The interaction between gaze and pointing towards remembered visual targets

M.A. Admiraal, N.L.W. Keijsers, C.C.A.M. Gielen
Dept. Biophysics, University of Nijmegen

Contact information:

M.A. Admiraal
Dept. Biophysics, University of Nijmegen
POBox 9101, 6500 HB Nijmegen
The Netherlands
e-mail: marjana@mbfys.kun.nl
phone: +31 (24) 3613326
fax: +31 (24) 3541435
ABSTRACT

We examined the role of gaze in a matching task, where subjects had to reproduce the position of remembered visual target with the tip of the index finger. For brevity, we will refer to this task as pointing. Subjects were tested in three visual feedback conditions: complete darkness (DARK), complete darkness with visual feedback of the finger position (FINGER), and with vision of a well-defined environment and feedback of the finger position (FRAME). Pointing accuracy increases with feedback about the finger or visual environment. In the FINGER and FRAME conditions, the 95% confidence regions of the variable errors have an ellipsoidal distribution with the main axis oriented towards the subjects' head.

During the one-second period, when the target is visible, gaze is almost on target. However, gaze drifts away from the target relative to the subject in the delay period after target disappearance. In the FINGER and FRAME conditions, gaze returns towards the remembered target during pointing. In all three feedback conditions, the correlations between the variable errors of gaze and pointing position increase during the delay period, reaching highly significant values at the time of pointing.

Our results demonstrate that gaze affects the accuracy of pointing. We conclude that the covariance between gaze and pointing position reflects a common drive for gaze and arm movements and an effect of gaze on pointing accuracy at the time of pointing.

Previous studies interpreted the orientation of variable errors as indicative for a frame of reference used for pointing. Our results suggest that the orientation of the error ellipses towards the head is at least partly the result of gaze drift in the delay period.

keywords:

Reaching, proprioception, reference frame, pointing, vision
INTRODUCTION

One of the main problems in understanding human motor control is related to the frames of reference, which are used for the preparation, planning, and execution of movements. Visual information about targets in 3-D space is initially coded in retinal coordinates. The visual information in retinal coordinates has to be combined with information of eye and head position in order to determine target position relative to the body. Finally, a pointing or grasping movement to a target requires a specific muscle activation pattern that brings the hand to the target. Obviously, both visual information about target position and proprioceptive information about arm position contribute to the final finger position, each in a different frame of reference. In order to gain more insight in the frames of reference that may be used, many studies have focused on movements of the hand to match the position of a visible or remembered visual targets. In previous literature, such movements are commonly referred to as pointing movements. In this paper we will use this term accordingly, although the term pointing may be somewhat misleading for the actual matching task.

Most studies on pointing towards remembered visual targets have in common that they showed that the distribution of finger positions for pointing to a remembered target is characterized by an ellipsoid with the long axis of the distribution oriented towards the subject. This has been interpreted as evidence for the hypothesis that the central nervous system (CNS) specifies the parameters of the endpoint of the movement separately for distance and direction (Georgopoulos et al. 1984, Flanders et al. 1992, Gordon et al. 1994), that the movement is planned in terms of displacement from the initial position (Messier and Kalaska 1997, Vindras et al. 1998), or as evidence for movement planning in a viewer-centered frame of reference (e.g. McIntyre et al. 1997). Other studies, however, interpreted errors in pointing as evidence for movement planning in a shoulder-centered frame of reference (Soechting and Flanders 1989, Berkinblit et al. 1995), or in both a shoulder-centered and a head-centered frame of reference (Soechting et al. 1990,
McIntyre et al. 1998). The different results regarding the variability in the orientation of the ellipses in these studies are most likely explained by differences in the experimental conditions with which different studies approached the issue: in some studies, subjects had no visual information whatsoever on the environment, nor on their arm (McIntyre et al. 1997), whereas in other studies subjects had feedback on their arm (Berkinblit, et al. 1995) or on both the arm and the visual environment (Soechting and Flanders 1989 and McIntyre et al. 1998). However, since none of these studies tested subjects for pointing in all these different experimental conditions, it is not clear to what extent the different contributions of information about finger position and about target position relative to the visual environment could explain the different results. In order to investigate the effect of vision of the finger and the environment on the pointing errors and in particular the effect of visual feedback of the finger, we have measured pointing movements in three visual conditions: 1) Complete darkness with visual feedback of the finger position; 2) Complete darkness without visual feedback on the position of the index finger; and 3) Vision of the index finger along with vision of a well-defined visual environment. A previous study by Van Beers et al. (2002) has demonstrated that the contributions of visual and proprioceptive information may vary, depending on the experimental conditions. The first aim of this study was therefore, to compare the distributions of pointing errors in each of the three conditions, in order to see whether the different results of previous studies could be explained by different contributions of visual and proprioceptive information in different experimental conditions. In particular, we tested whether the error distributions in each condition were oriented towards a single point, which might be interpreted as the center of some frame of reference for pointing (e.g. viewer-centered or shoulder-centered) and whether differences in the experimental conditions would change the location of any such point.

Several studies (Bock 1986, Enright 1995, Henriques et al. 1998 and Medendorp and Crawford 2002a) have stressed the importance of fixation of gaze to a target on pointing accuracy by demonstrating that pointing errors increase when gaze deviates from the target position. This
observation adds another complicating factor to the interpretation of the error ellipses that have been obtained in previous studies. Eye position has not been measured in most of the studies on pointing to remembered targets and fixation might very well have been different in conditions with and without visual feedback about the environment. If pointing errors depend on gaze and if gaze is different with and without visual feedback on the finger and the environment, different orientations of error ellipses might also be due to differences in gaze.

Therefore, the second aim of this study was to compare the errors in pointing and in gaze as a function of time in the period from target onset until completion of the pointing movement, and to look for the presence (or absence) of a correlation between constant and variable errors of 3-D gaze and of 3-D finger position during pointing.
MATERIALS AND METHODS

In this study we performed two experiments. In the first experiment we measured movements of
the arm during pointing towards remembered visual targets. In the second, we simultaneously
measured pointing movements of the arm and binocular eye movements to determine 3-D gaze.
Fifteen subjects (aged 21-49 years) participated in these experiments. Ten subjects participated in
the first experiment and six subjects participated in the second experiment. One subject (RK)
participated in both experiments. All subjects had normal or corrected to normal vision and all
gave informed consent. The experimental protocol was approved by the Medical Ethical
Committee of the University of Nijmegen. None of the subjects had any known history of
neurological sensory or motor disorders. All subjects were right handed, except for subject MA,
who participated only in the second experiment. Pointing movements were performed with the
right hand, unless explicitly indicated otherwise. Three subjects (MA, NK and SG) were familiar
with the aim of this study. Their results were not different from those of the other subjects.

Experimental paradigm

Subjects were standing in a completely dark room. An L-shaped obstacle was attached on the
gfloor to offer the subject a reference to maintain the correct location in the otherwise dark room in
all three test conditions. Seven red light-emitting diodes (leds) were attached on the vertices of
two 30 x 30 x 30 cm³ cubes, next to each other, about 25 cm in front of the subject (see Figure 1).
Each of these seven leds served as a target for pointing movements in the first experiment.
Targets 1, 4 and 5 were used in the second experiment.
In the first experiment, the subject was standing about 40 cm from the center of the 30x60x30cm frame, i.e. 25 cm from the front of the frame with the seven targets. Targets (black dots) were located 15 cm above and below the shoulder, and 15 cm to the right (targets 2, 3, and 7), 15 cm to the left (targets 1, 4 and 6) and 45 cm to the left (target 5) of the shoulder, such that the workspace of the right shoulder ranged from -30 to +60 degrees in azimuth, from -30 to +30 degrees in elevation, and from about 30 to 55 cm in distance. After target disappearance the frame was canted away (arrow). In the second experiment the subject stood right in front of target 4 and the framework was elevated, such that the upper targets were at eye level. Solid grey lines indicate the 90x90x90cm frame of optic fibers, which was illuminated in the FRAME condition.
The onset of the target marked the start of a trial. After one second the target led was switched off and the cubes with targets were canted away. Two seconds after target disappearance an auditory signal notified the subject to start the pointing movement to the remembered target. Subjects were instructed to wait for the auditory signal before positioning their index finger at the remembered target position, and to keep it at the position of the remembered target for at least half a second. No instruction was given about where to direct gaze, and subjects could freely move their head and eyes.

Three visual feedback conditions were tested: pointing in complete darkness (DARK), pointing with feedback by means of a red light emitting diode (led) on the tip of the index finger which was visible at all times (FINGER), and pointing in the presence of an illuminated cubic frame with a continuously lit red led attached on the tip of the index finger (FRAME). In the latter condition, a well-defined visual environment was shown to the subject by means of illuminated optic fibers along the edges of a cubic frame of 90 x 90 x 90 cm³ (see Figure 1). The surface of the optic fibers was roughened by sandpaper and red leds at the long ends of the optic fibers gave the optic fibers a red color. The frame was visible at all times in the FRAME-condition. All targets were within this illuminated cubic frame, well within reaching distance from the subject.

In the first experiment, we investigated pointing movements without measuring gaze, and tested pointing movements to the seven targets in each of the three visual feedback conditions. The visual feedback conditions were tested in pairs (DARK-FINGER and FINGER-FRAME). Measuring two feedback conditions took about an hour, and measuring all three feedback conditions in one experimental session would exceed the maximum amount of time that the subjects could remain concentrated. Four subjects participated in both pairs of conditions, and four other subjects participated in either the DARK-FINGER pair or the FINGER-FRAME pair.
As a result, we tested six subjects in the DARK and in the FRAME condition and all eight subjects were tested in the FINGER condition.

<table>
<thead>
<tr>
<th>Subject</th>
<th>FINGER(right)</th>
<th>DARK(left)</th>
<th>FRAME(right)</th>
<th>FINGER(left)</th>
<th>COIL (FRAME-FINGER-DARK)</th>
</tr>
</thead>
<tbody>
<tr>
<td>at</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dl</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>jl</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>mv</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nk</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>wv</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fh</td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>mk</td>
<td>2</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>fw</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>rk</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>ma</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>bb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>hn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>sg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>ba</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>10 (16)</strong></td>
<td><strong>6</strong></td>
<td><strong>6</strong></td>
<td><strong>6</strong></td>
<td><strong>6</strong></td>
</tr>
</tbody>
</table>

Table 1: Number of times of participation per condition by each subject.

At least six subjects participated in each condition. Due to pair-wise testing of the three feedback conditions, six subjects participated twice in the FINGER condition for pointing with the right hand.

In this experiment, we also examined the influence of the effector arm (left or right) for pointing to targets in the same workspace relative to the shoulder of the pointing arm. Therefore, we also had subjects to point with the left arm in the FINGER condition. Targets were presented at mirror-symmetric locations relative to the left shoulder. Four subjects performed pointing movements with the left and right arm (LEFT-RIGHT pair) and two subjects performed pointing
movements with the left arm only (LEFT ONLY). This resulted in a total of six subjects performing pointing movements with the left arm in the FINGER condition. Two of the subjects who pointed in the LEFT-RIGHT condition pair had not participated in the DARK-FINGER or FINGER-FRAME pairs. Therefore, the total number of subjects that performed pointing movements with the right arm in the FINGER condition increased to ten. Six subjects were tested twice in the FINGER condition for pointing with the right arm (See Table 1).

The targets were presented in a randomized order in sixteen blocks of twenty trials each. In each block, only one visual feedback condition was tested. For each of the visual feedback conditions, each target was presented 20 times, except for the target that was closest to the subject's eyes (target 4), which was presented 40 times. A block with 20 trials typically lasted about three minutes, and after each block, room lights were switched on for about one minute to avoid dark-adaptation.

In the second experiment we measured pointing movements and binocular eye-movements using the search-coil technique. In this experiment, subjects were tested in each of the three conditions (DARK, FINGER, and FRAME) in one experimental session. Since the duration of these experiments had to be restricted to 45 minutes (because of the limited time available to wear the search coils without discomfort), only targets 1, 4 and 5 were used, which were presented at least 13 times each in each condition. All three targets were at eye level. Targets were presented in a randomized order in six blocks of twenty trials each. In each block, only one visual feedback condition was tested. Blocks with different visual feedback conditions were tested in randomized order.

Experimental set-up
The position of several segments of the subject's body and the position of the targets were measured with an OPTOTRAK 3020 system (Northern Digital), which measures the three-
dimensional position of infrared-light-emitting-diodes (ireds) with a resolution better than 0.2 mm within a range of about 1.5 m³. The OPTOTRAK system was mounted on the ceiling above the subject at a distance of approximately 2.5 m to the right of the subject, tilted downward at an angle of 30 degrees relative to the ceiling. When pointing movements with the left arm were measured, the subject and the framework with targets was rotated 180 degrees, for better visibility of the pointing arm to the OPTOTRAK system. The position of ireds was measured with a sampling frequency of 100Hz.

Ireds were placed on the subject’s shoulder (acromion) and elbow (epicondylus lateralis). The position of the tip of the index finger was measured by means of an ired attached on a thimble on the index finger. This thimble also contained a visible red led that provided the subject with feedback on finger position in the FINGER and FRAME conditions. When gaze was measured, subjects were wearing a helmet with six ireds, which were attached in such a way that at least three ireds were visible for the OPTOTRAK system at all possible head orientations. This was necessary to calculate 3-D head orientation at all times, such that eye position could be reconstructed from head position (see below).

Gaze was measured using the scleral search coil technique (Collewijn et al. 1975) in a large magnetic field system (Remmel Labs). This system consists of a cubic frame of welded aluminum of 3 x 3 x 3 m³, which produced three orthogonal magnetic fields at frequencies of 48, 60, and 80 kHz. During these experiments subjects were tested such that the search coils were close to the center of the large magnetic field system. Care was taken that the calibration of the eye coil signals was performed in the same region of the magnetic field where the actual measurements took place. During the calibration procedure subjects fixated a series of red light-emitting diodes (leds) attached to a board at a distance of 90 cm in front of the subject. The leds were arranged at three circles of different radius (15, 27.5 and 37.5 degrees), concentric around the straight-ahead direction. With this setup, calibration errors -defined as twice the standard deviation - were typically about 0.5° in azimuth and 1° in elevation on average; resolution
was<0.04°. As a result, the errors in 3D-gaze position -resulting from calibration errors in the orientation of the two eyes in space - were on average about 0.6° and 1.1° in azimuth and elevation, respectively, and 3 cm in radial distance from the cyclopean eye.

Two PC’s controlled the experiment, one of which was equipped with hardware and software for the collection of the search-coil data and with software for the stimulus presentation. The second PC contained hardware and software to collect the ired data from the OPTOTRAK system, and was controlled by the first PC in order to synchronize the ired data collection (second PC) with the collection of the search-coil data (first PC). Coil signals were sampled at 500Hz. The OPTOTRAK system collected position data from the irreds at a sample frequency of 100Hz. In offline analyses, the coil signals were resampled at a 100Hz frequency by cubic spline interpolation.

Data analysis

We distinguish two types of pointing errors: the constant error, which is the distance between the led position of a target and the average of all pointing positions towards that target, and the variable error, which reflects the distribution of the pointing positions towards a target relative to the average pointing position to that target. Pointing position is defined as the position of the ired on the tip of the index finger at the end of the pointing movement towards the target. The distribution of the pointing positions for a target \(i\) is described by the 3-D covariance matrix \(S_i\). The three orthogonal eigenvectors of the covariance matrix \(S_i\) describe the orientations of the variable errors. The corresponding eigenvalues of the matrix give the size of the variable error along the eigenvectors. These eigenvalues of the covariance matrix \(S_i\) can be scaled to compute the limits that contain 95% of the data (see McIntyre 1998 and Morrison 1976). A \(\chi^2\) test was used to decide whether the three eigenvalues of the covariance matrix were statistically different (see Barlow 1989). In all figures we display only contours of the 95% confidence ellipses when one of the eigenvalues is significantly larger than the other two \((p<0.05)\). The eigenvector that corresponds to the largest eigenvalue will be referred to as the main axis of the distribution. We
derive the accuracy of the main axis by means of a bootstrap method (see e.g. Mooney and Duval 1993): From the 20 data points for each target, we drew a random sample of 1000 data points. For this artificial data set of 1000 points we calculate the 3-D covariance matrix, and determined the corresponding main axis. This procedure was repeated 500 times for each target, which resulted in 500 main axes per target. From the distribution of the 500 main axes we estimated the accuracy of the orientation of the ellipsoid.

**Intersection point of confidence ellipsoids**

Assuming that the ellipsoids for all targets are oriented towards one single point in space, one can find this point by estimating the intersection point of the main axes of the ellipsoids. We have determined the accuracy of the orientation of the main axis of each ellipse by a bootstrap method. Since the orientation of the main axes can be determined only up to certain accuracy, there will hardly ever be one single position in 3-D at which all main axes intersect exactly. Therefore, we have used a maximum log-likelihood method to determine the most likely position of the hypothetical intersection point. Finding the most probable intersection point in 3-D space given the pointing data set \( D \) corresponds to maximizing \( \log p(\hat{x}|D) \), where \( p(\hat{x}|D) \) represents the probability that the intersection point is at position \( \hat{x} \), given the data set \( D \). Finding the most probable intersection point for multiple data sets \( D_i \), corresponding to the targets \( i \), is equivalent to maximizing the product of the probabilities, according to:

\[
\max_{x} \log \left( \prod_{i} p(\hat{x}|D_i) \right) = \max_{\hat{x}} \sum_{i} \log(p(\hat{x}|D_i))
\]

We verified that the pointing responses could be considered to be normally distributed, using a Jarque-Bera test for goodness of fit (\( p<0.01 \), see Judge et al. 1988). For a normal distribution, \( p(\hat{x}|D) \) is proportional to:

\[
p(\hat{x}|D) \propto e^{-\frac{(\hat{x}-\mu_x)^2}{2\sigma_x^2}}
\]

and thus
\[ \log(p(x|D)) \propto -\sum_i (\bar{x} - \bar{\mu}_i) \Sigma_i^{-1} (x - \bar{\mu}_i) \]

where \( (\bar{x} - \bar{\mu}_i) \) corresponds to the distance of position \( \bar{x} \) relative to the main axis of the \( i \)-th distribution \( \bar{\mu}_i \). \( \Sigma_i \) is the covariance matrix describing the standard deviation of the data distribution for target \( i \). When the data sets \( D_i \) are normally distributed, maximizing the product of probabilities thus corresponds to minimizing:

\[ \min \sum_i (\bar{x} - \bar{\mu}_i) \Sigma_i^{-1} (x - \bar{\mu}_i) \]

The expression \( r_i = (\bar{x} - \bar{\mu}_i) \Sigma_i^{-1} (x - \bar{\mu}_i) \) is known as the Mahalanobis distance (see Duda and Hart 1973). The most probable intersection point \( \bar{x} \) given the data corresponds to the minimum of the sum of Mahalanobis distances. The basic idea behind this method is schematically displayed in 2D in Figure 2: For three targets we show hypothetical 2-D error ellipses. For each error ellipse, the probability of finding an intersection point decreases with distance relative to the main axis for that error ellipse. This probability distribution, orthogonal to the main axis, corresponds to a normal distribution. The standard deviation of this normal distribution depends on the ratio between the largest eigenvalue of the covariance matrix \( S_i \) and the smaller eigenvalue, and on the distance along the main axis relative to the center of the error ellipse. The most probable location for the intersection point corresponding to the minimum of the sum of Mahalanobis distances is indicated by a pentagon.

In order to study the relation between gaze and pointing we use the covariance between pointing position at the end of the pointing movement and gaze, which changes as a function of time, which is defined as:

\[ \rho(t) = \frac{\sum_i (g_i(t) - \bar{g}_i)(p_i(t) - \bar{p}_i)}{\sqrt{\sum_i (g_i(t) - \bar{g}_i)^2} \sqrt{\sum_i (p_i(t) - \bar{p}_i)^2}} \]
where $g_i^j(t)$ represents gaze in trial $i$ for target $j$ as a function of time. $p_i^j$ represents pointing position for trial $i$ for target $j$. $\overline{g^j(t)}$ and $\overline{p^j}$ represent the mean gaze as a function of time for target $j$ and the mean pointing position for target $j$, respectively. Note, that gaze is a function of time and pointing position $p$ is not. Therefore, any variations in the covariation between gaze and pointing in time are due to changes in gaze as a function of time.

Figure 2: Determining the most likely intersection point
Three ellipses represent the hypothetical pointing error distributions for three targets. Accuracy of the main axis direction is indicated with a cone of confidence along the main axis. The probability of finding an intersection point of the three main axes is represented in grey-scale. The star indicates the position of the most likely intersection point of the main axes of these three error ellipses.
RESULTS

In the analysis of pointing movements to remembered target positions we will mainly focus on the constant and variable errors of the pointing movements and on the relation between these errors and 3-D gaze position as a function of time after target onset. First we will focus on the pointing movements for the three visual conditions (FINGER, DARK and FRAME).

Figure 3 shows the main results for a typical subject (MV) when pointing with visual feedback of the fingertip (FINGER condition, Figures 3A and B), in the absence of visual feedback (DARK condition, Figures 3C and D), and with both vision of the environment and feedback of the fingertip (FRAME condition, Figures 3E and F). The top panels show top views on the 3-D position of the fingertip (Figures 3A, C and E), and the lower panels show side views (Figures 3B, D and F). In each panel we have also drawn a fictive subject to indicate the position of the subject relative to the targets.

As described in Methods, we distinguish between a constant error and a variable error. Figures 3A and B show that with vision of the tip of the index finger (FINGER condition) the constant errors are on average about 5 cm (range 3 to 8 cm for the different target positions). These constant errors are representative for all subjects: averaged over all subjects and all targets, the constant error is 5 cm (SD=2 cm).

Figures 3A and B show that the variable error is in general largest along an axis that is oriented towards the subject. Variable errors along the two minor axes of the ellipse are about the same in size, and are much smaller than errors along the main axis. As a result, the distribution of pointing positions has a significant orientation for most targets, and the distribution is indicated by an ellipse (see Methods).
Figure 3: Typical example of pointing responses

Pointing responses for subject MV in the FINGER (panels A and B), DARK (panels C and D) and FRAME condition (panels E and F). Top panels show a top view of the subject and the data, lower panels show a side view. Large black dots represent the target positions, small dots show the individual pointing responses towards the targets. Ellipses show the 95% confidence distribution of the pointing responses, and are drawn only when the distribution has a significant orientation. Lines emerging from the ellipses indicate the direction of the main axis.

Lack of vision of the tip of the index finger (DARK condition) leads to larger pointing errors than in the FINGER condition. Both constant and variable errors increase compared to that in the FINGER condition (compare Figures 3C and D and Figures 3A and B, respectively). The average constant error in Figures 3C and D is about 7.5 cm (range 4 to 11 cm for different target positions). The constant error is mainly in the radial-direction from the subject towards the target.
Over all subjects, the average constant error is about 9 cm (SD=4 cm), but pointing errors up to about 15 cm were observed. Averaged over all subjects, the constant error is significantly larger in the DARK condition than in the FINGER condition (t=5.79, p<0.05).

Similarly, the variable error is significantly larger in the DARK condition than in the FINGER condition (t=6.05, p<0.05). This is mainly caused by a large increase of errors in azimuth and elevation direction. Errors in distance are not significantly different from those in the FINGER condition. Due to the increased error in azimuth and elevation, the variable error is approximately the same in all directions for most target positions in the DARK condition. Fitting an ellipse to the data did not usually produce an ellipse with a clear orientation. For the data in Figures 3C and D the orientation for the variable error is only significant for the leftmost target.

Figures 3E and F show data for pointing towards remembered targets with vision of the index finger and with vision of an external frame. In this condition (FRAME) the average constant error for this subject is about 4 cm, ranging from 3 to 6 cm. Averaged over all subjects, the constant error in the FRAME condition is 4 cm (SD=2 cm). This is significantly smaller than in the other two conditions (t=7.08 and t=2.48, p<0.05, for DARK and FINGER, respectively).

The variable error is also smaller than in any of the other two conditions. This effect is significant across subjects (t=5.91 and t=3.00, p<0.05, compared to DARK and FINGER, respectively). The decrease in variable error is found especially along the axes in which the variability was already smallest (azimuth and elevation). Since the variability along the long-axis of the distribution decreases relatively little, this results in a more pronounced orientation tuning of the 95% confidence ellipses.

Thus: constant errors decrease when visual feedback of the finger is provided, and decrease even further, when additional feedback of the environment is presented by means of the illuminating frame. Providing visual feedback also results in a decrease in variable errors, mainly in azimuth and elevation direction, and only very slightly in radial distance.
Frames of reference for pointing movements

Figure 3 shows pointing responses towards seven targets for one subject in three conditions. For the DARK condition most ellipses do not deviate significantly from a spherical distribution. For the FINGER and FRAME conditions, most ellipses do have a long axis with a clear orientation, usually oriented towards the subject. Therefore, the analysis to test various hypotheses regarding the frame of reference for pointing, based on the search for a common origin of the main orientations of the variable error distributions, was limited to the FINGER and FRAME condition.

Using the Maximum Likelihood Estimation procedure (see Methods) we have determined the location of the most likely intersection point of all long axes of the error ellipses for each subject. The results of this analysis are shown in Figure 4. For each of the subjects, the point that most likely serves as the intersection point of all ellipses is indicated by a pentagon. The solid pentagon indicates the most likely intersection point for the data shown in Figure 3. We tested six subjects in the FRAME condition (Figures 4A and B), and ten in the FINGER condition (Figures 4C and D). Although the inter-subject variability in the location of the intersection point is rather large, all subjects seem to show a most likely intersection point close to or in front of the eyes for both conditions.
Figure 4: Most likely intersection points

Top views and side views (panels A, C and E, and panels B D and F, respectively) on the most likely intersection points (stars) for all subjects, for the FRAME condition (left hand panels) and FINGER condition for pointing with the right hand (central panels), and for pointing with the left hand (right hand panels). The most likely intersection point for the subject from Figure 3 is indicated in panels A-D with a filled circle. Six subjects participated in the FRAME condition and six in the FINGER condition for pointing with the left hand. Ten subjects participated in the FINGER condition for pointing with the right hand. Six of these subjects participated twice in this condition, and for these subjects the average of the two intersection points is displayed in this Figure.

Some studies in the past have suggested that the distributions of pointing errors are directed towards the shoulder (Soechting and Flanders 1989) or towards a position between the head and
the shoulder (Soechting et al. 1990). The left hand and central panels of Figure 4 show that the most likely intersection points of the subjects in our study do not lie close to the right shoulder in either condition. In order to test whether the location of the intersection point is affected by the pointing arm, we asked six subjects to point with their left arm instead of the right arm in the FINGER condition. The right-hand panels in Figure 4 show a top view and a side view of the most likely intersection points for these six subjects. Clearly, the intersection points lie close to the subjects’ head and the locations of the intersection points for pointing with the right vs. the left arm are not very different.

The relation between gaze and pointing position
In order to investigate to what extent pointing and gaze are related, binocular gaze was measured. In these experiments only a subset of the targets was tested in each of the three feedback conditions, because of the limited time available to wear the search coils. Nevertheless, the results provide evidence for a relation between pointing and gaze as a function of time, as will be illustrated below.

Table 2 shows the coordinates of the most likely intersection points relative to the cyclopean eye for all subjects, that pointed with either the left or the right arm in the FINGER condition. A statistical analysis showed that the sideward location of the intersection point was not significantly different for pointing with the left and the right arm (t=1.36, p=0.19). For both arms the intersection point seems to lie just in front of the cyclopean eye.
<table>
<thead>
<tr>
<th>subject</th>
<th>X (cm)</th>
<th>Y (cm)</th>
<th>Z (cm)</th>
<th>X (cm)</th>
<th>Y (cm)</th>
<th>Z (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FH</td>
<td>7</td>
<td>8</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JL</td>
<td>-1</td>
<td>-4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FW</td>
<td>7</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>13</td>
<td>-8</td>
</tr>
<tr>
<td>MK</td>
<td>7</td>
<td>-7</td>
<td>2</td>
<td>-1</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>NK</td>
<td>7</td>
<td>13</td>
<td>2</td>
<td>-1</td>
<td>13</td>
<td>-4</td>
</tr>
<tr>
<td>RK</td>
<td>8</td>
<td>-9</td>
<td>2</td>
<td>2</td>
<td>-1</td>
<td>-2</td>
</tr>
<tr>
<td>AT</td>
<td>5</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>-9</td>
<td>2</td>
</tr>
<tr>
<td>AT</td>
<td>5</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>-9</td>
<td>2</td>
</tr>
<tr>
<td>DL</td>
<td>-11</td>
<td>-13</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FH</td>
<td>-3</td>
<td>-8</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JL</td>
<td>15</td>
<td>10</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JL</td>
<td>27</td>
<td>-10</td>
<td>-3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MK</td>
<td>1</td>
<td>11</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MV</td>
<td>-1</td>
<td>-5</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MV</td>
<td>-3</td>
<td>-15</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NK</td>
<td>-3</td>
<td>-15</td>
<td>-8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WV</td>
<td>8</td>
<td>-20</td>
<td>-7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WV</td>
<td>6</td>
<td>-8</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**mean (sd)**

A: -4(9) -7(10) 3(7)

B: 1(3) 6(7) -2(4)

**Table 2: Position of most likely intersection points**

Coordinates of the most likely intersection point for pointing with the right arm (A) and with the left arm (B), relative to the cyclopean eye. Positive x-direction: leftward, positive y-direction: backward, positive z-direction: upward. When subjects were tested twice in the FINGER condition for pointing movements with the right arm, both data were included.
Figure 5: Gaze in time

Top views of gaze at different moments in time for subject BB: at the time of target presentation (panels A, D and G), at the end of the delay time (panels B, E and H) and at the time of pointing (panels C, F and I), for the FINGER (top row), DARK (middle row) and FRAME condition (bottom row). Pentagons indicate the target locations and small dots represent the gaze positions for all trials to that target. In the panels at the right hand side, ellipses indicate the 95% confidence levels of the distribution of the corresponding pointing positions. Lines emerging from the ellipses indicate the orientation of the distribution and are only plotted for elliptical distributions that have a significant orientation.

Figure 5 shows gaze at three different moments in time relative to target onset for the three visual conditions FINGER (Figures 5A-C), DARK (Figures 5D-F) and FRAME (Figures 5G-I), for
subject BB. The data are shown at 0.9 s after target onset (when the subject fixates the visible target, left panels); 2.9 s after target onset, which corresponds to 1.9 s after target offset (just before movement onset, middle panels); and 4.9 s after target onset (when the fingertip points at the remembered target position, right panels).

Figures 5A, D and G show that subjects fixate close to the target in almost all trials, when the target has been visible for 0.9 seconds. In the FRAME condition, subjects sometimes fixate at a point between the target and the back plane of the visible frame. At the end of the delay period, 1.9 s after target disappearance (Figures 5B, E and H), gaze has drifted away from the target in radial direction to a larger distance from the subject in all three conditions, but most clearly in the FRAME condition. During this drift period, the direction of gaze remains almost the same, i.e. any changes in azimuth and elevation are small. The amplitude of the drift at the end of the delay period depends on visual feedback of the environment. The amount of drift is approximately the same in the FINGER and DARK condition, but is considerably larger in the FRAME condition (compare Figures 5B and E and Figure 5H).

The panels at the right-hand side in Figure 5 (Figures 5C, F and I) show gaze at the time when the subject is pointing to the remembered target position. In these panels, the variability and location of the corresponding pointing positions is indicated by the error ellipses, which capture 95% of the pointing positions. Figures 5C, F and I show that gaze is much more variable at the time of pointing than at the time of fixation to the visible target (Figures 5A, D and G). Gaze and pointing seem to overlap quite well for the FINGER condition, but less so for the DARK and FRAME conditions (Figure 5F and I), where gaze locations at the time of pointing are farther away from the subject than the corresponding pointing positions.
Constant errors are displayed for the average pointing position (solid, horizontal lines) and gaze in time for three targets: target 5 (bold lines), target 1 (medium lines) and target 4 (thin lines). The constant errors are displayed for radial distance ($R$, left column), azimuth ($\phi$, middle column) and elevation ($\theta$, right column) separately, for the FINGER (top row), DARK (middle row) and FRAME conditions (bottom row). The dashed line corresponds to a perfect reproduction of the target position. Vertical grey bars represent the interval of target presentation ($0.0s < t < 1.0s$) and the interval of pointing ($4.3s < t < 5.4s$). The vertical line at $t=3s$ indicates the auditory tone that indicated the end of the delay period.

In Figure 5 we showed that gaze does not remain fixated to the target position throughout the trial, but changes during the delay period. In order to test whether errors in gaze (at some period
in time) and pointing position are related, we have analyzed the constant and variable errors in gaze as a function of time in relation to the constant and variable errors in pointing.

Figure 6 shows the constant errors in gaze averaged over all subjects, for each of the three targets. Similar to the constant error in pointing position (see Methods), we define the constant error in gaze as the deviation of mean gaze position from the target position. Since the average gaze position changes during a trial, the constant error in gaze also changes in time. The constant error in pointing position does not change during one trial, since -by definition- the pointing position is the mean position of the tip of the index finger at the end of each pointing movement. The constant pointing error is indicated by a horizontal line in order to simplify a comparison with the constant error in gaze.

Figure 6 shows that the constant gaze errors in elevation (right-hand column) show a more or less constant offset, slightly above (target 1) or below (targets 4 and 5) the target. These small deviations of about 2 deg or less may represent incorrect fixation to the target by the subjects, but they could also be due to small errors in calibration of 3-D gaze position (see Methods). The most interesting effects are found for radial distance relative to the cyclopean eye (R) and azimuth angle (φ), which are displayed in the left and middle columns, respectively.

In all conditions, the average radial distance of gaze at the end of target presentation lies within 2 cm from the target, except for the most distant target (target 1), for which gaze falls short by about 5 cm. After target disappearance gaze distance increases for all three targets and in all conditions, compatible with the drift in gaze described earlier in Figure 5. The amount of drift away from the subject and the duration of this gaze drift are different in the three conditions.

In complete darkness (FINGER and DARK conditions, Figures 6A and D) gaze distance at the end of target presentation matches the target's radial distance well. During the 2-second delay period, gaze slightly drifts away from the subject by about 3 to 10 cm. When the finger is visible during pointing (FINGER condition), gaze returns back to the radial distance of the initial gaze
position when the pointing movement starts, such that gaze and pointing position match quite closely when the subject points at the target (compare the traces for gaze error and the corresponding horizontal line for pointing error in Figure 6A).

In the DARK condition when the finger is not visible during pointing (Figure 6D), gaze distance relative to the cyclopean eye does not return to the target, nor to the pointing distance (flat lines in Figure 6D). The average pointing error in radial distance in the DARK condition is much smaller than the average error in gaze at the time of pointing (on average about 2 cm versus 7 cm, respectively).

Figure 6G shows the constant errors in radial distance for the FRAME condition: Gaze distance relative to the cyclopean eye corresponds well to the target distance at the end of target presentation. When the target disappears, gaze rapidly drifts away from the subject in radial direction. Figure 6G shows that gaze moves back towards the target position after the cue to start the pointing movement. However, at the time of pointing the decrease in gaze distance does not completely compensate for the drift in the delay period (as it did for the FINGER condition), and thus gaze does not correspond to the target distance, nor to the pointing distance, which is close to that of the target.

For constant errors in azimuth (middle columns), there is a clear distinction between targets 1 and 4 and target 5. For targets 1 and 4, which lie at eye level almost straight in front of the subject (see Figure 1), gaze errors in azimuth direction are small (less than 2 degrees from the target) in all three conditions (Figures 6B, E and H). Gaze azimuth remains almost constant from target offset until the end of pointing. For target 5, however, which lies about 30 degrees eccentric to the left at eye level (see Figure 1) there are clear differences between the three feedback conditions: At the end of target presentation, gaze azimuth corresponds within 2 degrees from the target position in all three conditions. After target disappearance, however, gaze drifts by about 3 degrees to the right of target 5 (decrease in azimuth, corresponding to a more straight ahead direction) in the FINGER and DARK condition (Figures 6B and E). Similar to the effects
described above for gaze distance, drifts in azimuth are compensated in the FINGER condition at the time of pointing, such that gaze direction returns to the target direction. In the DARK condition, gaze direction remains to the right of the target. In the FRAME condition, the visual feedback of the environment seems to prevent large drifts in azimuth for all targets (see Figure 6H).

In summary: as long as the target is visible, gaze is directed towards the target. After target disappearance gaze tends to drift away from the subject in radial direction. In the DARK and FRAME condition gaze remains too far from the cyclopean eye, relative to the target, whereas in the FINGER condition gaze almost completely returns at the time of pointing. In the FINGER condition, pointing errors correspond closely to gaze errors, which is not surprising since the finger is visible during pointing. In the DARK, pointing errors are smaller than gaze errors, mainly because gaze errors are large due to drift from the target position. In the FRAME condition pointing errors are small, but gaze is directed to a position between the pointing position and the visual background.

In order to study the effect of variable errors in gaze on the variability of the pointing positions or visa versa, we compared gaze in time with the corresponding pointing position. There is only one pointing position per trial, but gaze may vary in time. Therefore we tested whether there is a moment in time when the variability in gaze is closest related to the variability in pointing position.

Figure 7 shows the covariance between the pointing position and gaze position as a function of time, averaged over all six subjects. The average covariance for radial distance (R), azimuth (φ) and elevation (θ) is shown in the left, middle and right panels, respectively. We tested whether the time when the highest covariance was reached, was related to a specific stage in the delayed pointing task. Therefore, we focus on two time intervals: the interval of target presentation (from 0 to 1 second) and the interval during which the finger points to the target (on average from 4.3 to 5.4 seconds after target onset). These intervals are indicated by gray bars in Figure 7.
Figure 7: Covariance of pointing position and gaze in time

Covariance of pointing position and gaze in time -averaged over all subjects- is represented by a bold line. The thin lines indicate the standard deviation around the mean. Correlations are calculated for radial distance (R, left columns), azimuth (θ, middle columns) and elevation (θ, right columns), for the FINGER (top row), DARK (middle row) and FRAME condition (bottom row). Thin lines represent the standard deviation (1 SD). Vertical grey bars represent the interval of target presentation (0s < t < 1s) and the interval of pointing (4.3s < t < 5.4s). The vertical line at t=3s indicates the auditory tone that indicated the end of the delay period. The numbers in the grey bars indicate the number of subjects that show a significant covariance (p<0.05) within the interval.
All three feedback conditions show a similar increase in covariance from target offset towards the time of pointing (see the bold lines in Figure 7, indicating the values of the covariance averaged over all subjects). This increase towards the time of pointing was found for all subjects. The peak value of the covariance was reached at slightly different times in the interval between 4.3 and 5.4 seconds after target onset for different subjects. Therefore, the peak value of the average covariance shown in Figure 7 is about 25% smaller than the average of the peak values of all subjects. In order to overcome this problem of inter-subject timing differences, we counted the number of subjects that show a significant covariance (p<0.05) somewhere within the interval of target presentation (0-1 s), and the number of subjects having a significant covariance within the pointing interval (4.3 to 5.4 s). These numbers are displayed in the gray bars for each of the specific time intervals. Note that the number of subjects showing a significant covariance in the pointing interval is always at least equal to, but in general larger, than the number of subjects with a significant covariance during target presentation. We found that for almost all subjects, the largest covariance is reached in the pointing interval.

In the FINGER condition (Figures 7A-C), the average covariance exceeds a 5% significance level for all three coordinates (radial distance, azimuth and elevation) at the time of pointing. Moreover, Figures 7A-C show that the covariance increases gradually towards the time of pointing, indicating that the variability in pointing resembles the variability in gaze at a time, well before the time of pointing. This means that the variability of the pointing position in radial, azimuth and elevation direction can be explained (at least partly) from the variability in gaze at the time of pointing. When the subjects are considered individually, all subjects show a significant covariance in all three coordinates in the FINGER condition at the time of pointing (p<0.05), except for one subject, who shows a significant covariance for the two directional components (azimuth and elevation, p< 0.05), but not for radial distance (p=0.18).

For the DARK and FRAME conditions (Figures 7D-F and 7G-I, respectively), the covariance between pointing position and gaze is less pronounced: at the time of pointing the
average covariance is significant for azimuth and elevation (p<0.05). The average covariance for radial distance, however, increases towards the time of pointing, but does not reach a significant value (p=0.06 and p=0.11 for DARK and FRAME, respectively).

As indicated by the numbers in the gray bars in Figures 7D-F (DARK condition), all six subjects show a significant covariance at the time of pointing for the azimuth direction, but only four subjects also have a significant covariance for radial distance and elevation at the time of pointing. The high correlation between gaze and pointing in the FINGER condition is not surprising because of visual feedback. However, all subjects also show a significant covariance between pointing and gaze at the time of pointing in the DARK condition in at least two of the three spatial parameters. Thus, covariance at the time of pointing is also present without visual feedback of the finger during pointing.

For the FRAME condition (Figures 7G-I) three subjects show a significant covariance for all three coordinates at the time of pointing (p<0.05). The remaining three subjects show a significant covariance in azimuth (p<0.05), but only one of them also shows a significant covariance in radial distance (p<0.05).

In all conditions, we found subjects that have a significant covariance between pointing and gaze at the time of target presentation (see the numbers in the gray bars that indicate the time of target presentation). In most of these cases, the covariance was smaller at the time of target presentation than it was at the time of pointing.

In summary, Figure 7 shows that the variability in pointing position is related to the variability in gaze at the time of pointing and often already at the time of target presentation. This relation is most prominent for the FINGER condition, but is also present in the DARK and FRAME condition.
DISCUSSION

In this study we have investigated the accuracy of gaze and pointing movements towards remembered visual targets in 3-D space. The results demonstrate that the presence or absence of visual feedback of the finger and of the environment has a large effect on the variable and constant errors of pointing. Moreover, we found that the variability in the final position for the pointing movement is to a large extent related to gaze. Like pointing position, gaze is not always directed to the target position but may differ quite considerably from the target position depending on visual feedback conditions. We will first discuss the relation between pointing and gaze and its implications for pointing accuracy. After that, we will discuss the interpretation of the condition-specific constant and variable errors in pointing.

Gaze vs. pointing

We found a significant correlation between the variable errors in pointing and in gaze at the time of pointing. This covariance between gaze and pointing could suggest three possible explanations. The first explanation could be a common command signal to drive gaze and pointing towards the same target position. Other explanations might be that gaze affects pointing or that pointing serves as a target for gaze. Obviously, these explanations do not exclude each other. We will consider the implications of each of these hypothetical explanations in the context of our experimental results to investigate which explanations are consistent with the data.

The three possible explanations are schematically illustrated in Figure 8. At the time of target presentation, the orientations of the head in space (Hs), the eyes in the head (Eh) and the target on the retina (Tr) are available to calculate the perceived target position, which is stored during the delay period (Internal Target Representation, ITR). This ITR is used as target for the pointing movement (path A), and can also be used to guide the eyes to keep gaze at the remembered target position (path B). Evidence for such a common command signal for the eyes and the hand has
been presented before by several studies, which reported that eye and hand movements show similar characteristics in tasks like choosing between two targets (Gielen et al. 1984) or anticipating target displacements (Frens and Erkelens 1991). Our data provide additional evidence in favor of a common drive of gaze and pointing because of the -for many subjects- significant correlation between pointing position and gaze, when the target is visible (see Figure 7).

![Figure 8: Schematic overview of possible pathways](image)

Possible pathways to describe the transformation of retinal information to the pointing position. Eye-in-Head and Head-in-Space determine gaze during fixation (Gaze(0)). Target position on the retina and gaze during fixation determine the Internal Representation of Target position (ITR). The ITR provides a common signal to drive the arm towards the remembered target position (path A) and to drive gaze (path B). Path C represents a hypothetical pointing signal, which is used to direct gaze, and path D represents a hypothetical gaze signal, which is used to adjust the accuracy of pointing.

If the covariance between gaze and pointing were due only to a common command signal to the motor systems for gaze and pointing, one would expect that the gradual drift in gaze in the delay period should deteriorate the covariance between pointing and gaze. This is obviously not the case, as is shown in Figure 7, which shows that the covariance increases, rather than decreases with time. A possible explanation might be, that the gradual drift in gaze reflects a drift of the ITR, which then should result in a constant error in pointing, proportional to the drift of gaze.
Figure 6 clearly shows that this effect is not found at the end of the delay period: In general, the constant error in gaze is much larger than the constant error in pointing. Therefore, a common drive cannot be the only explanation for the results reported in this study.

Since various studies have shown that deviations of gaze from the target will affect pointing accuracy (e.g. Biguer et al. 1984, Bock 1986, Enright 1995, Van Donkelaar and Staub 2000, Henriques et al. 1998) one could argue that gaze accuracy affects the accuracy of pointing. Recently, Neggers and Bekkering (2001) have demonstrated a strong linkage between eye movements and pointing movements. Subjects were instructed to make a pointing movement and a saccade towards the same target. When the saccade had reached the target (but when the corresponding pointing movement was not yet completed), a new saccade target was presented. Subjects had to initiate a second saccade toward this new target, but the pointing movement had to stay directed towards the initial target. Neggers and Bekkering showed that the second saccade, away from the pointing target, was delayed until the pointing movement was nearly completed. These results were interpreted as evidence that gaze is used to improve accuracy of the ongoing pointing movement, in addition to a common command signal for eye and arm movements. This is in line with the conclusion obtained by Soechting et al. (2001), who reported that gaze position provides the target signal for hand movements to targets moving behind a moving background (Duncker Illusion). The third alternative explanation, that gaze depends on pointing, is less likely: subjects can quite well look at remembered targets irrespective of finger position.

Evidence in favor of a role of gaze at the time of pointing on pointing accuracy was found in various studies, in which gaze and pointing position were dissociated. Bock (1986) and Enright (1995) have shown that the pointing movement tends to overshoot the target distance relative to the gaze location, when gaze is not directed towards the location of the remembered target for the reaching movement. A similar finding was reported by Henriques et al. (1998, 2000), who studied pointing errors towards a remembered target, which was presented while gaze was directed in
various horizontal and vertical peripheral positions not coinciding with the target. Subjects were found to overshoot the magnitude of the retinal eccentricity of the target, both in horizontal and vertical directions. Medendorp and Crawford (2002a) have shown that a similar overshoot occurs when subjects are allowed to foveate a visual target but have to make a saccade in the delay period before pointing. In this paradigm, the pointing responses still indicated an overestimation of the retinal eccentricity of the target relative to gaze at the time of pointing. Therefore, the authors concluded that the target position is updated for gaze at the time of pointing, which was interpreted as evidence that target position is stored in retinocentric coordinates. A similar question was addressed by Pouget et al. (2002) who investigated whether the remembered position of a reach target is also stored in a retinocentric frame of reference for targets of other modalities (i.e. auditory, proprioceptive and imaginary targets). They found that when gaze was not directed to the target position, subjects largely overshot the retinal eccentricity of the target, irrespective of target modality.

If fixation away from the target leads to overshoot of the retinal position of the target, one would expect a negative correlation between variability in gaze and pointing in the present study, where gaze drifts away from the target in the delay period. This is not in agreement with our data, which clearly shows a positive correlation between pointing and gaze. Positive correlations have also been reported in other studies (see e.g. Flanders et al. 1999 and Soechting et al. 2001). One of the main differences between these studies and the ones that report overshoot of the target and a corresponding negative correlation is that in the latter studies, subjects deliberately fixated away from the target, whereas in our study (like in the study by Flanders et al. 1999 and Soechting et al. 2001), subjects were not aware of the off-target fixation. None of our subjects was consciously aware of the drift in the delay period. One subject explicitly mentioned that he tried to use the strategy to rigidly fixate the remembered target in the DARK and FINGER conditions, in order to "anchor" the visual target in the otherwise dark environment. The results of this subject were not different from those of the other subjects, which indicates that the drift in gaze occurs
unconsciously, even when a strategy is adopted to maintain gaze at the remembered target. When the remembered target position is stored relative to gaze, as suggested by e.g. Henriques et al. (1998), and when subjects are not consciously aware of the drift in gaze, the remembered target position is presumably not updated for the drifted gaze position at the time of pointing. Consequently, one will find a positive correlation between gaze and pointing, which is what we found.

**Gaze drifts during delay period**

From previous studies it is known that when subjects are left in complete darkness for a few minutes, gaze tends to shift towards a preferred distance, which varies between subjects from about 40 to 80 cm relative to the cyclopean eye, and slightly changes with gaze direction (*dark vergence*, Heuer et al. 1989). Gaze shifts in complete darkness, which occur already a few seconds after the disappearance of a visual target, were also described by Medendorp et al. (2002b). These authors tested the stability of gaze to a remembered target during active head movements. They found that the correspondence between target position and gaze gradually deteriorates after the visual target disappears, both for direction and for radial distance. Medendorp et al. presented a target at 20 cm from the cyclopean eye, and found that gaze distance starts to increase almost immediately after disappearance of the visual target, which is in agreement with our observations. We conclude that subjects do try to maintain gaze on the target, but fail to do so.

In the FRAME condition we found larger gaze drifts in radial distance than in the DARK and FINGER condition, and hardly any drift in direction (see middle columns in Figure 5 and Figure 6). For the FRAME conditions one might expect that it would be easier to maintain fixation at the remembered target position: the visual environment provides a reference that might assist the subject to correct for any unintended gaze drift. On the other hand, one might argue that since the visual environment provides a reference frame to store the remembered target position, precise
fixation may not be that important, since the visual frame serves as an “anchoring point” for the remembered target. Our results provide evidence for the latter, since at the time of pointing, gaze has not returned completely to the target position, and nevertheless, the pointing performance is more accurate than without the frame. More evidence for the latter interpretation was provided by Blouin et al. (2002), who showed that the definition of gaze direction after several saccades in the dark is more accurate when there is visual stimulation of the retina, than when there is no visual information whatsoever. This effect is irrespective of whether the visual stimulation carries spatial information or not. In the FINGER condition we found that vision of the index finger resulted in a correction of the radial drift in gaze. Following the reasoning of Blouin et al., this corrective movement of gaze may be the response to a more accurate definition of gaze direction due to visual stimulation of the retina.

Frames of reference

Many authors have studied pointing movements towards remembered targets. However, the conclusions of these studies have not always been congruent. Some studies concluded that subjects make pointing movements to remembered visual targets in an illuminated environment in a viewer-centered frame of reference, usually with respect to the head or the cyclopean eye (e.g. McIntyre et al. 1997 and Soechting et al. 1990). Other studies suggested that pointing movements are executed in a shoulder-centered frame of reference (e.g. Soechting and Flanders 1989), or both a shoulder-centered and a head-centered frame of reference, for pointing movements in the dark (Soechting et al. 1990 and McIntyre et al. 1998). These studies all describe pointing movements to remembered visual targets, but they tested subjects under different visual conditions. The question arises, whether these different conditions can explain the different observations.

Other work on the effect of vision and delay after target offset on the accuracy of pointing movements on a 2-D table was performed by Carlton 1981 and Elliott and Madalena 1987. Carlton showed that vision of the pointer (a hand held stylus) was the most important requirement
for accurate pointing, irrespective of visibility of the environment or the target. When the hand is not visible, continuous vision of the target improves the pointing performance of the hand, indicating that the CNS is able to correct pointing movements of an unseen hand during the execution (Prablanc et al. 1986).

Elliott et al. have shown that subjects are able to accurately use a visual representation of the target in the control of aiming movements only shortly after visual occlusion. For pointing movements after delays of $2\,s$ or longer, the pointing errors increase significantly, due to the decay of such visual representation.

Soechting et al. (1990) tested pointing movements towards remembered targets in the Dark. Unlike in our study, the room lights were on when the target was presented in the dark condition of their study. They tested the average orientation of the total error (constant and variable error taken together) between pointing position and target position, under the assumption that errors are largest in radial distance, and much smaller in direction. For pointing in the dark, they found that the total errors are best described relative to an origin located between the head and the shoulder. They also analyzed pointing movements when the room lights remained on after target presentation. In this Light condition they found that total pointing errors were best described relative to a position close to the subject’s eyes. From these results they concluded that there exists both a head-centered and a shoulder-centered representation of target location in the CNS, and that for pointing in the dark, the CNS includes the shoulder-centered representation.

The FRAME condition in the present study is similar to the Light condition in the study by Soechting and colleagues. In this condition we find that orientations of the variable errors are best described as originating from a center close to the subjects’ head. However, in the DARK setup, we did not find that the variable errors originate from a distinct origin at or near the head or the shoulder or in between. Differences between the results of Soechting and Flanders and the findings of our study may well result from differences in the approach to estimate the orientation of the pointing errors. The analysis by Soechting and Flanders was based on the total error.
whereas we analyzed the variable and constant errors separately. The two approaches will lead to the same conclusions, when the constant errors are mainly in radial distance. In our study however, we found moderate constant errors also in direction, especially in the DARK condition (2 degrees in the FINGER and FRAME condition and 4 degrees in the DARK condition), and these may result in different origins for the variable errors and for the total errors.

McIntyre et al. (1998) tested movements towards remembered targets performed in the dark and in a dimly lit room, apparently similar to our DARK and FRAME conditions. They found that -with vision of the environment and the arm- the distributions of variable pointing errors are oriented towards the subject’s head. McIntyre and colleagues also tested the local distortion, which described the fidelity with which the relative spatial organization of targets within a small workspace region (on a sphere of 22 mm radius) is maintained in the configuration of final pointing positions. They found that the pointing positions reflected a contraction of the target configuration along an axis that was oriented towards a position between the subject's head and the shoulder. They interpreted the orientation of the variable errors and that of the local contraction as indicative for the use of both a viewer-centered and a shoulder-centered frame of reference. However, there may be alternative explanations.

Carrozzo et al. (2002) have shown that the configuration of the targets influences the orientation of variable pointing errors, even though the targets were never presented simultaneously. Gentilucci et al. (1996) showed that the manifestation of such allocentric effects is strengthened by a delay between the visual stimulus and the motor response. In our study we used a delay period of 2 s, which suffices for allocentric effects to occur. Nevertheless, we found no effects on the variable errors related to our target configuration. This may be due to the fact that our target configuration was more complex than the configuration used by Carrozzo and colleagues, in which the targets were located on a straight line. In the setup in which the subjects were provided with the most allocentric information (FRAME), one might expect finding variable errors to be related to the illuminating frame. However, we found that the orientation of the pointing variability in the FRAME condition was largely related to gaze, instead of to the
illuminated frame. A significant correlation between the variability in pointing and gaze was also found in the DARK and FINGER conditions. This may seem in contradiction to earlier findings of Prablanc et al. 1979, who tested gaze and pointing movements towards targets in 2D, in a similar setup as our DARK condition. In their study, the target disappeared immediately after onset of the goal directed saccade. Therefore, the target was not foveated and moreover, the delay between target offset and the start of the pointing movement was much shorter than in our study. Probably, the use of gaze for the definition of the pointing target, as is indicated in Figure 8, takes place only when the target is actually foveated or when the target position has to be remembered for some time, as was the case in our study.

Another explanation for the distribution of variable errors was presented recently by Van Beers et al. (2002), who proposed an optimal integration model for the perception of position, which integrates information from different sensory modalities, weighted by their accuracy for each direction. They suggested that vision is more accurate for target direction than proprioception, but less so for radial distance. Therefore, errors in azimuth and elevation are thought to be due to errors in vision and errors in radial distance supposedly result from errors in proprioception, which are larger than directional errors in vision. This might explain why the pointing errors in the FINGER and FRAME condition are smaller in azimuth and elevation than in radial distance.

According to the ideas by Van Beers et al. (2002) the CNS will use visual information for direction and proprioceptive information for distance, when the finger is visible (FINGER and FRAME condition). When the finger is not visible (DARK condition), proprioceptive information will be used both for distance and direction. As a result, the variable errors in direction will be larger in the dark, since the CNS has to rely on proprioceptive information, which is less accurate than the visual information. This corresponds to almost spherical distributions of variable errors, just as we found for the DARK condition in the present study, and as reported previously by Desmurget et al. 1998. We found that providing visual feedback of the finger position (FINGER)
decreases variable errors in direction. The large directional variability in the DARK condition resulted in almost spherical distributions, thus making it impossible to decide on the orientation of the variable errors. Moreover, when the accuracy of the visual information is increased by providing visual information about the environment (FRAME condition), we find that variable errors in direction decrease accordingly, which is also in agreement with the model proposed by Van Beers et al. (2002).

We have shown that the variability in pointing positions correlates highly to gaze at the time of pointing. This can be explained by a combination of (1) a common drive to the motor systems for gaze and pointing, and (2) by an effect of gaze on pointing accuracy at the time of pointing. The eye-centered orientation of the distribution of pointing positions found in previous studies, may therefore reflect the effect of the variability in gaze (which is less stable in distance than in direction), in addition to possible internal reference frames used in processing and storage of the remembered target position.
ACKNOWLEDGEMENTS

We would like to thank Ton van Dreumel and Hans Kleijnen for their assistance in building the experimental setup, and we are grateful to Ger van Lingen for designing the software to drive the experiment. We thank Ingrid van der Ven and Wietske Kievit for their contributions in preliminary experiments and Lex Cools for useful discussions about the interpretation of the data.
REFERENCES


