Gaze-Centered Updating of Remembered Visual Space During Active Whole-Body Translations

Stan Van Pelt and W. Pieter Medendorp

You might find this additional information useful...

This article cites 46 articles, 23 of which you can access free at:
http://jn.physiology.org/cgi/content/full/97/2/1209#BIBL

Updated information and services including high-resolution figures, can be found at:
http://jn.physiology.org/cgi/content/full/97/2/1209

Additional material and information about *Journal of Neurophysiology* can be found at:
http://www.the-aps.org/publications/jn

This information is current as of March 9, 2007 .

Gaze-Centered Updating of Remembered Visual Space During Active Whole-Body Translations

Stan Van Pelt[1] and W. Pieter Medendorp[1,2]

[1]*Nijmegen Institute for Cognition and Information,* [2]*FC Donders Centre for Cognitive Neuroimaging. Radboud University Nijmegen, Nijmegen, The Netherlands*

Van Pelt S, Medendorp WP. Gaze-centered updating of remembered visual space during active whole-body translations. *J Neurophysiol* 97: 1209–1220, 2007. First published November 29, 2006; doi:10.1152/jn.00882.2006. Various cortical and sub-cortical brain structures update the gaze-centered coordinates of remembered stimuli to maintain an accurate representation of visual space across eyes rotations and to produce suitable motor plans. A major challenge for the computations by these structures is updating across eye translations. When the eyes translate, objects in front of and behind the eyes' fixation point shift in opposite directions on the retina due to motion parallax. It is not known if the brain uses gaze coordinates to compute parallax in the translational updating of remembered space or if it uses gaze-independent coordinates to maintain spatial constancy across translational motion. We tested this by having subjects view targets, flashed in darkness in front of or behind fixation, then translate their body sideways, and subsequently reach to the memorized target. Reach responses showed parallax-sensitive updating errors: errors increased with depth from fixation and reversed in lateral direction for targets presented at opposite depths from fixation. In a series of control experiments, we ruled out possible biasing factors such as the presence of a fixation light during the translation, the eyes accompanying the hand to the target, and the presence of visual feedback about hand position. Quantitative geometrical analysis confirmed that updating errors were better described by using gaze-centered than gaze-independent coordinates. We conclude that spatial updating for translational motion operates in gaze-centered coordinates. Neural network simulations are presented suggesting that the brain relies on ego-velocity signals and stereoscopic depth and direction information in spatial updating during self-motion.

INTRODUCTION

In daily life, we appear to be perfectly aware of objects in our surroundings. Even when we move, we seem to have no difficulty in keeping track of objects and reach or look at their locations whenever necessary. This seemingly automatic behavior, called spatial updating, works even in darkness and for targets that are otherwise no longer in view (Hallet and Lightstone 1976; Li and Angelaki 2005; Medendorp et al. 2002). But despite extensive investigations, the computational basis of spatial updating has remained an issue of great controversy (Andersen et al. 1985; Baker et al. 2003; Duhamel et al. 1992; Van Pelt et al. 2005).

A critical aspect of this issue is the reference frame with respect to which object locations for actions are encoded. A reference frame is characterized by a coordinate system, which represents locations using a set of coordinate axes fixed relative to some origin, like the eyes, head, body, or earth. Obviously, in theoretical terms, spatial updating could work in any coordinate frame as long as the correct updating signals and computational operations are used (Medendorp et al. 2003b). Adding to this notion, various studies have argued that the reference frame used to encode a spatial memory is not fixed but depends on several factors, including the sensory inputs, task constraints, the visual background, memory interval, and the cognitive context (Battaglia-Mayer et al. 2003; Bridgeman et al. 1997; Carrozzo et al. 2002; Hayhoe et al. 2003 Snyder et al. 1998; Van Pelt et al. 2005). Within this view, psychophysical evidence obtained in neutral open-loop testing situations has suggested that the early feedforward mechanisms for internal spatial updating operate in gaze-centered coordinates (Baker et al. 2003; Henriques et al. 1998; Medendorp and Crawford 2002). In further support of this evidence, many brain regions in parietal and frontal cortex have been shown to update their activity patterns relative to the new gaze direction after an eye movement has occurred (Batista et al. 1999; Duhamel et al. 1992; Medendorp et al. 2003a; Merriam et al. 2003; Sommer and Wurz 2002).

It is important to point out though that most of the actual evidence for gaze-centered updating was obtained using simple eye rotations only with the head and body restrained, ignoring the fact that in natural situations our eyes also translate through space, as for example when we walk. When the body translates, correct updating in a gaze-centered frame seems computationally much more demanding because the required updating varies from object to object, depending nonlinearly on their depth and direction as in motion parallax (Li et al. 2005; Medendorp et al. 2003b). In this respect, updating for translational motion seems much simpler if object locations were stored in, say, Cartesian body-centered coordinates because then the required updating would be the same for each object: the opposite of the amount of body displacement (Medendorp et al. 1999).

At present, it is unknown which reference frame is involved in the computations for the translational updating of remembered visual space. Here we address this question by characterizing the pattern of errors in manual reaching movements toward briefly flashed targets presented prior to a whole-body translation. Our goal is not to merely characterize a subject's ability to update spatial information for intervening translations. In fact, recent studies have already shown that humans and monkeys can look to remembered locations in near space,

---

Address for reprint requests and other correspondence: S. Van Pelt, Nijmegen Institute for Cognition and Information, Radboud University Nijmegen, P.O. Box 9104, NL-6500 HE, Nijmegen, The Netherlands (E-mail: s.vanpelt@nici.ru.nl).

compensating for intervening eye translation induced by head or body motion (Israel et al. 1999; Li et al. 2005; Medendorp et al. 2003b). However, the computational principles underlying the spatial constancy in this behavior, whether gaze-related or not, remain to be revealed.

We designed a novel experiment to discriminate between a gaze-dependent and gaze-independent model of visuospatial memory updating during translations. In our test, subjects fixate centrally at fixation point (FP) while a far or near target (Tf, Tn) is flashed onto the retinal periphery (Fig. 1, *middle*). Subjects then translate sideways (by making an active whole-body step displacement) while keeping their gaze at FP and
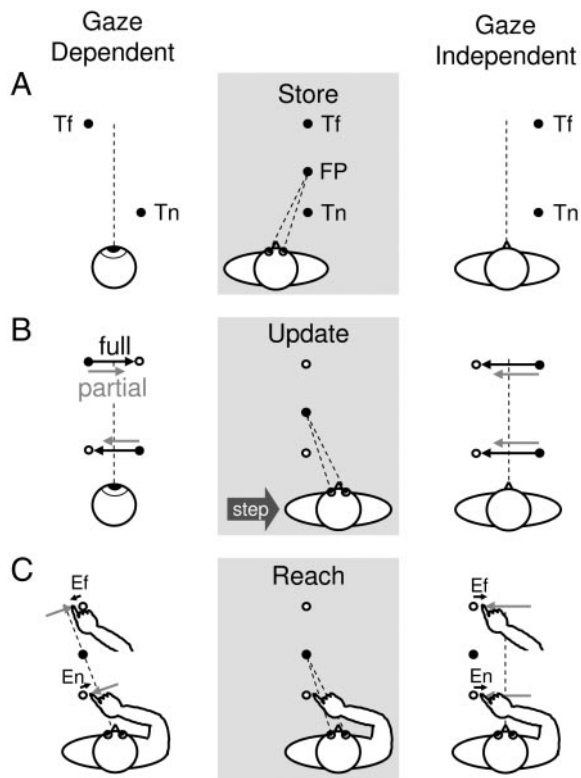

FIG. 1. Predictions of the gaze-dependent and gaze-independent models of internal spatial updating during whole-body translations. The basic assumption in the test is that subjects generally misestimate the amount of self-motion when the body translates. *A*: subject looks at a central fixation point (FP) while a target is flashed, either in front of (near T = Tn) or behind FP (far T = Tf). An internal representation of this target is coded in either a gaze-dependent frame (*A, left*) or in a gaze-independent frame (*right*). Thus in the gaze-dependent frame, near and far targets are stored as memories coding opposite locations relative to the gaze-line. In the gaze-independent frame, say a body-frame, they are transformed and stored as memories reflecting positions at the same side from the body midline. *B*: after viewing and storing the target, the subject translates the body, e.g., in rightward direction, while keeping fixation at FP, and then reaches toward the remembered location of the target (*C*). If the target is stored in a gaze-dependent frame (*B, left*), the subject should compensate for the induced change of gaze by updating the gaze-dependent memory trace. That is, the near-target memory should be shifted to the left while the far-target memory should be shifted to the right. If compensation is only partial as a result of an erroneous estimation of step size, memory traces will be shifted partially and hence will not match the actual location of the targets. This will result in reach errors, denoted by Ef and En, which reverse in direction for remembered targets at opposite depths from fixation (*C, left*). Alternatively, if targets are stored in a gaze-independent body frame, the subject should compensate for the induced changes of the body. In effect, when the body translates to the right, all memory traces should be shifted to the left, by equal amounts (*B, right*). If then updating is only partial, this will result in reach errors in the same direction for all remembered target locations (*C, right*).

subsequently reach to the remembered target location. The logic behind the test is the following. Suppose that the targets were visible at all times, including when the body translates sideways. Then parallax geometry dictates that targets in front and behind the eyes' fixation point (FP) shift in opposite directions on the retinas. Thus if the brain is to simulate motion parallax also in the active updating of *memorized* targets (*left*, black arrows), it can be predicted that if the body translation is not correctly taken into account (Glasauer et al. 1994; Medendorp et al. 1999), the updated locations (gray arrows) will deviate from the actual locations, leading to reach errors (Ef, En) in opposite directions for targets in front of and behind the FP (*hypothesis A*: gaze-dependent updating). Alternatively, parallax geometry plays no role if the brain codes locations in a gaze-independent reference frame, e.g., in a body-fixed frame (*right*). If then translations are misjudged, the updated locations will also deviate from the actual locations, but with updating errors (as probed by the reach) in the same direction for all targets (*hypothesis B*: gaze-independent updating).

Our results demonstrate that translational updating follows the predictions of the gaze-dependent scheme. To obtain further insights in the putative computations in this process, we trained a simple three-layer recurrent neural network to perform gaze-centered updating in these translation conditions. The network learnt correctly the geometric computations involved, and preferred velocity, rather than position signals for updating remembered visual space during self-motion (White and Snyder 2004).

METHODS

Subjects

Fifteen human subjects (4 female, 11 male, mean age of 26 ± 4 yr) were tested in four different task conditions as described in the following text. The main experiment involved 10 naïve subjects and the 2 authors. Each of the three additional control experiments tested five subjects (3 naïve). All subjects signed informed consent to participate in the experiment. All subjects were right-handed, and all were free of any sensory, perceptual, or motor disorders. All pointing movements were made using the right arm.

Experimental setup

Subjects were standing in a completely darkened room, within a designated area of 60 cm width, which we will refer to as the "translation zone." A U-shaped ridge of 6 cm height was attached to the floor indicating the outer borders of the translation zone to the left, right, and back of the subject. During the experiments, this ridge served as a reference for subjects to position their feet to accurately control their own positions and self-induced translations. This configuration led to lateral body translations with an amplitude of 30 ± 7 (SD) cm averaged over all subjects. Within subjects, positions and translations were reproduced with an accuracy >3 cm.

We used an OPTOTRAK 3020 digitizing and motion analysis system (Northern Digital) to record the position and orientation of various body parts in three dimensions (3-D). This system tracks the 3-D position of infrared-emitting diodes (ireds) with an accuracy >0.2 mm. We determined head position and orientation by means of four ireds attached to the eye tracking helmet worn by the subject (see following text). Prior to the experiment, we calibrated the locations of the eyes and ears with respect to the ireds on the helmet. During this calibration procedure, the subject faced the OPTOTRAK camera while wearing the helmet with three additional temporary ireds, one

near the right auditory meatus and one on each closed eyelid. The three-dimensional (3-D) locations of these ireds, which uniquely defined the location of the right ear and both eyes relative to the helmet, were recorded together with the ireds on the helmet. With this information, we were able, during the subsequent experiment, to compute the positions of the eyes and ear in space on the basis of the helmet ireds alone. The actual location of each eye, defined as its rotation center, was assumed to be 1.3 cm behind its cornea. In a similar fashion, we calibrated the position of the tip of the right index finger relative to four ireds attached to the middle phalanx of this finger. We further used the OPTOTRAK system to record the position of the shoulder (acromion) as well as the positions of the stimulus targets. OPTOTRAK data were sampled at 125 Hz. The ired coordinates were transformed to a right-handed space-fixed coordinate system. The x-y plane was aligned with the subject's horizontal plane. The positive x axis was pointing forward, perpendicular to the subjects' shoulder line; the positive y axis was pointing leftward along the shoulder line, seen from the subject; and the z axis was pointing upward. The position of the central light-emitting diode (LED) on the stimulus array (see following text) served as the origin of the coordinate system. The orientation of the head was determined with respect to a reference position adopted when the subject faced straight ahead. Orientation and location measurements were accurate to within $0.2°$ and 0.2 mm.

We used an Eyelink II eyetracker (SR Research) to record binocular eye movements. We ensured that its camera system, which was mounted to the helmet, remained stable on the head during the entire experiment. Stable recording of eye position was further warranted by measuring corneal reflections in combination with pupil tracking, which reduces the errors caused by any helmet slip and vibration. As a further precaution, subjects were also instructed to minimize speaking during the experiments. Eye movements were calibrated before the experiment by having subjects face straight ahead and fixate the stimulus LEDs two times each, in complete darkness, both when standing left and right within in the translation zone. Eye recordings were calibrated in the head-fixed coordinate system of the eye tracker. By combining the locations of the stimuli and the reconstructed locations of both eyes (using the helmet calibration data) as well as current head orientation, we computed the direction of the stimulus LEDs with respect to the subject's eyes in head-fixed coordinates. In this way, the eye-tracker data of both eyes could be matched to the corresponding vertical and horizontal stimulus directions and expressed as eye-in-head orientation signals. During the actual experiments, eye-in-space orientation was calculated by combining head orientation and calibrated eye-in-head orientation signals. The eye-calibration procedure resulted in a directional accuracy of the eye-in-head orientation $>1.5°$. Version and vergence positions were calculated from the left (L) and right (R) eye positions as $(R + L)/2$ and $L - R$, respectively.

Two PCs controlled the experiment. A master PC was equipped with hardware for data acquisition of the OPTOTRAK and Eyelink measurements, as well as visual stimulus control, while a slave PC contained the hardware from the Eyelink system.

Stimuli

Nine red LEDs (luminance <20 mcd/m2) served as stimuli. They were attached to a frame in the shape of a cross that was mounted on a two-link robot arm. This robot arm, equipped with stepping motors (type Animatics SmartMotors; Servo Systems), could rapidly position the center of the frame to virtually any desired position within a hemisphere (radius: 1 m) centered at its base. The frame was positioned with an accuracy of >0.2 mm, as confirmed by OPTOTRAK recordings. During the experiment, the stimuli were presented at space-fixed locations, at eye level in the subject's transverse plane (Fig. 2A). The location of the central LED, which served as fixation point (FP), corresponded to our space-fixed coordinate system's
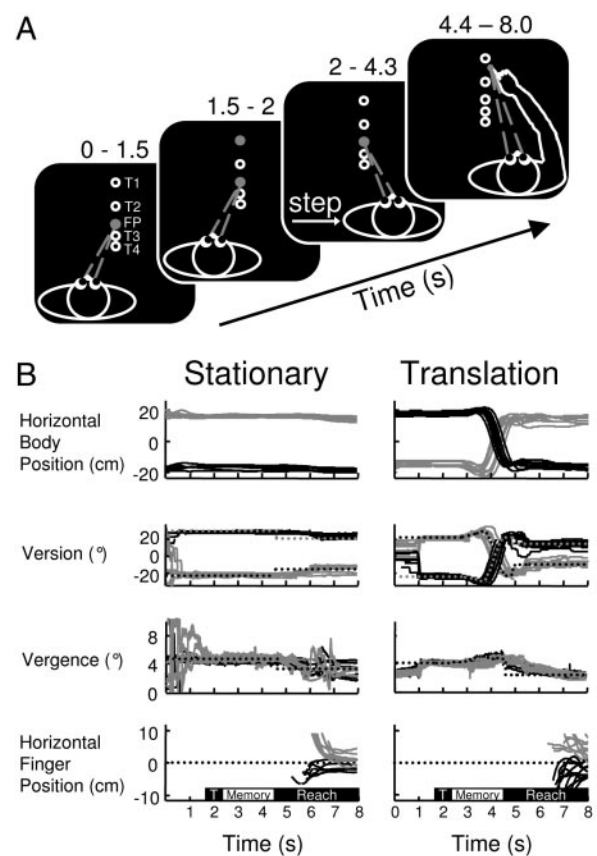

FIG. 2. *A*: sequence of stimuli and the subject's instructions during the translation trials. Subjects start by fixating a space-fixed target (FP) for 1.5 s. Next, a second space-fixed target was presented for 500 ms, either in front of or behind the fixation point. Subjects translated their body within a 2.3-s memory period while they maintained fixation on FP. After another 100 ms, an auditory cue signaled the subjects to reach toward the remembered location of the target. Stationary trials (not shown) differed from translation trials by the absence of subject translation during the memory period. Four space-fixed targets (1–4) served as potential target locations, presented such that their mutual distance in terms of retinal eccentricity was 4°, when the subject was standing at opposite ends within the translation zone. Figure not to scale. *B*: typical performance of one subject (*S1*). Body position, eye position (version and vergence), and finger tip position (horizontal component) plotted against time for 16 stationary (*left*) and 16 translation trials (*right*). The target for memory was T1 (i.e., behind the fixation point). Black traces, leftward final position; gray traces, rightward final positions. Dotted traces, geometrically ideal signals. Thin boxes, time intervals of the different trial stages [target presentation (T), memory period, reach interval].

origin, which was straight in front of the center of the translation zone, at a distance of ~35 cm. Four other LEDs were lined up with the x axis of the coordinate system and served as visual targets for task conditions described in the following text. Two of these targets were behind the central LED (from the subject's perspective) at distances of 7 and 17 cm (T1, T2), and two were in front of the central LED at distances of 6 and 10 cm (T3, T4). Using this configuration, we ensured that the target flashes stimulated both retinas during the experiments, at equal intervals of ~4°. We further positioned four other LEDS along the y axis of the coordinate system, at either side of the central light at 6 and 12 cm (not shown in Fig. 2A). These targets were used in catch trials to ensure that subjects did not simply make repeated stereotypic responses. Data for these catch targets were excluded from further analysis. We also made sure that subjects never saw the target configuration when the room lights were on by positioning it to an elevated level using the robot.

Main task

The experiments were designed to test between a gaze-dependent and gaze-independent model of visuospatial updating for translational motion. In our test, subjects were instructed to perform memory-guided reaching movements under two conditions, which will be referred to as "stationary" and "translation" tasks. The experimental paradigm of the translation task is illustrated in detail in Fig. 2A. Before the start of each trial, subjects positioned their feet on either the left or right end of the translation zone to certify a fixed starting position. A trial started with the onset of FP which was illuminated for 4.3 s and had to be fixated by the subject for its entire duration. At 1,500 ms after the onset of FP, a target for memory (here T1), closer or farther than FP, appeared in the visual periphery for 500 ms. Then a 2.3-s time interval followed in which subjects were instructed to either remain stationary (stationary task) or to make a sideward step to the opposite side of the translation zone (translation task) while still fixating FP. Then FP was extinguished, the stimulus frame was retracted, and 100 ms later an auditory signal cued the subject to conjointly look and reach at the remembered location of T, keeping the body and head still. Subjects had to hold that reaching position until another auditory signal was presented 3.6 s later. Then the next trial started. Targets were randomly chosen from the four locations. Each target location was tested 20 times for both starting positions, resulting in a total of 160 trials for each of the two tasks. Test trials were randomly interspersed with 32 catch trials. Subjects never received any tactile feedback during their reach. In all trials, subjects had to keep their head and body aligned in the straight ahead direction. In the translation trials, the starting position of a trial was the end position of the previous trial, whereas in the stationary task, the subject first moved to the other end within the translation zone before testing the next trial. Thus in the stationary task, response data were gathered at positions that also served either as initial or as final position in the translation task (F-test, $P > 0.05$). This allowed direct comparison of response behavior when updating was necessary (translation task) with behavior where no updating was needed (stationary task). For both test conditions, the total duration of each trial was 8.0 s.

During the reaching movement, visual feedback about hand position was provided by means of an LED attached to the fingertip. This way we tried to minimize the error attributable to an erroneous estimate of fingertip position during pointing (Beurze et al. 2006). We also allowed subjects to look where they were reaching to eliminate contributions of errors occurring otherwise, i.e., when gaze would be off the reach location (Henriques et al. 2003; see control experiment 2 in the following text).

The total experiment was divided into two sessions tested on different days. In each experimental session, half of the translation trials were tested first, followed by half of the stationary trials. Subjects performed blocks of 12 consecutive trials between which a brief rest was provided with the room lights on to prevent dark adaptation. During these periods, the stimulus frame was out of view. Each session lasted for ~60 min. One subject was tested over three sessions. During the experiments, subjects never received feedback about their performance. Before the actual experiment began, subjects practiced a few blocks to become familiar with the two task conditions.

Control tasks

We also performed three control experiments in which we varied a number of task parameters to test their implications for updating behavior. All controls were performed with the same timing and stimulus durations as in the main experiment, unless indicated otherwise. First, we tested updating performance in the absence of visual feedback about fingertip position during the reaching movement (control 1: reaching without feedback). This clarified whether the

results in the main experiment were not critically dependent on a visually monitored hand position during the reach. The next control experiment was inspired by the fact that reaching while looking where you reach is generally more accurate than reaching to a retinally peripheral location (Henriques et al. 2003). Therefore in contrast with the main experiment, subjects performed the reaching movement in this task by keeping gaze fixed at the remembered location of FP (control 2: reaching without looking). This tested whether the results of the main experiments were not mainly driven by one of the two motor systems (eye vs. arm). The final control was designed to test the effect of a visual fixation point (FP) during the updating task (control 3: updating without FP). Therefore in this task, FP was turned off immediately after the target flash, and subjects were instructed to make their body translation by keeping their gaze fixed on remembered FP. Reaching was performed under visual feedback of the fingertip, which had to be fixated. As the eyes may diverge from the remembered FP during the translation in darkness (Medendorp et al. 2003b), updating was tested for the two outermost targets only because these were most discriminative in terms of the models outlined in Fig. 1.

Data analysis

Data were analyzed off-line using Matlab (The Mathworks). We excluded trials in which subjects did not keep their eyes directed at FP within a 3° interval or made a saccade during target presentation. We also discarded trials in which the subject had not correctly followed other instructions of the paradigm, e.g., when stepping or reaching too early, or not making a step when this was required. Typically, 23 ± 11 trials (~7%) were discarded based on the arm and eye movement criteria. For each of the remaining trials, final reaching positions were selected manually at the time when the arm had the greatest degree of stability within the last 2 s of the response interval. For each trial, an average position was computed over a six-sample interval (48 ms) centered at this point in time. After categorizing the stationary and translation trials by starting position and translation direction, respectively, we computed the mean reach endpoint separately for each of the targets within these categories. Starting and final body positions were defined by the location of the center of the two eyes at the time of target presentation and reach response, respectively. The difference between these two positions determined the amplitude of the translation (step size). We tested between gaze-dependent and gaze-independent updating models by comparing the horizontal components of the updating errors of reaches toward the targets flashed in front of and behind FP in the translation trials. Because both variables are subject to natural variation and measurement error, a model 2 regression (also referred to as a major-axis regression) was used to determine their relationship, with slope and confidence limits estimated by the bootstrap method (Press et al. 1992). We used the results of the stationary paradigm as a measure for errors attributable to perception or motor effects assuming that both contributed equally. A further 2-D vectorial analysis was performed to entail how the interaction between initial target position, translational motion, and reach response can be described in both gaze-dependent and gaze-independent coordinate frames (see later). Statistical tests were performed at the 0.05 level ($P < 0.05$).

Neural network model

To understand our findings in neurophysiological terms, we trained a simple recurrent three-layer Elman-type neural network using back-propagation to perform gaze-centered updating for both intervening rotations and translations of the eye. We used a similar type of network architecture as White and Snyder (2004). who modeled the updating process for (conjugate) eye rotations only. The predictions of this model will be discussed in the DISCUSSION. In the present model, the input layer of the network includes a map of neurons with similar

spatial tuning properties as those observed in parietal region LIP: Gaussian-like receptive fields for the eye-centered direction of a stimulus and its relative depth from the plane of fixation (retinal disparity) (Gnadt and Mays 1995). For simplicity, we used a 2-D horizontal-disparity map of 121 units (11×11 units; horizontal range $-50–50°$ disparity range $-25–25°$). Each unit within the map had a 2-D Gaussian tuning curve, with a $10 \times 5°$ horizontal-disparity receptive field (1/e2 width), so that receptive fields of units at neighboring locations overlap considerably. Stimulus direction and disparity input to the network were limited to <20 and $<9°$, respectively. The network also received four eye-position units: one pair of units represented binocular gaze (version); another pair encoded binocular depth (vergence). For each unit, the activity was linearly scaled within the range -1 to $+1$, corresponding to -40 to $+40°$ version angle and 0 to $+10°$ vergence angle, respectively. In each pair, the second unit had the opposite activity of the first (push-pull arrangement). Another two pairs of push-pull input units coded for version velocity between -250 and $250°/s$ and vergence velocity between -10 and $10°/s$, respectively. Finally, two push-pull units encoded translation velocity of the eye between -250 and 250 cm/s; another unit pair represented the integrated velocity between -50 and 50 cm (translational path) of the eyes. The output layer was modeled corresponding to the input map. All units in the network were fully connected with each input unit connected to all hidden units and each hidden unit connected to all output units. The hidden layer had recurrent connections to enable the network to remember past events. Both the hidden layer units and the output neurons were characterized by a logarithmic sigmoid activation function of the form $A(x) = 1/[1+\exp(-x)]$. We simulated a trial as a series of 11 consecutive time steps with each step defined as a 200-ms interval. We tested the network with different numbers of units (25, 50, and 100) in the hidden layer. Each type of network was trained four times with random initial weights to validate reproducibility of behavior. The analysis presented in this paper was performed with 50 hidden units.

During training, targets were presented at one of five locations in space, at 25, 29, 35, 42, and 52 cm in front of the subject when viewing them from straight ahead (translation position 0). The other translational positions of the eyes at the start of the trial were 5, 10, 15, and 18 cm to the left or right from position 0. The binocular point of fixation was at the location of the 25-, 35-, or 52-cm target. The simulated translational motion was 0 (no translation), ±10, ±20, ±30, and ±36 cm. To simulate trial conditions with only rotational motion of the eyes (without translational motion), the fixation spot was moved by either 0, 5, 10, 15, or 18 cm to the left or right. Targets were presented for one time step, i.e., 200 ms, at the onset of a trial. Translation of the subject, or translation of binocular fixation point, which followed a bell-shaped velocity profile, was initiated 400 ms after the target disappeared, and lasted for 1 s. The network's output, the direction and disparity of the target in eye-centered coordinates, was read at the final time step of the trial. Trial types which moved the horizontal target direction $>20°$ in the output map were excluded to minimize edge effects at the boundaries of the workspace. Together, this led to 1,129 different types of trials in the training set.

Network testing included all combinations that comprise the binocular fixation position at 33 cm, targets presented at 27, 35, or 48 cm, the translational offset of the eyes -16, -6, 0, 3, or 9 cm, translation motion of 25, 12, 8, 0, and 14 cm, and movements of the fixation point of -13, -7, 0, 4, and 15 cm. The network was built, trained, and tested using the Matlab Neural Network Toolbox with a training function that updates weight and bias values according to gradient descent momentum and an adaptive learning rate. For training, individual weights were initially set to random values between -0.1 and $+0.1$.

RESULTS

We exploited the geometry of motion parallax to address the question whether the location of a space-fixed target, briefly presented before an intervening whole-body translation, is stored and updated in a gaze-dependent or gaze-independent coordinate frame (see Fig. 1). A gaze-dependent coding predicts that if the translation is not correctly taken into account, the updated locations will deviate from the actual locations with updating errors in opposite directions for targets in front and behind the FP. Alternatively, updating within a gaze-independent framework requires the readouts of the memories of such targets after the body translation to be affected by errors in the same direction. We tested between these hypotheses using memory-guided reaching movements in stationary and translation trials.

Task performance

Twelve subjects participated in the main experiment, outlined in Fig. 2A. Using the stationary trials, we first tested the ability of stationary subjects to look and reach to memorized locations of space-fixed targets flashed at different distances from the fixation point. Figure 2B, left, shows the performance of a typical subject over the time course of sixteen trials, either when standing at the leftward position (black traces) or at the rightward position (gray traces) within the translation zone with a target that was flashed 17 cm behind the eyes' fixation point (T1, see Fig. 2A). The top panel depicts the horizontal component of the subject's body position during the entire trial. Both within and across trials, this position remained constant, as instructed, also during reaching at about 17 cm left or right of the center of the translation zone. The second panel displays binocular gaze direction superimposed on the average signals for ideal performance (dotted lines) that were computed on the basis of the Optotrak data. Binocular gaze showed steady fixation when the target was presented and during the memory interval (as required to meet the 3° accuracy range of the trial inclusion criteria, see METHODS), and small saccades at the time of pointing. These saccades direct the eyes toward the finger tip, which is to point at the remembered location of the stimulus flash. The third panel shows a similar pattern for binocular fixation depth (in degrees, as indicated by the vergence component of the eye positions). The decline in vergence during the reach seems to match the requirements (dotted lines) to look at the remembered location of the flash, which is farther away than the fixation point. Finally, the bottom panel demonstrates the horizontal position of the finger tip (in cm), showing that the subject reached fairly accurately to the remembered location of the stimulus flash, with errors <3 cm. These few trials are exemplary for the performance of all subjects in the stationary trials, showing that they can localize a nonfoveated flashed target fairly well.

The question is how well are these subjects able to localize these flashed targets when they have translated after viewing the flash? This was tested using the translation task. Recall that a whole-body translation effectively disturbs the spatial registry of the location of the flash relative to any reference frame attached to the body. Hence, in any egocentric reference frame, whether gaze-dependent or gaze-independent, the location of the reach goal after the translation is different from the location of the flash before the translation.

Figure 2*B, right,* shows the typical performance of the same subject over the time course of 16 translation trials in which the translation was either rightward (gray traces) or leftward (black traces). As in the stationary examples (*left*), the target for updating was T1, flashed 17 cm behind the fixation point. As instructed, the subject only began moving after the target had flashed, and reached his final position before FP offset (*top*). Kinematics of the self-induced translation were highly reproducible across trials, with a mean displacement of 32 ± 2 (SD) cm. During the translation, changes in binocular fixation direction and depth matched the geometrically required modulations (dotted lines) to keep gaze fixed at FP quite well (*2nd* and *3rd panels*). In other words, the body translation had negligible influence on the ability to keep fixation at a lit fixation target. In accordance with the instructions, the changes of these signals during the reach period indicate a change in the binocular fixation point toward the remembered location of the target. The accuracy of the respective reaching movement reflects the accuracy of the spatial memory update as well as the perceptual and motor deficits involved. The reaching movements here show clearly larger errors than in the stationary condition, ranging up to ~7 cm.

To demonstrate the differences in performance in both tasks more clearly, Fig. 3 compares the reach endpoints in the stationary (*left*) and the translation task (*right*), in separate top-view panels for the four targets, ordered by their location from FP, for one subject. In both conditions, a general underestimation of target distance seems to be present. In the
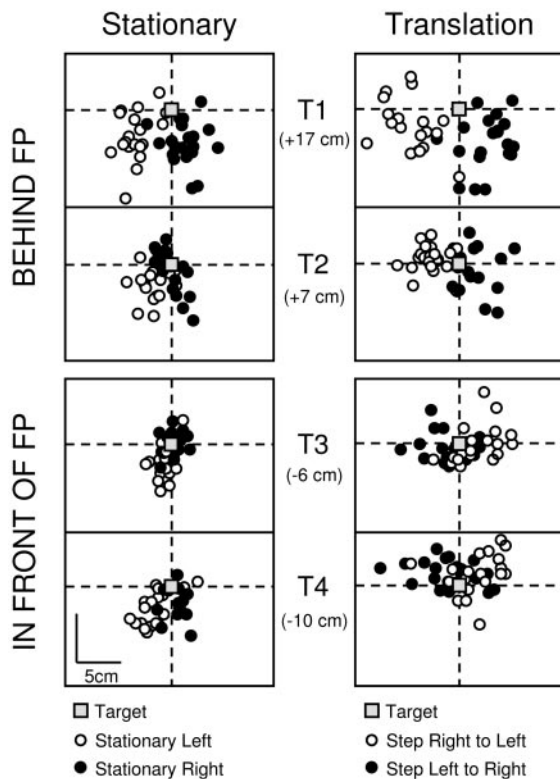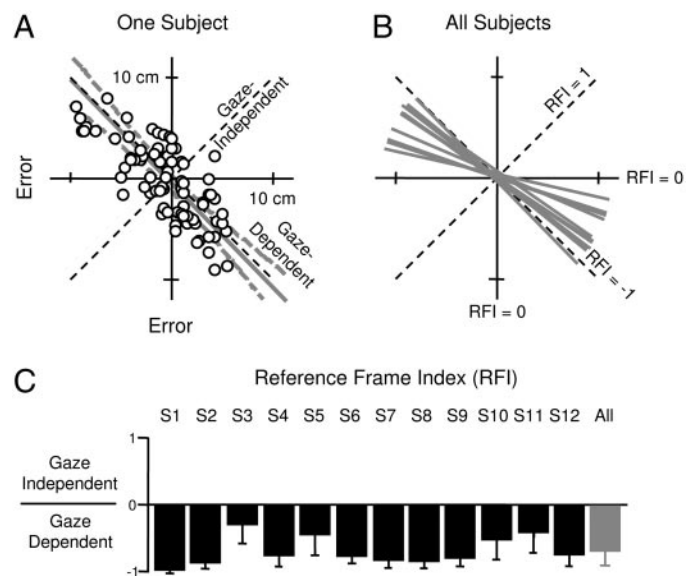

FIG. 4. Reach errors to targets at opposite but equiangular distances from fixation plotted vs. each other. Data would fall along the negative diagonal if subjects had updated remembered target locations in a gaze-dependent frame (reference frame index, RFI = −1). Data would scatter along the positive diagonal if subjects had employed a gaze-independent updating mechanism (RFI = 1). A: subject (*S1*) favoring the gaze-dependent model. The best-fit line (in gray) which characterizes the distribution of the data points, has a clear orientation to the negative diagonal. (gray dashed lines: ±95% confidence intervals), B. Best-fit lines from all subjects. *C:* RFI values (with bootstrap confidence intervals) from all subjects, with RFI −1 supporting the gaze-dependent model and RFI +1 the gaze-independent model. Subjects typically support the gaze-dependent updating model.

stationary task, errors are only small with a slight dependence on the subject's body position. Undeniably, errors in the translation trials exceed those in the stationary trials irrespective of step direction. Both size and horizontal direction of this error seem to depend on the direction of the intervened translation and on the location of the target. For rightward translations, the subject reached too far to the right for the farthest target, whereas there was a leftward bias for the nearest target. The opposite pattern is observed for a leftward translation. There is also a tendency for errors to increase for the targets flashed at farther distances from the fixation point despite the same amount of intervened translation. Thus for this one subject, the pattern of errors in the translation trials seem to follow the prediction by the gaze-dependent updating model: pointing positions deviate in opposite directions for targets in front and behind the FP, with a nearly mirror-symmetric pattern of errors for leftward and rightward translations.

Error analysis

To analyze these findings quantitatively, we assumed that the reach errors in the static trials reflect a sensorimotor deficit, whereas the reach errors in the translation trials reflect sensorimotor deficits as well as deficits in the spatial memory update (see METHODS). Therefore to compute the latter, i.e., the updating errors, we subtracted the mean horizontal reach error observed in the static trials from the horizontal reach errors that occur in the translation trials, for each target separately. Figure 4A plots these horizontal updating errors for targets behind FP, versus the errors for their corresponding equiangular counter-

FIG. 3. Reaching positions (circles) of one subject in the stationary (*left*) and translation task (*right*). Data from *subject S1.* Data presented in separate top-view panels for the 4 targets (□), ordered by their location from FP. Errors in the translation trials appear to depend on the direction of the intervened translation and on the depth of the target from fixation, which is most consistent with the predictions of the gaze-dependent updating model.

parts in front of FP, for each translation direction. Thus updating errors of target T1 were plotted versus the updating errors of target T4 and errors from target T2 with target T3. This pair-wise comparison was performed by picking, without return, the errors randomly from the respective trials, yielding a maximum of 80 data points. The gaze-dependent updating hypothesis predicts that these errors have equal size but opposite signs (Fig. 1). Accordingly, data points should fall in the even quadrants, ideally along the dashed line with slope −1. In contrast, the gaze-independent updating hypothesis predicts that these errors have equal size and signs, which would be indicated by data points along the positive diagonal (slope +1). Any other slope values, whether 0 (the data scatter around the *x* axis), infinity (the data scatter about the *y* axis) or any other value reflect a measure intermediate these two models. To deal with this in further analysis, we converted all slope values to a reference frame index (RFI) between −1 (perfect gaze-dependent coding) and +1 (perfect gaze-independent coding). For example, slopes of +2 and −2 correspond to a reference frame index of 0.5 and −0.5, respectively. Figure 4*A* presents the results of this analysis for the same subject as in Fig. 3, showing that the majority of the data points fall in the even quadrants. According to a model 2 regression, the best-fit line that characterized the direction of the data point clustering was closely directed along the line with slope −1. The reference frame index of this subject had a value of −0.93 ± 0.06 (mean ± SD), which is illustrative for a data distribution that best supports the gaze-dependent updating model. The best-fit lines of all 12 subjects are superimposed in Fig. 4*B,* generally indicating an orientation in the direction predicted by the gaze-dependent model. Figure 4*C* summarizes the corresponding reference frame indices (± SD) for all subjects (black bars), showing a clear bias toward the gaze-dependent model. Averaged across subjects, the reference frame index was −0.68 ± 0.23, which was significantly different from zero (*t*-test, $P < 0.05$), indicating that our data are most supportive for a gaze-centered coding and updating of spatial memory.

For completeness, Table 1 provides further statistical information about the data distribution of each subject, showing the mean correlation coefficient (*r*), RFI, and a variance ratio (VR), defined as the ratio between variance of the data along the main axis of the distribution and the variance in the direction orthogonal to it.

TABLE 1. *Results of the horizontal error analysis in each subject*

	r	RFI	VR	*n*
S1	−0.65 ± 0.05	−0.93 ± 0.06	4.17 ± 1.22	80
S2	−0.49 ± 0.06	−0.89 ± 0.08	3.03 ± 1.29	78
S3	−0.22 ± 0.04	−0.26 ± 0.28	1.61 ± 0.91	75
S4	−0.35 ± 0.08	−0.76 ± 0.17	2.04 ± 0.79	63
S5	−0.12 ± 0.09	−0.42 ± 0.32	1.56 ± 1.22	61
S6	−0.41 ± 0.08	−0.78 ± 0.13	2.33 ± 0.76	69
S7	−0.50 ± 0.14	−0.83 ± 0.13	2.30 ± 0.75	53
S8	−0.45 ± 0.04	−0.84 ± 0.11	2.56 ± 0.99	77
S9	−0.43 ± 0.08	−0.81 ± 0.13	2.37 ± 0.81	63
S10	−0.20 ± 0.08	−0.49 ± 0.30	1.59 ± 1.13	64
S11	−0.20 ± 0.06	−0.38 ± 0.31	1.56 ± 1.17	80
S12	−0.30 ± 0.10	−0.76 ± 0.20	1.85 ± 0.62	59
Mean ± SD	−0.37 ± 0.19	−0.68 ± 0.23	2.25 ± 0.73	12

r, correlation coefficient of model regression; RFI, mean reference frame index; VR, variance ratio. All values, bootstrap estimates (mean ± SD). *n,* number of data points.
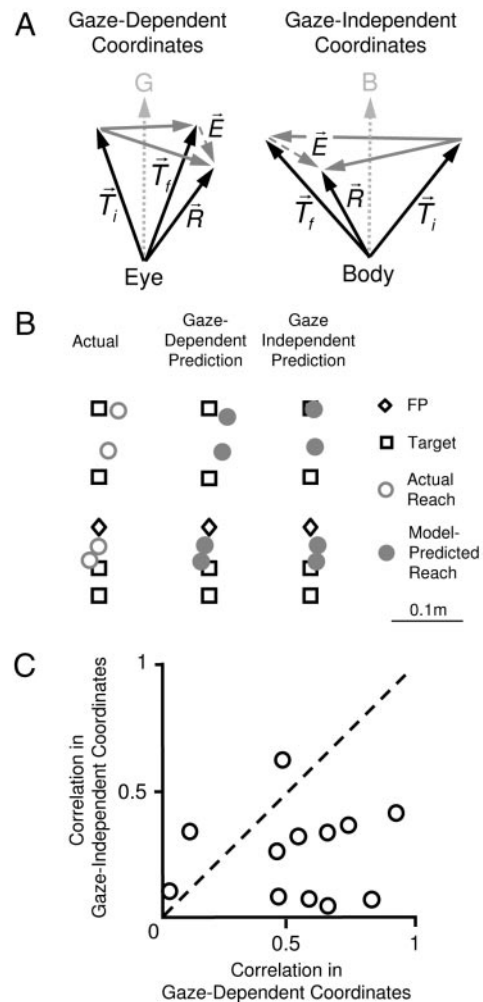

FIG. 5. Two-dimensional (2-D) vectorial analysis of updating performance. *A*: target location before ($\vec{T}_i$) and after ($\vec{T}_f$) the body translation, and reach location ($\vec{R}$) expressed as 2-D vectors in a coordinate frames fixed to either the gaze line, G, (*left*) or the line perpendicular to the shoulder line, B, (*right*). *Eq. 1* was fitted in terms of the predicted updating error $\vec{E}$, in both coordinate frames, with ($\vec{R} - \vec{T}_i$) representing the actual amount and ($\vec{T}_f - \vec{T}_i$) the ideal amount of updating. *B*: actual reach endpoints (○) of 1 subject (*subject S1*) flanked by reach endpoints (●) based on model fits (□) for the 4 targets (□) for rightward translation trials. The gaze-dependent model predicts the actual pattern of endpoints best. *C*: correlation coefficient for the fit in gaze-dependent vs. in gaze-independent coordinates for all subjects. The gaze-dependent model made the best description of the data in 9 of 12 subjects. Two subjects showed very low correlations for both models.

Vectorial analysis

Although the data of most of our subjects lend support for the gaze-dependent updating hypothesis, it should be pointed out that this conclusion is based on an (1-D) analysis of the horizontal reach errors. Because subjects also make updating errors in depth (see Figs. 2), it is desirable to validate this conclusion in a 2-D analysis. Therefore we investigated how the position of the target before the translation ($\vec{T}_i$, estimated by the average response in the stationary task), the position of the same target after the translation ($\vec{T}_f$), the actual translational motion ($\vec{T}_f - \vec{T}_i$) and reach response ($\vec{R}$), expressed as Cartesian 2-D vectors, are related in the coordinate frames of the two updating models (see Fig. 5*A*). The two coordinate axes of the gaze-dependent model were chosen to be aligned with and

orthogonal to the gaze line, respectively, with the origin at the center of the two eyes (cyclopean eye). At the same origin, the coordinate axes of the gaze-independent model were arranged to be aligned with and orthogonal to the shoulder line, respectively. Note that the same (space-fixed) target T_i in this example is described by quite different vectors in each coordinate system. In both coordinate frames, the following updating relationship can be specified

$$\vec{R} - \vec{T}_i = a(\vec{T}_f - \vec{T}_i) + \vec{b} \qquad (1)$$

in which $\vec{T}_f - \vec{T}_i$ represents the ideally required updating vector, $\vec{R} - \vec{T}_i$ the actual updating vector, fit parameter a the updating gain, and vector $\vec{b}$ the bias in the updating process. If a subject had a correct percept of $\vec{T}_i$, but did not account for the intervening translation, reach vector $\vec{R}$ would be equivalent to target vector $\vec{T}_i$, and hence the internal updating vector $\vec{R} - \vec{T}_i$ would equal zero, thus $a = 0$, $\vec{b} = 0$. In contrast, if translational updating were flawless, reach vector $\vec{R}$ would be identical to the new target vector $\vec{T}_f$, and thus $a = 1$, and $\vec{b} = 0$.

We fitted *Eq. 1* in terms of the predicted updating error $\vec{E}$ (see dashed gray vector in Fig. 5*A*). The results of this analysis are shown in Fig. 5*B* for one subject for the rightward translation trials. The actual average endpoints (*left*) are compared with those predicted by each of the two models on basis of the fit parameters of *Eq. 1*. Close scrutiny indicates that the predictions of gaze-dependent model (*middle*) better match the observed reach endpoints than the gaze-independent model (*right*). The gaze-dependent model seems to capture the observed pattern of opposite errors for targets behind and in front of the fixation point, whereas the gaze-independent model shows only a small rightward shift of each of the reach endpoints. On a population level (Fig. 5*C*), *Eq. 1* gave a better description (higher correlation coefficients) of the updating errors when expressed in gaze-dependent coordinates than in gaze-independent coordinates (*t*-test, $P < 0.01$), which is consistent with the 1-D analysis described in the preceding text. Within individual subjects, the gaze-dependent model produced the best description for 9 of 12 subjects. The gaze-independent model performed slightly better in three subjects although its performance remained at a rather low level in two of them (see also Table 2).

Table 2 lists the best-fit coefficients of *Eq. 1*, showing the updating gain, a, and bias vector, $\vec{b}$, for both models for each

subject separately. Across the population, the bias vector was not significantly different from a zero vector (*t*-test, $P > 0.05$ for all components) for both of the two models.

In the gaze-dependent model, the updating gain, *a,* specifies how well the translational-depth geometry is taken into account in the updating of remembered visual space. Averaged across subjects, its value was 1.16 ± 0.15 (SD), which was significantly different from 1 (*t*-test, $P < 0.05$). This suggests that this model takes the systematic reach errors into account in fitting the data or, in other words, that subjects generally overestimated the amount of self-motion when updating targets in 3-D space during active whole-body translations. In contrast, the gaze-independent model yielded an average updating gain that was statistically not distinguishable from 1 (*t*-test, $P = 0.62$), which essentially indicates that this model has no provision to account for the systematic errors observed in the data.

Control experiments

To determine the robustness of these findings, we performed three control experiments (see METHODS). The task designs of these controls were kept identical to that of the main experiment as much as possible. In the analysis of these experiments, each performed on five subjects, we focused on the horizontal reaching errors, investigating the relationship between the errors for targets in front of FP and errors to targets behind FP. As in the preceding text (see Fig. 4), a negative relationship would confirm gaze-dependent coding (ideal slope -1); a positive relationship would be suggestive of a gaze-independent coding scheme (ideal slope $+1$). We first asked whether the same results would be obtained if the reaching movement toward the updated target locations were not accompanied by any visual feedback about hand position (control I: reaching without feedback). The results show that the absence of hand feedback does not alter our main conclusion. All subjects performing the task without hand feedback produced data consistent with the gaze-centered updating hypothesis (see Fig. 6*A*). This is reflected by the average reference frame index, which was -0.70 ± 0.18 and significantly different from a value of 0 (*t*-test, $P > 0.05$).

Next we investigated if the effects were mainly specific to moving the eyes to the updated target locations rather than to moving the hand (control 2: reaching without looking). For eye
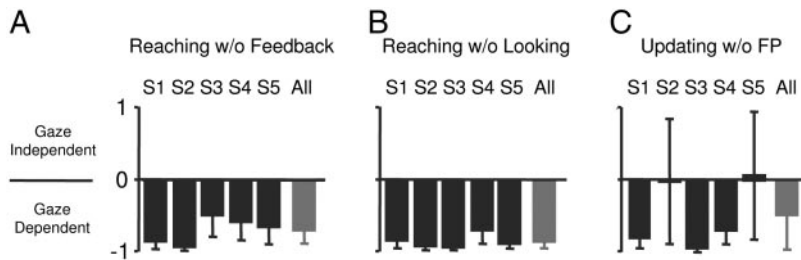
TABLE 2. *Fit performance of Eq. 1 ($\vec{R} - \vec{T}_i = a(\vec{T}_i - \vec{T}_i) + \vec{b}$) in gaze-dependent and gaze-independent coordinates in each subject*

	Gaze-Dependent Model			Gaze-Independent Model		
	r	*a*	$\vec{b}$, cm	*r*	*a*	$\vec{b}$, cm
S1	0.83	1.38	[−0.84, 1.90]	0.06	0.99	[−0.48, 2.34]
S2	0.65	1.26	[0.15, 0.87]	0.32	0.97	[0.46, 1.52]
S3	0.47	1.28	[0.06, 1.00]	0.61	0.95	[0.29, 2.37]
S4	0.58	1.11	[0.73, −0.88]	0.07	1.00	[0.87, −0.88]
S5	0.45	0.89	[1.09, 0.08]	0.07	1.00	[0.77, −0.20]
S6	0.45	1.12	[0.28, 1.10]	0.25	1.02	[0.22, 0.69]
S7	0.65	1.19	[0.31, 2.13]	0.04	1.00	[0.98, 2.24]
S8	0.93	1.10	[0.22, 1.88]	0.4	1.06	[0.21, 1.05]
S9	0.74	1.37	[0.60, 2.39]	0.35	0.95	[0.75, 2.10]
S10	0.09	1.03	[0.36, −1.25]	0.33	0.97	[0.66, −0.96]
S11	0.01	1.00	[−0.02, 0.57]	0.09	0.99	[0.15, 0.65]
S12	0.53	1.15	[−0.26, 1.27]	0.31	1.03	[−0.13, 1.31]
Mean ± SD	0.59 ± 0.42	1.16 ± 0.15	[0.23 ± 0.49, 0.92 ± 1.15]	0.25 ± 0.20	1.00 ± 0.03	[0.39 ± 0.44, 1.02 ± 1.20]

r, correlation coefficients (values also shown in Fig. 5*C*). Best-fit values of *a* and $\vec{b}$ refer to updating gain and bias vector, respectively.

FIG. 6. Results of 3 control experiments, each performed on 5 subjects. Reference frame indices according to Fig. 4*C*: +1 reflects a gaze-independent scheme; −1 the gaze-dependent scheme. *Left*: control I, reaching without visual feedback of the fingertip. All subjects support the gaze-dependent model. *Middle*: control II, reaching without looking at the finger tip provides unanimous support for the gaze-dependent model. *Right*: control III, updating without FP. Subject translated their body keeping gaze fixed on a remembered fixation point. Although the reference frame test may be less discriminative due the vergence drift, clear support for the gaze-dependent model can be seen in most subjects. Error bars, bootstrap confidence intervals.

movements, the sensory frame of reference imposed by the retina and oculomotor reference frames for the eyes are quite similar (Snyder 2000). Hence for the eyes to look at the remembered target locations, saccadic amplitude must depend nonlinearly on target depth and direction. If saccadic amplitude was not scaled appropriately (Medendorp et al. 2003b) and the eyes lead the arm, the errors that appeared reflected an eye-centered motor representation rather than information about the spatial representation that codes the target. Arm movements do not suffer from this drawback: the sensory frame of the retina is quite distinct from the motor frame of reference imposed by the joints and muscles of the arm (Snyder 2000). Therefore in this control experiment, subjects were instructed to keep gaze fixed at FP at all times during the trials as well as when probing the remembered target by the reach. These results show once again clear evidence for the gaze-dependent coding scheme (Fig. 6*B*). All subjects had reference frame indices significantly smaller than zero. Moreover, the average RFI across subjects was −0.87 ± 0.11, which was significantly different from zero (*t*-test, $P < 0.05$) but not from −1 (*t*-test, $P > 0.05$).

Finally we asked whether the visual FP, available during the main experiments, was a biasing factor for the gaze-centered updating hypothesis. To test this, we conducted an experiment in which subjects had to keep their eyes fixated on the remembered FP during the self-motion and then looked and reached to the remembered location of the flashed target (control 3: updating without FP, Fig. 6*C*). It is important to realize that in this situation, our test has less discriminative capabilities. Because of possible vergence drift caused by the absence of a visual FP during translation in this paradigm, updating vectors in gaze-coordinates will not be of equal size for targets in front and behind FP (compare Fig. 1). In spite of that, across the five subjects that participated here, three followed the gaze-dependent model. The RFIs in the other two subjects had values around zero. Averaged across subjects, we found a RFI of −0.48 ± 0.47—a clear bias in favor of the gaze-dependent updating model.

Taken together, the results of all our experiments lead to the conclusion that the brain uses a gaze-dependent reference frame to store and update visuospatial memories during self-generated whole-body translations.

DISCUSSION

Inspired by the work of Von Helmholtz, investigators have made abundantly clear over the last decades that humans can remember visual direction across rotary eye and head movements (Blouin et al. 1998; Hallet and Lightstone 1976; Herter and Guitton 1998; Medendorp et al. 2002; Schlag et al. 1990; Von Helmholtz 1867; Wexler 2005). Since Gibson, vision

scientists have also become aware of the complexity of motion parallax for seeing in depth when the eyes translate through space (Gibson et al. 1955; Rogers and Graham 1979). Here we have exploited a paradigm based on the conjunction of these two challenges for visual stability testing if the brain internally simulates motion parallax when updating remembered visual space during active whole-body translations. We called this the gaze-dependent hypothesis as it predicts a systematic pattern of updating errors depending on gaze fixation if the intervening translation is not correctly taken into account with the errors reversing in direction for targets at opposite depths from gaze fixation. As a contrasting hypothesis, we set up the predictions of a gaze-independent coding scheme. According to this hypothesis, the brain codes remembered space irrespective of gaze fixation and therefore predicts no such reversal of updating errors if translations are misjudged. We emphasize that the central premise behind our test was that subjects misestimate their traveled distance during self-generation motion as shown by many studies (Glasauer et al. 1994; Israel et al. 1993; Kudoh 2005; Medendorp et al. 1999; Philbeck and Loomis 1997), although the exact explanation for why this occurs is not directly relevant (but see later). Our results show that target updating for translational motion is compromised by small errors, which increase with depth from fixation and reverse in direction for opposite depths from fixation. This is consistent with the gaze-dependent prediction, so we conclude that the brain employs a gaze-centered mechanism to internally update remembered visual space during whole-body translations.

We will now list a number of observations that further support this conclusion. First, reaching errors were larger in translation trials (with intervening body translation) than in the stationary trials (without body translation), suggesting that the differences indeed arose during the updating of spatial information (Fig. 3). Second, a quantitative analysis of these errors revealed that they were opposite for targets in front of and behind FP (Fig. 4). Third, a two-dimensional vectorial analysis of the translational-depth geometry in the transversal plane showed that the interaction among target location, translational motion, and reaching response is much better described in a gaze-centered than in a gaze-independent coordinate system (Fig. 5). Fourth, the gaze-centered updating errors were quite robust and invariable among various task constraints (Fig. 6). More specifically, the same error pattern was found irrespective of whether the eyes and hand moved to the memorized target location or the hand alone. Neither did the pattern of errors change when subjects performed the reaching movement with or without visual feedback of hand position. Even the presence or absence of a visual fixation point during the translations was not essential for a gaze-centered description of updating errors.

Although our data provide support for the gaze-dependent model across subjects, it is important not to overstate this. The results are not perfect, and our conclusions follow from relatively small systematic errors. As a matter of fact, three of our subjects did not show support for the gaze-dependent hypothesis in all conditions and analyses (see Figs. 5*C* and 6*C*). It is also important to note that our test was based on relative simple geometry, whereas the brain may actually represent visual space in a more complex manner (Cuijpers et al. 2002). Furthermore, we should emphasize that we have focused on only one important signal, the central representation of body translation, as an underlying basis for the updating errors, which is but one of a myriad of variables which might lead to errors. In this respect, further experiments are needed to isolate the various signals related to overall performance of the present task. Nevertheless, despite these reservations, we think that our behavioral tests provide evidence that the brain possesses a geometrically complete, dynamic map of remembered space, the spatial accuracy of which is maintained by internally simulating motion parallax during volitional translatory body movements.

It is true that even when you walk around normally in the environment, it is difficult to experience motion parallax even if you try (Palmer 1999). And without doubt it is even harder to imagine motion parallax with locations of remembered objects or objects that are out of view. Nevertheless, this cannot be taken to imply that the neural mechanism for spatial coding cannot act by simulating the parallax geometry to maintain spatial constancy as we have shown here.

Recently various studies have shown that both human and non-human primates can adjust the amplitude of memory-guided eye movements after intervening translation, taking into account the amount of translation and distance of the memorized target (Israel and Berthoz 1989; Li and Angelaki 2005; Li et al. 2005; Medendorp et al. 2003b). None of these studies, however, explicitly assessed the exact nature of the representation of remembered visual space during these tasks. Here for the first time, we were able to establish that targets in such tasks are stored in a gaze-centered reference frame, an inference based on the assessment of the operational errors in the system.

Our evidence for gaze-centered updating during translational motion agrees well with recent studies showing gaze-centered updating for rotational motion (Baker et al. 2003; Henriques et al. 1998; Medendorp and Crawford 2002; Pouget et al. 2002). The first three showed that subjects overshoot the direction of a previously seen but foveally viewed target when reaching toward it after an intervening eye rotation. Interestingly, here we show a similar type of overshoot for translation-induced changes of gaze, corroborating these gaze-centered results. Baker et al. (2003) investigated updating behavior during horizontal whole-body rotations using a memory-guided saccade task. Based on the assumption of noise propagation at various processing stages in the brain, they found their results most consistent with a gaze-centered representational system for storing the spatial locations of memorized objects.

Which signals are needed in the updating process? In the present study, the updating mechanism may have received information about the self-motion through efference copy and proprioceptive signals (available in the context of active motion), and by vestibular inputs (Klier et al. 2005; Li and

Angelaki 2005; Li et al. 2005; Medendorp et al. 2003b; Van Pelt et al. 2005). Li et al. (2005) found updating during passive translation to be compromised after bilateral labyrinthectomy, attributing an important role of the vestibular system. Also, Israel and Berthoz (1989) have provided evidence for spatial updating with the vestibular system as the main extraretinal source of motion-related information. Furthermore, in the present study, the changes in eye position to keep the eyes fixed at FP during the translation—the version and vergence eye movements—are essential for a well-functioning updating system. All of this information must be must be integrated at a central level within the brain and unified with retinal information about target direction and depth to mediate the computations for gaze-centered spatial updating, as outlined in detail in Medendorp et al. (2003b).

In line with our findings, many brain regions have been demonstrated to store and update target locations within an eye-fixed, gaze-centered reference frame (Batista et al. 1999; Duhamel et al. 1992; Gnadt and Andersen 1988; Medendorp et al. 2003a; Merriam et al. 2003; Sommer and Wurtz 2002). However, the majority of these studies have focused on direc-
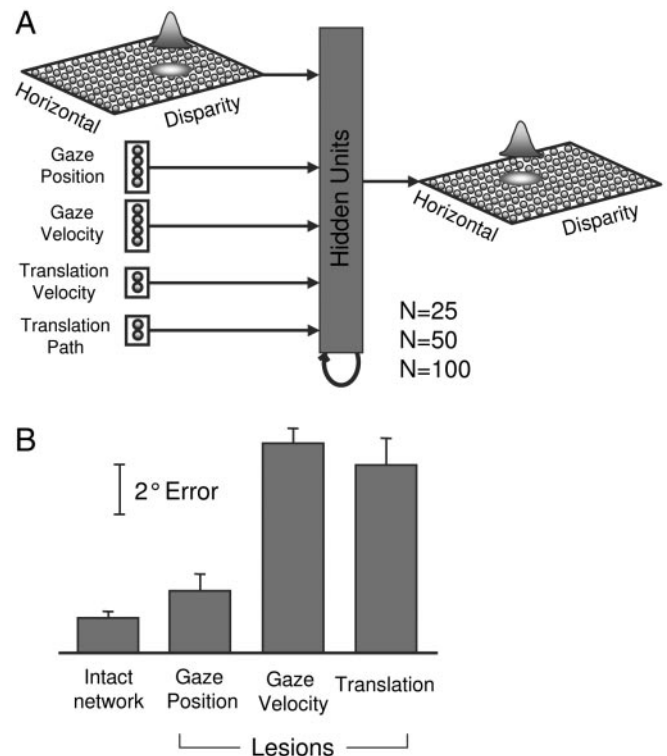

FIG. 7. *A*: diagram of the 3-layer recurrent network model trained to perform gaze-centered updating during both eye rotations and translations. Inputs: target location as a 2-D Gaussian hill of activity within a 11×11 units horizontal-disparity (retinal eccentricity-depth) map and extraretinal signals, including gaze position (version/vergence), gaze velocity signals (version/vergence), and the eyes' translation velocity and path signals. Hidden layer contained 25, 50, or 100 units. Output layer encodes the memory of the target in terms of direction and disparity relative to the binocular fixation point, i.e., in gaze-centered coordinates. *B*: performance of the network ($n = 50$) when particular input signals are removed. Updating error in the direction of the target (in degrees) is shown for the intact network, the network with gaze position inputs removed, the network with gaze velocity inputs eliminated, and the network with translational inputs removed. The network has a strong preference for gaze velocity inputs over gaze position inputs. Similar results were obtained for the networks trained with 25 or 100 hidden units. Error bars denote SD for 4 networks.

tional updating of target location in the frontal parallel plane. For example, the lateral intraparietal area and superior colliculus have been shown to update its retinotopic map of target directions for each eye movement (Duhamel et al. 1992; Walker et al. 1995). On the other hand, it also known that the activity of LIP neurons is modulated by retinal disparity information, providing them with three-dimensional receptive fields (Genovesio and Ferraina 2004; Gnadt and Mays 1995). Moreover, Cumming and DeAngelis (2001) indicated that the updating of target distance may be expressed by changes in retinal disparity representations.

To obtain further insights in the interactions between self-motion information and retinal signals at the level of the parietal cortex, we designed a simple recurrent neural network performing gaze-centered target updating during translations and rotations (see Fig. 7A and METHODS). The input to the network was a transient distributed representation of target direction and disparity in a 2-D retinotopic map (as a hill of activity) as well as a variety of extraretinal signals, including angular gaze position and velocity signals (version/vergence), and translational velocity and path signals of the eyes. The network was trained to store the memory of the target for successive time intervals and update its representation for any intervening rotational or translational eye motion.

Figure 8 shows the simulation results for updating a target in front of ("near" target) and target behind ("far" target) the eyes' fixation point during a translational motion of the eyes. The extraretinal signals involved are the same in both situations (Fig. 8A), for which the geometrical relationships are depicted in Fig. 8B. As shown in Fig. 8C, the near target appeared at 8° to the right of the center of gaze, at −2° disparity, and shifted to an 8° deg leftward, −2° disparity

position after the rightward translational motion. The activity pattern of the far target evolves in the opposite direction of the map during the translation (Fig. 8D). In other words, the updating network must have used information about target depth to determine how the hill of activity should move over the map. The exact location of the target was decoded from the map by means of a weighted average of the activity of all neurons (see Fig. 8, C and D, bottom, ○), which closely follow the geometrically required changes for ideal updating over time (thin lines). Likewise, the network also incorporated the geometrically-required properties of updating targets in the same direction on the map, irrespective of their depth, when the eyes rotate only (not shown). Using 25 neurons in the hidden layer was already sufficient to learn the task acceptably, but performance improved for the 50 and 100 hidden units networks.

Because the network was trained to perform these tasks under the provision of extraretinal position and velocity signals, an interesting question to ask is whether one input is more relied on than another (White and Snyder 2004). To this end, we removed one of the inputs after training ("artificial lesion") and looked at the performance of the network in terms of its updating errors (Fig. 7B). As the figure shows, the network has a clear preference for gaze velocity over gaze position inputs, which is consistent with findings by White and Snyder (2004) for rotational updating. The use of velocity signals may give the network a benefit to update continuously, irrespective of initial or final gaze position. Thus our simulation results provide good evidence for the idea that the brain synthesizes ego-velocity signals and stereoscopic depth and direction information to update the internal representation of 3-D space during self-motion. This integration may occur in parietal area LIP using the computations that we have described or in any
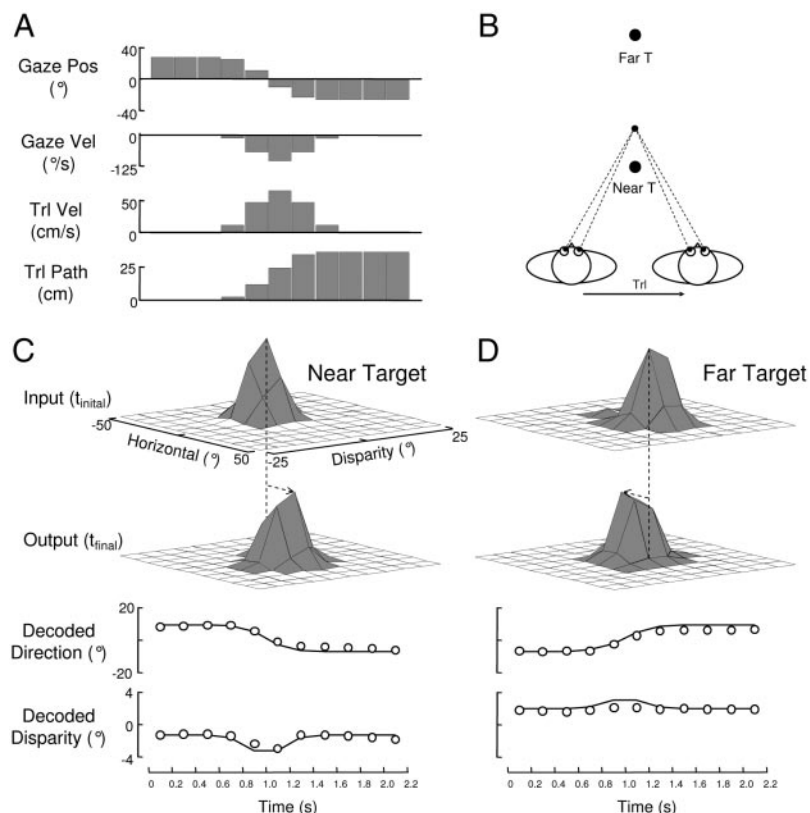

FIG. 8. Network performance for target updating during a rightward translation trial. A: activation of the units representing the eyes' rotational and translational kinematics at each time step. B: geometry that has been simulated. C: updating of a target flashed in front of the eyes' fixation point. The hill of activity coding the target memory shifts across the horizontal-disparity map. Bottom: target representation encoded by the output layer, showing that a near target shifts from right to left relative to the gaze line (○), as geometrically required (—). The network also matches the required changes in disparity. D: target farther away than the fixation point shifts in the opposite direction on the map. Thus activity patterns evolve during translation in a way that depends on target depth.

other cortical or subcortical structures involved in updating as long as they have the necessary signals at their disposal.

R E F E R E N C E S

Andersen RA, Essick GK, Siegel RM. Encoding of spatial location by posterior parietal neurons. *Science* 230: 456–458, 1985.

Baker JT, Harper TM, Snyder LH. Spatial memory following shifts of gaze. I. Saccades to memorized world-fixed and gaze-fixed targets. *J Neurophysiol* 89: 2564–2576, 2003.

Batista AP, Buneo CA, Snyder LH, Andersen RA. Reach plans in eye-centered coordinates. *Science* 285: 257–260, 1999.

Battaglia-Mayer A, Caminiti R, Lacquaniti F, Zago M. Multiple levels of representation of reaching in the parieto-frontal network. *Cereb Cortex* 13: 1009–1022, 2003.

Beurze SM, Van Pelt S, Medendorp WP. Behavioral reference frames for planning human reaching movements. *J Neurophysiol* 96: 352–362, 2006.

Blouin J, Labrousse L, Simoneau M, Vercher JL, Gauthier GM. Updating visual space during passive and voluntary head-in-space movements. *Exp Brain Res* 122: 93–101, 1998.

Bridgeman B, Peery S, Anand S. Interaction of cognitive and sensorimotor maps of visual space. *Percept Psychophys* 59: 456–469, 1997.

Carrozzo M, Stratta F, McIntyre J, Lacquaniti F. Cognitive allocentric representations of visual space shape pointing errors. *Exp Brain Res* 147: 426–437, 2002.

Cumming BG, DeAngelis GC. The physiology of stereopsis. *Annu Rev Neurosci* 24: 203–238, 2001.

Cuijpers RH, Kappers AM, Koenderink JJ. Visual perception of collinearity. *Percept Psychophys* 64: 392–404, 2002.

Duhamel JR, Colby CL, Goldberg ME. The updating of the representation of visual space in parietal cortex by intended eye movements. *Science* 255: 90–92, 1992.

Genovesio A, Ferraina S. Integration of retinal disparity and fixation-distance related signals toward an egocentric coding of distance in the posterior parietal cortex of primates. *J Neurophysiol* 91: 2670–2684, 2004.

Gibson JJ, Olum P, Rosenblatt F. Parallax and perspective during aircraft landings. *Am J Psychol* 68: 372–385, 1955.

Glasauer S, Amorim MA, Vitte E, Berthoz A. Goal-directed linear locomotion in normal and labyrinthine-defective subjects. *Exp Brain Res* 98: 323–335, 1994.

Gnadt JW, Andersen RA. Memory related motor planning activity in posterior parietal cortex of macaque. *Exp Brain Res* 70: 216–220, 1988.

Gnadt JW, Mays LE. Neurons in monkey parietal area LIP are tuned for eye-movement parameters in three-dimensional space. *J Neurophysiol* 73: 280–297, 1995.

Hallett PE, Lightstone AD. Saccadic eye movements towards stimuli triggered by prior saccades. *Vision Res* 16: 99–106, 1976.

Hayhoe MM, Shrivastava A, Mruczek R, Pelz JB. Visual memory and motor planning in a natural task. *J Vis* 3: 49–63, 2003.

Henriques DYP, Klier EM, Smith MA, Lowy D, Crawford JD. Gaze-centered remapping of remembered visual space in an open-loop pointing task. *J Neurosci* 18: 1583–1594, 1998.

Henriques DYP, Medendorp WP, Gielen CCAM, Crawford JD. Geometric computations underlying eye-hand coordination: orientations of the two eyes and the head. *Exp Brain Res* 152: 70–79, 2003.

Herter TM, Guitton D. Human head-free gaze saccades to targets flashed before gaze-pursuit are spatially accurate. *J Neurophysiol* 80: 2785–2789, 1998.

Israel I, Berthoz A. Contribution of the otoliths to the calculation of linear displacement. *J Neurophysiol* 62: 247–263, 1989.

Israel I, Chapuis N, Glasauer S, Charade O, Berthoz A. Estimation of passive horizontal linear whole-body displacement in humans. *J Neurophysiol* 70: 1270–1273, 1993.

Israel I, Ventre-Dominey J, Denise P. Vestibular information contributes to update retinotopic maps. *Neuroreport* 10: 3479–3483, 1999.

Klier EM, Angelaki DE, Hess BJM. Roles of gravitational cues and efference copy signals in the rotational updating of memory saccades. *J Neurophysiol* 94: 468–478, 2005.

Kudoh N. Dissociation between visual perception of allocentric distance and visually directed walking of its extent. *Perception* 34: 1399–1416, 2005.

Li N, Angelaki DE. Updating visual space during motion in depth. *Neuron* 48: 149–158, 2005.

Li N, Wei M, Angelaki DE. Primate memory saccade amplitude after intervened motion depends on target distance. *J Neurophysiol* 94: 722–733, 2005.

Medendorp WP, Crawford JD. Visuospatial updating of reaching targets in near and far space. *Neuroreport* 13: 633–636, 2002.

Medendorp WP, Goltz HC, Vilis T, Crawford JD. Gaze-centered updating of visual space in human parietal cortex. *J Neurosci* 23: 6209–6214, 2003a.

Medendorp WP, Smith MA, Tweed DB, Crawford JD. Rotational remapping in human spatial memory during eye and head motion. *J Neurosci* 22: 196RC, 2002.

Medendorp WP, Tweed DB, Crawford JD. Motion parallax is computed in the updating of human spatial memory. *J Neurosci* 23: 8135–8142, 2003b.

Medendorp WP, Van Asselt S, Gielen CCAM. Pointing to remembered visual targets after active one-step self-displacements within reaching space. *Exp Brain Res* 125: 50–61, 1999.

Merriam EP, Genovese CR, Colby CL. Spatial updating in human parietal cortex. *Neuron* 39: 361–373, 2003.

Palmer SE. *Vision Science: Photons to Phenomenology.* Cambridge MA: MIT Press, 1999.

Philbeck JW, Loomis JM. Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions. *J Exp Psychol Hum Percept Perform* 23: 72–85, 1997.

Poggio GF, Poggio T. The analysis of stereopsis. *Annu Rev Neurosci* 7: 379–412, 1984.

Pouget A, Ducom JC, Torri J, Bavelier D. Multisensory spatial representations in eye-centered coordinates for reaching. *Cognition* 83: 1–11, 2002.

Press WH, Flannery BP, Teukolsky SA, Vettering WT. *Numerical Recipes in C* (2nd ed.). Cambridge, MA: Cambridge, 1992.

Rogers B, Graham M. Motion parallax as an independent cue for depth perception. *Perception* 8: 125–134, 1979.

Schlag J, Schlag-Rey M, Dassonville P. Saccades can be aimed at the spatial location of targets flashed during pursuit. *J Neurophysiol* 64: 575–581, 1990.

Sommer MA, Wurtz RH. A pathway in primate brain for internal monitoring of movements. *Science* 296: 1480–1482, 2002.

Snyder LH. Coordinate transformations for eye and arm movements in the brain. *Curr Opin Neurobiol* 10: 747–754, 2000.

Snyder LH, Grieve KL, Brotchie P, Anderson RA. Separate body- and world-referenced representations of visual space in parietal cortex. *Nature* 394: 887–891, 1998.

Van Pelt S, Van Gisbergen JA, Medendorp WP. Visuospatial memory computations during whole-body rotations in roll. *J Neurophysiol* 94: 1432–1442, 2005.

Von Helmholtz H. *Treatise on Physiological Optics.* New York: Dover, 1962, vol 3; English translation by Southall JPC for the Optical Society of America 1925 from *Handbuch der Physiologischen Optik* (3rd ed.). 1867.

Walker MF, Fitzgibbon J, Goldberg ME. Neurons of the monkey superior colliculus predict the visual result of impeding saccadic eye movements. *J Neurophysiol* 73: 1988–2003, 1995.

Wexler M. Anticipating the three-dimensional consequences of eye movements. *Proc Natl Acad Sci USA* 102: 1246–1251, 2005.

White RL 3rd, Snyder LH. A neural network model of flexible spatial updating. *J Neurophysiol* 91: 1608–1619, 2004.