

RESEARCH NOTE

J.J. Marotta · A. Kruyer · M.A. Goodale

The role of head movements in the control of manual prehension

Received: 22 December 1997 / Accepted: 4 February 1998

Abstract Binocular information has been shown to be important for the programming and control of reaching and grasping. But even without binocular vision, people are still able to reach out and pick up objects accurately – albeit less efficiently. As part of a continuing investigation into the role that monocular cues play in visuomotor control, we examined whether or not subjects could use retinal motion information, derived from movements of the head, to help program and control reaching and grasping movements when binocular vision is denied. Subjects reached out in the dark to an illuminated sphere presented at eye-level, under both monocular and binocular viewing conditions with their head either free to move or restrained. When subjects viewed the display monocularly, they showed fewer on-line corrections when they were allowed to move their head. No such difference in performance was seen when subjects were allowed a full binocular view. This study, combined with previous work with neurological patients, confirms that the visuomotor system “prefers” to use binocular vision but, when this information is not available, can fall back on other monocular depth cues, such as information produced by motion of the object (and the scene) on the retina, to help program and control manual prehension.

Key words Monocular · Binocular · Manual prehension · Retinal motion · Head movements

Introduction

When we reach out to pick up an object, our motor system must have access to information about the exact location of the object in egocentric space, as well as information

about its actual size. The absolute distance of the object (within a particular frame of reference) must be computed in order to program both the trajectory of the reach and the aperture of the grasp. One reliable source of absolute distance information for the calibration of reaching and grasping is binocular vision (Servos et al. 1992; Marotta et al. 1995b; Dijkerman et al. 1996; Jackson et al. 1997). Servos et al. (1992) demonstrated that grasping movements made under monocular viewing were less “efficient” than those performed under binocular viewing conditions, achieving lower peak velocities and showing prolonged periods of deceleration during the closing phase of the grasp. One of the most striking differences between binocular and monocular reaches is the number of on-line adjustments made by the subject during the execution of the movement, particularly in the closing phases of the movement (Kruyer et al. 1997; Marotta et al. 1997). These adjustments appear to arise as a consequence of errors in the subject’s initial estimate of the target’s distance. But even though denying binocular vision has significant effects on performance, the subjects were still able to reach out and pick up objects reasonably well. They must therefore be relying on monocular depth cues – but which ones?

One source of monocular information about distance (and thus object size) that is potentially useful, is the motion of the object (and the scene) on the retina – particularly motion generated by head movements. It is possible to generate an accurate estimate of absolute distance from the movement of a point on the retina if one “knows” the magnitude of one’s head movement (and/or the movements of the eyes in the orbit as one fixates that point). There have been several motion perception studies that have found that subjects make more accurate judgements of depth when their head is free to move (Ferris 1972; Biguer et al. 1984).

Our laboratory has been involved in several investigations into the role of head movements in the control of action. For example, Mongolian gerbils execute a series of vertical head movements prior to jumping a gap, the amplitude and velocity of which were strongly correlated

J.J. Marotta · A. Kruyer · M.A. Goodale (✉)
Vision and Motor Control Laboratory,
Department of Psychology,
The University of Western Ontario, London,
Ontario, Canada N6A-5C2
e-mail: goodale@uwo.ca,
Tel.: +1-519-661-2070, Fax: +1-519-661-3961

with gap distance. These gerbils were found to bob their head more often before jumping a gap when using monocular vision than when using binocular vision (Ellard et al. 1984). This suggests that the gerbils were employing retinal motion, derived from self-generated head movements, to judge distance, and that when animals were tested using monocular vision they generated more head bobs to better utilize retinal motion.

As a first step in investigating the use of retinal motion cues in the programming and control of reaching and grasping in human subjects, we examined grasping in individuals who had been deprived of binocular vision for a long time – namely, individuals who had lost an eye. We found that such individuals made larger and faster head movements during the execution of their reaching movement, and that the tendency to make these movements increased as a function of time since enucleation (Marotta et al. 1995a). Subjects with normal vision, however, did not appear to use this learned strategy when one eye was temporarily covered (Marotta et al. 1995b). Although subjects with normal vision do not generate larger head movements when wearing an eye-patch, they still move and show head movements. The aim of the current study was to determine whether these movements generate useful retinal motion cues to depth under monocular viewing conditions.

Method

The experiment was carried out at the University of Western Ontario in compliance with the Social Sciences and Humanities Research Council (Canada) Guidelines (1981).

Subjects

Eight right-handed subjects (4 males, 4 females; mean age 24.5 years) with normal or corrected-to-normal vision participated in the experiment, for which they were paid. Subjects were strongly right-handed, as determined by a modified version of the Edinburgh handedness inventory (Oldfield 1971). All subjects had stereoscopic vision in the normal range with assessed stereoacuity of 40 seconds of arc or better as determined by the Randot Stereotest (Stereo Optical, Chicago, Ill.).

Apparatus

In this study, we utilized a cue-deprived test environment developed in our laboratory (Kruyer et al. 1997). In this test environment a wide range of depth cues can be removed by presenting lit spheres in three-dimensional grasping space to subjects who are in the dark and viewing the scene with only one eye. By systematically reintroducing depth cues into this severely cue-deprived environment, we are able to examine the contribution of individual depth cues to the programming and control of manual prehension.

Three sizes of Styrofoam sphere (6.25, 7.5 and 10 cm in diameter) were presented at one of three different distances (20, 30, 40 cm from the subject), on a rod that was positioned at a height of 127 cm in a matte black vertical presentation board (183×120 cm). The centre of each sphere contained four light-emitting diodes controlled by computer. The voltage sent to each sphere was controlled so that the surface luminance levels for each size of sphere were equivalent (10 candelas/m² as measured by a light meter). (It should be noted that perfect spheres would offer no retinal disparity cues to depth or distance. Of course, the Styrofoam spheres we used were not per-

fect and in addition they had a textured surface. Moreover, even with perfect spheres, other binocular cues such as convergence could provide depth information.)

Subjects sat in a height-adjustable chair at a table that was positioned in front of the presentation board. A start button and a bite bar were mounted on the table. The bite bar was used to restrict subjects' head movements but also allowed for translational and rotational head movements when a central rod was removed and replaced with a flexible spring. Subjects wore PLATO spectacles (Translucent Technologies, Toronto, Canada) throughout the testing sessions. These liquid-crystal shutter spectacles permitted monocular or binocular viewing and when both shutters were closed prevented subjects from viewing the spheres being put into position. Subjects also wore earphones that emitted white noise between trials, to prevent them from using any audible cues from the sphere being put into position. The room was dark and subjects reached for the sphere, which remained lit for 2.5 s.

Three infrared light-emitting diodes (IREDs) 8 mm in diameter were attached to the subject's right hand. One IRED was mounted on the end of an aluminum extension tab 3 cm long attached to a watchband worn at the radius at the wrist; a second IRED was positioned at the end of a 2 cm aluminum tab attached to the ulnar border of the thumbnail. A final IRED was placed on the distal portion of the index fingertip. The aluminum extensions were used to allow the camera system an optimal view of each IRED.

Two additional IREDs were attached to the earphones the subjects wore, in order to measure the amplitude and direction of the subjects' head movements. Unfortunately, equipment failure prevented the collection of useful head movement data. While subjects were free to move their head when the bite bar was fitted with a flexible spring instead of the rigid bar, which prevented head movement (see below), the exact nature of these movements cannot be reported. Nevertheless, it is well known that the action of reaching towards a visual goal involves a combination of movements. In normal conditions, subjects will first orient their gaze, then their head, and finally their arm in the proper direction (Biguer et al. 1982, 1984). Moreover, movements of the head and torso are required to deal with the forces generated by movements of the limb and to extend the range of the limb.

The IREDs were monitored by an infrared-sensitive camera system (Optotrak) positioned approximately 2 m from the subject. The three-dimensional coordinates of the IREDs were stored by the Optotrak's data acquisition unit and later filtered off line (with a low-pass second-order Butterworth filter with a 7-Hz cut-off).

Procedure

At the beginning of the test session, subjects were given the handedness questionnaire and tested for eye dominance (viewing preference). The seat was adjusted so that the sphere height (127 cm high) would be at eye-level. Subjects were instructed to place the tips of their index finger and thumb of their right hand on the start button. Subjects were further instructed that as soon as they saw the lit sphere, they were to reach out quickly and accurately but as "naturally" as possible and grab hold of the sphere with their whole hand. Subjects were asked to hold onto the sphere until they heard a tone signalling the end of the trial. The experimenter initiated the start of a trial by signalling the computer to simultaneously open the shutter-spectacles and illuminate the spheres for a period of 2.5 s.

Subjects were administered four separate blocks of 36 experimental trials, each consisting of four instances of each of the three sphere sizes at the three presentation distances. Each subject performed two blocks of trials using binocular vision and two blocks of trials using monocular vision. Under each viewing condition the subject performed one block of trials with their head free to move (using the bite bar fitted with a flexible spring) and one with their head restrained (using the bite bar fitted with a rigid rod). The order of presentation of these blocks of trials was counterbalanced across subjects. Trial presentation was random and each testing block was preceded by five practice trials. The testing session lasted approximately 90 min.

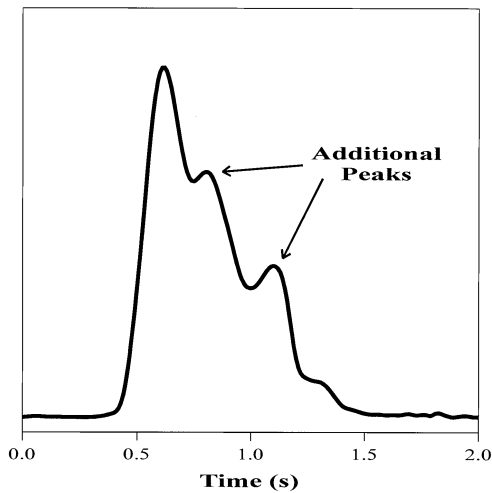


Fig. 1 Representation of additional on-line correction peaks

Dependent measures

If subjects program their grasp on the basis of an incorrect estimate of target distance, then they will have to make an on-line correction in order to acquire the target. If they overestimate the distance, then they will sometimes collide with the target. If they underestimate the distance, then they will have to adjust the trajectory (and grasp) during the closing phase in order to make successful contact. These latter movements in particular have been observed in a number of different experiments in our laboratory where cues to distance (and thus object size) were either ambiguous or absent (Marotta and Goodale 1996; Kruyer et al. 1997; Marotta et al. 1997). The methods we have developed for quantifying these adjustments are outlined below.

On-line velocity corrections

In a typical reach, subjects accelerate smoothly to a peak (or maximum) velocity and then decelerate as their hand approaches the object to be grasped. Occasionally, however, subjects show on-line adjustments in the reach, which are evident as additional "peaks" in the velocity profile (Fig. 1). A peak is defined as a sharp increase in velocity followed by a decrease. The number of these additional velocity peaks was recorded for each trial.

On-line aperture corrections

In a typical grasp, subjects open their hand smoothly to a peak (or maximum aperture) and close it as their hand approaches the object. As with their reach, occasionally subjects adjust their grasp on line. Again these adjustments are reflected as additional peaks in the aperture profile. The number of these additional aperture peaks was recorded for each trial.

These measures provide a more accurate representation of the "efficiency" of a manual prehension movement than do more "traditional" kinematic measures (e.g., maximum velocity, maximum grip aperture). In an earlier study on the relative contributions of binocular vision and pictorial cues, we showed that on-line corrections revealed differences in performance that were not evident in the pattern of traditional kinematic measures (Marotta et al. 1997).

Results

For each of the subjects, mean values of the two dependent measures were calculated for each viewing condition and

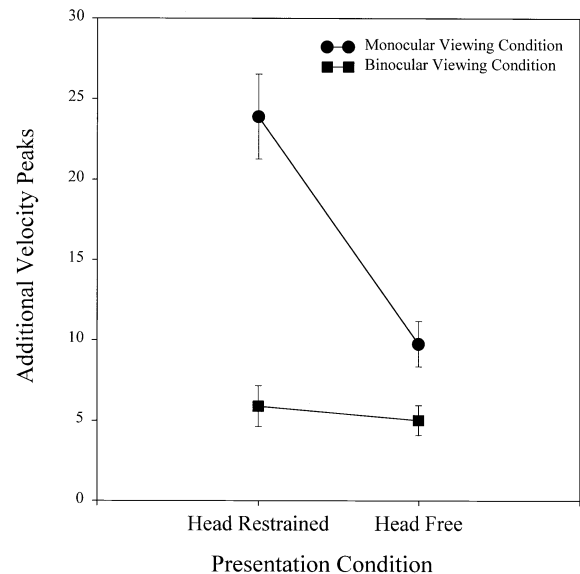


Fig. 2 The effects of presentation array and viewing condition on additional velocity peaks (*error bars SEMs, filled circles monocular viewing condition, filled squares binocular viewing condition*)

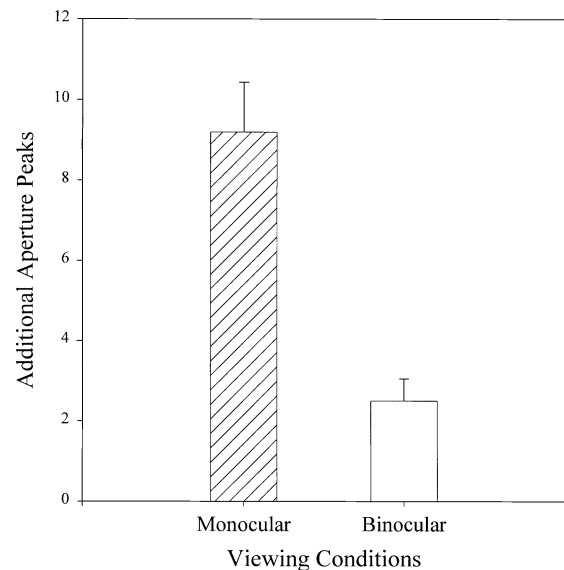


Fig. 3 The effects of viewing condition on additional aperture peaks (*error bars SEMs, hatched column monocular viewing condition, empty column binocular viewing condition*)

each head movement condition. (Equipment failure resulted in some loss of data, but this constituted less than 4.5% of the trials.) The values were entered into separate 2x2 (viewing condition x head movement condition) repeated-measures analyses of variance. All test of significance were based upon an alpha level of 0.05. Post-hoc Newman-Keuls analysis was performed where necessary.

As was seen in previous studies (Kruyer et al. 1997; Marotta et al. 1997), under monocular viewing conditions subjects produced more on-line corrections in their reaching and grasping movements than they did when binocular vision was available. As can be seen in Fig. 2, monoc-

ular reaches exhibited significantly more peaks ($F(1,7)=32.88$, $P<0.001$) in their velocity profiles than did binocular reaches.

When the spheres were presented in the head-restrained condition, subjects showed significantly more peaks ($P<0.01$) in their velocity profiles than when they were allowed to move their heads freely – but only when they viewed the spheres with one eye covered. Under binocular conditions, the number of additional velocity peaks did not vary as a function of head movements ($P>0.05$). This interaction between the head movement conditions and the viewing conditions is evident in Fig. 2 for the additional velocity peaks ($F(1,7)=73.64$, $P<0.0005$).

As can be seen in Fig. 3, the only effect on grip aperture was an increase in the number of peaks in the aperture profile during monocular grasping movements when compared with binocular grasping movements ($F(1,7)=33.63$, $P<0.001$).

Discussion

As in previous studies (Servos et al. 1992; Marotta et al. 1995b), restricting vision to only one eye clearly disrupted reach performance, which resulted in an increased number of on-line corrections. In addition, when head movements were also restricted, thereby preventing the use of retinal motion cues to distance as well as binocular cues, subjects produced even more on-line corrections in their reaching and grasping movements, corrections that were evident as additional peaks in their velocity profile. In contrast, restricting head movements under binocular viewing conditions had no effect on performance. These results suggest that retinal motion is an important source of depth information that can be used for the programming and control of reaching and grasping movements – but only when binocular cues are not available. Finally, it is worth emphasizing that the difference in performance between head-fixed and head-free conditions cannot be attributed to any biomechanical constraint that might have been present in the head-fixed condition. If subjects simply could not reach out as well when their head was fixed, then they should have performed equally poorly under both monocular and binocular viewing conditions.

The presence of on-line corrections can be used as an indication that subjects were underestimating the distance of the spheres on certain trials. In other words, if the subjects decelerated and began to close their grasp too early, they would be forced to adjust their trajectory to acquire the target. Of course, when they overestimated the distance of the target, they would encounter the target sooner than anticipated, would be unable to decelerate properly, and would collide with it with some force. Such collisions were occasionally observed and, when they occurred, they were obvious to both the subject and the experimenter. Unfortunately, such collisions were difficult to measure unambiguously. Nevertheless, like the adjustments that occurred when subjects presumably underestimated the

sphere's distance, these occasional collisions occurred only during the monocular head-fixed condition.

One paradox in the results remains to be discussed. Although the calibration of the reach was affected by the removal of binocular cues and by restricting head movements, the calibration of grip aperture was largely unaffected by restricting head movements. In other words, even when their heads were restricted, subjects were able to use information that was left in the visual scene to calibrate their grasp. It seems that while the subjects are able to use retinal motion information derived from head movements to compute distance for transporting the limb to the correct location, they do not rely on this information to calibrate their grasp to the true size of the object. Because only three sphere sizes were used, subjects could simply have learned to use three different motor routines for scaling their grasp, each of which would be evoked by the appropriate sphere size. Such a strategy would not require that the sphere's distance be accurately computed, just reasonably "ball-parked" so that it could be categorized as small, medium-sized, or large. In other words, the actual retinal image size would have to be only roughly calibrated for distance in order for the visuomotor system to settle on one of the three possible sizes. But this of course is just speculation. Another possibility is that the visuomotor systems mediating grasp simply do not access information about distance provided by retinal motion. This possibility was first discussed in a recent study in which patients with visual form agnosia were tested on a grasping task under monocular and binocular viewing conditions (Marotta et al. 1997). In that experiment, removing binocular cues affected grasp scaling but not distance. At the time, we suggested (but did not demonstrate) that retinal motion may be used to program and control the reaching movement itself, but not the grasp.

This study provides clear evidence that intact subjects wearing an eye-patch are able to use the available retinal motion information generated from head movements to help program and control manual prehension. Nevertheless, as we observed in earlier studies (Marotta et al. 1995a, b), subjects seem reluctant to make better use of this information by actively generating larger head movements – in contrast to enucleated patients who are used to living in a monocular world. These results, combined with the neuropsychological work by Marotta et al. (1997), also suggest that under normal viewing conditions the visuomotor system "prefers" to use binocular information but can fall back on retinal motion information when binocular vision is not available.

Acknowledgements This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) grant to M.A.G., a postgraduate scholarship from the Ontario Graduate Scholarship Program (OGS) to J.J.M. and a postgraduate scholarship from NSERC to A.K.. We would like to thank J. Orphan and D. Pulham for their technical support. We would also like to thank the subjects for the cooperation they showed throughout the testing.

References

- Biguer B, Jeannerod M, Prablanc C (1982) The coordination of eye, head, and arm movements during reaching at a single visual target. *Exp Brain Res* 46:301–304
- Biguer B, Jeannerod M, Prablanc C (1984) The contribution of coordinated eye and head movements in hand pointing accuracy. *Exp Brain Res* 55:462–469
- Dijkerman HC, Milner AD, Carey DP (1996) The perception and prehension of objects oriented in the depth plane. 1. Effects of visual form agnosia. *Exp Brain Res* 112:442–451
- Ellard CG, Goodale MA, Timney B (1984) Distance estimation in the Mongolian gerbil: the role of dynamic depth cue. *Behav Brain Res* 14:29–39
- Ferris SH (1972) Motion parallax and absolute distance. *J Exp Psychol* 95:258–263
- Huynh H, Feldt LS (1976) Estimation of the Box correlation for degrees of freedom from sample data in randomized block and split-plot designs. *J Educ Stat* 1:69–82
- Jackson SR, Jones CA, Newport R, Pritchard C (1997) A kinematic analysis of goal-directed prehension movements executed under binocular, monocular, and memory-guided viewing conditions. *Vis Cogn* 4:113–142
- Krayer A, Marotta JJ, Goodale MA (1997) Scene-based binocular cues improve grip trajectory accuracy. *Invest Ophthalmol Vis Sci* 38:988
- Marotta JJ, Goodale MA (1996) Height and distance in the visual field: calibrating a monocularly guided reach. *Soc Neurosci Abstr* 22:885
- Marotta JJ, Goodale MA (1997) Restricting head movements hampers reaching under monocular but not binocular viewing conditions. *Soc Neurosci Abstr* 23:885
- Marotta JJ, Perrot TS, Nicolle D, Goodale MA (1995a) The development of adaptive head movements following enucleation. *Eye* 9:333–336
- Marotta JJ, Perrot TS, Nicolle D, Servos P, Goodale MA (1995b) Adapting to monocular vision: grasping with one eye. *Exp Brain Res* 104:107–114
- Marotta JJ, Behrmann M, Goodale MA (1997) The removal of binocular cues disrupts the calibration of grasping in patients with visual form agnosia. *Exp Brain Res* 116:113–121
- Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh Inventory. *Neuropsychologia* 9:97–112
- Servos P, Goodale MA, Jakobson LS (1992) The role of binocular vision in prehension: a kinematic analysis. *Vision Res* 32: 1513–1521