between conditions with and without knowledge about the actual target when participants encode the breakfast scene. This suggests that allocentric information is utilized, even though a purely egocentric reference frame would be sufficient to code the target location. This indicates that the paradigm in my studies does not overrate the use of the allocentric information.

One further option for future research could be to address the question whether allocentric information is also used for performing visually-guided reaches compared to memory-guided reaches. In chapter 6.3, based on the findings by Seth and Shimojo (2004), I predicted a similar use of the allocentric information in closed-loop compared to open-loop reaches for the present paradigm. However, the stimuli they used in their experiment contained only one bar as landmark next to a target dot and thus, results might not be fully transferable to a naturalistic setting. To overcome this limitation, utilizing the present paradigm but performing closed-loop instead of open-loop reaches would reveal insights in the use of allocentric information for visually-guided reaches under more realistic conditions. This would be a promising approach to increase ecological validity of the outcomes because humans mostly look at the object that they want to act on.

7 References

Ambrosini, E., Ciavarro, M., Pelle, G., Perrucci, M. G., Galati, G., Fattori, P., Galletti, C., & Committeri, G. (2012). Behavioral investigation on the frames of reference involved in visuomotor transformations during peripheral arm reaching. *PLoS One, 7*(12), e51856, doi: 10.1371/journal.pone.0051856.

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, doi: 10.1089/cpb.2007.9935.

Avillac, M., Deneve, S., Olivier, E., Pouget, A., & Duhamel, J. R. (2005). Reference frames for representing visual and tactile locations in parietal cortex. *Nature Neuroscience, 8*(7), 941–949, doi: 10.1038/nn1480.

Baldauf, D. & Deubel, H. (2010). Attentional landscapes in reaching and grasping. *Vision Research, 50*(11), 999-1013.

Ballard, D. H. & Hayhoe, M. M. (2009). Modelling the role of task in the control of gaze. *Visual Cognition, 17*, 1185–1204, doi: 10.1080/13506280902978477.

Batista, A. P., Buneo, C. A., Snyder, L. H., & Andersen, R. A. (1999). Reach plans in eye-centered coordinates. *Science, 285*(5425), 257–260.

Battaglia, P. W., Jacobs, R. A., & Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *JOSA A, 20*(7), 1391–1397.

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.

Bernier, P. M. & Grafton, S. T. (2010). Human posterior parietal cortex flexibly determines reference frames for reaching based on sensory context. *Neuron, 68*(4), 776–788, doi: 10.1016/j.neuron.2010.11.002.

Beurze, S. M., Van Pelt, S., & Medendorp, W. P. (2006). Behavioral reference frames for planning human reaching movements. *Journal of Neurophysiology, 96*(1), 352–362, doi: 10.1152/jn.01362.2005.

Biegstraaten, M., Smeets, J. B., & Brenner, E. (2003). The influence of obstacles on the speed of grasping. *Experimental Brain Research, 149*(4), 530–534.

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.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433–436.

Bridgeman, B., Gemmer, A., Forsman, T., & Huemer, V. (2000). Processing spatial information in the sensorimotor branch of the visual system. *Vision Research, 40*(25), 3539–3552.

Bridgeman, B., Peery, S., & Anand, S. (1997). Interaction of cognitive and sensorimotor maps of visual space. *Perception & Psychophysics, 59*(3), 456–469.

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

Buneo, C. A., Jarvis, M. R., Batista, A. P., & Andersen, R. A. (2002). Direct visuomotor transformations for reaching. *Nature, 416*(6881), 632–636.

Burns, J. K. & Blohm, G. (2010) Multi-sensory weights depend on contextual noise in reference frame transformations. *Frontiers in Human Neuroscience, 4*, 221.

Byrne, P. A., Cappadocia, D. C., & Crawford, J. D. (2010). Interactions between gaze-centered and allocentric representations of reach target location in the presence of spatial updating. *Vision Research, 50*(24), 2661–2670, doi: 10.1016/j.visres.2010.08.038.

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, doi: 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, doi: 10.1016/j.visres.2015.02.018.

Carrozzo, M., Stratta, F., McIntyre, J., & Lacquaniti, F. (2002). Cognitive allocentric representations of visual space shape pointing errors. *Experimental Brain Research, 147*(4), 426–436, doi: 10.1007/s00221-002-1232-4.

Chapman, C. S. & Goodale, M. A. (2010). Obstacle avoidance during online corrections. *Journal of Vision, 10*(11), 1–14.

Chen, Y., Byrne, P., & Crawford, J. D. (2011). Time course of allocentric decay, egocentric decay, and allocentric-to-egocentric conversion in memory-guided reach. *Neuropsychologia, 49*(1), 49–60.

Chieffi, S. & Allport, D. A. (1997). Independent coding of target distance and direction in visuo-spatial working memory. *Psychological Research, 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, doi: 10.1038/nrn873.

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

Crawford, J. D., Henriques, D. Y., & Medendorp, W. P. (2011). Three-dimensional transformations for goal-directed action. *Annual Review of Neuroscience, 34*, 309–331, doi: 10.1146/annurev-neuro-061010-113749.

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

DeAngelus, M., & Pelz, J. B. (2009). Top-down control of eye movements: Yarbus revisited. *Visual Cognition, 17*(6–7), 790–811, doi:10.1080/13506280902793843.

Deubel, H., Koch, C., & Bridgeman, B. (2010). Landmarks facilitate visual space constancy across saccades and during fixation. *Vision Research, 50*(2), 249–259.

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

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*(6870), 429–433.

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, doi: 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, doi: 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.

Hay, L. & Redon, C. (2006). Response delay and spatial representation in pointing movements. *Neuroscience Letters, 408*(3), 194–198, doi: 10.1016/j.neulet.2006.08.080.

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. (2007). Visual saliency does not account for eye movements during visual search in real-world scenes. In Van Gompel, R. P., Fischer, M. H., Murray, W. S., & Hill, R. L. (Eds.), *Eye movements: A window on mind and brain* (pp. 537-562). Oxford, UK: Elsevier Ltd.

Henriques, D. Y., Klier, E. M., Smith, M. A., Lowy, D., & Crawford, J. D. (1998). Gaze-centered remapping of remembered visual space in an open-loop pointing task. *Journal of Neuroscience, 18*, 1583–1594.

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, doi: 10.1007/s00221-002-1295-2.

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(3), 683–702.

Khan, A. Z., Pisella, L., Rossetti, Y., Vighetto, A., & Crawford, J. D. (2005). Impairment of gaze-centered updating of reach targets in bilateral parietal-occipital damaged patients. *Cerebral Cortex, 15*, 1547–1560, doi: 10.1093/cercor/bhi033.

Kingstone, A., Smilek, D., Ristic, J., Friesen, C. K., & Eastwood, J. D. (2003). Attention, researchers! It is time to take a look at the real world. *Current Directions in Psychological Science, 12*(5), 176–180.

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

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, 1–13, doi: 10.1167/15.13.24.

Klinghammer, M., Blohm, G., & Fiehler, K. (submitted). Scene configuration and object reliability affect the use of allocentric information for memory-guided reaching.

Klinghammer, M., Schütz, I., Blohm, G., & Fiehler K. (under review). Allocentric information is used for memory-guided reaching in depth: a virtual reality study.

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.

Knill, D. C. (2007). Robust cue integration: A Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant. *Journal of Vision, 7*(7), 1–24, doi: 10.1167/7.7.5.

Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research, 43*(24), 2539–2558, doi: op10.1016/S0042-6989(03)00458-9.

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS ONE, 2*(9), e943, doi: 10.1371/journal.pone.0000943.

Körding, K. P. & Tenenbaum, J. B. (2007). Casual inference in sensorimotor integration. NIPS 2006 conference proceedings, in Advances in Neural Information Processing Systems, Vol. 1, eds. B. Schölkopf, J. Platt, and T. Hoffman (Cambridge, MA: MIT Press), 641–647.

Krigolson, O., Clark, N., Heath, M., & Binsted, G. (2007). The proximity of visual landmarks impacts reaching performance. *Spatial Vision, 20*(4), 317–336, doi: 10.1163/156856807780919028.

Krigolson, O. & Heath, M. (2004). Background visual cues and memory-guided reaching. *Human Movement Science, 23*(6), 861–877, doi:10.1016/j.humov.2004.10.011.

Lacquaniti, F. & Caminiti, R. (1998). Visuo-motor transformations for arm reaching. *European Journal of Neuroscience, 10*(1), 195–203.

Land, M., & Hayhoe, M. M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research, 41*, 3559–3565, doi:10.1016/s0042-6989(01)00102-x.

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.

Leclercq, G., Lefèvre, P., & Blohm, G. (2013). 3D kinematics using dual quaternions: theory and applications in neuroscience. *Frontiers in Behavioral Neuroscience, 7*, 7.

Lemay, M., Bertram, C. P., & Stelmach, G. E. (2004). Pointing to an allocentric and egocentric remembered target. *Motor Control, 8*(1), 16–32.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279–281.

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

Maxcey-Richard, A. M. & Hollingworth, A. (2013). The strategic retention of task-relevant objects in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(3), 760, doi: 10.1037/a0029496.

McGuire, L. M. M. & Sabes, P. N. (2009). Sensory transformations and the use of multiple reference frames for reach planning. *Nature Neuroscience, 12*(8), 1056–1061.

Medendorp, W. P., Beurze, S. M., Van Pelt, S., & Van Der Werf, J. (2008). Behavioral and cortical mechanisms for spatial coding and action planning. *Cortex, 44*(5), 587–597, doi: 10.1016/j.cortex.2007.06.001. Epub 2007 Dec 23.

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

Medendorp, W. P., Goltz, H. C., Vilis, T., & Crawford, J. D. (2003). Gaze-centered updating of visual space in human parietal cortex. *Journal of Neuroscience, 23*(15), 6209–6214.

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.

Mills, M., Hollingworth, A., Van der Stigchel, S., Hoffman, L., & Dodd, M. D. (2011). Examining the influence of task set on eye movements and fixations. *Journal of Vision, 11*(8):0, 1–15, doi:10.1167/11.8.17.

Mon-Williams, M. (1999). Some recent studies on the extraretinal contribution of distance perception. *Perception, 28*(2), 167–181, doi: 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, doi: 10.1007/s002210000410.

Mullette-Gillman, O. A., Cohen, Y. E., & Groh, J. M. (2009). Motor-related signals in the intraparietal cortex encode locations in a hybrid, rather than eye-centered reference frame. *Cerebral Cortex, 19*(8), 1761–1775, doi: 10.1093/cercor/bhn207.

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, doi: 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, doi: 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, doi: 10.3200/JMBR.40.3.203-213.

Neely, K. A., Tessmer, A., Binsted, G., & Heath, M. (2008). Goal-directed reaching: movement strategies influence the weighting of allocentric and egocentric visual cues. *Experimental Brain Research,186*(3), 375–384, doi: 10.1007/s00221-007-1238-z.

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*(3), 335–344, doi: 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.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3–25.

Ramsey, J. O. & Silverman, B. W. (2005). *Functional data analysis* (2nd ed.). New York: Springer.

Saling, M., Alberts, J., Stelmach, G. E., & Bloedel, J. R. (1998). Reach-to-grasp movements during obstacle avoidance. *Experimental Brain Research, 118*(2), 153–160.

Sato, Y., Toyoizumi, T., & Aihara, K. (2007). Bayesian inference explains perception of unity and ventriloquism aftereffect: identification of common sources of audiovisual stimuli. *Neural Computation, 19*(12), 3335–3355.

Schenk, T. (2006). An allocentric rather than perceptual deficit in patient D.F. *Nature Neuroscience, 9*(11), 1369 – 1370, doi: 10.1038/nn1784.

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, 19*(87), 46–52, doi: 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*, doi: 10.1007/s00221-015-4197-9.

Selen, L. P. & Medendorp, W. P. (2011). Saccadic updating of object orientation for grasping movements. *Vision Research, 51*(8), 898–907, doi: 10.1016/j.visres.2011.01.004.

Sheth, B. R. & Shimojo, S. (2004). Extrinsic cues suppress the encoding of intrinsic cues. *Journal of Cognitive Neuroscience, 16*(2), 339–350, doi: 10.1162/089892904322984616.

Snow, J. C., Skiba, R. M., Coleman, T. L., & Berryhill, M. E. (2014). *Frontiers in Human Neuroscience*, 8:837, doi: 10.3389/fnhum.2014.00837.

Sober, S. J. & Sabes, P. N. (2003). Multisensory integration during motor planning. *Journal of Neuroscience, 23*(18), 6982–6992.

Sober, S. J. & Sabes, P. N. (2005). Flexible strategies for sensory integration during motor planning. *Nature Neuroscience, 8*(4), 490–497, doi: 10.1038/nn1427.

Soechting, J. F. & Flanders, M. (1992). Moving in three-dimensional space : frames of reference, vectors, and coordinate systems. *Annual Review of Neuroscience, 15*, 167–191.

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, 1–6, doi: 10.1167/11.9.10.

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

Spetch, M. L. (1995). Overshadowing in landmark learning: touch-screen studies with pigeons and humans. *Journal of Experimental Psychology Animal Behavior Processes, 21*(2), 166–181.

Taghizadeh, B. & Gail, A. (2014). Spatial task context makes short-latency reaches prone to induced Roelofs illusion. *Frontiers in Human Neuroscience,* 8:643, doi: 10.3389/fnhum.2014.00673.

Thaler, L. & Goodale, M. A. (2011a). Reaction times for allocentric movements are 35 ms slower than reaction times for target-directed movements. *Experimental Brain Research, 211*(2), 313–328, doi: 10.1007/s00221-011-2691-2.

Thaler, L. & Goodale, M. A. (2011b). The role of online visual feedback for the control of target-directed and allocentric hand movements. *Journal of Neurophysiology, 105*(2), 846–859, doi: 10.1152/jn.00743.2010.

Thompson, A. A. & Henriques D. Y. (2008). Updating visual memory across eye movements for ocular and arm motor control. *Journal of Neurophysiology, 100*(5), 2507–2014, doi: 10.1152/jn.90599.2008.

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

Tipper, S. P., Howard, L. A., & Jackson, S. R. (1997). Selective reaching to grasp: Evidence for distractor interference effects. *Visual Cognition, 4*(1), 1–38.

Treisman, A. M. & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97–136.

Tresilian, J. R. (1998). Attention in action or obstruction of movement? A kinematic analysis of avoidance behavior in prehension. *Experimental Brain Research, 120*(3), 352–368.

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.

Triesch, J., Ballard D. H., Hayhoe, M. M., & Sullivan, B. T. (2003). What you see is what you need. *Journal of Vision, 3*(1), 86–94, doi: 10.1167/3.1.9.

Van Pelt, S. & Medendorp, W. P. (2008). Updating distance across eye movements in depth. *Journal of Neurophysiology, 99*(5), 2281–2290, doi: 10.1152/jn.01281.2007.

Vaziri, S., Diedrichsen, J., & Shadmehr, R. (2006). Why does the brain predict sensory consequences of oculomotor commands? Optimal integration of the predicted and the actual sensory feedback. *The Journal of Neuroscience, 26*(16), 4188–4197, doi: 10.1523/JNEUROSCI.4747-05.2006.

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

Walter, E. & Dassonville, P. (2006). Fragments of the Roelofs effect: A bottom-up effect equal to the sum of its parts. *Perception & Psychophysics, 68*(8), 1243–1253, doi: 10.3758/bf03193724.

Westwood, D. A. & Goodale, M. A. (2003). Perceptual illusion and the real-time control of action. *Spatial Vision, 16*(3), 243–254.

Westwood, D. A., McEachern, T., & Roy, E. A. (2001). Delayed grasping of a Müller-Lyer figure. *Experimental Brain Research, 141*(2), 166–173.

Wolfe, J. M. (1994). Visual search in continuous, naturalistic stimuli. *Vision Research*, *34*(9), 1187–1195.

Wolfe, J. M., Võ, M. L., Evans, K. K., & Greene M. R. (2011). Visual search in scenes involves selective and nonselective pathways. Trends in cognitive sciences, 15(2), 77–84, doi: 10.1016/j.tics.2010.12.001.

World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. Retrieved from http://www.wma.net/en/30publications/10policies/b3/17c.pdf.

Zaehle, T., Jordan, K., Wüstenberg, T., Baudewig, J., Dechent, P., & Mast, F. W. (2007). The neural basis of the egocentric and allocentric spatial frame of reference. *Brain Research*, *1137*(1), 92–103, doi: 10.1016/j.brainres.2006.12.044.

Selbstständigkeitserklärung

Ich erkläre: Ich habe die vorgelegte Dissertation selbstständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis" niedergelegt sind, eingehalten.

Gießen, im September 2016 Mathias Klinghammer