Visuomotor sensitivity to visual information about surface orientation

David C. Knill[1] and Daniel Kersten[2]

[1] Center for Visual Science, University of Rochester

[2] Dept. of Psychology. University of Minnesota

Running Head: Visuomotor sensitivity to surface orientation

Correspondence to:

David Knill

Center for Visual Science

274 Meliora Hall

University of Rochester

Rochester, NY 14627

e-mail: knill@cvs.rochester.edu

To appear in *Journal of Neurophysiology*

FINAL ACCEPTED VERSION

Abstract

We measured human visuomotor sensitivity to visual information about three-dimensional surface orientation by analyzing movements made to place an object on a slanted surface. We applied linear discriminant analysis to the kinematics of subjects' movements to surfaces with differing slants (angle away form the fronto-parallel) to derive visuomotor d's for discriminating surfaces differing in slant by 5 degrees. Subjects' visuomotor sensitivity to information about surface orientation was very high, with discrimination "thresholds" ranging from 2 to 3 degrees. In a first experiment, we found that subjects performed only slightly better using binocular cues alone than monocular texture cues and that they showed only weak evidence for combining the cues when both were available, suggesting that monocular cues can be just as effective in guiding motor behavior in depth as binocular cues. In a second experiment, we measured subjects' perceptual discrimination and visuomotor thresholds in equivalent stimulus conditions in order to decompose visuomotor sensitivity into perceptual and motor components. Subjects' visuomotor thresholds were found to be slightly greater than their perceptual thresholds for a range of memory delays, from 1 to 3 seconds. The data were consistent with a model in which perceptual noise increases with increasing delay between stimulus presentation and movement initiation, but motor noise remains constant. This result suggests that visuomotor and perceptual systems rely on the same visual estimates of surface slant for memory delays ranging from 1 to 3 seconds.

1 Introduction

A principle function of three dimensional vision is to provide animals with the information needed to guide goal directed motor behaviors; for example, picking up, manipulating and placing objects. Despite this, most vision research has relied on experimental paradigms based on explicit perceptual report to study the processes underlying three-dimensional perception. A number of recent studies, however, suggest some degree of functional dissociation between the perceptual estimation of scene attributes and the use of visual information about those attributes to guide motor behavior [Goodale and Milner 1992, Milner and Goodale 1995, Bridgeman 1999, Haffenden and Goodale 1998]. While several authors have called into question the interpretation of some of these results [Franz et. al. 2001, Bruno 2001], the full body of research raise questions about whether one can unquestioningly use results from perceptual studies to make inferences about visuomotor transformations. In the current paper, we use a natural, visually guided object placement task to measure humans' abilities to transform visual information about a surface's orientation in three-dimensional space to motor behavior. We further compare subjects' performance on matched perceptual and motor tasks to parcel their visuomotor error into independent perceptual and motor components.

In order to address the main goals of the experiment, we applied discriminant analysis to the analysis of movement kinematics to accurately measure the visuomotor system's sensitivity to the visual information used to control a simple goal-directed hand movement in three-dimensional space. Such methods have been previously applied to study the time evolution of hand grip formation for grasping complex objects [Santello and Soechting 1998]. We applied the analysis to motion trajectories in an object placement task to derive 'd-prime' measures (d's) of subjects' visuomotor sensitivity to visual information about the three-dimensional orientation of flat surfaces. This

allowed us to address two main questions. First, we measured visuomotor sensitivity to different sources of information about a surface's orientation in three-dimensional space. We measured sensitivity to binocular and texture information in isolation and looked for evidence that visuomotor sensitivity improves significantly when texture information is added to binocular information about orientation in depth. Second, by comparing perceptual and motor performance over a range of memory delays, we tested the hypothesis that visuomotor and perceptual performance derive from a common visual representation of surface slant. In particular, we tested the hypothesis that errors in visuomotor performance could be decomposed into independent perceptual and motor noise sources.

1.1 Background

The visual scene contains many cues to the three dimensional layout of objects within it. Of these cues, binocular information (provided by retinal disparities and vergence angle) is often considered dominant within a person's immediate workspace - the space relevant to normal object manipulation movements. Consistent with this assumption, a number of researchers have found that reaching performance degrades in monocular viewing conditions - movement times increase, the proportion of time spent in the deceleration phase increases, the number of re-accelerations in hand movements increases and the number of secondary re-openings of finger grip increase [Servos 2000, Moll and Kuypers, 1980, Kruyer et. al. 1996, Watt and Bradshaw 2003]. Marotta, et. al. have also argued for a special role of binocular information based on a functional dissociation between monocular and binocular performance in an apperceptive agnosic (patient DF) [Marotta et. al. 1997]. In some of these studies, however, monocular cues were kept sparse (for example an illuminated ball in the dark). In such conditions, it is not surprising to find degradations in motor performance in monocular conditions. Perhaps more importantly, the tasks studied required reaching to tar-

gets at different depths, for which accurate information about absolute egocentric depth is critical. Monocular cues generally provide poor information about absolute depth, so that it makes sense for the visuomotor system (or, for that matter, the perceptual system) to rely more heavily on binocular information to estimate absolute depth.[1].

In order to place binocular and monocular cues on more equal footing, we studied a task whose performance requires visual information about planar surface orientation, a geometric property determined by relative depth rather than absolute depth. Cue integration studies have shown that at low slants (angles away from the fronto-parallel), stereo information dominates perceptual judgments, while at larger slants monocular cues like texture can dominate [Knill and Saunders 2003]. In the two experiments reported here, we studied a simple object placement task in which subjects were required to place a cylindrical object flush onto a flat surface oriented at different slants away from the viewer. We fixed the axis of rotation (tilt) of the surface in space to be horizontal, and the target location for cylinder placement to be at the center of that axis. This effectively removed uncertainty about the location at which subjects had to place the cylinder. The visual information relevant to accurately performing the task, therefore, was principally that specifying the orientation of the surface.

The first experiment measured subjects' visuomotor sensitivity to binocular and texture cues to three-dimensional surface orientation. By comparing sensitivity measures in single cue stimulus conditions with those found when both cues were available to subjects, we looked for evidence that subjects' performance improves when monocular cues like texture are added to binocular cues. Experiment 2 was designed to separate the contributions of perceptual and motor "noise" to variability in subjects' visuomotor performance. We did this by comparing subjects' perceptual

[1]Blur and accommodation provide information about absolute depth, but this has been found to be relatively weak [Mon-Williams and Tresilian 2000]

discrimination thresholds and visuomotor discrimination "thresholds" for the same stimuli at a range of memory delays. This allowed us to test the additivity prediction of the hypothesis that the two sources of noise are independent. It also allowed us to more accurately test the cue additivity hypothesis, since the hypothesis makes predictions specifically about perceptual uncertainty.

2 Experiment 1: visuomotor sensitivity

Place figure 1 here.

Figure 1 illustrates the task that subjects performed in the experiments. Subjects placed a cylindrical object flush onto a flat surface positioned by a robot arm at different orientations in space. Subjects viewed the surface through field-limiting tubes placed in front of each eye, so that the information about surface slant was limited to that available in a region within the bounds of the surface - stereo information provided by retinal disparities and vergence angle, and texture information provided by the texture projected from the surface. We used linear discriminant analysis to measure the discriminability of the movements generated for surfaces with slightly different slants (5 degrees in our experiment). In effect, the analysis measured how reliably one can estimate the orientation of the target surface using the motion of the cylinder that a subject placed on the surface.

The first experiment measured the sensitivity of the visuomotor system to binocular and texture cues to three-dimensional surface orientation. We created stimuli with, respectively, stereo information alone (binocular views of the white noise texture)[2], texture information alone (monocular views of the regular texture pattern), both cues combined (binocular views of the regular texture

[2]Technically, the white noise textures contained texture gradients; however, because these gradients were contained in the very high frequency components of the patterns, we expected them to be perceptually unreliable. Subjects performance with monocular views of the white noise textures bore this prediction out (see figure 7).

pattern) or no reliable visual information (monocular views of the white noise texture) (see figure 3). The last condition served as a control for the efficacy of information not independently controlled in the experiment, such as the brightness of the target surface and auditory cues provided by the movement of the robot arm.

2.1 Methods

2.1.1 Subjects

Three undergraduates at the University of Minnesota served as subjects in the experiment. All three subjects had corrected to normal vision and were paid for their participation. The subjects were naive to the purpose of the experiment.

2.1.2 Apparatus and stimuli

Figure 1 illustrates the apparatus used in the experiment. Figure 2 shows the physical dimensions of the experimental setup. A robot arm (PUMA 260) positioned a flat surface 45 cm from the subject's eyes (30.5 cm in front of the subject and 32.5 cm below the level of their eyes). The arm was used to rotate the surface by different angles around a horizontal axis through a fixed point in space. The angle of the rotation axis in the horizontal plane was chosen to be horizontal from the viewpoint of the observer. Textures were printed on 8-1/2 x 11" paper and slipped into slots on the edges of the target surface. Different textures could be displayed by changing the textured paper placed on the target surface. The target position for cylinder placement was generated using a laser pointer illuminating a spot at the center of the target surface. The horizontal starting platform for the cylinder was positioned to the right of and above the test surface. In a coordinate frame centered on the target location for placement, with the y-axis taken parallel to gravity (positive up), the x-axis parallel to the rotation axis of the surface (positive to the right of the subject)

and the z-axis parallel to the cross-product between those two (positive towards the subject), the starting position for the cylinder was at ($35\,cm$, $4\,cm$, $0\,cm$). The starting platform had lips on two sides into which subjects could slot the cylinder at the end of each trial, guaranteeing that the starting position was the same on each trial.

Place figure 2 here.

The cylinder was hollow, 1 cm thick, with an outer diameter of 5.65 cm and a height of 12.6 cm. It weighed 134 grams. Three infrared emitting diodes were positioned on a circular disk mounted on the side of the cylinder facing toward the Optotrak system. The diodes formed a triangle on the disk. In relation to the subject, the disk was on the left side of the cylinder, while subjects grasped the cylinder from the right. An Optotrak 3020 system was used to measure the positions of the markers at a sampling rate of 100 Hz. Prior calibration of the system revealed that it measured marker positions with an error less than .1 mm. The triad of marker positions was used to estimate the 3-D orientation of the cylinder in space as a function of time. Two pieces of rough tape were placed on either side, half-way up the cylinder for subjects to position their forefingers and thumbs. Subjects used five finger precision grips to grasp the cylinder.

Place figure 3 here.

Two different types of texture patterns could be mounted on the target surface, a white noise texture and a texture composed of a regular array of dots (see figure 3). Subjects viewed these either monocularly or binocularly, making four stimulus conditions for the experiment. Seven different surface orientations (target slants) were tested, ranging from 70° to 100°, where 90° was level (perpendicular to gravity) and 47° would have been fronto-parallel to the subject.

Subjects, with heads fixed in a chin rest, viewed the target surfaces through translucent cylinders mounted in front of each eye in such a way that each eye's view formed a circle centered on the target location for cylinder placement. Subjects' field of view on the surface subtended a visual angle of 13.2° , which, in all conditions, was within the bounds of the target surface. The left eye was occluded in the monocular conditions.

Subjects wore headphones through which auditory signals were given to begin a movement, to return to the starting position and to close eyes between trials.

2.1.3 Procedure

Subjects ran in eight sessions on separate days. Each session was further sub-divided into four blocks, one for each stimulus condition (monocular vs. binocular viewing crossed with regular vs. noise textures). Stimulus blocks were randomized across sessions. Data from the first session were discarded as practice. Each block consisted of ten trials per test slant, making 70 trials per block. After every 10 trials, there was a brief break of about 20 seconds during which the experimenter changed the texture pattern mounted on the target surface. Within a block the different textures represented different samples of the same type of texture (different white noise patterns, large or small dot patterns). Subjects were given a break of several minutes after each block of trials. Subjects finished each session in approximately 40 minutes. The order of blocks was randomized between sessions and counter-balanced so that the different cue conditions appeared in each temporal position within a session twice.

Each trial began with the subject holding the cylinder stationary in the starting position. A trial was initiated by a "go" signal given over the headphones. One second later, a "stop" signal was given over the headphones. Subjects were instructed to place the cylinder flush onto the surface

before hearing the stop signal. Movements lasted for between 300 and 600 msec., reflecting the fact that the one-second window given to complete the movement was well within the natural limits imposed by the task. One second after the "stop" signal, a "return" signal was given instructing the subject to return the cylinder to the start position. After the cylinder was stably placed at the start position, a "close" signal was given instructing subjects to close their eyes, during which interval the robot arm rotated the surface to a new test orientation. This was followed by an "open" signal, a one second delay and the beginning of a new trial with the "go" signal.

We recorded 2 seconds (200 frames) of data from the Optotrak on each trial, beginning at the time of the "go" signal.

2.2 Results

2.2.1 Parsing the trajectories

Detecting movement start and stop times

For purposes of analyzing visuomotor performance, we defined the trajectory of the cylinder to extend from the time a movement started to the time of initial contact with the surface. We defined the start time to be the time at which the cylinder's orientation first deviated from the starting orientation (vertical) by more than .5°. We used the cylinder's acceleration profile to determine the initial contact time. Figure 4 shows an example acceleration profile for the point at the center of mass of the three markers on the cylinder (calculated using discrete differences). As clearly evident in figure 4, contact with the target surface was marked by a sharp negative peak in acceleration. Often, a second peak was also evident, reflecting final contact with the surface following a secondary rotation to bring the cylinder flush with the surface. In order to mark the initial contact time, we first found the two local minima in acceleration with the largest negative values in the movement.

When these occurred in brief succession (within 50 msec. of each other), the first was selected as the contact time; otherwise the minimum with the largest negative value was selected.

Place figure 4 here.

In order to check the automatic parsing method, we used the kinematic data and 3D graphics (OpenGL) to simulate a virtual moving cylinder, marking the measured contact time with a change in cylinder color. Contact with the target surface was visually clear from the motion of the cylinder. Visual inspection of a large number of trials showed that the automatic method captured the contact slant, with occasional underestimates of contact time of one frame. Never did the method appear to over-estimate the time of contact.

Time normalization

Subjects' movements varied in duration from trial to trial with a standard deviation of approximately 10% of the mean. We used a cubic spline interpolator to normalize the trajectories for analysis. The marker positions on the side of the cylinder were used to calculate the orientation (slant and tilt) and position (at the center) of the cylinder at each sample point in time. For each trial, the sampled orientations and positions of the cylinder were interpolated to give 100 uniformly spaced measures between the detected start and contact times. The result was a set of 100-dimensional vectors characterizing the slant, tilt and three-dimensional positions of the cylinder as it moved from the starting surface to the target surface.

Outlier rejection

We defined outlier trials to be thrown out of the analysis as those which matched one of several criteria. We discarded trials with durations less than 250 msec., durations greater than 1 sec. or a cylinder orientation at the time of the "go" signal greater than .5° away from the vertical.

Histograms of final contact slants showed that occassionally, though rarely, subjects were very far off from the mean settings. These contact slants were far enough away from the mean that they showed up as isolated points well away from the mass of the histogram. To account for these outliers, we removed trials on which the final contact slant was more than three standard deviations away from the mean of the contact slants found for a given stimulus and target slant condition. Based on the criteria listed here, we rejected on average 10 trials out of 70 for each stimulus/slant condition. On average, less than one of those trials was rejected on the basis of being more than three standard deviations away from the mean.

2.2.2 Kinematic analysis

In order to assess the rotation kinematics of subjects' movements, we decomposed the orientation of the cylinder into three components: its slant (angle in the z-y plane as defined in Methods), its tilt (angle in the x-y plane defined in Methods) and its spin (angle around its central axis). Note that we have defined slant in a gravitational frame of reference so that 90° is upright relative to gravity. To obtain the slant of the target surface in the subjects' frame of reference when the cylinder was flush on the surface, one would subtract 47 degrees from the measured slant of the cylinder. Tilt was the same in both frames of reference. We will focus our analysis on the rotational motion of the vertical axis of the cylinder (its slant and tilt). The spin of the cylinder was irrelevant to the task. Analyses showed that the spin kinematics did not vary significantly between stimulus conditions.

Figure 5 shows average slant and tilt trajectories (expressed as functions of normalized time) for the seven different target slants in the full cue condition. The figures show data from the best stimulus condition (the binocular, good texture condition). The qualitative shapes of the paths and the trajectories were the same for all three "informative" stimulus conditions (excepting the

monocular noise texture condition). Not surprisingly, since only the slant of the target surface changed between stimulus conditions, most of the between condition variation in the rotation kinematics appears in the slant kinematics of the cylinder.

Place figure 5 here.

2.2.3 Sensitivity analysis

A standard way to analyze the sensitivity of the visuomotor system to the visual information in our task would be to measure constant and variable errors in the end-point slants of the cylinder as a function of target slant. Figure 6 plots the average slant of the cylinder at initial contact with the target surface as a function of the target slant for each of the four stimulus conditions. For the three information-rich stimulus conditions, there is little constant error, except at the highest slant. While the contact slants in the monocular/white noise stimulus condition show a regression toward a constant, intermediate slant, it clearly shows that some information for target slant is available in that condition. Figure 6 also shows the standard deviations of cylinder slants at initial contact for the four stimulus conditions. Note that the standard deviations increase dramatically for the monocular / white noise condition.

Place figure 6 here.

The data shown in the previous figures, while providing an initial insight into subjects' performance on the task, do not necessarily provide an accurate representation of the visuomotor system's sensitivity to visual information about slant. Earlier points in the trajectories could provide further information about the visuomotor system's estimate of target slant (e.g., because of cumulative effects of motor noise through the course of a motion). In order to accurately determine visuomotor

sensitivity to slant information, we must invert the forward mapping from target slant to cylinder trajectories and measure how much information the trajectories provide about target slant.

One way to do this would be to derive from the sample data an optimal, unbiased estimator of target slant from cylinder trajectory data. The variance of this estimator would reflect the visuomotor sensitivity of the system to target slant information. Doing so would require making assumptions about the global form of the estimator (e.g. linear) which might bias the results. Because the data from the experiment were derived from a small set of discrete target slants, we took a somewhat different tack, treating each target slant as a discrete stimulus category. We applied linear discriminant analysis[Duda and Hart, 1973] to derive d' measures specifying the average discriminability of trajectories generated for target slants that differ by 5 degrees (see Appendix A for details). This approach fits a different linear discriminant function to local pairs of slants (e.g. 75° vs. 80° and 85° vs. 90°), allowing for local deviations from global linearity. We refer to the resulting d-prime measures as the visuomotor d's for slant estimation. A d' of 1, for example, would indicate that an optimal linear model could correctly discriminate randomly drawn trajectories to surfaces differing in slant by 5 degrees 76% of the time. The d' measure provides a bias-free measure of the reliability of visuomotor slant "estimates" that uses all of the information provided by the output of the visuomotor system, the motion kinematics.

In applying discriminant analysis to our problem, we are faced with the problem of selecting the appropriate representation to use. Given the constraints on our data (approx. 65 trajectories per target slant per stimulus condition) and the total number of samples we have for each trajectory (approx. 35 - 65 per trajectory), we cannot use every sample point in the trajectory for the analysis (the number of free parameters in the discriminant function would approach or exceed the number of samples available for the analysis). Our first simplification was to analyze only the slant

trajectories of the cylinder; that is, the temporal trajectories of the angle of the cylinder in the same plane in which the target surface rotated. Including other kinematic parameters (e.g. the two other rotation angles and the transport parameters) did not increase the d' discriminability indices, justifying our choice. We then applied the discriminant analysis to three different representations of the slant trajectories. Two were derived by subsampling the trajectories and a third was derived from a principal components analysis of the trajectories.

Ten-point trajectories - These were ten-dimensional vectors specifying the slant of the cylinder at ten equally spaced times between movement start and initial contact with the target surface.

Contact slants - These were scalar values specifying the slant of the cylinder at the time of initial contact with the target surface.

Principal components - For each stimulus condition, we computed the principal components of the entire set of cylinder trajectories. Ten-dimensional vector representations of the trajectories were derived by projecting them into the space of the top ten principal components (which accounted for 99.9% of the variance in the trajectories).

d' values were computed for pairs of neighboring slants in each stimulus condition using covariance matrices and mean vectors estimated from the experimental data. Since d' values are non-negative, such a direct estimation method is inherently biased. We used a parametric bootstrap procedure to estimate both the bias in the d' estimates and the standard errors in the estimates [Efron and Tibshinari 1993][3]. The d' values reported here have been corrected for the estimated bias.

[3]In the parametric bootstrap, we used the covariance matrices and mean vectors calculated from the data to generate artificial trajectory vectors from which we estimated d'. Repeating the estimation many times allowed us to estimate the inherent bias in the estimation as well as the standard error of the estimates.

The d' measures did not show any consistent variation as a function of slant, so we only report the average d' measures here. Figure 7 shows the d' values computed for each stimulus condition averaged across target slants, using each of the three estimation methods described above. As can be clearly seen, the different trajectory representations used lead to the same estimates of d' (there were no interactions between the trajectory representation used for the analysis and target slant). Most notable is the fact that the contact slant of the cylinder captures all of the discriminable information from the trajectories. We explored the question of whether kinematic features derived from non-linear functions of the trajectories (e.g. maximum angular acceleration, time of maximum angular acceleration) add to the discriminability of the trajectories by combining them with the contact slant in the trajectory representation used to compute d' values. This gave no significant improvement, leaving us confident that the contact slant of the cylinder captures all of the information in the trajectories that reflects the different target slants used in the experiment.

Place figure 7 here.

A one way ANOVA on average d' values for the three informative stimulus conditions (using estimates of within condition variances derived from resampling) revealed a significant effect of stimulus condition for two of the subjects and a marginally significant effect for the third (Subject LES, $F(2, \infty) = 6.4$, $p < .01$, Subject MDY, $F(2, \infty) = 7.0$, $p < .01$, subject MEL, $F(2, \infty) = 2.6$, $p < .1$). Planned post-hoc comparisons revealed a significant difference between the monocular texture condition and the stereo noise condition for two subjects (subject LES, $Z = 2.5$, $p < .05$, subject MEL, $Z = 2.15$, $p < .05$) and a significant difference between stereo-noise and stereo texture conditions for only one subject (subject MDY, $Z = 4.6$, $p < .001$).

2.3 Discussion

The most significant feature of the results is how well subjects are able to use visual information about surface slant to control their orienting movements. In the full cue condition, subjects had visuomotor d's ranging from 1.5 to 2.5 for 5^o differences in target surface slant. These correspond to slant discrimination thresholds in a two-alternative forced-choice task of 1.9^o to 3.3^o (at 76% correct). By comparison, in a perceptual slant discrimination experiment using computer rendered displays, we have found thresholds ranging from, on average, 4^o to 10^o for the range of slants used in the current experiment[Saunders and Knill 2003]. This may reflect a higher efficiency for visuomotor processing or simply a difference in visual conditions (virtual vs. real surfaces). Experiment 2 was designed in part to resolve this question.

Two of the three subjects showed somewhat better performance in the stimulus condition containing only binocular information than in the condition containing only monocular, texture information. Only one of the three subjects showed a significant improvement in performance when both stereo and texture information were available. Two things are notable about the results. First, subjects' performance with binocular information is little better, on average, than it is with monocular texture information. This is consistent with earlier results using virtual stimuli, in which perceptual discrimination thresholds for surfaces slanted away from the viewer at 30° are, on average, equivalent for stimuli containing only stereo information and stimuli containing only texture information[Saunders and Knill 2003]. Second, only one of the three subjects shows evidence for efficient cue integration (subject MDY). The lack of improvement in the multiple cue condition (binocular views of regular textures) in the other two subjects may, however, simply reflect high levels of motor noise relative to noise in perceptual estimates of slant. Were this the

case, the high levels of motor noise would overwhelm the small improvements in d' predicted by efficient cue integration. Experiment 2 helps to resolve this issue.

3 Experiment 2: Comparison with perceptual sensitivity measures

The d' values measured in the first experiment are quite high. Translated to equivalent perceptual threshold measures, subjects' performance would reflect average 76% discrimination thresholds in a two alternative forced choice task of 1.9°, 2.6° and 3.3° for the three subjects in the full cue stimulus condition and slightly higher for the single cue conditions. Using virtual stimuli similar to the stereo views of white noise used here (random dots instead of white noise, however), Knill and Saunders [Knill and Saunders 2003] found slant discrimination thresholds no better than 10° for surfaces at 30° slant, which is within the range tested here. Even for stereo views of textured surfaces, they found thresholds averaged approcimately 7° for test surfaces at 30° slant. This runs counter to expectation, as visuomotor sensitivity is limited by the cumulative effects of sensory/perceptual uncertainty and motor noise. One would expect that visuomotor thresholds would be higher than visual discrimination thresholds. Of course, the results are not strictly comparable, as the current results were obtained with real surfaces while the others were obtained with virtual stimuli.

One explanation for the high levels of visuomotor sensitivity measured here is that visuomotor performance is driven by specialized transformations that are more efficient than those subserving perceptual judgments. This is the position argued by Goodale and Milner in their "two visual systems" hypothesis. Experiment 2 was designed to test whether this explanation accounts for the results of experiment 1 or if a simpler hypothesis, that both perceptual judgments and visuomotor performance derive from a common representation of slant, could account for the data.

The logic of the experiment follows from the claim that visuomotor transformations have a short memory. Goodale et. al. (1994) have argued that the visuomotor system's effective memory is less than 2 seconds and that for memory delays greater than that, the system reverts to stored perceptual representations of object properties to program grasping movements [Goodale et. al. 1994]. Previous studies have also shown that some perceptual illusions only begin to reflect themselves in visuomotor behavior after delays of two or more seconds between stimulus presentation of the initiation of movement. The result has been interpreted as reflecting a shift in relying on special-purpose visuomotor mechanisms to a reliance on biased perceptual representations at long delays [Bridgeman et. al. 2000], though this interpretation has been called into question.

Were subjects' visuomotor performance based on the same representation as perceptual judgments regardless of delay, one would expect similar changes in performance as a function of the delay between stimulus presentation and movement, on the one hand, or judgment, on the other. Previous studies of perceptual discrimination performance as a function of memory delay are consistent with a random walk model on the stored variable. This leads to a linear change in squared discrimination thresholds as a function of delay. The common representation hypothesis, therefore, predicts that visuomotor performance will decay in a similar manner. The "two systems" hypothesis, on the other hand, predicts a non-linear change in performance at the point where the system, switches from using special purpose visuomotor transformations to relying on stored perceptual representations of stimuli.

In experiment 2, we measured visuomotor slant discrimination "thresholds" as a function of the delay between stimulus presentation and the signal to initiate the object pacement movement used in experiment 1. For the same subjects, we measured perceptual discrimination thresholds

using the same stimuli used in the visuomotor task as a function of the delay between stmuli that subjects were required to compare to perform the task.

3.1 Methods

3.1.1 Motor task

The methods used to measure visuomotor sensitivity were similar to those used in experiment 1. Differences in the specifics are outlined below.

Stimuli

The viewing geometry for viewing stimuli duplicated that used in the first experiment. Stimuli consisted of field-limited, binocular views of real surfaces oriented at different slants around the horizontal by the robot arm. Target surfaces were planar white noise patterns mounted on the end of the robot arm. The depth of the target surface away from the viewer (measured at the center of the subjects' field of view) was randomized within a range from 44 to 46 cm. This was done so that, in the discrimination experiment (see below) subjects could not rely on absolute depth judgments at any individual point on a surface to make their judgments of slant (e.g. that the stimulus in a test pair that appeared further away at the top of subjects' field of view was the most slanted)[4]. The robot positioned the target surface randomly within the plane of the surface, so that different noise patterns appeared in the subject's field of view both within a trial and from trial to trial.

[4]For pairs of surfaces both at 45 cm from the observer, a 6° difference in slant (the largest used in the experiment) gave rise to a 1.5 cm depth difference at the top of the field of view. The 2 cm randomization in depth was designed to swamp this effect. Near subjects' thresholds, the largest depth difference for a fixed surface depth from the viewer was more on the order of .5 cm.

Subjects viewed stimuli at slants ranging between 13° and 37° in steps of 3° (specified in the subjects' frame of reference, where a slant of 0° represents a fronto-parallel surface). Nine slants in all were tested. The spacing between the target slants allowed us to calculate psychometric functions at different slants, and hence to calculate equivalent visuomotor discrimination "thresholds' for the test slants used in the discrimination experiment. Subjects viewed stimuli for 2 seconds before the shutter glasses closed.

Procedure

Four delay conditions were tested, corresponding to delays of 0, 1, 2 and 3 seconds between the extinction of target surface and the signal to initiate the movement to place the cylinder on the surface. Subjects ran in eight sessions on separate days. Each session was further sub-divided into four blocks, one for each delay condition. Stimulus blocks were randomized across sessions. Each block consisted of ten trials per test slant, making 90 trials per block. Subjects were given a 20 - 30 second break every 20 trials to avoid fatigue. They also had approximately a 2 minute break between blocks within a session. The order of blocks was randomized between sessions and counter-balanced so that the different delay conditions appeared in each temporal position within a session twice. Experimental sessions took less than 1 hour to complete. The first session was discarded as practice.

Figure 8a illustrates the time course of an experimental trial. The shutter glasses were initially closed. At a preset time, the shutters opened on the stimulus and remained open for 2 seconds, after which they closed again. Three delay conditions were tested, in which a 1, 2 or 3 second delay was imposed between shutter closing and the start signal to begin the motion. A fourth, no delay condition was tested in which the start signal was given 2 seconds after the shutter glasses first opened and the glasses remained open until the subject began moving the cylinder, at which point

the shutters were closed. Thus, in the no delay condition, subjects had vision of the target up to the point of movement initiation, but not during the movement. Subjects were asked to place the cylinder on a target projected onto the center of the target surface by a laser pointer.

The shutter glasses re-opened after the cylinder made contact with the target surface to give subjects feedback about the relative placement of the cylinder. After 1 second, a return signal was given and subjects returned the cylinder to the starting position.

3.1.2 Perceptual discrimination experiment

In the discrimination experiment, subjects judged which of two sequentially presented test surfaces was more slanted away from them. We measured discrimination thresholds around target slants of 22°, 25° and 28° (measured from the fronto-parallel as seen by the subject) as a function of the delay between presentations of the first and second stimulus in a trial. We varied the delay from 1 to 3 seconds.

Stimuli

Stimuli were equivalent to those used in the motor task. Subjects wore a pair of opaque liquid crystal shutter glasses (PLATO shutter glasses) that allowed us to automatically limit the duration over which they viewed the target surfaces. As in experiment 1, they viewed surfaces through translucent tubes placed in front of each eye to limit their field of view on the surface. The shutter glasses were closed between trials and between stimulus presentations within a trial.

Procedure

We used a temporal, 2-AFC task to measure discrimination thresholds. Figure 8b illustrates the sequence of events in a trial. A subject viewed the target surface at one slant for 2 seconds. The

view was then occluded while the robot repositioned the surface at a different slant (and depth) and the subject viewed the second surface for another 2 seconds. Four delay conditions were tested, corresponding to periods of 1, 1.5, 2 and 3 seconds between the extinction of the first surface and the display of the second surface. Because of the time it took the robot to move the surface from one slant to another and for the surface to stabilize after the transient surface motion (wobble) created by stopping the robot, 1 second was the minimal delay we could impose between views of the two surfaces being discriminated. Subjects used the computer mouse to indicate which of the two surfaces appeared more slanted (closer to a ground plane). Subjects ran in eight sessions on separate days. Each session was further sub-divided into four blocks, one for each delay condition. There was a small break of approximately 2 minutes between each block within a session. Stimulus blocks were randomized across sessions and counterbalanced so that each delay condition appeared in each of the four temporal positions within a session twice. Experimental sessions took less than 1 hour to complete. The first session was discarded as practice.

Place figure 8 here.

We used a method of constant stimuli to estimate discrimination thresholds. Test stimuli differed by 2, 4 or 6 degrees around each of the target slants for which we estimated discrimination thresholds (22°, 25° and 28°). Thus, for a target slant of 25°, test surface pairs were shown at 24° and 26°, 23° and 27° and 22° and 28°, corresponding to slant differences of 2°, 4° and 6°, respectively. Test slants thus varied between 19° and 31°, with an equal number of presentations at each slant. The range of slants corresponded to angles ranging from 15° to 27° away from the horizontal (relative to gravity).

Subjects were given feedback in the form of a tone when their judgment was incorrect.

Subjects

Three undergraduates from the University of Minnesota served as subjects in the experiment. All subjects were naive to the purposes of the experiment and had normal or corrected-to-normal vision.

3.2 Results

3.2.1 Visuomotor performance

Place figure 9 here.

Experiment 1 showed that contact slant carried all of the discriminative power of the motion trajectories. We therefore measured equivalent visuomotor discrimination thresholds by simulating an observer that made discrimination judgments based on the contact slants derived from a pair of cylinder orientation trajectories derived from randomly drawn trials in the visuomotor experiment. For each target slant, we simulated a discrimination experiment by randomly drawing trials from the target slant condition and each of the slant conditions within $\pm 9^\circ$ of the target slant condition. The average probability correct for each pair of slants was obtained by calculating the proportion of all combinations of trials on which the contact slant of the cylinder for the surface with the greater slant was greater than the contact slant for the surface with the smaller slant[5]. Figure 9 plots the visuomotor thresholds calculated from these psychometric functions for each of the three subjects.

[5]We threw away trials on which contact slants were more than three standard deviations from the mean for a particular delay/slant condition.

3.2.2 Perceptual performance

We assumed that subjects' judgments, when stimuli are properly attended to, are governed by a cumulative gaussian psychometric function, with a mean shifted away from 0. We modeled the effects of attentional lapses using two parameters, the probability of a lapse (leading to guessing) and the probability, given that the subject is guessing, that the subject will guess the second stimulus. Appendix B describes the psychometric model used to account for the bias and lapse effects. The thresholds reported here were derived from a model which assumed that the guessing parameters were fixed across all conditions of the experiment (see Appendix B for more discussion of this point).

In the experiment, the 25° slant condition was the only one that truly isolated memory effects. Thresholds for the other two slant conditions are contaminated by the fact that some stimulus slants in these conditions were either always the smallest within a pair (21°, 20° and 19°) or were always the largest largest within a pair(29°, 30° and 31°). When one of these stimuli was presented as the first stimulus in a pair, subjects could, in theory, have made their judgment accurately based only on the first stimulus and learned absolute thresholds. We will, therefore, focus our analysis and discussion on the data for the 25° target slant.

Place figure 10 here.

Figure 10 shows plots of the perceptual discrimination thresholds for the 25° target slant conditions at the 1, 2 and 3 second delay conditions at which we also measured visuomotor thresholds. Discrimination thresholds for subjects SS and JJ increase significantly with increasing delay. SS showed a 60% increase in threshold when th delay increased from 1 to 2 seconds ($Z = 2.4$, $p < .01$), while subject JJ showed a 76% increase ($Z = 3.9$, $p < .01$). Subject MEL only showed a

5% increase which was not significant (Z = .3, p > .35). These values clearly show a significant memory cost for two of the three subjects.

Shown for comparison in the same figures are equivalent visuomotor thresholds. These have been corrected for the fact that the discrimination thresholds were derived from a task requiring comparison of a remembered stimulus and a no-delay stimulus, whereas the visuomotor thresholds shown in figure 9 were computed from motor performance measured for a common delay. Assuming that the thresholds in each condition were proportional to the standard deviation of the noise in the internal representation of slant, we can correct for this difference using a non-linear average of the visuomotor thresholds at delay, Δ, and the thresholds at zero delay. The correction takes the form,

$$\hat{T}_m(\Delta) = \sqrt{\frac{1}{2}(T_m(\Delta)^2 + T_m(0)^2)}, \tag{1}$$

where $\hat{T}_m(\Delta)$ is the corrected visuomotor threshold for stimuli delayed by Δ seconds, $T_m(\Delta)$ is the visuomotor threshold computed for the same stimuli as described above and $T_m(0)$ is the visuomotor threshold calculated as above for the no delay condition.

3.3 Discussion

Visuomotor thresholds for subjects SS and JJ closely follow the increase as a function of memory delay that would be predicted by a constant temporal decay in perceptual uncertainty. Threshold functions for MEL flatten at large delays. In order to gain more quantitative insight into the results, we fitted a simple version of the "common representation" model to the data. According to this model, performance on both the perceptual discrimination task and the visuomotor task is determined by the same uncertainty in visually derived slant, which is expected to vary with

memory delay. The two tasks, however, are affected by different constant noise sources decision noise for the perceptual task and motor noise for the motor task.

According to the model, perceptual and visuomotor thresholds should be related by

$$\hat{T}_m(\Delta)^2 = \hat{T}_p(\Delta)^2 + K, \tag{2}$$

where K is a constant representing the difference between the variance of the motor noise and the variance of the decision noise. Assuming that the decision noise is significantly lower than the motor noise, we can treat the constant K as an estimate of the motor noise level.

Place figure 11 here.

Figure 11 shows the value of K derived for each of the three common delay conditions in the experiment. The common representation hypothesis predicts that the estimated values of K should remain fixed as a function of delay. The data shows no significant difference between the estimates of K, consistent with the model. Note that this is true for subject MEL, as well as subjects SS and JJ, despite the flattening of her visuomotor threshold function at high delays. This is because her perceptual thresholds follow a similar pattern as a function of delay.

While the data are consistent with the common representation hypothesis, we note that the differences between perceptual and visuomotor performance are low. It is possible that a putative shift from special-purpose visuomotor transformations to ones based on distinct perceptual representations when memory at large delays adds noise that is too low to pick up with the present technique. It is also possible that the shift occurs at delays less than one second. We do not have perceptual data at a zero second delay; however, a diffusion model for perceptual memory decay predicts a linear shift in squared thresholds as the delay is decreased to zero[Kinchla and Smyzer 1967]. Discrimi-

nation thresholds for a number of simple, two-dimensional geometric properties of visual stimuli (e.g. contour curvature) show a linear shift as a function of inter-stimulus delay[Hole 1996, Sakai 2003]. The small change in subjects SS and JJ's visuomotor thresholds between the zero delay and 1 second delay conditions is consistent with this model.

The data from experiment 2 are notable in two more regards. First, subjects are remarkably good at making the slant discrimination required in the experiment - much better than has been reported in psychophysical experiments using virtual stimuli[Saunders and Knill 2003]. It may derive from the fact that virtual stimuli contain cues that remain fixed, regardless of the simulated surface geometry, causing cue conflicts (e.g. accommodation and blur), while those same cues co-vary appropriately with surface geometry in real stimuli. It may also reflect the greater reliability of stereo information when cues to absolute surface depth, needed to calibrate the interpretation of disparity, are stronger in the stimulus, as they are for real surfaces. Second, assuming that the difference between perceptual and visuomotor performance is largely due to motor noise, the data suggest highly efficient transformations between visual estimates of slant and motor output. The variance in subjects' motor performance induced by motor noise had a slant discrimination equivalence of between 1° and 2.5° for the three subjects.

4 General Discussion

4.1 Sources of uncertainty in orienting movements

The results of experiment 2 give no support for the hypothesis that orienting movements of the hand to match the target orientation of a surface during a natural object placement movement rely on a visuomotor channel that is independent of a perceptual processing channel. Rather, they are consistent with the hypothesis that visuomotor variability is a result of independent perceptual

and motor noise sources. In order to test this more fully, however, conditions must be created in which the perceptual uncertainty varies over a larger range than followed from the memory delay manipulation used in this experiment. Nevertheless, the results interpreted according to the common representation hypothesis show very small levels of internal noise, both at the perceptual input and the motor output. Since visual feedback from the moving hand (or the cylinder) was not available for the task, the low motor noise suggests efficient transformations from the perceptual representation of slant to the motor output (equivalent motor noise on the order of 1 to 2 degrees, standard deviation). That end-point slant carries all of the information needed to discriminate trajectories toward one target slant from trajectories toward another further suggests that this efficiency derives both from motor planning and the efficient use of on-line control based either on internal feedback loops (using efferent copy information) and/or proprioceptive feedback.

4.2 The role of binocular and monocular cues for guiding orienting movements

Experiment 1 showed that subjects could accurately guide hand orienting movements to place a cylindrical object on a flat surface based on either binocular and texture information about surface orientation alone. The sensitivity analysis left open the question of whether or not performance improved significantly when both cues were available. One subject showed a statistically significant improvement, one showed a small, but not quite significant improvement and the other showed no improvement. The results of the second experiment may well explain the small amount of improvement obtained in the multiple cue condition. Since different subjects ran in the two experiment, we cannot compare results directly, however, we can do a simple back-of-the-envelope calculation to show why the motor noise that corrupts subjects' performance compresses any effect obtained to a very small one.

Experiment 2 showed that subjects' motor noise remains constant across memory delays. If we assume that it is independent of perceptual noise and hence equivalent across cue conditions, we can express the visuomotor d-primes measured in experiment 1 as a non-linear sum of perceptual and motor d-primes,

$$d' = \frac{d'_{percept} d'_{motor}}{\sqrt{(d'_{percept})^2 + (d'_{motor})^2}}, \tag{3}$$

where d' is the visuomotor d-prime measured in experiment 1. The lowest level of motor noise measured in experiment 2 was about equal to the perceptual noise. Using this conservative estimate, we have for visuomotor d-primes for the individual cue conditions in experiment 1,

$$d' = \frac{d'_{percept}}{\sqrt{2}}. \tag{4}$$

Assuming that perceptual uncertainty is equal in the texture and stereo conditions, optimal cue integration predicts that the perceptual d-prime for the combined cue condition will be greater than the perceptual d-prime for each individual cue condition by a factor of $\sqrt{2}$. After some algebraic manipulation, this predicts that a subjects' visuomotor d-prime for the combined cue condition cue condition in experiment 1, d'_{TS}, would be related to the vsuomotor d-primes in the individual cue conditions, d', by

$$d'_{TS} = 1.15d', \tag{5}$$

This is to be compared with the value of $d'_{TS} = 1.41d'$ predicted by optimal integration when no motor noise is present. Our calculation shows that the relatively small improvement shown by two of the three subjects in experiment 1 is, in fact, consistent with optimal cue integration, given the levels of motor noise that were likely present in the data.

4.3 Visuomotor control issues

Place figure 12 here.

While we have focused on visuomotor sensitivity in this paper, we also noticed patterns in the kinematic data charactering the transport of the cylinder that suggest an interesting synergy between hand rotation and hand transport. Examination of the transport kinematics revealed that the transport kinematics of the cylinder appeared to be invariant over target surface orientation when represented by the mid-point of the bottom of the cylinder (figure 12). This result suggests that the position of the bottom of the cylinder is one of the parameters, along with orientation, that is controlled by the motor system to perform the object placement. Such a strategy makes sense given the task demands, which were to center the cylinder on the target location. Since subjects grasped the cylinder at its mid-point, controlling the path of the bottom of the cylinder so as to maintain invariance over cylinder orientation required coordinated control of the transport of the wrist through space and of the movements used to orient the cylinder - rotation of the wrist and coordinated movements of the fingers to "roll" the cylinder between them. While a number of physiological and behavioral results support the hypothesis that proximal (e.g., hand transport) and distal components (e.g. grip formation) of prehension movements are controlled through independent visuomotor "channels" [Jeannerod 1981, Jeannerod et. al. 1995, Paulignon and Jeannerod 1996], other results strongly suggest some interaction between the two, particular as regards transport and hand orientation [Haggard 1991, Soechting and Flanders 1993, Gentilluci et. al. 1996, Mamamssian 1997]. The current results are consistent with the concept of a flexible control system that coordinates distal and proximal components of hand movements to successfully match the demands of specific motor tasks [Todorov and Jordan 2002].

5 Summary

Applying linear discriminant analysis to the motion trajectories of a cylinder measured during an object placement task, we were able to compute measures of visuomotor sensitivity to visual information about surface slant. Subjects' motor performance reflects a very high sensitivity to surface slant information. This sensitivity is nearly the same for binocular information and for monocular texture information, suggesting that, for the control of hand orientation, monocular cues can be just as useful for visuomotor control as binocular cues to 3D surface structure. Comparing visuomotor performance to perceptual discrimination measures in equivalent stimulus conditions reveals that motor and perceptual noise contribute, on average, similar amounts of variability to subjects' visuomotor performance. That the difference between visuomotor variance and perceptual variance measures (squared thresholds) did not differ significantly across different delay conditions is consistent with a common perceptual representation of slant underlying both tasks for the delays tested.

Appendix A: Discriminant analysis

Place figure 13 here.

Given an N-dimensional parametric representation of cylinder trajectories at each of two neighboring target slants (e.g. 90° and 85°), we computed the discriminant vector, $\vec{W}$, that supports maximum discriminability of the two sets of trajectories. This is given by the equation,

$$\vec{W} = (\vec{\Sigma}_1 + \vec{\Sigma}_2)^{-1}(\vec{\mu}_1 - \vec{\mu}_2), \tag{6}$$

where $\vec{\Sigma}_1$ and $\vec{\Sigma}_2$ are the covariance matrices and $\vec{\mu}_1$ and $\vec{\mu}_2$ are the means of the trajectories measured for each target slant. Given a sample trajectory, an ideal classifier would project the N-dimensional vector representation of the trajectory onto the discriminant vector and compare the result to a scalar decision criterion to decide which target slant had been presented to the subject. The discrimination performance of such an observer is characterized using a single d' value computed as

$$d' = \frac{| < \vec{W}, \vec{\mu}_1 - \vec{\mu}_2 > |}{\sqrt{\frac{1}{2}(\vec{W}^T\vec{\Sigma}_1\vec{W} + \vec{W}^T\vec{\Sigma}_2\vec{W})}}, \tag{7}$$

where $< \vec{x}, \vec{y} >$ is the inner product between $\vec{x}$ and $\vec{y}$. The numerator is the distance between the projections of the two mean trajectories onto the discriminant vector and the denominator is the square root of the average variance of each set of trajectories when projected onto the discriminant vector. One way to interpret the discriminant vector is that it is the vector that maximizes the d' value computed from equation 7; that is, it specifies the projection of the trajectory vectors that maximizes their discriminability. Figure 13 illustrates the analysis in two dimensions. d' also specifies the probability correct of the ideal classifier, with an unbiased observer having a probability

correct given by the value of the cumulative normal distribution at the value specified by d'; thus, for example, $d' = 1$ corresponds to approximately 76% correct classification performance.

Appendix B

We found that psychometric functions for some subjects in the discrimination experiment leveled off at points below 1.0 and above 0.0. We modeled these effects by assuming that subjects suffered from occasional lapses in attention, leading to a percentage of trials in which they guessed the correct answer[Wichmann and Hill, 2001]. In order to correct for these lapses and the resultant guessing, we fit a modified cumulative Gaussian psychometric function to each subject's data in which the probability of selecting a comparison stimulus was assumed to be a mixture of an underlying Gaussian discrimination process and a random guessing process. Writing subjects' decision as

$$D = \begin{cases} 1; & \text{Second stimulus judged more slanted} \\ 0; & \text{First stimulus judged more slanted} \end{cases} \tag{8}$$

The psychometric model was

$$p(D = 1 \,|\, \Delta S) = (1 - p)\, \Gamma(\Delta S; \mu, \sigma) + pq, \tag{9}$$

where ΔS is the difference in slant between the first and second stimulus, μ is the mean of the cumulative Gaussian, σ is the standard deviation of the cumulative Gaussian, p is the probability that a subject guessed on any given trial and q is the probability that subject guessed the comparison stimulus, given that he or she guessed at all. The mean parameter, μ, is a measure of the point of subjective equality between first and second stimuli. It accommodates effects like perceptual drift in the remembered slant of the first stimulus. We computed a corrected 76% threshold from the standard deviation parameter σ of the model. The corrected threshold reflects the 76% threshold difference in slant between test and comparison stimuli that a subject would have in a 2-AFC choice without guessing and without a temporal order bias in slant judgments (reflected by the μ parameter).

The guessing parameters add two free parameters to the psychometric model, which, for an individual subject and condition, make model fits difficult. For example, they correlate strongly with both the threshold and the bias parameters. We hypothesized that the lapse and guessing parameters would be constant across different test slants and perhaps even across different delays. We therefore considered three qualitatively different models when fitting discrimination thresholds:

• Unconstrained model - subjects had different lapse rates and different guessing biases for every different condition in the experiment (different test slant and different memory delay).

• Delay-conditioned lapse model - subjects had the same lapse rate and guessing biases for different surface slants, but different values for these parameters for different memory delays (e.g. one might reasonably expect the lapse rate to increase with increased memory delay).

• Constant lapse model - subjects had constant lapse rates and guessing biases across all conditions of the experiment.

In terms of the model parameters, the three models correspond to, respectively, fitting the p and q parameters independently for each condition of the experiment, fitting one set of p and q parameters for each memory delay condition, and fitting only one set of p and q parameters across all conditions.

We used Bayesian model testing to find the model which best fit the data. In particular, we calculated $p(\text{Model}i|\,\text{Data})$, for each of the three models by integrating the likelihood functions for the free parameters in each model (the threshold and bias estimates and the free p and q parameters) [MacKay 1992]. This involves selecting priors on the model parameters. We assumed uniform priors on the lapse and guessing bias parameters between 0 and 1, a uniform prior on the thresholds between 0 and 10 degrees and a uniform prior on the bias parameter, μ between -10 and 10 degrees. Since the only difference between the models was in the number of p and

q parameters that were free to vary, the priors on the thresholds and bias parameters did not contribute significantly to the relative likelihoods of the three models. For all three subjects, the constant lapse model was significantly more likely than the other two, by factors greater than 1000. We therefore used this model to fit subjects' thresholds. Standard errors for the threshold fits were computed from the inverse of the Gaussian of the model likelihood function.

An alternative approach to Bayesian model testing is to used a standard nested-hypothesis test to determine if we can reject the simplest model fit to the data (constant lapse model). The relative likelihoods of the simplest and more complex models,

$$2log_e(\frac{L_{independent}}{L_{simple}}), \tag{10}$$

is distributed as χ^2 with either 4 degrees of freedom (when comparing the delay-conditioned lapse model to the constant lapse model) or 22 degrees of freedom (when comparing the unconstrained modell to the constant lapse model). In neither case did the Chi-square statistic reach the $p = .05$ level of significance. While this confirms the results of the Bayesian model testing, we prefer the former, as it places all three models on equal footing. The nested hypothesis test, on the other hand, is inherently biased to accept the simplified model (it is rather arbitrarily considered to be the null model) and hence is subject to Type 2 inference errors. Bayesian model testing has it's own difficulty - having to assume a fixed prior; however, in this case, a uniform prior makes sense and changing the prior (e.g. making the prior on squared values of p and q uniform) has little effect on the probabilities assigned to the different models.

Acknowledgements The authors would like to thank Bruce Hartung for his help running subjects.

This research was supported by NIH RO1 EY11507 and NIH R01 EY13319.

References

[Bridgeman 1999] Bridgeman, B. (1999) Separate representations of visual space for perception and visually guided behavior, in (Aschersleben, G. and Bachman, T. (eds.)) *Cognitive contributions to the perception of spatial and temporal events, Advances in Psychology 128*, 3-13.

[Bridgeman et. al. 2000] Bridgeman, B., Gemmer, A., Forsman, T. and Huemer, V. (2000) Processing spatial information in the sensorimotor branch of the visual system, *Vision Research* **40** (25), 3539-3552.

[Bridgeman, et. al., 2003] Bridgeman, B, Dassonville, P., Bala, J. and Thiem, P. (2003) What is stored in the sensorimotor visual system: map or egocentric calibration? *Journal of Vision*, **3** (9), 13.

[Bruno 2001] Bruno, N. (2001) When does action resist visual illusions, *Trends in Cog. Sci.*, **5** (9), 379 - 382.

[Desmurget et. al. 1997] Desmurget, M., Rossetti, Y., Jordan, M., Meckler, C., Prablanc, C. (1997). Viewing the hand prior to movement improves accuracy of pointing performed toward the unseen contralateral hand, *Experimental Brain Research*, **115** (1), 180-186.

[Duda and Hart, 1973] Duda, R. O. and Hart, P. E. (1973) Pattern classification and scene analysis, Wiley, New York.

[Efron and Tibshinari 1993] Efron, B. and Tibshirani, R. J. (1993) *An introduction to the bootstrap*, Chapman and Hall, New York.

[Franz et. al. 2001] Franz, V. H., Fahle, M., Buelthoff, H. and Gegenfurtner, K. R. (2001) Effects of visual illusions on grasping, *JEP: Human Perc. and Perf.*, **27** (5), 1124 - 1144.

[Gentilluci et. al. 1996] Gentilucci, M, Daprati, E, Gangitano, M, Saetti, M. C. and Toni, I. (1996) On orienting the hand to reach and grasp an object, *NeuroReport*, **7**, 589-592.

[Gharamani et. al. 1997] Gharamani, Z., Wolpert, D. M. and Jordan, M. I. (1997) Computational models of sensori-motor integration, in *Self-organization, Computational Maps, and Motor Control* (eds Morasso, P. G. and Sanguineti, V.) Elsevier Press, Amsterdam.

[Goodale and Milner 1992] Goodale, M. A. and Milner, A.D. (1992). "Separate visual pathways for perception and action." *Trends in Neuroscience*, **15**, 20-25.

[Goodale et. al. 1994] Goodale , M A, Jakobson, l S and Keillor, J M (1994) Differences in the visual control of pantomimed and natural grasping movements, *Neuropsychologia*, **32** (10), 1159-1178.

[Haffenden and Goodale 1998] Haffenden, A. M. and Goodale, M. A. (1998) The effect of pictorial illusions on prehension and perception, *J. Cog. Neuro.*, **10**(11), 122-136.

[Haggard 1991] Haggard, P. (1991) Task coordination in human prehension, *Journal of Motor Behavior*, **23** (1), 25 - 37.

[Hole 1996] Hole, G J (1996) decay and interference effects in visuospatial short-term memory, *Perception*, **25**(1), 53-64.

[Jeannerod 1981] Jeannerod, M (1981) Intersegmental coordination during reaching at natural visual objects, in (Long, J, Baeddely, A (eds.)) *Attention and Performance IX*, Lawrence Erlbaum, Hillsdale, NJ, 153-168.

[Jeannerod et. al. 1995] Jeannerod, M, Arbib, M. A., Rizzolati, G. and Sakata H. (1995) Grasping objects: the cortical mechanisms of visuomotor transformations, *Trends in Neuroscience*, **18**, 314-20.

[Kawato 1999] Kawato, M. (1999) Internal models for motor control and trajectory planning, *Current Opinion Neurobiology*, **9**, (6), 718-27.

[Kinchla and Smyzer 1967] Kinchla R A and Smyzer F (1967) A diffusion model of perceptual memory, *Perception and Psychophysics*, **2**, 219-229.

[Knill and Saunders 2003] Knill, D. C. and Saunders, J. (2003) Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research*.

[Saunders and Knill 2003] Saunders, J. and Knll, D. C. (In press) Humans use continuous visual feedback from the hand to control fast reaching movements? *Experimental Brain Research*.

[Kruyer et. al. 1996] Kruyer A, Marotta J J, Goodale M A (1996) Manual grasping in a cue-deprived environment, *Investigative Ophthalmology and Visual Science*, **37** (3), 2408.

[MacKay 1992] Mackay, D. J. C. (1992) Bayesian interpolation, *Neural Computing*, **4**, 415 - 447.

[Mamamssian 1997] Mamassian, P. (1997) Prehension of objects oriented in three-dimensional space, *Exp. Brain Res.*, **114**, 235-245.

[Marotta et. al. 1997] Marotta J J, Behrmann M, Goodale M A (1997) The removal of binocular cues disrupts the calibration of grasping in patients with visual form agnosia, *Exp. Brain Res.*, **116** (1), 113-121.

[Milner and Goodale 1995] Milner, D. and Goodale, M. A. (1995) *The Visual Brain in Action*, Oxford Univ. Press, Oxford, U.K.

[Moll and Kuypers, 1980] Moll, L and Kuypers, H G J M (1980) impaired visually guided arm, hand and finger movements, *Neurology*,

[Mon-Williams and Tresilian 2000] Mon-Williams, M. and Tresilian, J. R. (2000) Ordinal depth information from accommodation?, *Ergonomics*, **43** (3), 391-404.

[Paulignon and Jeannerod 1996] Paulignan, Y and Jeannerod, M. (1996) Prehension movements: the visuomotor channesl hypothesis revisited, in (Wing, A. M., Haggard, P. and Flanagan, R. (eds.)) *Hand and Brain: The Neurophysiology and Psychology of Hand Movement,* Academic Press, Orlando, FL, 265-282.

[Sakai 2003] Sakai, K (2003) Short-term visual memory for contour curvature in a delayed discrimination task, *Japanese Psychological Research*, **45** (2), 122-128.

[Santello and Soechting 1998] Santello, M. and Soechting, J. F. (1998) Gradual molding of the hand to object contours, *J. Neurophysiol*, **79** (3), 1307-20.

[Servos 2000] Servos, P. (2000) Distance estimation in the visual and visuomotor systems, *Exp. Brain Res.*, **130** (1), 35-47.

[Soechting and Flanders 1993] Soechting, J F and Flanders M (1993) Parallel, interdependent channels for location and orientation in sensorimotor transformations for reaching and grasping, *J. Neurophysiology*, **70**, 1137-1150.

[Todorov and Jordan 2002] Todorov, E. and Jordan, M. (2002) Optimal feedback control as a theory of motor coordination, *Nature Neuroscience*, **5** (11), 1226 - 1235.

[Watt and Bradshaw 2003] Watt S J and Bradshaw M F (2003) The visual control of reaching and grasping: Binocular disparity and motion parallax, *Journal of Exp. Psychology: Human Perception and Performance*, **29** (2), 404-415.

[Wichmann and Hill, 2001] Wichmann, F. A. and Hill, N. J. (2001) The psychometric function: I. Fitting, sampling, and goodness of fit, *Perception and Psychophysics*, **63** (8), 1293-1313.

[Wolpert et. al. 1995] Wolpert, D. M., Gharamani, Z. and Jordan, M. I. (1995) An internal model for sensorimotor integration, *Science*, **269**, 1880-82.

Figure Captions

Figure 1: The task subjects performed in the experiment. Starting from a horizontal platform (Panel a), subjects placed a cylinder flush onto a flat target surface of variable slant mounted on the end of a robot arm (Panel b). Stimulus information about surface orientation was isolated by fixing the location of the targeted placement in space. The robot arm was used to orient the surface at different slants across trials.

Figure 2: The dimensions of the experimental setup. The hatched area represents the back-projection of the round viewing apertures onto the target surface when the surface was level (perpendicular to gravity). The dashed line through the target surface in the bottom figure represents the axis around which the surface was rotated to create different target surface slants.

Figure 3: Examples of the textures mounted on the end of the robot arm. (a) A white noise texture that provides little if any texture cues to surface orientation when projected into the retinal image and (b) a regular pattern of dots whose projection provides very good information about surface orientation.

Figure 4: The acceleration profile for one trajectory. In this case, contact with the target surface was marked by two sharp negative peaks in acceleration following in quick succession. The first peak reflects the initial contact. The second peak reflects the point at which the cylinder rotates down flush to the surface.

Figure 5: Average slant and tilt trajectories for each subject for the full cue stimulus condition. Each curve reflects a different target slant.

Figure 6: Left column: Average contact slants as a function of target slant for each of the three subjects. Right column: Standard deviations of contact slants as a function of target slant.

Figure 7: d's averaged across target surface slants for each of the stimulus conditions. Three sets of bar plots are shown for each subject corresponding to the three methods used to derive d' estimates.

Figure 8: (a) The sequence of events for a trial in the visuomotor experiment. A tone sounded after the delay period instructing the subject to begin the movement to place the cylinder on the surface. The LCD glasses were closed during the delay period. (b) The sequence of events in a discrimination trial. The LCD glasses were closed during the inter-stimulus interval, preventing vision of the moving surface. Note that the surface randomly changed positioned in the inter-stimulus interval. Parts of the surface outside the viewing aperture were not visible to the subject.

Figure 9: Visuomotor discrimination thresholds, derived as described in the text, for the three subjects, plotted for each target slant as a function of the delay between stimulus presentations in the temporal 2-AFC task.

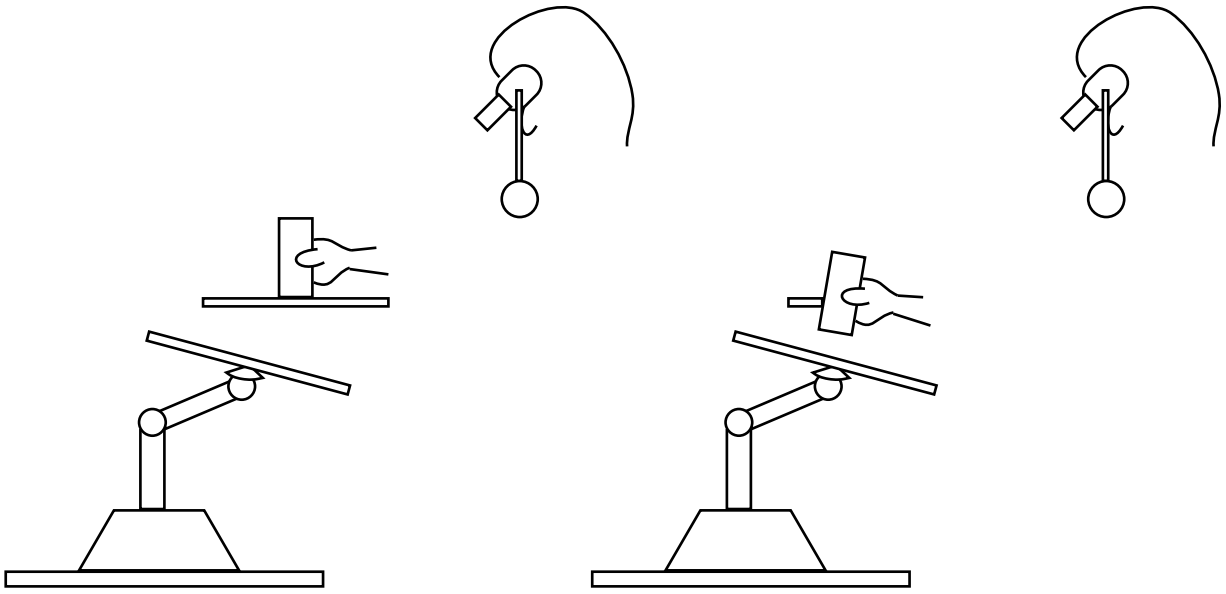
Figure 10: Perceptual discrimination thresholds for the 25° slant condition plotted for the three matching delays used in the visuomotor task (1,2 and 3 seconds). Also shown are the equivalent visuomotor thresholds.

Figure 11: Plots of the standard deviation of motor noise derived from the simple additive noise model. This is given by the square root of the difference of squared visuomotor and perceptual discrimination thresholds (assuming independent noise sources

and no effective decision noise in the perceptual task). Error bars were derived from the error bars of the visuomotor and perceptual discrimination thresholds.

Figure 12: Transport paths for the cylinder in the binocular, good texture condition. The left graphs show the average paths of the center of mass of the cylinder for each target slant. The right graphs show the average paths of the bottom of the cylinder.

Figure 13: (a) Scatter plot of cylinder slants at two points in a movement (mid-point and contact) derived from movements in one condition toward target slants of 90° (red) and 85° (blue). The ellipses represent the covariance ellipses of the data for the two target slants. The line, D represents the line of maximal discrimination. An ideal classifier would label movements falling below the line as being toward a target slant of 85° and movements falling above the line as being toward a target slant of 90°. Mathematically, this is done by projecting the point representing the movement onto a vector, W, perpendicular to the discriminant line and determining whether it is above or below some criterion (b). The discriminability of the two sets of trajectories is captured in a d' measure given by the distance between the means of the projected points normalized by their average standard deviations. In higher dimensions, the discriminant line, D, becomes a hyper-plane.



(a) (b)

Figure 1:

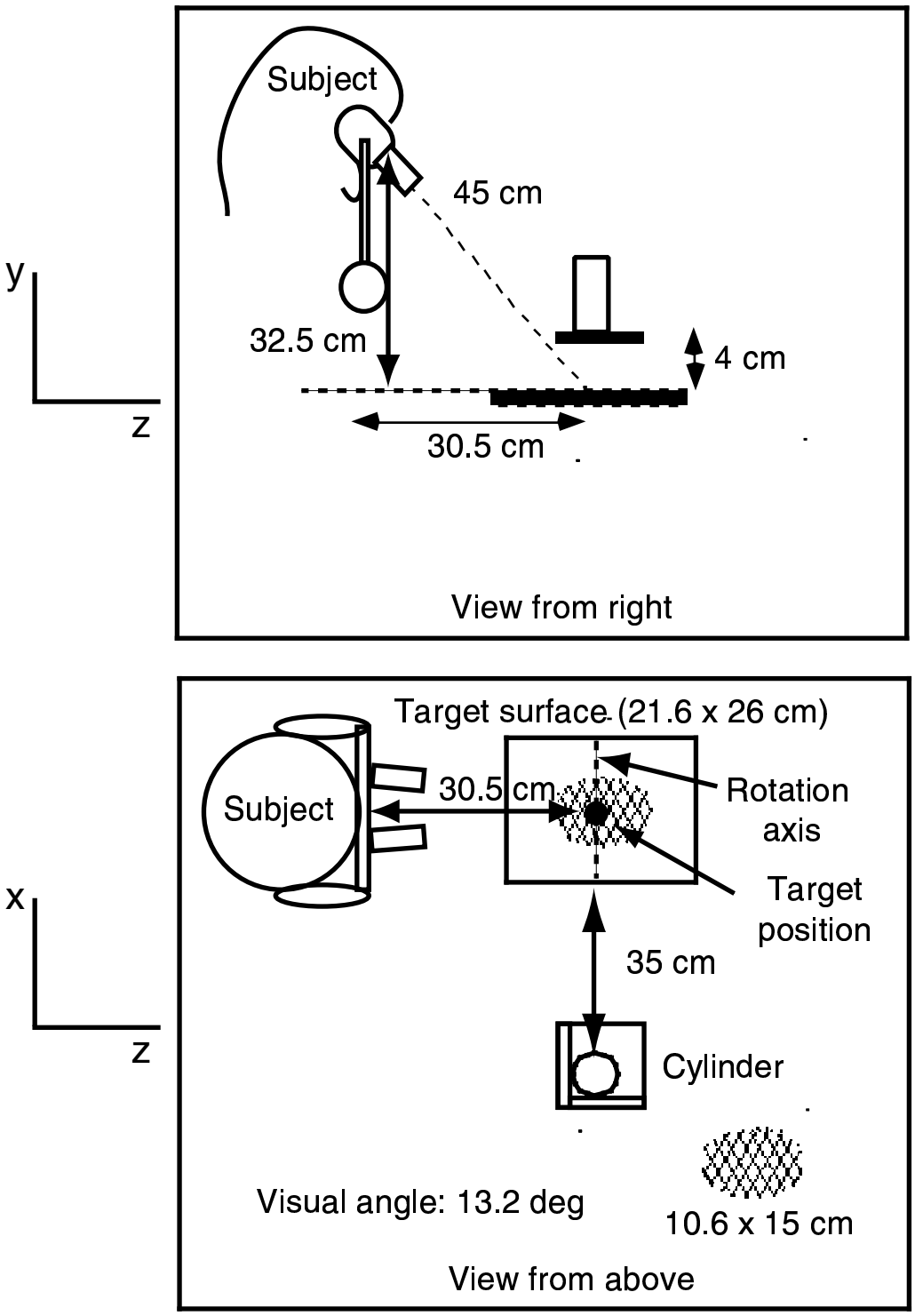
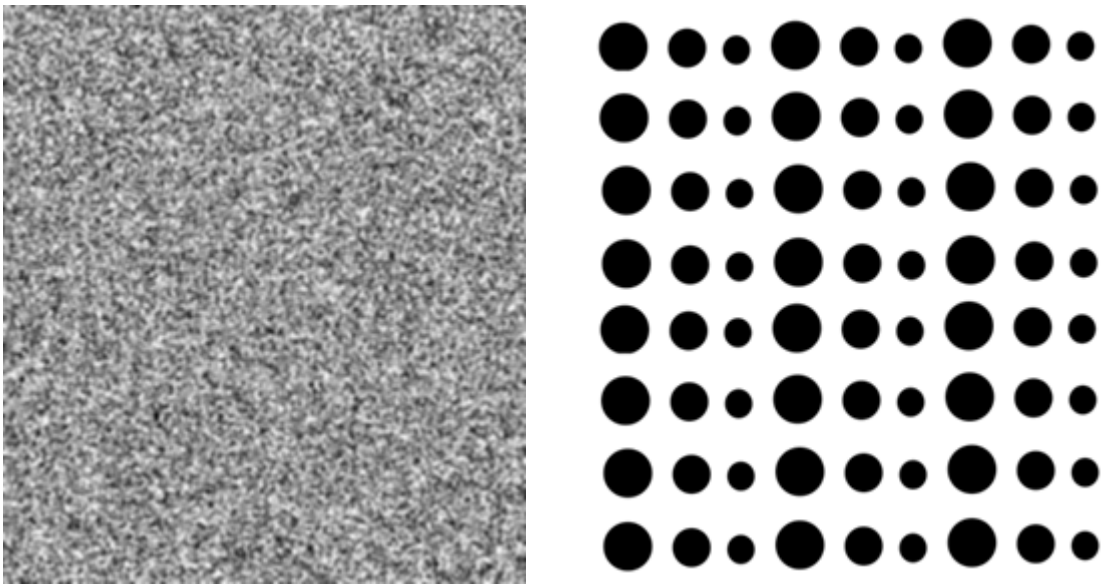

Subject
45 cm
y
32.5 cm
4 cm
z
30.5 cm
View from right
Target surface (21.6 x 26 cm)
Subject
30.5 cm
Rotation axis
Target position
x
35 cm
z
Cylinder
Visual angle: 13.2 deg
10.6 x 15 cm
View from above

Figure 2:



(a) (b)

Figure 3:

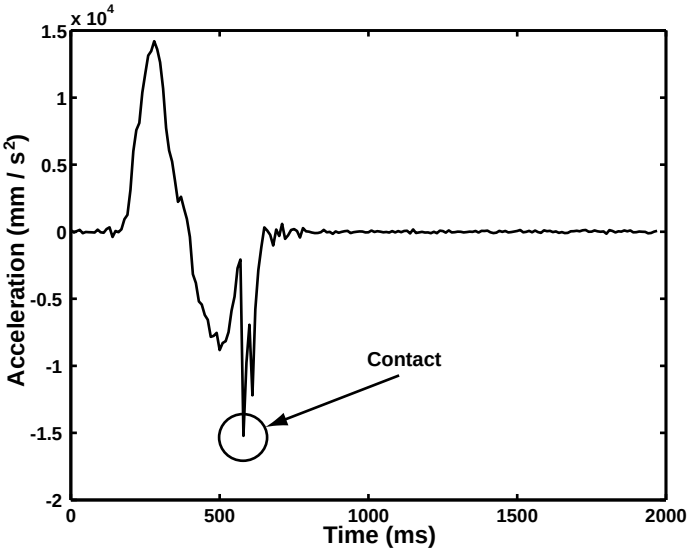
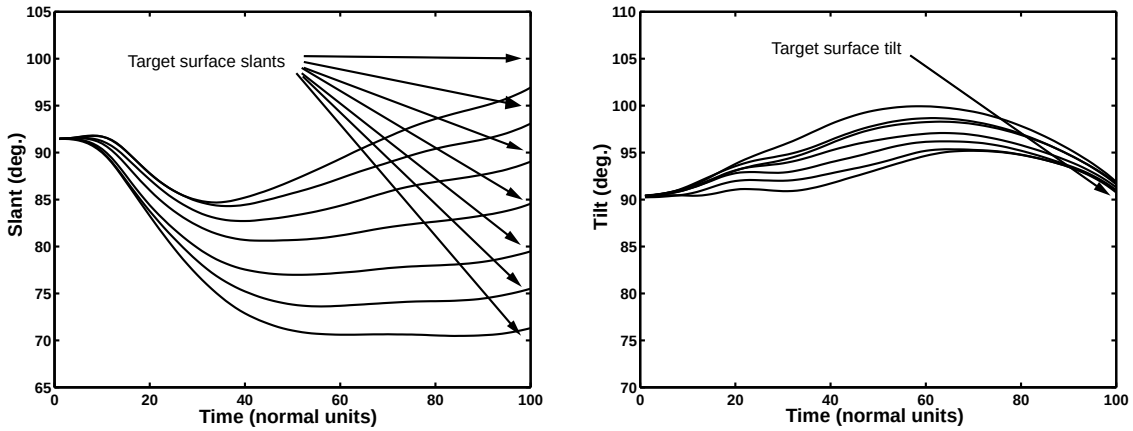
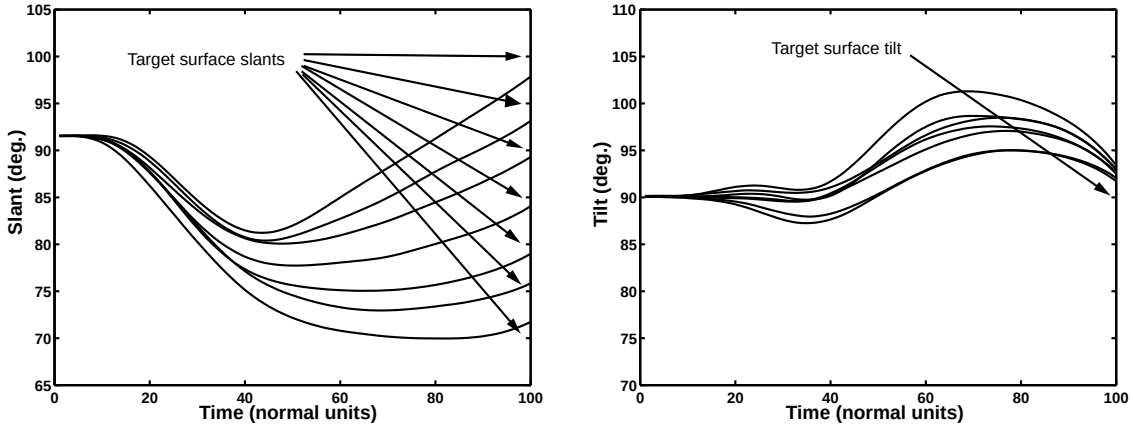

x 10^4
1.5
1
0.5
0
-0.5
-1
-1.5
-2
Acceleration (mm / s^2)
0
500
1000
1500
2000
Time (ms)
Contact

Figure 4:

Subject LES



Subject MDY



Subject MEL

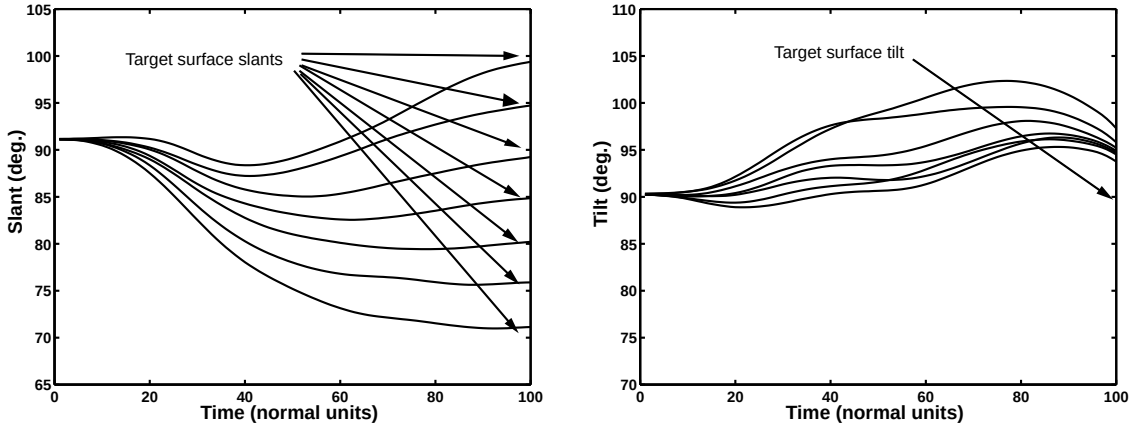
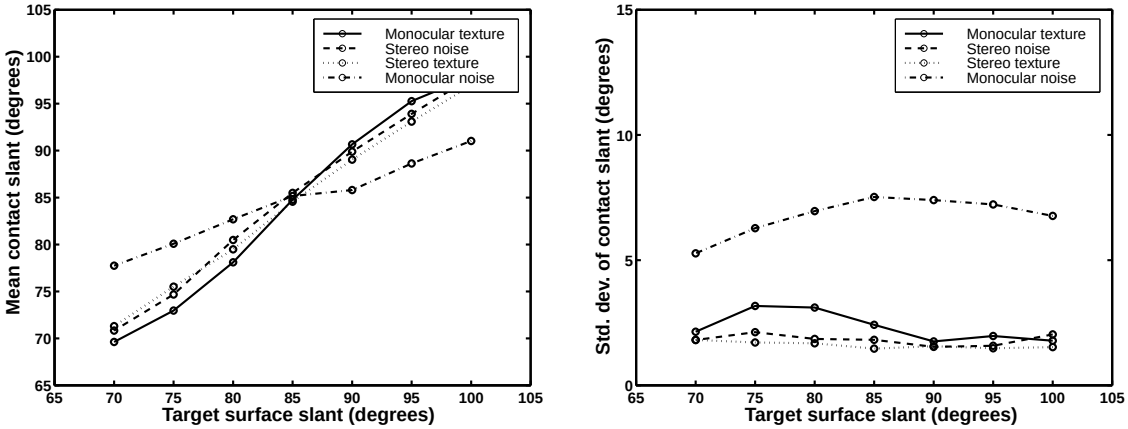
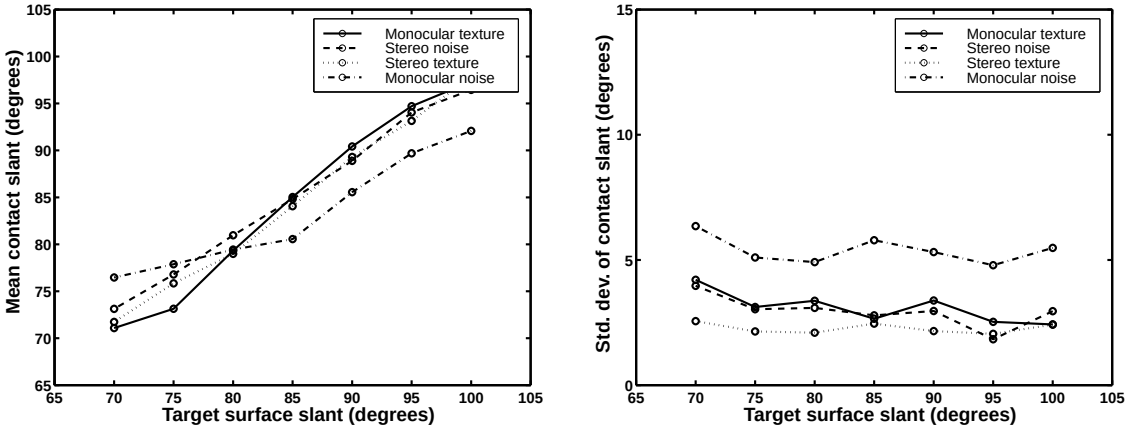


Figure 5:

Subject LES



Subject MDY



Subject MEL

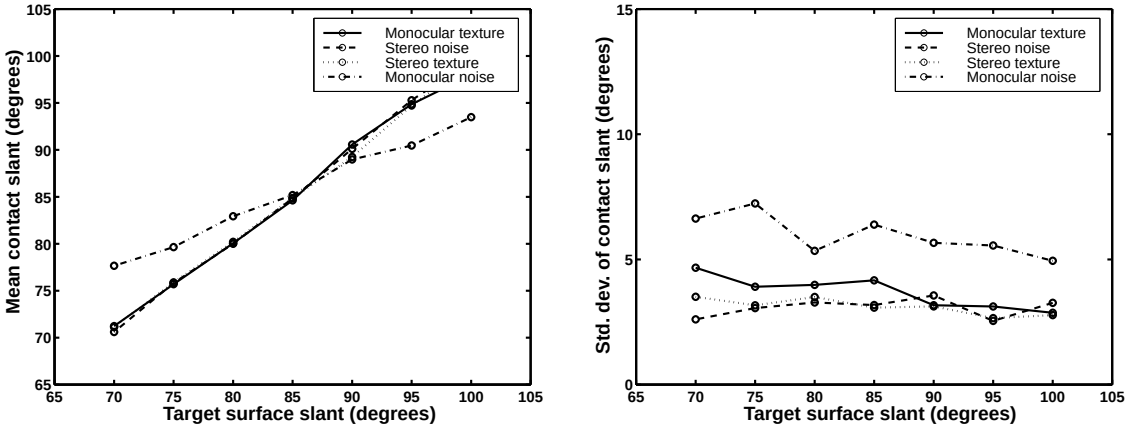
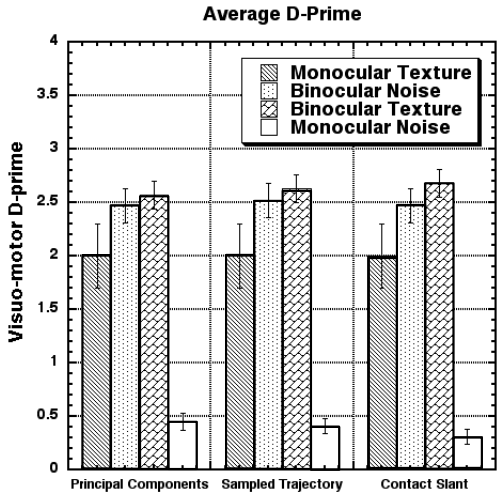
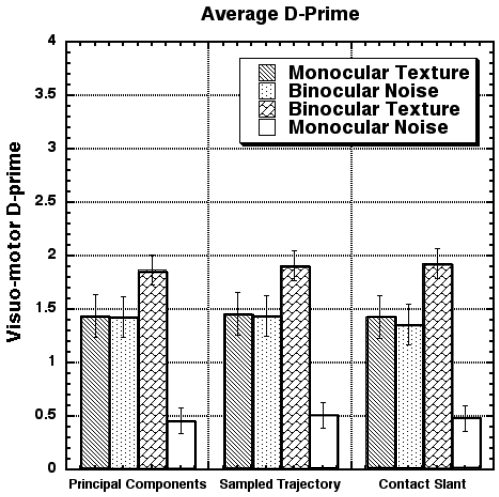


Figure 6:

Subject LES



Subject MDY



Subject MEL

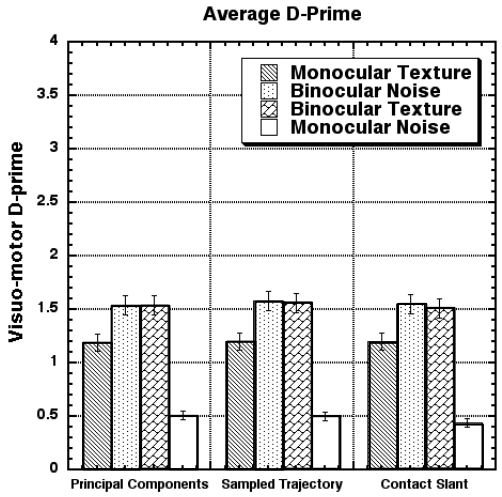
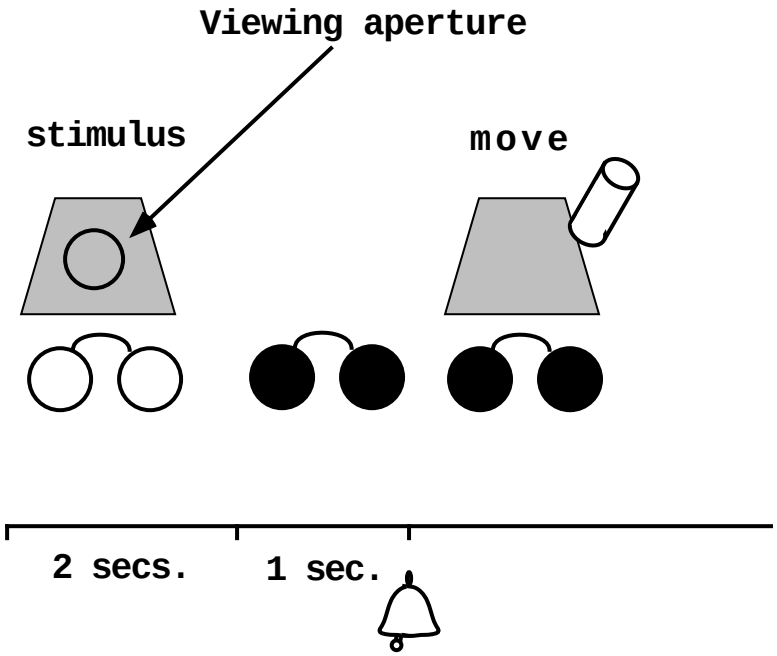
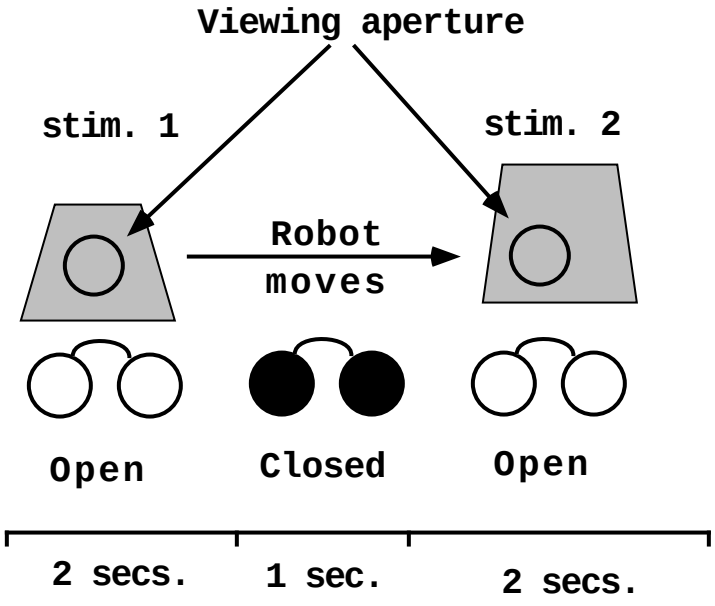


Figure 7:


Viewing aperture
stimulus
move
2 secs.
1 sec.

(a)


Viewing aperture
stim. 1
stim. 2
Robot
moves
Open
Closed
Open
2 secs.
1 sec.
2 secs.

(b)

Figure 8:

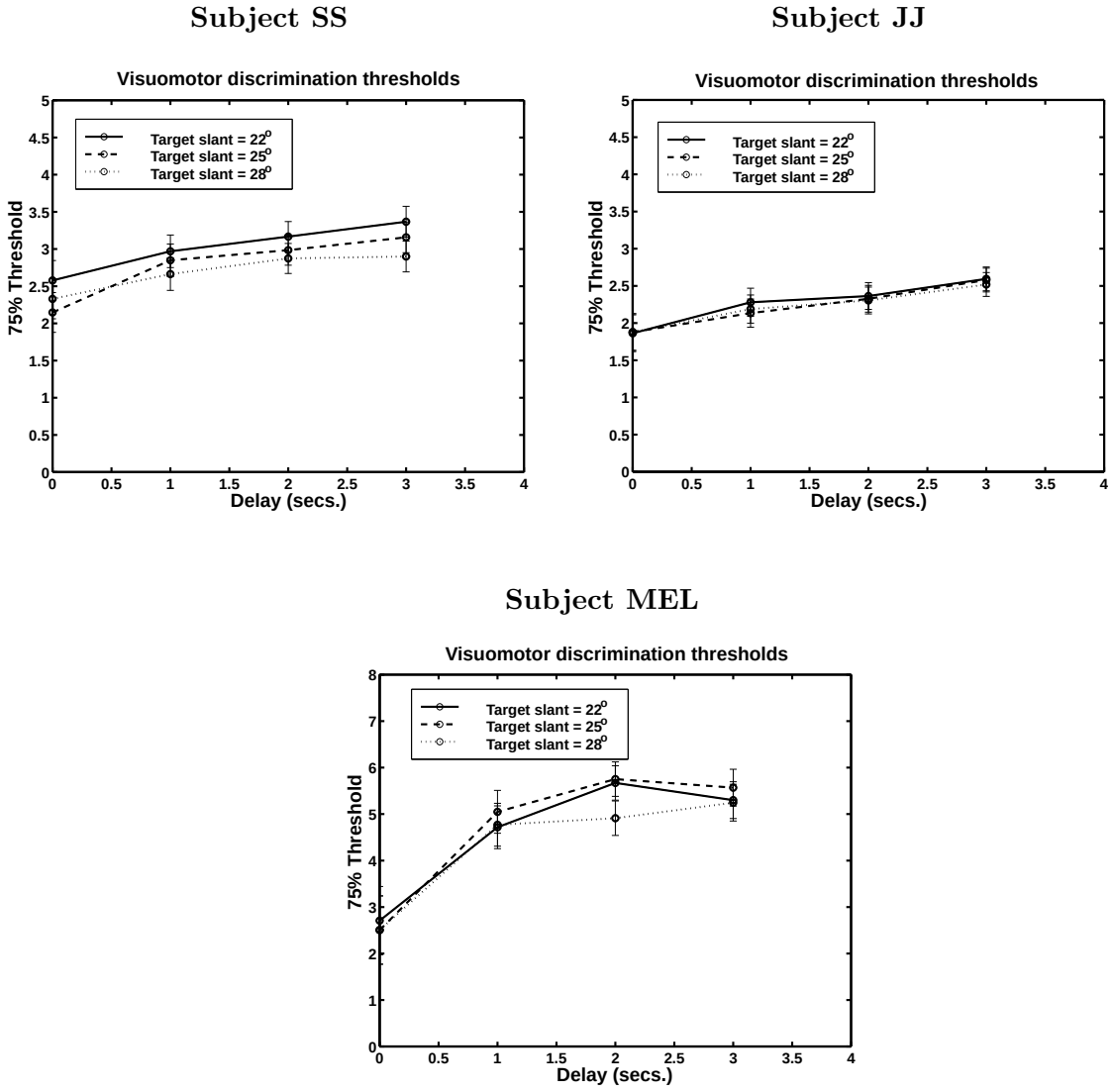

Subject SS
Visuomotor discrimination thresholds
Target slant = 22°
Target slant = 25°
Target slant = 28°
75% Threshold
Delay (secs.)
Subject JJ
Visuomotor discrimination thresholds
Target slant = 22°
Target slant = 25°
Target slant = 28°
75% Threshold
Delay (secs.)
Subject MEL
Visuomotor discrimination thresholds
Target slant = 22°
Target slant = 25°
Target slant = 28°
75% Threshold
Delay (secs.)

Figure 9:

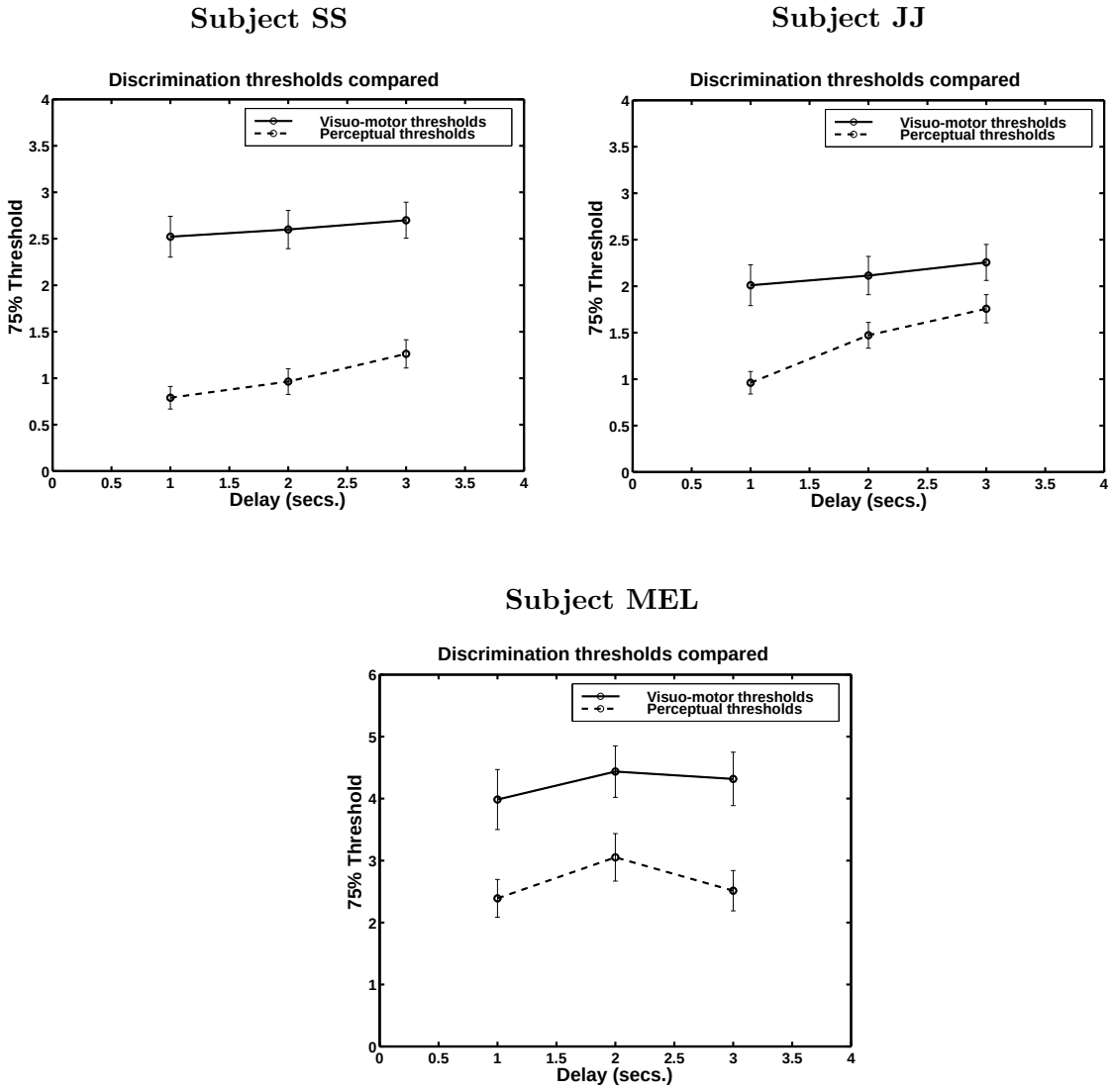

Subject SS
Discrimination thresholds compared
Visuo-motor thresholds
Perceptual thresholds
75% Threshold
Delay (secs.)
Subject JJ
Discrimination thresholds compared
Visuo-motor thresholds
Perceptual thresholds
75% Threshold
Delay (secs.)
Subject MEL
Discrimination thresholds compared
Visuo-motor thresholds
Perceptual thresholds
75% Threshold
Delay (secs.)

Figure 10:

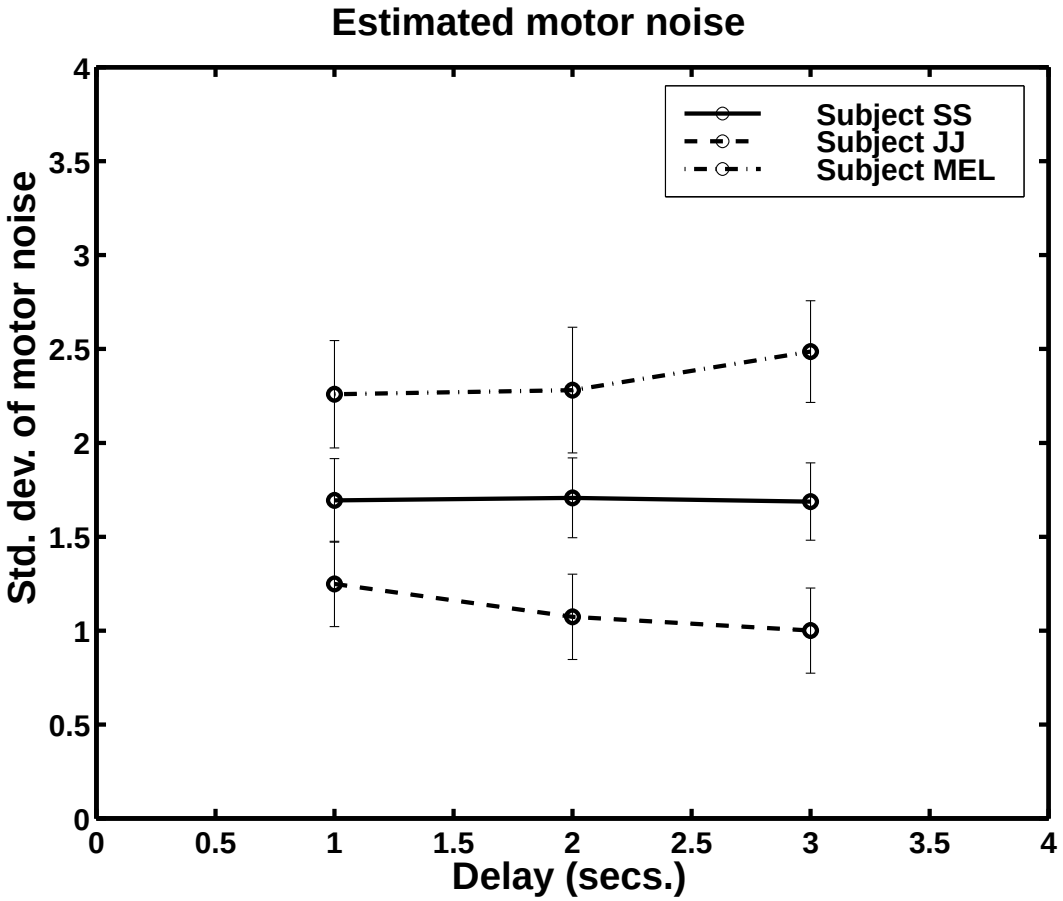
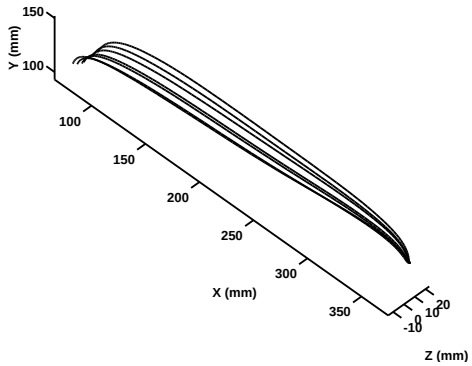
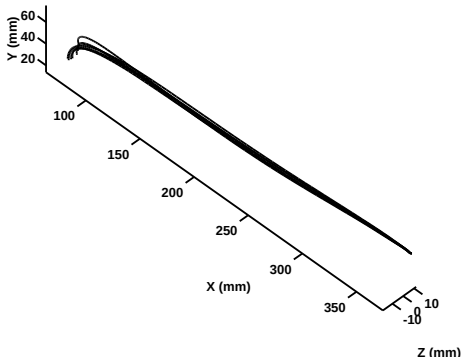

Estimated motor noise
Subject SS
Subject JJ
Subject MEL
Std. dev. of motor noise
Delay (secs.)
0
0.5
1
1.5
2
2.5
3
3.5
4

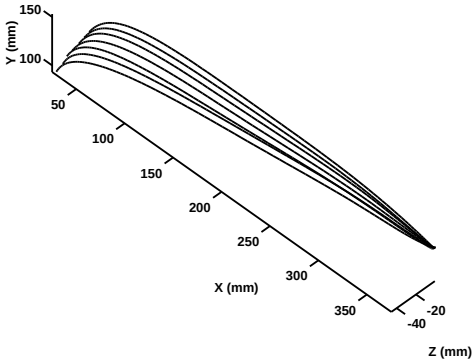
Figure 11:

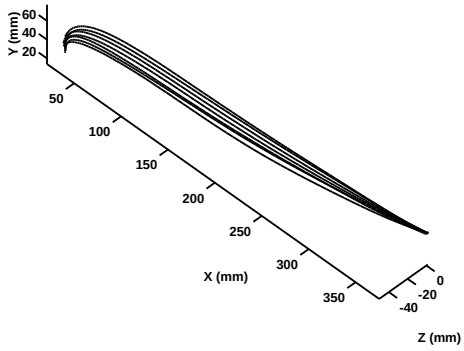
Subject LES



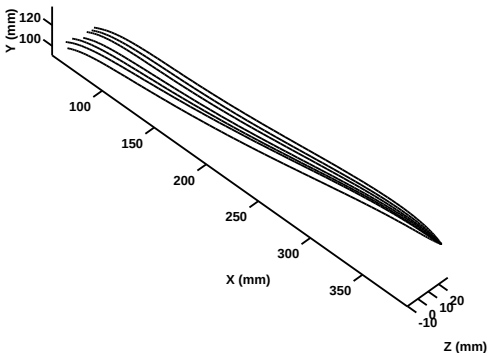


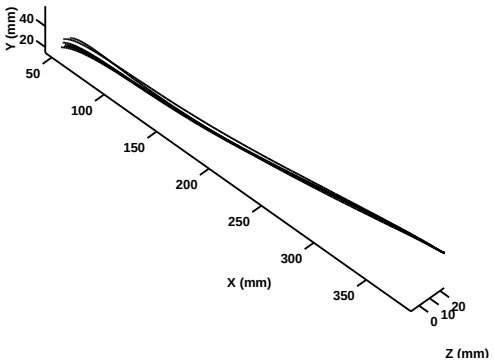
Subject MDY





Subject MEL





Trajectories of center of cylinder **Trajectories of bottom of cylinder**

Figure 12:

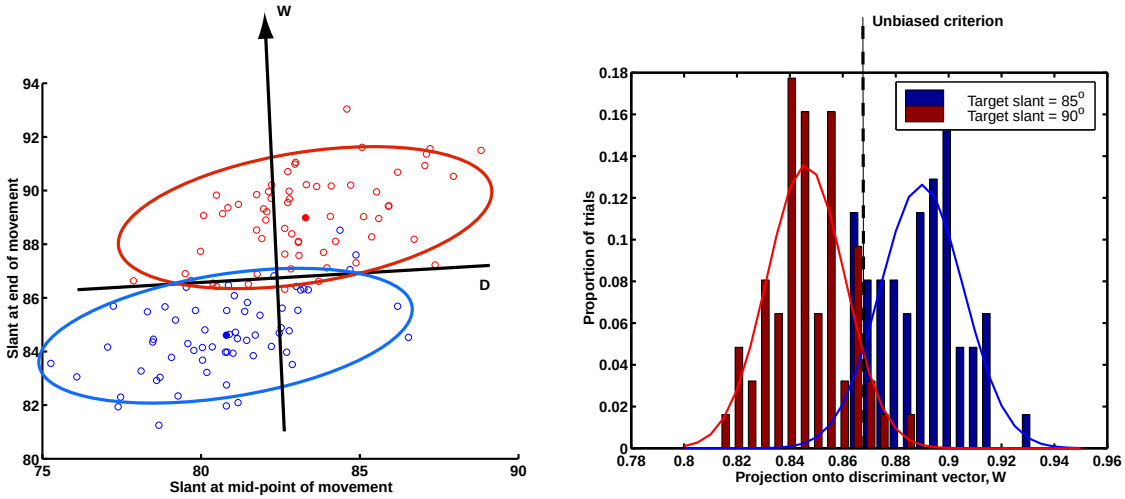

W
D
Slant at end of movement
Slant at mid-point of movement
94
92
90
88
86
84
82
80
75
80
85
90
Unbiased criterion
Target slant = 85°
Target slant = 90°
Proportion of trials
Projection onto discriminant vector, W
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0.78
0.8
0.82
0.84
0.86
0.88
0.9
0.92
0.94
0.96

Figure 13: