Bangor University

DOCTOR OF PHILOSOPHY

Depth-cue integration, and the role of uncertainty in grasping

Keefe, Bruce D.

Award date:
2010

Awarding institution:
Bangor University

Depth-cue integration, and the role of uncertainty in grasping

Bruce D. Keefe, BSc, MSc

This thesis is submitted in partial fulfilment of the requirement for the degree of Doctor of Philosophy, completed in the school of Psychology, Bangor University.

Acknowledgements

First and foremost I would like to thank Simon for the opportunity to undertake this PhD. Without his patience, guidance and expertise it would not have reached fruition. He has helped to develop both my skills and thinking as a scientist.

Thanks to both Kevin and Chie for being able to share an office with me. Kevin's addition to the lab early in my PhD proved to be invaluable. He has answered countless questions and has helped me to develop in all areas of my PhD.

Thanks to Raych Clayton and Matt Elsby for help in data collection in Chapters 4 and 5 respectively.

Thank you also to Simon, Kevin, Chris Saville, and Wendy Adams for their useful comments on sections of this Thesis.

Contents Page

Abstract

The aim of this thesis is to use the cue-integration framework as a normative model to study the role of binocular depth cues in grasping, and to explore how the visuomotor system decreases uncertainty in visual information. We explored whether the availability of different cue types, or visual uncertainty *per se* determines the margin-for-error in programmed grip apertures. Results showed that both monocular and binocular cues contributed to grasping performance, and that performance was improved when both cues were available compared to when either cue was available alone. We also examined whether cue type or the precision of feedback determines online grasping performance. Our findings suggested that the contribution of feedback from binocular and monocular cues was determined by the precision of information in a particular situation. Taken together, the above results suggested that the idea of a binocular specialism for grasping is incorrect, and that a cue-integration account better explains the data. We extended this idea to examine whether learnt, familiar size information is used alongside visual depth cues. The results suggested that familiar size is also integrated with binocular and monocular depth cues. The above experiments confirmed that the visuomotor system opens the grasp wider when visual uncertainty increases in order to increase the margin-of-error in the movement. We examined this behaviour further by exploring different movements, for which a different response to perceptual uncertainty is required. We found that the visuomotor system encodes both perceptual uncertainty, and the probability of making an error to programme grasping movements. Overall our results suggest that binocular information is not critical for the planning or online control of grasping. Rather, our findings suggest that all sources of information are integrated for grasp control,

minimising uncertainty in estimates of object properties, which allows for smaller margin-of-error in grasping movements.

Chapter 1: General Introduction

1.1. *Overview*

For centuries, man has held admiration for the exquisite design of the hand and the brain that governs it; as Aristotle noted, "The hand can become a claw, a fist, a horn or spear or sword or any other weapon or tool. It can be everything because it has the ability to grasp anything or hold anything". The versatility and precision of human grasping raises the question of how such a feat is achieved. Part of the answer lies in the visual information upon which grasping relies. Grasping has two main phases, planning and online control, that require visual information about the object and its position relative to the moving digits. Planning requires information concerning object properties such as size, shape, and distance in order to accurately programme the movement. Once the hand starts to move, online feedback is also available concerning the position of the digits relative to the object. Fine-tuning of the movement is achieved by correcting for planning errors and changes in the world using both feed-forward and feedback control loops to accurately guide the digits to the pre-selected contact points on the object (Bhushan and Shadmehr, 1999; Miall, Weir, Wolpert & Stein, 1993; Paulignan, Jeannerod, MacKenzie & Marteniuk, 1991; Paulignan, MacKenzie, Marteniuk & Jeannerod, 1991). Thus, accurate grasping performance relies on visual information that is able to specify the requirements of both the planning and online control of grasping. Accurate grasping could be achieved both by explicitly encoding the distance and size of the object, or the target locations at which the digits meet the object (Smeets & Brenner, 1999). Although it remains to be determined which of these two possibilities is encoded, it is likely that either method requires the computation of 3-D properties. The extraction of such information is clearly achieved given the seamless manner in which we can reach towards and pick up objects, such as our morning cup of coffee. In the next section we consider how such information may be recovered by the visual system.

1.2. *Combining multiple sources of information*

Different visual sources of information or 'depth cues' are available simultaneously that can, in principle, be used by the visual system to estimate information about the world, including its metric structure. For example, textured surfaces such as the bark on a tree, give rise to a stochastically uniform texture gradient, that can be used to recover the relative distance to each texture element, and therefore the orientation of the surface. Also, the fact that we have two eyes separated laterally by about 6 cm produces differences or 'disparities' in the two retinal images, which, in combination with an estimate of viewing distance, can be used to recover metric depth (Howard & Rogers, 2002). Overall, there are over a dozen depth cues that simultaneously provide the visual system with information. However, each source of information is noisy/variable and so provides estimates of the properties of the world that will change over space and time (Hillis, Watt, Landy, & Banks, 2004; Knill & Saunders, 2003). What then is the optimal use of these multiple, noisy, estimates from different visual depth cues?

One possibility is that the most reliable cue be relied upon to form the percept, effectively vetoing other cues (Bülthoff & Mallot, 1998, Turner, Braunstein, & Anderson 1997). However, such a strategy would be sub-optimal, as information from other cues would be discarded. Further, a solitary source of information will not always provide the best estimate of object properties; for example, changes in viewing geometry may create conditions in which it is uninformative (Hillis et al., 2004).

Another solution proposes that the visual system integrates information from all cues (Landy, Maloney & Johnston, 1995). Determining how the visual system uses different cues requires us to tease apart the underlying mechanisms. This is difficult to achieve under natural viewing conditions because the different estimates are highly correlated. By independently manipulating different cues so that they come into conflict, and observing how these conflicts affect visual estimates, researchers have begun to determine empirically whether the visual system integrates different sensory signals.

1.3. *Integrating information from multiple depth cues*

Maximum-likelihood estimation (MLE) provides a statistical framework to model the combination of different sensory sources and has allowed the mechanisms of cue combination to be tested empirically. Consider the case of estimating the size of an object specified by binocular disparity and monocular texture (perspective) cues. Size estimates from each cue ($\hat{S}_d$, $\hat{S}_t$) are noisy, and the most likely size can be computed from a weighted sum of estimates from each cue (assuming each estimate has independent, Gaussian noise and that all sizes in the world are equally likely)

$$\hat{S} = w_d\hat{S}_d + w_t\hat{S}_t \tag{1}$$

where

$$w_d = \frac{1/\sigma_d^2}{1/\sigma_d^2 + 1/\sigma_t^2} = 1 - w_t \tag{2}$$

In this model, the weight given to each estimate (w_i, normalised to sum to 1) is proportional to the reciprocal of the variance (σ_i^2) of the estimate. Thus, less variable (more precise or reliable) cues are given more weight (Backus & Banks, 1999; Ernst & Banks, 2002; Ghahramani, Wolpert, & Jordan, 1997; Jacobs, 1999; Oruç, Maloney, & Landy, 2003). The variance of the combined estimate $\hat{S}$ is

$$\sigma^2 = \frac{\sigma_d^2 \sigma_t^2}{\sigma_d^2 + \sigma_t^2} \tag{3}$$

The variance of the combined estimate is always lower than the variance of any single-cue estimate. So, theoretically, a key advantage of combining information from multiple depth cues is that the 3-D properties of objects can be estimated with greater precision than by relying on any one cue alone (Clarke & Yuille, 1990; Knill & Pouget, 2004; Landy, Maloney, Johnston, & Young, 1995; Oruç et al., 2003; Yuille & Bülthoff, 1996). Indeed, cue-integration of this kind is statistically optimal, in the sense that it gives the minimum variance unbiased estimate of object properties (Ghahramani et al., 1997). This cue-integration model has been shown to give a good, quantitative account of the contribution of different signals to a variety of perceptual tasks, both within and across sensory modalities (e.g., Alais & Burr, 2004; Ernst & Banks, 2002; Gepshtein & Banks, 2003; Hillis, et al., 2004; Jacobs, 1999; Knill & Saunders, 2003; Landy & Kojima, 2001).

Cue-integration provides an optimal solution to the problem of having multiple estimates. This is useful because the reliability of information changes due to both the sensory apparatus and 'geometrical' factors. For example, the reliability of information from binocular disparity falls off with increases in viewing distance as the disparities in the projected retinal images become increasingly small (Hillis et al., 2004; Howard & Rogers, 2002). And, when judging a slanted surface from texture, the reliability of information will be greater at higher slants than lower ones because the same change in slant will cause a greater change in the retinal image (Hillis et al., 2004; Knill, 1998; Knill & Saunders, 2003). Therefore, the system should ideally take account of moment-to-moment changes in cue reliability in a dynamic fashion. A good example of such behaviour has been shown by Hillis and colleagues who examined how binocular (disparity) and monocular (texture and perspective) are combined across variations in slant and distance (Hillis et al., 2004, see also Knill & Saunders, 2003). They first measured single-cue thresholds for stimuli defined only by binocular or monocular cues. This was achieved by using a two-interval, forced-choice (2IFC) task in which observes judged which of two sequentially presented slanted surfaces had the greater slant. Consistent with the predictions based on the viewing geometry of the task, monocular thresholds were shown to improve as the slant increased and binocular thresholds were found to improve as distance decreased. The single cue thresholds were next used to predict the improvement in thresholds that should be observed when both binocular and monocular cues were available together, and to predict the respective weights given to each cue in the two-cue stimulus. When both cues were available, the slant specified by disparity and texture was put into conflict during a 2IFC task to determine the empirical weights given to either cue. These were accurately predicted according to the reliability of the individual cues, suggesting that binocular and monocular cues were integrated in a weighted sum. Importantly two-cue discrimination performance improved by the predicted amount compared to when either cue was available in isolation, suggesting that different sources of information are integrated in a statistically optimal fashion.

These results demonstrate both the manner with which the reliability of information can change over space (i.e. disparity reliability falling off with increases in distance) and time, and that for perception the brain can integrate different visual depth cues in a statistically optimal fashion. To gain accurate estimates of the 3-D structure of the world therefore requires a dynamic system that can update its estimates of object properties both spatially and temporally, according to the reliability of information from individual cues. Indeed the findings of this, and other studies (Hillis, et al., 2004; Jacobs, 1999; Knill & Saunders, 2003; Landy & Kojima, 2001), suggest that this is how the visual system is designed, allowing it to optimise precision across a wide range of viewing conditions.

Recent fMRI findings have also provided evidence for the combination of different visual depth cues in the human brain. Welchman and colleagues examined whether there are brain areas that represent 3-D shape from combined estimates (Welchman, Deubelius, Conrad, Bülthoff & Kourtzi, 2005). To do so they compared fMRI responses evoked by changes in the perceived 3-D shape of a two-cue stimulus. The authors used an event-related adaptation procedure whereby neural activity is increasingly attenuated for previously viewed stimuli. To determine the sensitivity of the underlying substrate, adapted responses to the test stimulus, shown twice (test-test), were compared with responses to the test stimulus followed by a reference stimulus (test-reference). For example, if a greater fMRI response was observed ('rebound' effect) when the test-reference pair differed in perspective (compared to the adapted test-test response) this would indicate sensitivity of a particular area to the changes in the perspective cue. Observers viewed stimuli that were defined by both binocular disparity and monocular (perspective) cues. The authors first tested the weight each observer gave to each cue (using a cue conflict stimulus) in a manner similar to Hillis et al., (2004). The weights for each observer were used to create 'metamer' stimuli in which the consistent parts changed (disparity specified slant and perspective specified slant) but perceived slant did not. By comparing 'rebound' effects between the test-test, and the test-reference ('metamer') stimuli allowed the authors to delineate the underlying brain areas involved in cue combination. Early visual areas (V1, V2, V3, Vp, V4) showed a rebound effect but extrastriate areas (hMT+/V5 and lateral occipital complex) did not, suggesting that these extrastriate areas (hMT+/V5 and lateral occipital complex) were responsible for differences in perceived shape based on cue combination (Welchman et al., 2005). These results support the idea that different visual depth cues are integrated for perception.

The cue-integration model gives a good, quantitative account of the contribution of different signals to a variety of perceptual tasks, both within and across sensory modalities (e.g., Alais & Burr, 2004; Ernst & Banks, 2002; Gepshtein & Banks, 2003; Hillis et al., 2004; Jacobs, 1999; Knill & Saunders, 2003; Landy & Kojima, 2001). It remains unclear, however, whether the same model can explain the contribution of binocular and monocular depth cues to grasp control. Because one of the primary roles of vision is to provide precise information for the control of grasping it makes sense that the visuomotor system is also designed to minimise uncertainty in the estimates of object properties used for the control of grasping. Another long standing view, however, is that there is a particular special role for binocular depth cues in grasping (Goodale and Milner, 2004; Marotta, Behrmann and Goodale, 1997; Melmoth, Finlay, Morgan & Grant, 2009; Melmoth & Grant, 2006; Watt & Bradshaw, 2000). In the following sections we discuss the evidence for both views.

1.4. *Cue-integration in grasping?*

Knill and colleagues have carried out several studies exploring whether visuomotor control of the hand integrates information from binocular and monocular cues. Knill (2005) asked participants to place an object on a virtual surface defined by disparity and texture cues. Using a technique similar to perception studies the orientation specified by each cue was varied independently, to create 'cue-conflict' conditions (Hillis et al., 2004; Knill & Saunders, 2003). The weight given to each cue was then determined by measuring the orientation of the object shortly before contact with the surface. Motor responses varied with variations in the orientation specified by either cue (see also Greenwald, Knill & Saunders, 2005; van Mierlo, Louw, Smeets & Brenner, 2009). Moreover, the average weight given to disparity and texture varied systematically with variations in cue reliability (although disparity was weighted more highly for grasping than for perceptual judgments). More recently, Greenwald and Knill (2009a,b) found similar results when participants grasped an oriented object with their hand (rather than placing an object on a surface). Again using cue-conflict stimuli, the orientation of the 'opposition space' between thumb and finger was affected by variations in either the disparity- or texture-specified orientation of the target object.

The above results are broadly consistent with a cue-integration account, and inconsistent with a binocular-specialism account, because they suggest that the grasping system is sensitive to changes in monocularly specified depth even when binocular depth information is available concurrently. However, they provide only weak evidence that the visuomotor system routinely combines signals in the manner described by the above model, for two reasons. First, Knill and colleagues' stimuli varied only in terms of the slant of the target surface, and the monocular signal was slant from texture. This is arguably a special case, because the monocular texture gradient specifies the *absolute* orientation of a surface (Hillis et al., 2004; Knill, 1998). Static monocular signals typically specify only *relative* estimates of the various parameters needed to programme a natural grasp (e.g. object size, location), in which case they may not make the same contribution. Second, and more importantly, the above studies do not provide a critical test of the central feature of cue-integration, namely that the system exploits the redundancy from having two signals simultaneously available to improve performance beyond single-cue levels (Ernst 2007; Equation 3). Biases in response to variations in depth specified by one or other cue do not provide firm evidence of this integration, because they could result not only from participants basing their responses on an integrated estimate but also from 'switching' between cues, trial by trial (see Serwe, Drewing and Trommershäuser, 2009, for an example of such behaviour). In previous perception studies, a hard test of cue-integration has therefore been to show that discrimination performance with two cues is better (by the predicted amount) than with one cue (e.g. Ernst & Banks, 2002). There is no direct visuomotor analogy, but Knill and Kersten (2004) developed a measure that allowed them to compare visuomotor sensitivity to slant specified by disparity or texture alone, and by both cues together. Participants were required to place an object on a virtual surface that could be composed of either cue or both cues together. Movements to a range of target surface slants were analysed using linear discriminant analysis that provided a measure of the movements sensitivity to changes in the slant of the virtual surface. Essentially this analysis examined how well one could estimate the slant of the target surface from the object's movement trajectory. Two of their three participants showed no improvement in visuomotor sensitivity with the addition of an extra (reliable) cue. This result was inconclusive, however, because the authors found that the predicted improvement was likely not measurable, given the level of motor noise present.

1.5. *Neurophysiological evidence for cue-integration in grasping*

Recent single-unit physiology studies on macaques also provide evidence that is consistent with cue-integration in grasping. (see Castiello, 2005; Castiello & Begliomini, 2008 for reviews). Preliminary observations show similarities common to grasping behaviour in both macaques and humans. Reaches in both macaques and humans have a bell shaped velocity profile and their peak velocities vary with both the distance and difficulty of the task (Jeannerod 1984, 1988; Bootsma, Marteniuk, MacKenzie & Zaal, 1994; Roy, Paulignan, Farnè, Jouffrais & Boussaoud, 2000), in line with a speed vs. accuracy trade-off (Fitts, 1954, see General Methods). Grip aperture in monkeys and humans reaches a maximal opening prior to object contact, that is found to scale reliably with object size (Jeannerod 1984, 1988; Roy et al, 2000). For grasps to different objects, hand configurations in macaques and humans vary according to the object's shape (Mason, Gomez & Ebner, 2001; Mason, Theverapperuma, Hendrix, & Ebner, 2004), and online corrections are observed to changes in object size and location (Paulignan et al., 1991a, 1991b; Roy, Paulignan, Meunier & Boussaoud, 2006). These findings lend support to the existence of functional similarities between human and macaque brains and suggest that it is reasonable to make inferences about human brain functioning from findings in macaque cortex.

Neurons have been identified in the macaque anterior intraparietal area (AIP) that are selective for 3-D properties of visual objects (surface orientation, and axis orientation), and are involved in the sensorimotor transformations required for hand manipulation (Murata, Gallese, Luppino, Kaseda, & Sakata, 2000; Sakata, Taira, Murata & Mine, 1995; Sakata, Taira, Kusunoki, Murata, Tsutsui et al., 1999; Taira, Mine, Georgopoulos, Murata & Sakata, 1990; Taira, Tsutsui, Jiang, Yara & Sakata, 2000), including the type of grasp required (e.g. precision or power grip; Baumann, Fluet & Scherberger, 2009). Inactivation of macaque AIP by injecting muscimol (Gallese, Murata, Kaseda, Niki & Sakata, 1994) results in impaired grip aperture scaling, especially in precision grips, (but, interestingly does not affect the reach). Macaque area AIP receives input from the anatomically nearby caudal intraparietal area (CIP), which has been shown to contain neurons selective for 3-D surface orientation. Thus, these areas apparently code for information that could provide fundamental building blocks for programming the grasp. Recent studies indicate that their selectivity is derived from both binocular disparity and monocular (linear perspective and texture gradient) cues to slant (Tsutsui, Jiang, Yara, Sakata & Taira, 2001; Tsutsui, Sakata, Naganuma & Taira, 2002). This is consistent with the idea that depth from binocular and monocular cues contributes to grasp control. Overall, these findings suggest that grasping systems are not hard-wired to rely on binocular depth information and support the idea that binocular and monocular cues are combined for the control of grasping.

1.6. *Evidence for a binocular specialism for grasping*

The idea that binocular vision may play a key role in grasping comes partly from consideration of the nature and quality of depth information it provides. Humans are extremely sensitive to depth from binocular disparity (Banks, Gepshtein & Landy, 2004; Bradshaw & Rogers, 1999). Moreover, horizontal disparities, together with an estimate of viewing distance from the eyes' angle of convergence and/or vertical disparities, can, in principle, be used to recover the full metric properties of the visual scene (Bradshaw et al., 1996; Gårding, Porrill, Mayhew, & Frisby, 1995; Howard & Rogers, 2002; Rogers & Bradshaw, 1993). Unlike many perceptual tasks, which can be performed largely independently of spatial scale (face or object recognition, for example), programming actions such as grasping presumably requires accurate, absolute estimates of object properties (Desmurget, Pélisson, Rossetti & Prablanc, 1998; Milner & Goodale, 1995). The system could therefore be hard-wired to rely on signals that provide such information.

There is also experimental evidence consistent with a unique role for binocular vision in grasping. Patient DF, who has visual form agnosia, has been found to have largely intact grasping performance with binocular viewing. However, this preserved performance appears to depend almost entirely on information from binocular vision, because she shows striking failures to correctly open and orient her grasping hand under monocular viewing (Dijkerman, Milner and Carey, 1996; Marotta et al., 1997; see also Carey, Dijkerman & Milner, 1998). Similarly, differences in the efficacy of binocular and monocular cues to control grasping have been found in normal participants. Watt and Bradshaw (2003) found that binocular depth cues, in isolation, supported reliable grasp formation, whereas motion parallax did not. Given that perceptual sensitivity to depth from disparity and from motion parallax is similar (Rogers & Graham, 1979), and that motion parallax can also in principle specify the metric properties of the scene (Koenderink & van Doorn, 1991), this result, too, suggests grasping relies selectively on information from binocular vision.

As noted earlier there is convincing evidence for cue-integration in the perception of 3-D space. One might therefore assume, that the visuomotor system will rely on similar estimates, since it too requires estimates of the 3-D properties of the world. It is possible, however, that visual depth cues are processed differently for perception and for the control of grasping because different types of information are required for each task. Such a model is put forward in Milner and Goodale's (1995) 'two visual systems' hypothesis. In this model the dorsal stream, that projects from striate cortex to posterior parietal cortex, and the ventral stream that projects from striate cortex to inferotemporal cortex, are proposed to subserve vision for action and vision for perception respectively. Specifically, it is proposed that the dorsal stream processes absolute, egocentric information for the control of action and the ventral stream processes relative, allocentic information for perception. This idea is based mainly on findings of a double dissociation between these two types of functioning. Lesions of the ventral stream can result in visual form agnosia, whereby the perception of object properties is impaired but grasping remains intact. This contrasts with lesions to the dorsal stream that can cause optic ataxia in which the ability to grasp objects is impaired but the perception of the same object is not. Although there has been much evidence to support the two visual systems model (see Milner & Goodale, 2006) other findings suggest that there is not a strict dichotomy between the two streams (see Schenk & McIntosh, 2010 for a review). For example, optic ataxic patient AT shows recovered grasping performance when picking up familiar objects (Jeannerod, Decety & Michel, 1994), presumably using information from an intact ventral stream (object recognition is thought to involve the ventral pathway; Grill-Spector, 2003). These results do not support the idea of separate visual systems for the control of grasping and perception because both the dorsal and ventral streams appear to process information for the control of grasping. Based on these and other findings, it has been suggested that the functional specialisation of the two streams is relative, not absolute (Schenk & McIntosh, 2010).

By far the largest amount of empirical research on the role of binocular vision in grasping comes from studies comparing performance under normal binocular viewing, and with one eye covered (Jackson, Jones, Newport & Pritchard, 1997; Keefe & Watt, 2009; Loftus, Servos, Goodale, Mendarozqueta & Mon-Williams, 2004; Melmoth & Grant, 2006; Melmoth et al., 2009; Servos et al., 1992; Servos & Goodale, 1994; Watt & Bradshaw, 2000; for a review, see Melmoth & Grant, 2006). Removing or degrading binocular cues for grasp planning typically leads to wider grasp opening, and sometimes (but not always) slower movements. Studies interested in the contribution of binocular cues to the online control of grasping have compared performance when binocular cues are available throughout the movement to when they are removed at movement onset (Jackson et al., 1997; Servos & Goodale, 1994). Reaches performed without binocular feedback result in longer movement durations and deceleration phases (Jackson et al., 1997; Servos & Goodale, 1994) and has sometimes, though not always, resulted in larger grip apertures (see Jackson et al., 1997, Experiment 3). These results have led to the conclusion that binocular vision plays a particular, 'critical' role in grasping both for planning (Marotta et al., 1997) and online control (Melmoth & Grant, 2006; Watt & Bradshaw, 2000), or even, taken together with the neurological data described above, that there is a functional specialism for binocular vision in the control of grasping (e.g. Goodale and Milner, 2004; Marotta et al., 1997).

However, while this experimental design clearly indicates that information from binocular vision contributes to grasping performance in typical viewing situations, it tells us little about the nature of this contribution, because it confounds removal of a particular *source* of information (binocular cues) with an overall reduction in the available information *per se*. The results are thus consistent both with a binocular specialism for grasping, and with the cue-integration account. The change in performance when an eye is covered could reflect either the system switching from exclusively using binocular cues to using non-preferred monocular cues (Marotta et al., 1997), or degradation of a signal based on multiple cues, resulting in a less precise (higher variance; Equation 3) overall estimate of object properties. In order to be able to test between these two accounts of grasping performance it is first important to understand the effects of perceptual uncertainty on grasping performance.

1.7. *Visual uncertainty and grasp programming*

Grasping movements reflect a consideration of many factors, including biomechanical costs, movement variability, timing constraints and the probability of different outcomes of the movement, such as failing to grasp the object, or knocking it over (Christopoulos & Schrater, 2009). As noted above, a key advantage of combining information from multiple depth cues is that the 3-D properties of objects can be estimated as precisely as possible, across different circumstances. All else being equal, more precise estimates of object properties could support more efficient grasping movements. To do this, however, the *degree* of uncertainty must be encoded by the system (i.e. the variance of the combined estimate in Equation 3). Also, to factor this uncertainty into the movement appropriately, the system needs to know the probability of an error, given different possible movements. Grasps are opened wider when the hand is moved faster than normal (Wing, Turton & Fraser, 1986), and when vision of the hand and target are prevented during a movement (Jakobson and Goodale, 1991; Wing et al., 1986). These finding are consistent with the idea that the grasping system responds adaptively—adjusting the 'margin-for error' of the movement—to changes in the quality of information available (as well as to change in other factors such as motor variability) that affect the probability of success. Schlicht and Schrater (2007) systematically manipulated the reliability of visual information during visually closed-loop grasping by having participants look at various eccentric fixation points, thereby placing the hand and target object at different locations in the retinal periphery. Increasing the retinal eccentricity of visual information resulted in systematic increases in the opening of the grasp. This demonstrates that the system can encode visual uncertainty and that movements are altered accordingly. It seems likely, however, that Schlicht and Schrater's (2007) results reflect changes principally in the uncertainty of online visual feedback about the hand's position relative to the object, rather than in the *initial estimate* of object properties. Each participant made over a thousand grasps to the same object, at the same location (large trial numbers were required to carry out Principal Component Analysis), and so it is likely that the properties of the target object were learned, and that this information was used to programme their movements (indeed, there is evidence for such learning even with a far smaller number of repetitions, to several different objects; Keefe & Watt, 2009;

Marotta & Goodale, 2001). More recently, using visually open-loop grasping, Christopoulos and Schrater (2009) have shown that the orientation of the 'opposition space' between thumb and finger (and to some extent the opening of the grasp) is programmed to compensate for uncertainty in the orientation of the target object: the fingers approach the object in directions that increase the probability of achieving a stable grasp. Taken together, these results suggest it is likely that uncertainty in the initial estimate of object size is also factored into movement programming. Thus effects of removing binocular cues could be a strategic response to increase the margin-of-error when faced with reduced precision in the estimates of object properties.

1.8. *Studying grasping within a cue-integration framework*

The fact that changes in perceptual uncertainty have predictable consequences on grasping behaviour allows cue-integration theory to be used as a normative model in order to examine the relationship between the uncertainty of visual information and grasping performance. By manipulating the quality of information available from visual depth cues and examining changes in grasping performance with respect to predictions from cue-integration will help to determine the mechanisms by which visual depth cues are combined for the control of grasping.

The aim of this thesis is to use the cue-integration framework as a normative model to study the role of binocular depth cues in grasping, and the response of the visuomotor system to changes in perceptual uncertainty. In Chapter 3 we explore whether the availability of different cue types, or visual uncertainty *per se* determines the margin-for-error in programmed grip apertures. This allowed us to directly test between the binocular specialism and cue-integration accounts of grasp programming. In Experiment 3.1 uncertainty in object-size estimates was measured for binocular and monocular-only conditions, and when both cues were available together. Changes in viewing geometry were used to manipulate the reliability of information, creating conditions in which removing either binocular or monocular information had similar effects on the uncertainty of size estimates. In Experiment 3.2, grasping performance was compared when either cue alone, or both cues were available, allowing us to determine if changes in grasping performance result from a loss of critical information or an overall reduction in depth information. We were also able to test whether performance when both cues are available is improved over single-cue performance, as predicted by a cue-integration account of grasping that minimises uncertainty in its final estimate of object properties. A third experiment was also conducted to test whether increases in perceptual uncertainty *per se*, independent of changes in the available depth cues, result in increased grip apertures (Experiment 3.3). In Chapter 4 we extended this to examine whether binocular cues are specially relied upon for the online control of grasping. Changes in viewing geometry were used to manipulate the relative precision of binocular and monocular feedback in separate conditions of the experiment. This allowed us to compare the effects of removing binocular cues when binocular feedback was relatively informative to when it was relatively uninformative. The results of Chapters 3 and 4 support the idea that different visual depth cues are integrated for the planning and online control of grasping. Given this, we might expect other signals, not just disparity and texture, to be readily integrated for the control of grasping. One particular case is information from familiar size that is known to affect grasping performance (Marotta & Goodale, 2001; McInstosh & Lashley, 2008). Familiar size information is often available in natural scenes (presumably you know the size of your coffee cup) and can in principle provide an estimate of the metric size of objects. It is also a particularly important case because it could represent a confound in cue-removal studies of grasping. Most grasping studies have used repeated movements to a small number of object sizes and distances. Thus, the stimuli could themselves become familiar, adding another uncontrolled depth signal. We examined whether repeated grasps to a small stimulus-set (few object sizes and distances), results in learnt information that is used alongside visual depth cues for grasp programming. Using the cue-integration framework we were able to make general predictions about the presence of leant information on grasping performance. We compared grasps made to a small stimulus-set that could presumably be learnt against those to a randomised stimulus set that could not be learnt. The results allowed us to determine whether the contribution of binocular cues to grasping has previously been misestimated due to the presence of this uncontrolled, learnt information.

The results of Chapters 3 and 5 suggest that uncertainty is encoded and used to programme grasps, because the precision of information available predicted the margin-of-error programmed in grip apertures. This implies a 'strategic' system that can appropriately control the probability of errors under different circumstances. In Chapter 6 we examined how uncertainty in estimates of object properties is factored into grasping movements by exploring different movements, for which a different response to uncertainty might be expected. Performance was compared for three different grasping movements for which increasing, decreasing, or keeping grip aperture constant as perceptual uncertainty increased reduced the probability of making a grasping error (knocking into or missing the object). This reflects the different task demands of real world behaviour and was designed to see if the visuomotor system encodes both perceptual uncertainty and the probability of error for grasp planning.

Chapter 2: General Methods

2.1. *The kinematics of grasping*

In this work we planned to measure changes in grasping performance across different viewing conditions. When doing so it is important to consider that grasping movements are complex and vary according to a number of factors, including the size, shape, distance, weight, texture and fragility of the object that is to be grasped, as well as what the person intends to do with it (Jeannerod, 1984, 1988, Savelsbergh, Steenbergen & van der Kamp, 1996; Weir, MacKenzie, Marteniuk & Cargoe, 1991; Weir, MacKenzie, Marteniuk, Cargoe, & Fraser, 1991). However, grasping movements are also highly stereotypical. For example, the velocity of the movement generally follows a bell-shaped profile and there is a gradual maximum opening of the finger and thumb before the grasp is closed towards object contact. Various kinematic parameters have been identified which (i) can be readily extracted from grasp trajectory data, and (ii) have been found to be useful ways to describe and qualify changes in grasping performance. For example, many kinematic indices vary closely with object properties, which suggests that object properties are encoded by the visuomotor system in order to programme appropriate movements. For example, the peak velocity of the reach scales reliably with object distance – people move faster to further objects (Jeannerod, 1984, 1988) (Figure 2.1a,c) and maximum grip aperture (the maximum separation between the finger and thumb) is highly correlated with object size – the grasp is opened wider for larger objects (Jeannerod, 1984, 1988) (Figure 2.1b,d). Thus, one can infer the underlying estimates of object properties the visuomotor system uses to programme grasps from the kinematics of the movement.

A primary aim of this thesis was to study grasping under a cue-integration framework. To do so we tested whether the visuomotor system responds adaptively to changes in perceptual uncertainty by increasing grip apertures in order to adjust the 'margin-for error' of the movement (Christopoulos & Schrater, 2009; Schlicht and Schrater, 2007, see Experiment 3.3). If so, changes in grasp kinematics could be used as an indirect measure of the underlying precision of estimates used to programme the movement.

2.2. *Movement recording*

During our experiments, we captured grasping movements by recording the x, y, and z positions of markers attached to the nail of the thumb and index finger, sampling at 240 Hz, using a 3-camera ProReflex motion capture system (Qualisys AB). The ProReflex system is a passive system that emits infrared light. This infrared light is reflected directly back to the cameras via retroreflective markers, allowing the system to track the markers in real time. The cameras are calibrated to a known position in order to accurately reconstruct the 3-D trajectories of the markers. The 3-D co-ordinates for each trial were low-pass filtered (Butterworth filter, 12 Hz cut-off) prior to further analysis. Kinematic indices were computed from the 3-D trajectory data using custom written software in Matlab.

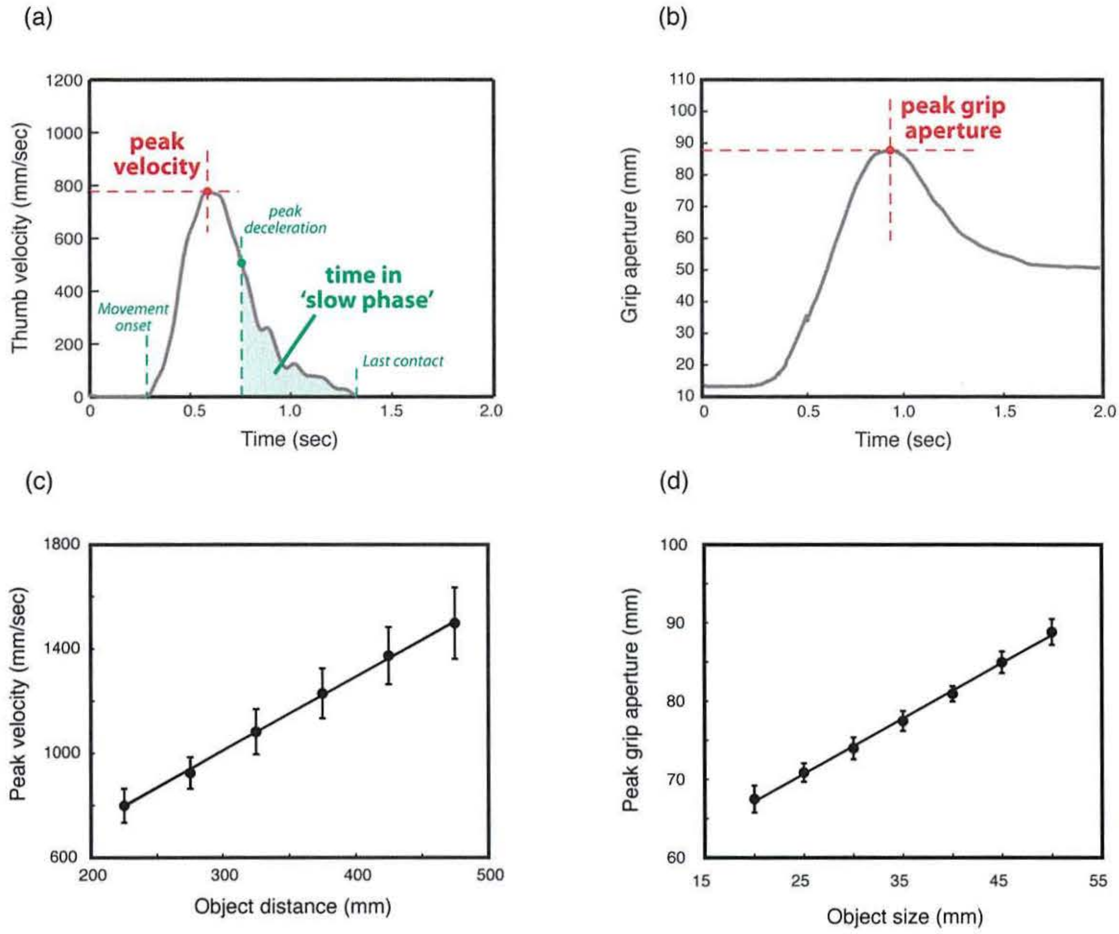


Figure 2.1. Graphs detailing the kinematic measures used. (a) A reach profile that plots the velocity of the thumb as a function of time. The points at which movement onset, peak velocity, peak deceleration and object contact occur are shown on the plot. (b) A grasp profile that plots the separation between the finger and thumb as a function of time. The point at which peak grip aperture occurs is shown on the plot. (c) Peak velocity of the thumb plotted as a function of object distance. (d) Peak grip aperture plotted as a function of object size. The solid lines show the best-fitting linear regressions to the data in each case. Error bars denote ±1 SEM.

2.3. *Dependent variables*

To characterise grasping movements we examined a number of kinematics measures derived from the three-dimensional coordinates of the movement. These are listed below and detailed in Figure 2.1. Different kinematic indices were used according to the aims of each experiment. When determining time in the slow phase and movement time in Chapter 4, last contact describes the time at which the second digit (be it the thumb or index finger) contacted the object. That is, the time at which both digits were in contact with the object. For determining the time spent decelerating in

Chapter 6, object contact refers to the time the thumb marker reached the distance to the front edge of the object (nearest the observer).

Peak velocity

Peak velocity is the maximum velocity of the reach and provides a measure of the speed of the movement. We extracted movement velocities from the 3-D coordinates of the thumb marker on each experimental trial and determined the point at which peak velocity occurred. Peak velocity is found to scale reliably with object distance (Figure 2.1a,c) (Jeannerod, 1984, 1988) suggesting that object distance is encoded by the visuomotor system. Slower reaches are sometimes observed when information is removed (Jackson et al., 1997; Loftus et al., 2004; Melmoth & Grant, 2006; Servos et al., 1992; Watt & Bradshaw, 2000), suggesting that the visuomotor system may compensate for increases in perceptual uncertainty by moving slower. Peak velocity is usually measured from the wrist, to provide a measure of hand speed that is biomechanically independent of grip formation. In Chapters 3, 5 and 6 we instead compute peak velocity from the thumb. This was necessary, because in our apparatus the wrist marker was often occluded from the cameras. We chose to use the thumb marker because the thumb moves only a small amount relative to the wrist during grasping. In control experiments we also confirmed that effects seen in wrist recordings were similar to those based on thumb recordings. We were confident, therefore, that the peak velocity of the thumb would provide a good measure of the maximum velocity of the reach. In Chapter 4 we compute peak velocity from the wrist.

Peak grip aperture

Peak grip aperture is measured as the maximum separation between the index finger and thumb, and provides a measure of the scaling of the grip to the object's size. We computed the separation between the finger and thumb from the 3-D coordinates of the index finger and thumb marker at each time frame for each experimental trial and determined the point at which peak grip aperture occurred. Peak grip aperture is found to scale reliably with object size (Figure 2.1b,d) (Jeannerod, 1984, 1988) suggesting that object size is encoded by the visuomotor system. Larger grip apertures are observed when information is removed (Jackson et al., 1997; Loftus et al., 2004; Keefe & Watt, 2009; Melmoth & Grant, 2006; Watt & Bradshaw, 2000). It has been suggested that this reflects the visuomotor system responding to increases in perceptual uncertainty by increasing grip aperture, thereby adding a 'margin of error' to the movement, to reduce the chances of knocking into or missing the object when its properties are not known precisely. We explore this idea in several experiments.

Time spent decelerating

The time spent decelerating is a measure of the final portion of the reach, as the hand approaches the object. It is measured as the time taken from peak velocity to object contact. Longer decelerations are observed when information is removed (Jackson et al., 1997; Servos & Goodale, 1994), suggesting that the visuomotor system may compensate for increases in perceptual uncertainty by taking longer to complete this phase of the movement. This makes sense, because longer decelerations allow more time to implement online corrections and gather additional information. Also, the end of the movement will be slower, reducing the chance of knocking into the object while allowing more time for haptic feedback.

Time in the slow phase

The time in the slow phase is also a measure of the final approach of the reach, as measured from peak deceleration to last contact. Longer slow phases are observed when information is removed (Melmoth & Grant 2006; Servos & Goodale, 1994), suggesting that the visuomotor system compensates for increases in perceptual uncertainty by increasing the time in the slow phase.

Time spent in contact with the object

The time spent in contact with the object provides a measure of the time spent manipulating the object in order to successfully grasp it. It is measured as the time between initial object contact and movement end (Melmoth & Grant, 2006). That is, the time between the first and second digit touching the object. It thus provides a measure of the portion of the grasp that requires fine adjustment of the digits onto the preselected object contact points. Removing information has been found to increase the time spent in contact with the object (Melmoth & Grant, 2006; Servos & Goodale, 1994), suggesting that the visuomotor system compensates for increases in perceptual uncertainty by increasing the time taken to complete this portion of the movement. This shows an adaptation, by allowing longer for online corrections in guiding the digits to their pre-selected contact points, when faced with imprecise estimates of object properties.

Movement time

Movement time is measured as the time between movement onset and last contact. Movement time increases with distance and the removal of information (Bradshaw & Elliott, 2003; Jackson et al., 1997; Loftus et al., 2004; Melmoth & Grant, 2006;

Servos & Goodale, 1994; Watt & Bradshaw, 2000) congruent with decreases in peak velocity and extended decelerations. This suggests that the visuomotor system programmes longer movement times in response to increases in perceptual uncertainty.

2.4. *Normalisation of grip aperture data*

To account for differences in hand sizes we normalised peak grip apertures across individuals by expressing each participant's grip apertures as a proportion of his or her average peak grip aperture across all trials. We then multiplied these normalised values by the average peak grip aperture across all participants to provide normalised peak grip apertures in millimetres.

2.5. *Interpreting changes in movement kinematics*

Many factors can affect grasp kinematics so one needs to take care when interpreting results. For example, increasing the difficulty of the task, by decreasing the width of the object, results in slower peak velocities and extended decelerations (Bootsma et al., 1994). Further, the grasping system can compensate for increased perceptual uncertainty not only by increasing grip apertures, but also by reaching more slowly (Wing, Turton & Fraser, 1986). In many experiments our principle measure is changes in grip aperture. A potential difficulty with interpreting experimental data, therefore, is that reach velocity and grip aperture could be 'traded-off' differently across different experimental conditions. Thus, changes in grip aperture could not be attributed unambiguously to perceptual uncertainty. We therefore examined both velocity and grip aperture, in order to be able to interpret changes in grip aperture unambiguously.

2.6. *Psychophysical methods*

In Chapter 3 we used psychophysical procedures to directly measure variability in underlying estimates of object size. To do this we computed just-noticeable differences (JNDs), namely the smallest detectable change in object depth. This was achieved by having participants complete a two-interval forced-choice task in which they judged which of two stimuli had the greater apparent depth. On each trial the standard stimulus had a fixed depth and the comparison varied around this size using staircase procedures. To accurately estimate the slope of the resulting psychometric function, while minimising the number of trials, we used an adaptive staircase technique with different reversal rules (1-up, 2-down; 2-up, 1-down; 1-up, 3-down; 3-up, 1-down) to distribute data along the psychometric function, with most data points at the most informative positions. The resulting data were fit with cumulative Gaussians using a maximum-likelihood criterion. Following previous research (Ernst & Banks, 2002; Hillis et al., 2004; Knill & Saunders, 2003) we assumed that the standard deviation of the fitted psychometric function is proportional to the standard deviation of the underlying size estimate in each case. The resulting standard deviations were divided by the square root of 2 to estimate the standard deviation of a single estimate (this provided us with a measure of the JND, equivalent to the 76% point of the psychometric function; Green & Swets, 1974). Larger JNDs (shallower slopes) indicate a less reliable (more uncertain) estimate of object size in a given condition. This is illustrated in Figure 2.2a which plots example psychometric functions for a disparity-defined stimulus at three different distances. Each psychometric function plots the proportion of times the comparison was judged as having greater depth as a function of the comparison depth. Figure 2.2a demonstrates how the slope of the function relates to the underlying precision of the size estimate; at further distances the precision of the size estimate from disparity decreases.

Perceived size in different viewing conditions was measured by computing the point of subjective equality (PSE). The PSE is the point at which the size of the comparison is judged to be the same size as the standard, and is measured from the 50% point of the psychometric function. Put another way, the PSE determines the point at which the comparison was indistinguishable in size compared to the standard, because 50% of the time it was judged as larger, and 50% of the time it was judged as smaller. To compare PSEs across different viewing conditions the stimulus from one viewing condition was the standard, and the stimulus from the other viewing condition was the comparison. If the comparison was perceived as physically smaller than the standard it follows that one would have to make it larger to reach the PSE, whereas if it were perceived as physically larger than the standard, it would have to be made smaller to reach the PSE. Example PSEs for judgements in two different viewing conditions (comparison stimuli) against a standard stimulus with a fixed size of 45 mm are shown in figure 2.2b. The first viewing condition is plotted as the psychometric function in brown, and shows a PSE of 45 mm, indicating that both the standard and comparison were perceived as being the same size. The second viewing condition is plotted as the blue psychometric function, and shows a PSE of 47.5 mm, which demonstrates that for the comparison to be perceived as the same size as standard stimulus it was made 2.5 mm larger. This indicates that the perceived size of the comparison was in fact 2.5 mm smaller than the standard stimulus, namely 42.5 mm.

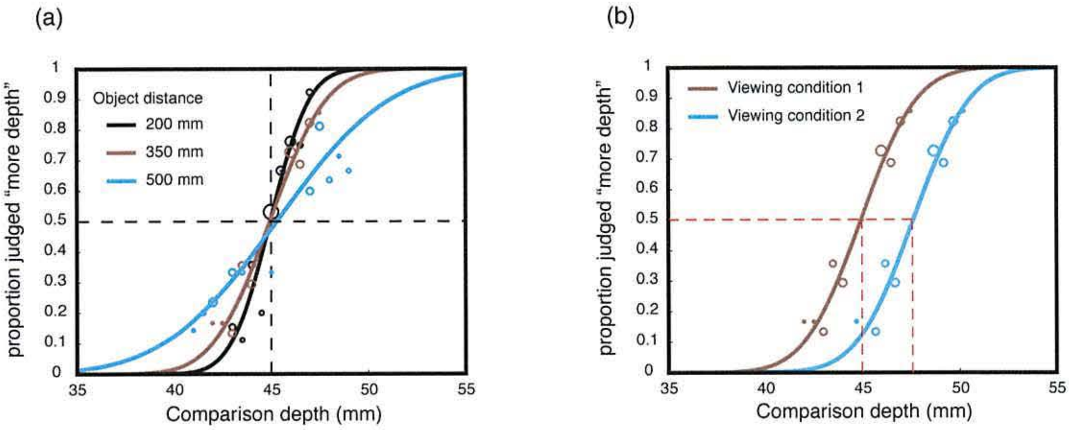


Figure 2.2. Graphs detailing the psychophysical measures used. (a) Psychometric functions for disparity-defined stimuli viewed at 200, 350 and 500 mm (see Chapter 3). Each psychometric function plots the proportion of times the comparison stimulus was judged as having greater depth as a function of the comparison depth. The slope of the function determines the precision of the underlying estimate in each case and is used to determine the just-noticeable difference (JND). The hollow circles show the data points used to fit the psychometric function; the size of the circles is proportional to the number of trials at that comparison size. (b) Psychometric functions to show how the point of subject equality (PSE) is computed from the 50% point of the function for a standard object of 45 mm.

2.7. *General experiment setup*

Although individual methods differed across experiments, all studies were carried out using the same basic setup of movement recording apparatus and custom written c-code that controlled various inputs and outputs, and trial order. We used the apparatus schematised in Figure 2.3. A LabJack (digital input/output interface) was used to control a series of relays that triggered responses from the various pieces of apparatus (LabJack^{TM}, Lakewood, Colorado). We also used the LabJack to record digital inputs from the experiment (e.g. the start button).

The LabJack was controlled by custom-written c-code running on a Power Mac to which the LabJack was connected via a Universal Serial Port. When the start button was pressed, a circuit was closed allowing 5V to pass to the digital input on the LabJack. The LabJack monitored this voltage and the point at which it ceased, marking the start button circuit being broken and the time at which movement onset occurred. In Chapter 4 both the front and rear contact surfaces of the objects were embossed with an open circuit constructed from a conductive fabric (woven nylon coated with silver; Schlegel Electronic Material, Inc., Belgium). Each circuit was created from two intersecting grids that covered the surface of the object without overlapping. One grid was attached to a 5V connection and the other to a digital input connection. The circuit was closed when the finger or thumb came into contact with, and covered the separate grids, allowing for an accurate reading of object contact time, for each digit, from the LabJack. This was possible due to the conductance of the skin allowing a small current to pass through the circuit. Both of the object circuits were monitored via the LabJack, and the times at which the circuits were closed marked the times at which first contact and second contact occurred. In some experiments we used a plastic liquid crystal film to control visual feedback (PolyVision^{TM}, Witham, Essex), This film, which we refer to as a privacy screen, is able to change state from opaque to transparent when a current is passed through it. The experimental code controlled the privacy screen via the LabJack. The code also turned a desk lamp on and off via the LabJack to allow the experimenter to see where to place the objects between trials. Recording by the ProReflex motion tracking system was also triggered by the experimental code via the LabJack. Movement recordings were saved to a PC running Microsoft Windows for off-line analysis.

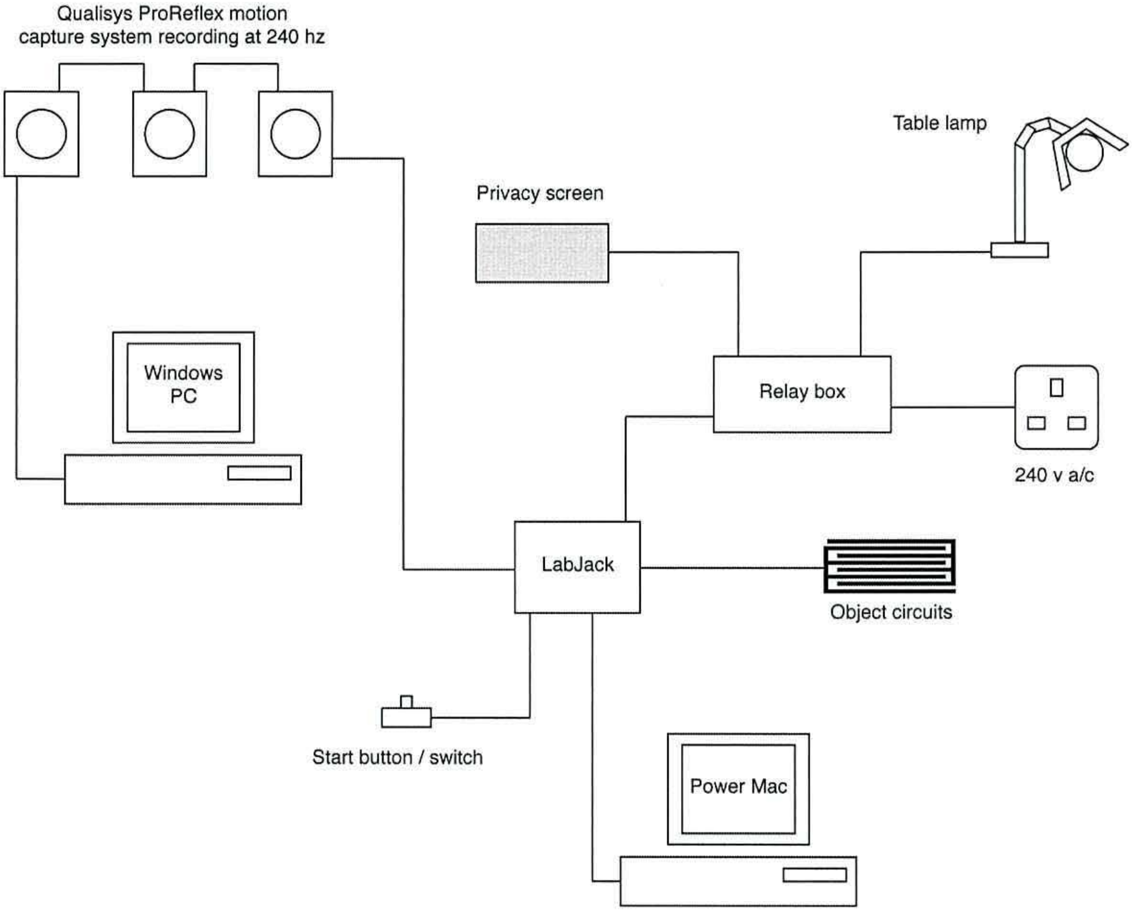


Figure 2.3. A schematic of the set up used to control the various apparatus used in the experiments of this thesis.

Chapter 3: Depth-cue integration in grasp programming: visual uncertainty, not cue type, determines the margin-for-error in grip apertures

3.1. Introduction

One critical contribution of vision to complex tasks such as grasping is to provide initial estimates of the three-dimensional (3-D) properties of objects, which are used to programme movements. Recent studies have confirmed the intuitive idea that the visuo-motor system encodes the degree of uncertainty in these estimates, and programmes a 'safety margin' into the movement accordingly (Schlicht & Schrater, 2007; Christopoulos & Schrater, 2009). Thus, minimising visual uncertainty is an important goal for the system: it allows a smaller margin-for-error, and therefore more efficient movements. How might this be achieved? In a typical visual scene there are multiple depth cues that provide useful, but uncertain, information about the three-dimensional (3-D) properties of the environment. Recent statistical models show how these different depth cues can be combined optimally, to minimise perceptual uncertainty. These models have been shown to give a good quantitative account of performance on a range of perceptual tasks (e.g., Alais & Burr, 2004; Ernst & Banks, 2002; Gepshtein & Banks, 2003; Hillis, Watt, Landy, & Banks, 2004; Jacobs, 1999; Knill & Saunders, 2003; Landy & Kojima, 2001), and they also provide a potentially useful framework for analysing and understanding visually guided movements, such as reaching and grasping (Knill 2005; Knill & Kersten, 2004). These models make specific predictions about the certainty with which object properties are known under particular viewing conditions, and specifically, they predict that to minimise perceptual uncertainty the visuo-motor system will use all the available information, independent of the particular depth cues by which it is specified. This contrasts markedly with the commonly expressed view that there is a particular, special role of binocular vision in grasping (Goodale and Milner, 2004; Marotta, Behrmann and

Goodale, 1997; Melmoth, Finlay, Morgan & Grant, 2009; Melmoth & Grant, 2006; Watt & Bradshaw, 2000). Here we used cue-integration theory as a normative model to examine the relationship between the uncertainty of visual information and grasping performance. Specifically, we explored whether effects on grasping of removing depth cues are best explained by the loss of particular 'critical' types of information, or increases in uncertainty *per se,* and in so doing, we determined whether the visuo-motor system integrates information from different depth cues to reduce uncertainty.

3.2. *Integrating information from multiple depth cues*

As discussed in the General Introduction (see section on *Integrating information from multiple depth cues*) empirical findings have shown that different visual depth cues are combined, according to their reliability in a proportionally weighted sum, for perception (Hillis et al., 2004; Jacobs, 1999; Knill & Saunders, 2003; Landy & Kojima, 2001). Importantly the variance of the combined estimate is lower than the variance of any single-cue estimate, meaning that the 3-D properties of objects can be estimated with greater precision than relying on any one cue alone (Clarke & Yuille, 1990; Knill & Pouget, 2004; Landy et al., 1995; Oruç et al., 2003; Yuille & Bülthoff, 1996). Indeed, cue-integration of this kind is statistically optimal, in the sense that it gives the minimum variance unbiased estimate of object properties (Ghahramini et al., 1997).

3.3. *Cue-integration in grasping?*

Assuming that minimising uncertainty in estimates of object properties results in better grasping performance (see General Introduction: *Visual uncertainty and grasp* *programming*), the visuo-motor system will perform better, over a wider range of circumstances, by integrating cues in the manner described above, rather than by relying on one cue alone. One might assume that within near (i.e. reaching) space binocular cues will generally be far more reliable than other sources of information, and so the system could perform quite well relying solely on this source of information. This is not the case, however. Psychophysical experiments have shown that the reliability of information from binocular disparity and monocular texture (perspective), for example, depend differently on a number of 'geometrical' factors including the orientation of surfaces relative to the observer (Hillis et al., 2004; Knill, 1998; Knill & Saunders, 2003;) and viewing distance (Hillis et al., 2004). They also depend differently on retinal eccentricity, and the position of the stimulus relative to the horopter (Greenwald & Knill, 2009a). Consequently, for example, texture can often be a significantly more reliable cue to slant than disparity within reaching distance (Hillis et al., 2004; Greenwald & Knill, 2009a). Moreover, these data refer to reduced-cue laboratory situations, and natural scenes generally contain more monocular depth cues, including familiar-size information (McIntosh & Lashley, 2008), further reducing the predicted contribution of binocular cues. Thus, natural variations in the availability and informativeness of different signals suggest that a grasping system reliant solely on binocular vision would perform significantly more poorly, under a range of commonly encountered circumstances, than a system that integrates all available signals. (Note also that to best take advantage of these various signals—to be statistically optimal—cue weights must be updated moment-to-moment to reflect changes in viewing position etc.)

Empirical studies have provided some evidence for reliability-based cue weighting in the visuo-motor system (see General Introduction: *Cue-integration in grasping?*). However, the central prediction of cue-integration—that the grasping system exploits the redundancy from having multiple signals simultaneously available to improve performance beyond single-cue levels (Ernst 2007; see General Introduction: Equation 3)—remains to be demonstrated. This represents a hard test of cue-integration because biases in mean responses to cue-conflict stimuli can result from the system switching between cues, on a trial-by-trial basis (Serwe, Drewing and Trommershäuser, 2009).

3.4. *Visual uncertainty and grasp programming*

Empirical findings suggest that the visuomotor system is able to encode visual uncertainty of object properties and alter movements accordingly (see General Introduction: *Visual uncertainty and grasp programming*). Specifically larger grip apertures appear to be programmed in response to increases in perceptual uncertainty (Christopoulos and Schrater, 2009; Schlicht and Schrater, 2007). Overall the findings suggest it is likely that uncertainty in the initial estimate of object size is factored in to movement programming. If so, increasing uncertainty by removing cues should have predictable consequences.

3.5. *The role of binocular vision in grasping*

One important aspect of the cue-integration framework is that all available information is used, irrespective of the depth cue by which it is specified. This contrasts with the main approach in the grasping literature, which has been to concentrate on the role of information from binocular vision. This focus on binocular cues perhaps stems from the fact that they are traditionally considered the most important source of depth information (Bishop, 1989; Foley 1980; Previc, 1990; Servos, Goodale & Jakobson, 1992; Sheedy, Bailey, Buri & Bass, 1986). But, as described above, it is not generally true that binocular vision is a better source of depth information than other (monocular) cues.

There is both behavioural and neurophysiological evidence to suggest that the grasping system relies selectively on binocular vision (see General Introduction: *Evidence for a binocular specialism for grasping*). Behavioral studies interested in the role of binocular vision in grasping have shown that the removal of binocular cues typically leads to a wider grasp opening, and sometimes (but not always) slower movements (Servos et al., 1992; Servos & Goodale, 1994; Jackson et al., 1997; Watt & Bradshaw, 2000; Loftus et al., 2004; Melmoth & Grant, 2006; Keefe & Watt, 2009; Melmoth et al., 2009; for a review, see Melmoth & Grant, 2006).

These results have led to the conclusion that binocular vision plays a particular, 'critical' role in grasping, or even, taken together with the neurological data described in the General Introduction (see section on *Evidence for a binocular specialism for grasping*) that there is a functional specialism for binocular vision in the control of grasping (e.g. Marotta et al., 1997; Goodale and Milner, 2004). However, while this experiment design clearly indicates that information from binocular vision contributes to grasping performance in typical viewing situations, it tells us little about the nature of this contribution, because it confounds removal of a particular *source* of information (binocular cues) with an overall reduction in the available information *per se*. The results are thus consistent both with a binocular specialism for grasping, and with the cue-integration account. The change in performance when an eye is covered could reflect either the system switching from exclusively using binocular cues to using non-preferred monocular cues (Marotta et al., 1997), or degradation of a signal based on multiple cues, resulting in a less precise (higher variance; See General Introduction: Equation 3) overall estimate of object properties.

3.6. *Aims of the current study*

Using the cue-integration framework as a normative model, we examined the relationship between perceptual uncertainty and the margin-of-error programmed into grasping movements. In particular we explored whether changes in grasp programming are best explained by the availability of binocular (disparity-plus-vergence) and monocular (texture) depth cues or by changes in perceptual uncertainty *per se.* In Experiment 3.1 we measured the relative uncertainty of size estimates of objects specified by binocular *and* monocular cues, monocular cues alone, and binocular cues alone. We established conditions in which estimates from binocular and monocular cues alone had equal uncertainty (that is, selective removal of either depth cue resulted in the same increase in perceptual uncertainty). In Experiment 3.2 we measured visually open-loop grasps to the same stimuli, allowing us to examine the relationship between variations in perceptual uncertainty and grasp opening. In so doing, we could also examine whether the binocular specialism or cue-integration account best explains grasp programming. According to the binocular specialism view, selective removal of monocular texture cues will have no effect on grasping, because the binocular input on which the system depends remains intact. According to the cue-integration model, in this situation, removing monocular *or* binocular depth cues will reduce the precision of the available depth information by the same amount (see General Introduction: Equation 3), and so should result in similar increases in the margin-for-error programmed into the grasp. This provides a hard test of the cue-integration model: if estimates of object properties are more reliable with binocular *and* monocular cues than with either signal alone, grip apertures should be smaller when both cues are available than with either signal alone. In Experiment 3.3 we used stereoscopic transparency to vary perceptual uncertainty without varying the available depth cues. We again measured uncertainty of size estimates psychophysically, and examined open-loop grasping performance to the same stimuli. This allowed us to determine whether increased uncertainty *per se* in estimates of object properties results in the programming of increased grip apertures.

3.7. Experiment 3.1: Uncertainty of size estimates from binocular and monocular cues

We determined the uncertainty of estimates of object size when only binocular cues were available (disparity and vergence), when only monocular cues were available, (texture/perspective) and when both were available together, by measuring size-discrimination thresholds in each condition, using a two-interval, forced-choice (2-IFC) psychophysical experiment. We manipulated the relative reliabilities of binocular and monocular cues by varying viewing distance. The reliability of depth information from binocular disparity should fall off particularly rapidly with viewing distance because, for a given depth, the magnitude of the projected disparities is proportional to the reciprocal of the square of viewing distance (Hillis et al., 2004; Howard & Rogers, 2002; Ogle, 1950).

Methods

Participants

Eight participants took part in Experiment 3.1 (five female and three male, aged 21-42 years). All had normal or corrected to normal vision and stereoacuity better than 40 arcsec. Participants were paid, and all procedures were in accordance with the Declaration of Helsinki.

Apparatus

The experiment apparatus is shown in Figure 3.1. Participants' eyes were positioned 400 mm above a table surface, directly above a start button (used only in the grasping experiments) located on the mid-sagittal plane. Virtual objects were displayed on a TFT monitor, placed face down, and viewed via a horizontal first–surface mirror. Stereoscopic presentation was achieved using red-green anaglyph glasses. The surface of the monitor was optically coincident with the table surface, and the mirror occluded the participants' hands. Eye position was stabilised using individually adjusted bite bars. We used a sighting device to adjust each participant's eye position to ensure that the stimulus information was always geometrically correct (Hillis & Banks, 1999). Each observer's inter-ocular distance was taken into account when generating the images.

Stimuli

The stimuli were virtual rectangular blocks positioned on a ground plane coincident with the table surface (Figure 3.1), and presented along the mid-sagittal plane. Examples of the stimuli are shown in Figure 3.2. There were three depth-cue conditions. In the *binocular-plus-monocular* condition, the ground plane and target object were defined by the perspective projection of Voronoi textures (de Berg, van Kreveld, Overmars & Schwarzkopf, 2000), viewed stereoscopically (Figure 3.2a). Such stimuli have been shown to provide reliable binocular and monocular depth cues (Knill and Saunders 2003; Hillis et al. 2004). In the *monocular-only* condition the stimuli were defined using the same technique, but we selectively removed binocular information by presenting only the right eye's image. A patch was worn over the left eye. Lastly, in the *binocular-only* condition the ground plane and objects were defined by a random-dot stereogram (Figure 3.2b), thereby selectively removing monocular depth cues from the scene. The dots were correctly perspective projected, so that there was no geometrically incorrect information in the stimulus that could otherwise give rise to biased estimates of object properties. The dot pattern was sparse, however, to minimise the reliability of any residual texture cue (Hillis et al., 2004). To ensure object size was always accurately specified, in each stimulus 10 dots were constrained to lie on the near and far edges of the top of the object. In all conditions we used anti-aliasing to position the stimulus elements with subpixel accuracy. The experiment was performed in the dark, so only the stimulus was visible.

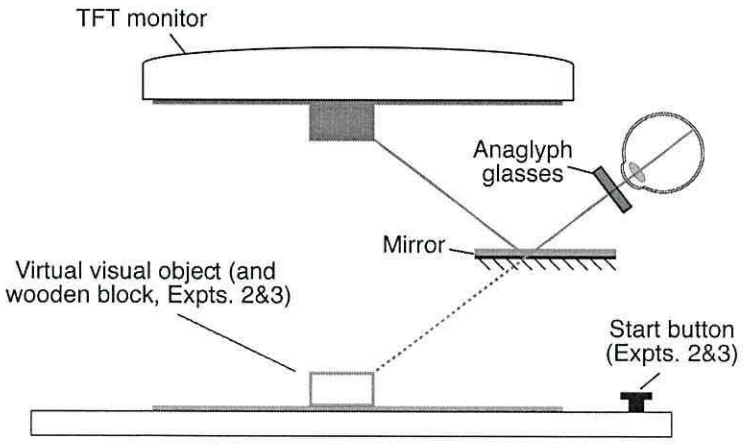


Figure 3.1. A schematic of the experiment setup.

Procedure

We measured size-discrimination thresholds (just-noticeable differences, or JNDs) using a 2-IFC procedure. Here “size” refers to the front-to-back depth of the objects. In each depth-cue condition we measured thresholds at three distances (200, 350, and 500 mm), measured along the table surface from the start button. At the start of each trial a fixation square was shown for 750 ms, at the approximate position of the upcoming stimulus. This was followed by the two target objects, which were presented for 2000 ms each, with a 300 ms inter-stimulus interval. The screen then went blank, and participants indicated which interval contained the larger object (front-to-back) using a key press. The size of the standard object was always 45 mm and the comparison size was varied using adaptive staircase procedures. Participants completed each condition with each of four staircase reversal rules (1-up, 2-down; 2-up, 1-down; 1-up, 3-down, and 3-up 1-down) to distribute data along the psychometric function. Step size was initially 8 mm, and was halved on each of the first 4 reversals. The order of standard and comparison within each trial was randomised. The staircases quit after 12 reversals, typically resulting in ~40 trials per staircase type (~160 trials per psychometric function). Trials were blocked by depth-cue condition and staircase reversal rule. Within a block, distance was randomised trial-by-trial (participants completed three interleaved staircases). Each participant completed the blocks in a random order. On each stimulus interval we randomly varied the width (90 ± 10 mm) height (25 ± 3 mm), and distance (+/- .03 × viewing distance) of the stimuli, so participants had to rely on their estimates of size to complete the task, and could not use simple cues such as the angular size of the object’s top surface, or whether the front surface of the object was nearer or farther in interval two.

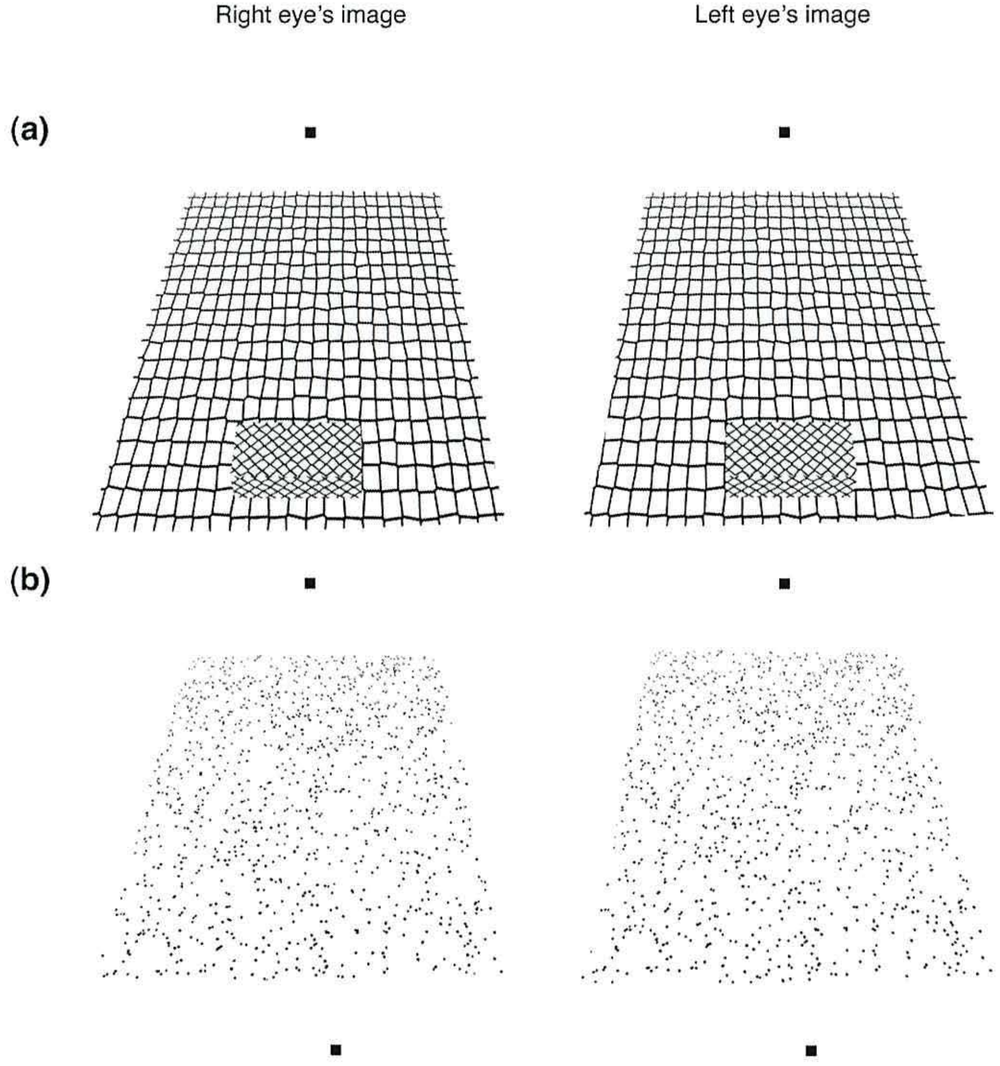


Figure 3.2. Stereograms showing examples representative of the stimuli in the *binocular-plus-monocular* condition (a), and *binocular-only* condition (b). Cross-fuse to view. The *monocular-only* condition was achieved by presenting only the right eye's image from (a).

Results and Discussion

Size JNDs were computed by fitting cumulative Gaussians to the data using a maximum-likelihood criterion. Following previous research (Ernst & Banks, 2002; Hillis et al., 2004; Knill & Saunders, 2003) we assumed that the standard deviation of the fitted psychometric function is proportional to the standard deviation of the underlying size estimate in each case. The resulting standard deviations were divided by the square root of 2 to estimate the standard deviation of a single estimate (Green

& Swets, 1974) (this provided us with a measure of the JND, equivalent to the 76% point of the psychometric function). Smaller JNDs therefore indicate a more reliable (less uncertain) estimate of object size in a given condition.

Figure 3.3a plots JNDs in each depth-cue condition (averaged across all observers) as a function of viewing distance. The uncertainty of size estimates was clearly dependent both on the available depth cues and on viewing distance. A 3 × 3 (depth-cue condition × distance) analysis of variance showed there were significant main effects of depth-cue condition ($F_{(2,14)} = 45.50$, $p < 0.001$) and of object distance ($F_{(2,14)} = 38.70$, $p < 0.001$), as well as a significant depth-cue × distance interaction ($F_{(4,28)} = 7.64$, $p < 0.001$). We conducted post-hoc pair-wise comparisons (Tukey's tests) to specify the exact nature of these effects. Figure 3.3a shows that, as expected, increasing object distance resulted in increased uncertainty of size estimates (increased JNDs), particularly in the binocular-only condition. In the binocular-plus-monocular condition the small increase in uncertainty with increasing distance was not statistically significant. In the monocular-only condition size estimates were significantly more uncertain at 500 mm than at 200 mm ($p < 0.01$). As expected, the uncertainty of binocular-only size estimates increased rapidly with distance (Hillis et al., 2004). Uncertainty did not differ significantly between 200 and 350 mm but at 500 mm it was significantly higher than at either of the two nearer distances ($p < 0.001$). These effects show that varying distance successfully modulated the relative uncertainties of size estimates in each depth-cue condition.

Figure 3.3a also plots predicted discrimination performance in the binocular-plus-monocular condition. These data are the average of each participant's predicted 'two-cue' performance, computed using his or her JNDs from the monocular- and binocular-only conditions, and Equation 3 (see General Introduction). Consistent with cue-integration theory, uncertainty was generally lower (lower JNDs) when both cues were available. There is reasonable (though not perfect) quantitative agreement between predicted and empirically determined uncertainty (the grey shaded area shows ± one standard error of the predicted two-cue JNDs). *t*-tests at each object distance showed that there were no statistically significant differences between predicted and observed binocular-plus-monocular performance ($p > 0.05$, Bonferroni corrected).

Specifying the predictions for grasping movements

If the overall uncertainty of size estimates is factored into grasp apertures (to control the margin-for-error), cue-integration theory predicts that the pattern of changes in the uncertainty of size estimates under different cue conditions, and changes in grip apertures (Experiment 3.2) will be similar. Grip apertures typically scale with movement amplitude, presumably to compensate for increased movement variability (Harris and Wolpert 1998; Loftus et al. 2004; Watt and Bradshaw 2000; Wing et al., 1986). We will therefore consider the effect of cue availability *at each distance,* holding movement amplitude constant.

Figure 3.3b shows the statistical significance of the effects of cue condition on uncertainty of size estimates at each distance (Tukey's pair-wise comparisons). At 200 mm, removing binocular information (monocular-only condition) resulted in a significant increase in uncertainty of size estimates. Thus, the cue-integration account predicts that this stimulus manipulation will result in a significant increase in grip apertures in Experiment 3.2. There was no significant effect of removing monocular information. However, binocular- and monocular-only conditions were not reliably different either. Therefore, at this distance, we cannot precisely determine the effect of removing monocular information. The same pattern of statistical effects (and therefore predictions for grasping) was observed at 350 mm. At 500 mm, however, removing either binocular or monocular cues resulted in a quantitatively similar, statistically significant increase in uncertainty of size estimates. According to the cue-integration account, therefore, removing either signal should have a reliable effect on grip apertures at this distance.

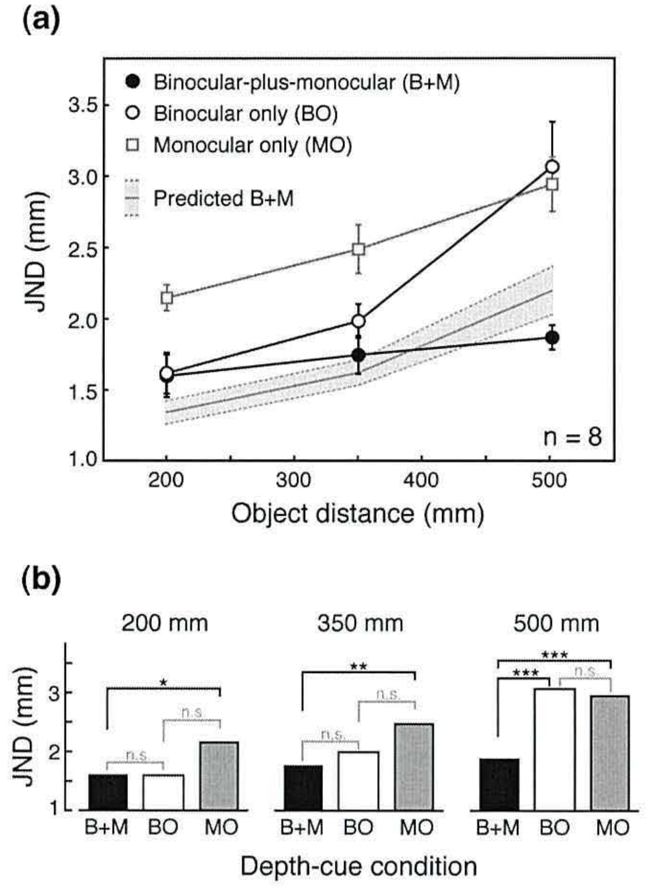


Figure 3.3. Size-discrimination performance in Experiment 3.1. (a) Mean just-noticeable differences (JNDs) in each depth-cue condition plotted as a function of object distance. The black circles denote the binocular-plus-monocular condition, the hollow black circles denote the binocular only condition, and the hollow grey squares denote the monocular only condition. The solid lines show point-by-point fits to the data in each case. Error bars denote ±1 SEM. The solid grey line shows predicted two-cue JNDs, calculated using the single-cue reliabilities at each distance and Equation 3 (see General Introduction). The shaded area shows +/- 1 standard error around the mean predicted value. (b) The results of post-hoc Tukey's tests examining the effect of depth-cue condition on JNDs at each viewing distance. The data are replotted from Figure 3.3a; each subplot shows a different object distance. The connecting lines between bars denote statistical significance (Tukey's test). Grey lines show non-significant differences ($p > 0.05$). Black lines denote significant differences at the $p < 0.05$ (*), $p < 0.01$ (**), and $p < 0.001$ (***) levels.

3.8. Experiment 3.2: Grasping objects defined by monocular and/or binocular cues

Here we examined grasping movements made to similar stimuli as those used in Experiment 3.1, under the same depth-cue conditions.

Methods

Participants

Five of the participants from Experiment 3.1 took part in Experiment 3.2, along with seven new participants (overall 4 female and 8 male, aged 21-42 years). All were right handed, had normal or corrected to normal vision, and stereoacuity better than 40 arcsec. None reported any motor deficits.

Apparatus and Stimuli

We used the same display setup as in Experiment 3.1 (Figure 3.1). The stimuli and depth-cue conditions in Experiment 3.2 were near-identical to those used in Experiment 3.1, with the following exceptions. First, we varied object size, so that the motor demands of the task were different on each trial. Second, participants grasped a real wooden block, the size and distance of which exactly matched the visual stimulus. This ensured that reaches were natural and not "pantomimed" (Cuijpers, Brenner & Smeets, 2008; Goodale, Jackobson & Keillor, 1994). The wooden block was occluded by the mirror, and so could not be seen by the participants.

Procedure

Participants began each trial by pressing down on the start button with the thumb and index finger of their right hand. The visual stimulus was presented for 2 sec followed by an audible beep, which was the cue to initiate the movement. Participants were instructed to grasp the object, front-to-back, using their thumb and index finger, and to make quick, natural movements. Releasing the start button extinguished the visual stimulus, and the hand was not visible at any time. Movements that began before the start signal, or > 600 ms after it, were discarded and repeated at the end of the block.

In each depth-cue condition there were again three object distances (200, 350, 500 mm from the start button) and also three object sizes (30, 45 and 60 mm). The objects were always presented on the table surface, along the mid-sagittal plane. Trials were blocked by depth-cue condition. One block consisted of six repetitions of each object × distance combination, giving a total of 54 trials per depth-cue condition. Trial order was randomised within blocks and the block order was counterbalanced across participants.

Results and Discussion

Movement velocity analysis

The grasping system could compensate for increased perceptual uncertainty not only by increasing grip apertures, but also by reaching more slowly (Harris & Wolpert, 1998; Wing et al., 1986). If reach velocity and grip aperture are 'traded-off' differently across different depth-cue conditions changes in grip aperture (the margin-for-error) could not be attributed unambiguously to perceptual uncertainty. Before analysing grip apertures, we tested for this by comparing peak velocity of the movements (computed from the thumb marker) in each cue condition. Figure 3.4 plots the average peak velocity across all participants in each depth-cue condition, as a function of object distance. It can be seen that movement velocities scaled in the normal manner with object distance in all conditions (Jeannerod, 1984, 1988). Moreover, velocities were unaffected by depth-cue condition. A 3 × 3 (depth-cue condition × distance) analysis of variance confirmed that there was a significant main effect of distance ($F_{(1.3,14.4)} = 1008.26$, $p < 0.001$; Greenhouse-Geisser corrected), but no main effect of depth-cue condition ($F_{(2,22)} = 1.37$, $p > 0.05$), and no significant depth cue × distance interaction ($F_{(4,44)} = 1.01$, $p > 0.05$). This finding is consistent with previous results from our lab for grasps to similar stimuli (Keefe & Watt, 2009), and confirms that it is meaningful to compare grip-aperture effects across depth-cue conditions.

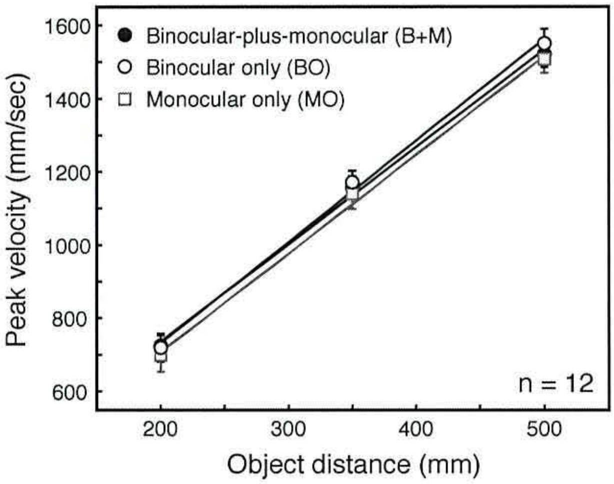


Figure 3.4. Mean peak velocity of grasping movements in each depth-cue condition in Experiment 3.2, as a function of object distance. The data points show mean velocity of the thumb marker at each distance, collapsed across all object sizes. Solid lines show the best fitting linear regressions to the data in each case. Error bars show ±1 SEM.

Grip aperture analysis

We computed a grip aperture profile for each trial by computing the 3-D distance between the thumb and finger markers at each time frame. For each participant we measured the separation of the markers when the finger and thumb were pressed together, and subtracted this measurement from the marker separation to give the actual grip aperture. We then extracted the peak grip aperture for each trial. To account for differences in hand sizes we normalised peak grip apertures across individuals by expressing each participant's grip apertures as a proportion of his or her average peak grip aperture across all trials. We then multiplied these normalised values by the average peak grip aperture across all participants to provide normalised peak grip apertures in millimetres.

Figure 3.5a plots mean peak grip aperture as a function of object distance for each depth-cue condition. If grip apertures were programmed to factor in uncertainty in size estimates (Figure 3.5), at the 500 mm object distance the removal of either binocular or monocular cues should result in similar increases in grip apertures (see *Specifying the predictions for grasping movements* section). Figure 3.5a shows that this was the case. Planned pair-wise comparisons (linear contrasts) confirmed that removing either cue had a statistically significant effect (Figure 3.5b), and showed there was no reliable difference between grip apertures in the monocular- and binocular-only conditions. Based on the measured perceptual uncertainties in Experiment 3.1, cue-integration theory also predicts that at the 200 mm and 350 mm distances removing binocular cues (monocular-only condition) should have an effect on grasping. Again, it can be seen in Figure 3.5 that this was the case. At these two distances, the predictions for the effect of removing monocular signals were not well specified, because JNDs in the binocular-only condition did not differ reliably from those in either of the other two conditions. It can be seen in Figure 3.5 that in both cases removing monocular cues resulted in quantitatively similar increases in grip apertures as removing binocular cues. Thus, there is overall good agreement between the predicted and observed effects of uncertainty on grip aperture, although there is far-from-perfect 'quantitative' agreement between effects in the perception and grasping experiments.

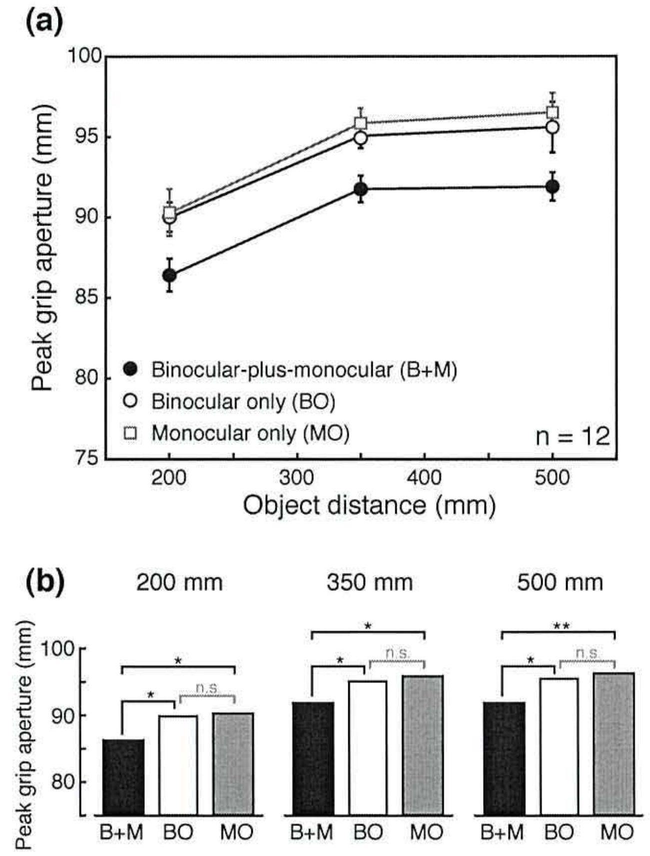


Figure 3.5. Grip aperture results from Experiment 3.2. (a) Mean peak grip aperture in each depth-cue condition as a function of object distance (collapsed across object size). Error bars denote ±1 SEM. (b) The results of planned comparisons (linear contrasts) examining the effect of depth-cue condition on peak grip aperture at each viewing distance. The data are replotted from Figure 3.5a; each subplot shows a different object distance. The connecting lines between bars denote statistical significance. Grey lines show non-significant differences ($p > 0.05$). Black lines denote significant differences at the $p < 0.05$ (*), and $p < 0.01$ (**) levels.

The pattern of grip aperture results in Figure 3.5 is inconsistent with the idea of a binocular-specialism for grasp programming. If the system relied selectively on binocular cues when both monocular and binocular information is available (Marotta et al., 1997), removing monocular information should have no effect. However, removing either signal resulted in a reliable increase in grip apertures, indicating that the system was relying on both information sources. Thus, the data are qualitatively consistent with a central tenet of cue-integration theory: grip apertures were significantly smaller when both binocular and monocular cues were available, compared with either signal alone, providing evidence that the grasping system does exploit the redundancy in multiple depth signals to improve performance (Knill & Kersten, 2004; see General Introduction: Equation 3). As noted earlier, however, the agreement between the detailed pattern of changes in perceptual uncertainty and grip apertures is not perfect because, at the 200 mm distance in particular, removing monocular information had a greater-than-predicted effect on grasping. We explore one possible explanation for this in the next section.

Isolating the effect of uncertainty on grip apertures

There are natural constraints on opening the grasp: as grip apertures increase due to increasing object size it presumably requires increased effort to open them yet wider to compensate for increased uncertainty. Consistent with this, removing binocular information has been found to have a larger effect when grasping small objects than when grasping large objects (Watt & Bradshaw, 2000), particularly during visually open-loop grasping, when grip apertures are larger overall (Jakobson and Goodale, 1991; Keefe & Watt, 2009). This effect was evident in Experiment 3.2. Figure 3.6 plots grip apertures in each depth-cue condition as a function of object size, and shows that removing binocular information resulted in more than three times the increase in grip aperture for the 30 mm object than for the 60 mm object (6.9 mm vs. 2.2 mm).

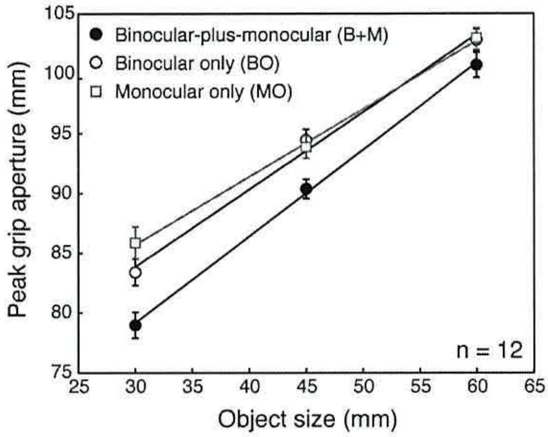


Figure 3.6. Mean peak grip apertures as a function of object size in each depth-cue condition (collapsed across object distance). The solid lines show the best-fitting linear regressions to the data in each case. In all cases peak grip apertures scaled reliably with object size, in the normal manner (Jeannerod, 1984, 1988), indicating that sufficient visual information was available to estimate object size. Error bars denote ±1 SEM.

The experiment could not be run with only small objects, because variability in motor demands is necessary to prevent pantomimed movements. However, by considering grasps only to the subset of smallest objects in our experiment we can minimise the influence of grip aperture ceiling effects in the data, yielding a more accurate measure of the effects of removing binocular or monocular information *per se* (Keefe & Watt, 2009). Figure 3.7a plots grip apertures for grasps to the 30 mm object only, as a function of object distance. It can be seen that the pattern of grip aperture changes for the small objects is more similar to the changes in perceptual uncertainty (Figure 3.3). The overall pattern of statistical effects remains the same (Figure 3.7b).

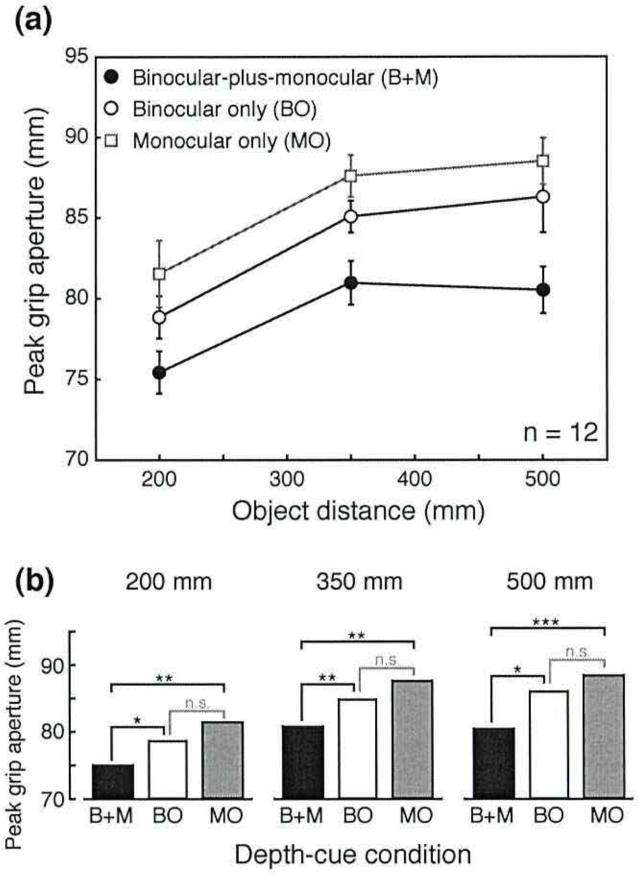


Figure 3.7. Grip aperture results from Experiment 3.2 for grasps to the 30 mm object only. (a) Mean peak grip aperture in each depth-cue condition as a function of object distance (collapsed across object size). Error bars denote ±1 SEM. (b) The results of planned comparisons (linear contrasts) examining the effect of depth-cue condition on peak grip aperture at each viewing distance. Grey lines show non-significant differences ($p > 0.05$). Black lines denote significant differences at the $p < 0.05$ (*), $p < 0.01$ (**), and $p < 0.001$ (***) levels.

3.9. Experiment 3.3: The effect of uncertainty *per se* on grip apertures

The results of Experiment 3.2 are broadly consistent with the predictions of cue-integration theory, provided one assumes that wider grip apertures are programmed in response to increased perceptual uncertainty in the estimate of object properties. In Experiment 3.3 we tested this assumption directly by varying perceptual uncertainty of visual-size information while keeping the available depth cues constant. We again measured perceptual uncertainty using a size-discrimination experiment, and then compared the results to open-loop grasping movements to the same stimuli.

Methods

Participants

The same seven participants took part in the size-discrimination task and the grasping task (3 female, 4 male, aged 21-41 years). All had normal or corrected to normal vision, stereoacuity better than 40 arcsec, and no reported motor deficits.

Apparatus and Stimuli

We used the same apparatus as in the previous two experiments. In both the size-discrimination and grasping tasks we presented two different stimuli designed to have low and high perceptual uncertainty. In one condition the stimuli were random-dot-stereogram objects, identical to the binocular-only conditions in Experiments 3.1 and 3.2. Here the object's surfaces occluded the ground plane, and so we refer to this as the *opaque condition.* In the other condition we increased the dot density from 1 to 2 dots per cm^2, and made the object transparent, so that the ground-plane dots were visible through the object. The uncertainty of the visual information should be higher in this *transparent condition,* because the problem of binocular matching—determining corresponding image points in the two eyes' images—is more difficult (Akerstrom & Todd, 1988; Banks et al., 2004; van Ee & Anderson, 2001; Tsirlin, Allison, & Wilcox, 2008; Wallace & Mamassian, 2004).

Procedure

Size-discrimination task

Participants completed a 2-IFC size-discrimination task identical to that used in Experiment 3.1, using the same standard object size and distances. Trials were blocked by viewing condition (opaque or transparent), and the order of blocks was counterbalanced across participants.

Grasping task

The grasping task was similar to that used in Experiment 3.2. Again, vision was extinguished at movement onset, to isolate the role of vision in programming the movement. In an attempt to minimise learning of the object properties, we used seven object sizes (30 – 60 mm, in 5 mm increments), at eight distances (200 – 500 mm in ~43 mm increments) (Keefe and Watt, 2009). Participants completed one repetition of each object × distance combination, making a total of 56 trials per viewing condition. Trials were blocked by viewing condition (opaque or transparent) and trial order was randomised within each block. The order in which each participant completed the viewing conditions was counterbalanced.

Results and Discussion

Figure 3.8 plots mean peak velocity of the thumb marker for the opaque and transparent conditions as a function of object distance. It can be seen that movement speed scaled reliably with object distance in the normal way (Jeannerod, 1984, 1988) in both conditions. Velocities were slightly slower in the transparent condition. However a 2 x 8 (viewing condition x distance) repeated measures analysis of variance showed that that there was no significant main effect of viewing condition ($F_{(1,6)} = 1.25$, $p > 0.05$) or depth cue × distance interaction ($F_{(7,42)} = 1.34$, $p > 0.05$). There was a significant main effect of distance ($F_{(7,42)} = 173.90$, $p < 0.001$). These findings confirm that participants did not compensate for increased uncertainty by moving slower. It is meaningful, therfore to compare grip-aperture effects across different levels of perceptual uncertainty.

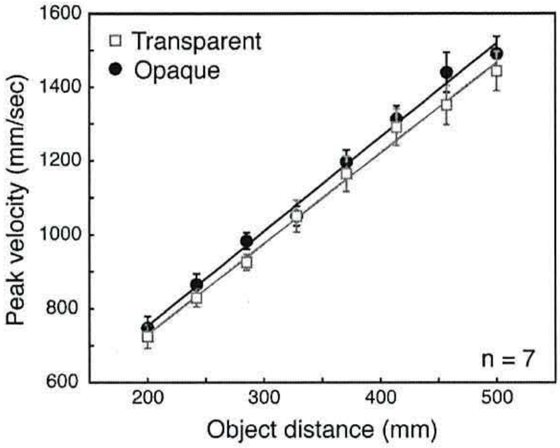


Figure 3.8. Mean peak movement velocity as a function of object distance in Experiment 3.3. Filled black circles denote the opaque condition and hollow grey squares denote the transparent condition. The data are the average of reaches to the subset of small object sizes (30, 35 and 40 mm). Solid lines show the best fitting linear regressions to the data in each case. Error bars denote ±1 SEM

The perceptual and grasping data were analysed in the same way as in the previous experiments. We again analysed grasps to a subset of small object sizes (here we could include 30, 35 and 40 mm objects, because we used smaller increments in object size than in Experiment 3.2) to minimise grasp ceiling effects (see *Isolating the effect of uncertainty* section).

The results of both the perceptual and grasping experiments are shown in Figure 3.9. Figure 3.9a plots perceptual uncertainty (size JNDs) as a function of object distance for the opaque and transparent conditions. Performance in the opaque condition was quantitatively very similar to the identical binocular-only condition in Experiment 3.1 (Figure 3.3). Uncertainty again increased with object distance, with mean JNDs approximately doubling from 200 to 500 mm object distance. Uncertainty also increased with distance in the transparent condition, but did so much more rapidly, more than tripling across the distance range tested. Thus, the transparency manipulation had a small effect at the nearest and middle distances, but caused JNDs to nearly double at the farthest distance. These effects were confirmed by statistical analysis. A 2 × 3 (stimulus condition × distance) analysis of variance showed there were significant main effects of distance ($F_{(2,12)} = 19.52$, $p < 0.001$) and stimulus condition ($F_{(1,6)} = 20.89$, $p < 0.01$), and also a significant stimulus condition × distance interaction ($F_{(2,12)} = 6.78$, $p < 0.05$). Tukey's pair-wise comparisons showed that JNDs were not significantly higher in the transparent condition at 200 or 350 mm viewing distances ($p > 0.05$), but they were at 500 mm distance ($p < 0.01$). Our stimulus transparency manipulation was therefore successful in modulating the uncertainty of size estimates.

Figure 3.9b shows grip apertures for grasps to similarly defined stimuli, as a function of distance. If increased uncertainty in the initial estimate of object properties *per se* causes increases in grip apertures we would expect making the stimulus transparent to have a qualitatively similar effect on size-discrimination performance and grip apertures (as before, we consider these effects separately at each distance, to avoid confounding changes in perceptual uncertainty and motor variability; Harris and Wolpert 1998). Figure 3.9b shows that this is the case. At the near to mid distances (< 350 mm) there was no appreciable difference between grip apertures in the opaque and transparent conditions. At farther distances, however, grip apertures were substantially larger in the transparent condition. Planned pair-wise comparisons confirmed that there was no significant difference between grip apertures in the two stimulus conditions at 200 mm ($F_{(1,6)} = 0.09$, $p > 0.05$), but they were reliably larger at 500 mm ($F_{(1,6)} = 6.56$, $p < 0.05$). To confirm that grip aperture increases were caused by increased uncertainty, and not simply by increases in the estimate of object size, we ran a control experiment in which we measured the perceived size of the objects in each condition. This experiment is described in the Control Experiments section (Experiment C3). We found that at the far distance the apparent size of the stimuli in the two conditions differed by less than 0.1 mm, indicating that perceived size changes did not cause the observed increase in grip apertures.

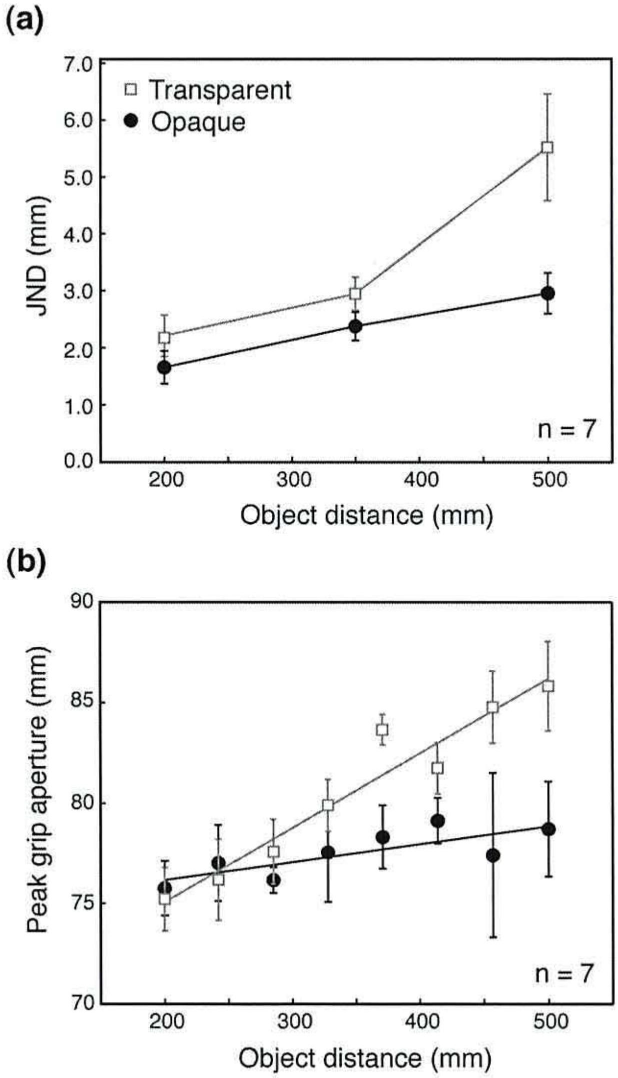


Figure 3.9. Perceptual uncertainty and grasping results from Experiment 3.3. (a) Size-discrimination performance (JNDs) for the transparent (black circles) and opaque (grey squares) objects as a function of viewing distance. Error bars show ±1 SEM. (b) Normalised peak grip apertures for grasps to the small (30, 35 & 40 mm) objects in each condition, as a function of viewing distance. The solid lines show best fitting linear regressions to the data in each case. Error bars again show ±1 SEM.

3.10. General discussion

3.11. *Summary of results*

In Experiment 3.1 we measured the uncertainty in object-size estimates as a function of the availability of depth information from binocular cues (disparity and vergence) and/or static monocular cues (texture). We varied viewing distance in order to manipulate the relative uncertainty of each signal, creating conditions in which removing either binocular or monocular information had similar effects on the uncertainty of size estimates. In Experiment 3.2, we recorded visually open-loop grasping movements to similarly defined stimuli. We found that wider grip apertures were programmed in response to increased uncertainty, largely independently of the particular depth cue (binocular or monocular) that was removed. Thus, grip apertures were smaller (smaller margin-for-error) when both cues were available, than with either signal alone. In Experiment 3.3 we increased uncertainty without changing the available depth cues, and found that this too caused an increase in grip apertures. Overall, when we accounted for ceiling effects in grip apertures, there was reasonable qualitative agreement between the patterns of variation in perceptual uncertainty and grip aperture increases. Taken together, these results suggest that the visuo-motor system integrates information from multiple depth cues, thereby reducing uncertainty in estimates of object properties. Moreover, the degree of uncertainty in this combined estimate appears to be encoded, and factored into the movement programme to adjust the margin-for-error in the movement.

3.12. *No evidence for a binocular specialism for grasp programming*

One clear implication of our findings is that they rule out the idea that grasp programming depends critically on depth information from binocular vision. As found previously, grip apertures increased when binocular cues were removed (Jackson et al., 1997; Watt & Bradshaw, 2000; Mon-Williams & Dijkerman. 1999; Loftus et al., 2004; Melmoth & Grant, 2006; Keefe & Watt, 2009). However, grip apertures also increased when we selectively removed monocular texture cues to depth, but left the binocular signal intact. This indicates that monocular depth information contributes to grasping performance even when binocular information is available simultaneously, and so the system does not rely selectively on binocular signals. This finding is consistent with the results of Knill and colleagues, who showed that orientation of the hand during grasping is affected by variations in the slant specified by either texture or disparity cues (Greenwald and Knill, 2009a,b; Greenwald et al., 2005; Knill, 2005; Knill & Kersten, 2004; van Mierlo et al., 2009; see General Introduction: *Cue-integration in grasping?*).

In light of these findings, it is unclear why patient DF, who has visual form agnosia, showed such a clear impairment in her (otherwise relatively normal) grasping performance when binocular cues were removed (Carey et al., 1998; Dijkerman et al., 1996; Marotta et al., 1997). Our results suggest her 'intact' grasping performance is based on a subset of the information relied on by normal participants, and so she may have particular deficits in her ability to process static monocular depth information for grasp control. One significant source of monocular information in our stimuli was provided by the visible outline of their edges (Figure 3.2a), and DF has been shown to have difficulty using this type of outline information, whether for perceptual or action tasks (Goodale and Milner, 2004).

If, as our data suggest, we can reject the idea of a general binocular specialism for grasping, it is perhaps surprising that motion parallax, alone, has been found not to support reliable grip aperture scaling, whereas binocular cues do (Watt & Bradshaw, 2003). It is possible that the visuo-motor system is constructed so as to not utilise the motion parallax signal, which in itself would represent a specialisation of information use. A more straightforward explanation, however, is that depth information in Watt & Bradshaw (2003)'s motion parallax condition was less precise (higher uncertainty) than in their binocular condition, resulting in poorer grasping performance in the former case. They did not empirically match the precision of the binocular and motion signals *in their stimuli*, but instead used head-motion parameters that had previously been found to result in similar sensitivity to sinusoidal corrugations in depth (Rogers & Graham, 1982; Bradshaw & Rogers, 1996)—rather different stimuli and task to estimating the size of the random-dot-covered boxes used by Watt and Bradshaw (2003). We speculate that, if the precision of information from motion parallax and disparity were matched, motion parallax will be found to support equivalent grasping performance (indeed this seems likely in light of other findings that motion parallax can be useful in guiding hand movements under some conditions; Bingham, Bradley, Bailey, & Vinner, 2001; Bingham, Zaal, Robin, & Shull, 2000; Marotta, Kruyer, & Goodale, 1998; Marotta, Perrot, Nicolle, & Goodale, 1995; Marotta, Perrot, Nicolle, Servos, & Goodale, 1995; Wickelgren, McConnell & Bingham, 2000).

The binocular-specialism account of grasping implies that different control processes are employed during 'normal' binocularly guided movements, and when the system must 'switch' to relying on monocular input (Marotta et al., 1997). In contrast, our data (and the results of the studies by Knill and colleagues, reviewed earlier) suggest that it is the amount of *information* available, not the particular cues, that determines grasping performance. This is a fundamentally different account, because only the available information changes in different viewing conditions, and not the underlying control processes.

3.13. *Evidence for cue-integration in grasp programming*

A central tenet of cue-integration theory is that, by exploiting the redundancy inherent in multiple estimates of the same property, the system can in principle estimate object properties more precisely than with any one estimate alone (see General Introduction: Equation 3). A critical empirical indicator of cue-integration in grasping is therefore to show that performance is improved by adding sources of information (Knill & Kersten, 2004). In Experiment 3.2 we found that grip apertures were reliably smaller when object properties were specified by both binocular and monocular (texture) cues, compared to when they were specified by either signal alone. Based on the assumption that more precise estimates of object properties allow a smaller margin-for-error in grip apertures to be programmed (see General Introduction: Experiment 3.3 and *Visual uncertainty and grasp programming*), this finding indicates that the visuo-motor system was able to estimate object properties with less uncertainty with both signals than with either one alone. Thus, our data indicate that information from binocular and monocular depth cues is *integrated* for grasp programming.

3.14. *Comparison of observed and expected effects of uncertainty on grasping*

In Experiment 3.1 there was quite close agreement between the measured uncertainty of perceptual size estimates from binocular and monocular cues together, and the predicted uncertainty, based on the model outlined in the General Introduction (see section on *Integrating information from multiple depth cues*). This is consistent with previous psychophysical studies showing that this model accounts well for integration of disparity and texture cues to slant (Hillis et al., 2004; Knill & Saunders, 2003), and extends this finding to the more complex (and arguably more naturalistic) task of estimating the size of multi-surface objects. We then made specific predictions about how grasp control would be affected by the loss of binocular and monocular information, based on the measured uncertainty of size estimates in Experiment 3.1, and assuming that increased uncertainty would result in larger grip apertures. There was quite good agreement between the pattern of effects on grasping and the effects predicted on the basis of the statistical analysis in Experiment 3.1 and, comparing Figures 3.3 and 3.7, it can be seen that there was reasonable qualitative agreement between the effects of removing each signal on the uncertainty of perceptual size estimates (Experiment 3.1) and on grip apertures (Experiment 3.2). In general, therefore, the cue-integration model gives a good account of grasping performance, and suggests that depth cues are similarly combined for perception and for grasp programming.

Clearly there are quantitative differences between the exact patterns of effects in the two domains. Most notably, at the nearest distance (200 mm) removing monocular information had a greater effect than we might have expected based on the perceptual uncertainty in Figure 3.3. It is useful to explore possible reasons for this discrepancy, because it illustrates several difficult issues in measuring perceptual uncertainty, and in using these data to predict motor effects.

The first possible explanations is that the JND measurements do not accurately reflect perceptual uncertainty in size estimates. It could simply be that our predictions were based on unrepresentative measurements of perceptual uncertainty, because we did not measure size JNDs for all of the participants who completed the grasping task. We think this is unlikely because, as the error bars on Figure 3.3 show, discrimination performance was highly consistent across different observers. For the purposes of our analysis we also assumed that discrimination performance in our 2-IFC task reflects only the uncertainty in the sensory estimates in each interval. This is unrealistic, because responses were presumably also affected by noise in the decision-making process (Hillis et al., 2004; Knill & Saunders, 2003). If decision noise is low (and consistent across viewing conditions), and sensory noise is high, this approximation may hold reasonably well. However, with decreasing sensory noise, the performance will increasingly reflect the influence of decision noise, in effect introducing a floor effect in JNDs. The most visually apparent difference between the effects in Experiments 3.1 and 3.2 is the relatively high JND for the binocular-plus-monocular condition, at the 200 mm object distance. Since sensory estimates were presumably the most precise in this condition, the JND would be most affected by decision noise here, and therefore most overestimated.

A second explanation is that the perceptual size JNDs are themselves accurate, but they do not accurately reflect the uncertainty in signals required to programme the grasps. This could arise from the particular methods employed. For example, we have previously shown that participants learn the properties of relatively small stimulus sets such as that used in Experiment 3.2, presumably through haptic feedback, and that this information contributes to grasp programming (Keefe & Watt, 2009; see also Marotta & Goodale, 2001). Thus, participants may have had access to additional information from familiar size to programme the grasp, which was unavailable in the size-discrimination task in Experiment 3.1. This is unlikely to explain the discrepancy between our predicted and observed results because, assuming that the reliability of learnt information was constant across conditions, it would be expected to reduce the overall magnitude of all effects of removing information in Experiment 3.2, but would not change the pattern of effects. Nonetheless, we did not control for this possibility. There could also be inherent differences in the information required to carry out the two tasks. For instance, grasping movements require estimates of distance, and in Experiment 3.1 we made the implicit assumption that noise in distance estimates propagates into size estimates, and so measuring a size JND captures all the relevant uncertainty for programming a grasping movement. This assumption may be incorrect, given that there is some evidence that variability in perceived distance and shape are independent (Brenner & van Damme, 1999). Moreover, some researchers have claimed that grasping does not rely on an explicit estimate of size, but is instead based on programming independent finger and thumb movements to the relevant object surfaces (Smeets & Brenner, 1999). If so, our perceptual reliability measurements in Experiment 3.1 may be only indirectly related to the variance of the estimates of object parameters required for grasping.

A third possible explanation for the quantitative discrepancy between JND effects and grasp effects is that cues are weighted differently for perception and for visuo-motor control. Using an object-placement task, Knill (2005) found that although visuo-motor weights for disparity and texture cues to slant reflected variations in their relative reliabilities, disparity was weighted slightly more than in perceptual slant judgements. Similar results have been found for a task requiring participants to grasp an oriented disk (Greenwald and Knill, 2009a,b, although see Van Mierlo et al., 2009), and

Greenwald and Knill (2009b) also found that binocular cues were weighted more highly for grasping than for object placement. A task-dependent increase in the weight given to binocular vision for grasping cannot account for the differences between our predicted and observed grasping effects, however. At 500 mm viewing distance we did see a slight trend towards extra weight for disparity (we predicted the same magnitude of effect of removing either binocular or monocular cues, but removing binocular cues resulted in a greater increase in grip apertures (8.0 vs. 5.8 mm) than removing monocular information. At the 200 mm distance, however, we found the opposite pattern.

A fourth possibility is that the changes in grip aperture reflect not only changes in uncertainty, but also changes in size estimates when information was removed. Changes in size estimates of a few millimetres would have little effect on discrimination thresholds, but could have a significant effect on grip apertures. To test this we ran a control experiment, in which we measured changes in perceived size when monocular and binocular signals were selectively removed (see Control Experiments; Experiment C1). To do this we used the same method as Experiment 3.1, except participants judged the relative size of a two-cue (binocular plus monocular) stimulus, and a single-cue (binocular- or monocular-only) stimulus, allowing us to determine the point of subjective equality (PSE) for two-cue versus single-cue stimuli. We found only small biases in perceived size when cues were removed (< 1.5 mm), which cannot explain the discrepancy between the predicted and observed grip aperture effects.

Overall, given the above difficulties, the fit between the pattern of changes in perceptual uncertainty across viewing conditions, and changes in grip apertures, is quite good, and is broadly consistent with the idea that binocular and monocular signals are integrated for grasp programming as well as for perception.

3.15. *Implications for the 'role' of binocular cues in grasp programming*

As noted earlier, the literature has typically focused on the role of binocular depth cues in grasp control. However, our results, taken together with work by Knill and colleagues, suggests that binocular and (static) monocular cues are integrated in the control of grasping, and weighted according to their reliability (precision, or uncertainty). In this context, questions regarding the particular role or 'specific contributions' of binocular information to grasp programming are ill posed. Instead, it seems likely that the contribution of information from binocular vision depends on its informativeness in any given situation, relative to other available information: its role will be large when it is relatively reliable, and small when it is relatively unreliable. Moreover, the fact there are significant variations in cue reliability even within reaching space (see Introduction, and our Experiment 3.1) suggests that the effects of removing different visual cues on grasping can only be understood by considering (and ideally controlling) their relative informativeness for the particular experimental situation under study.

It has also been suggested that binocular disparity may make a critical contribution to grasping during online control, particularly towards the end of the movement, as the fingers approach the object's surfaces (Anderson & Bingham, 2010; Bingham et al., 2001; Melmoth, Storoni, Todd, Finlay & Grant, 2007; Mon-Williams & Dijkerman, 1999; Morgan, 1989). Consistent with this, it has been found that significantly more time is spent decelerating when binocular information is removed (Bradshaw & Elliott, 2003; Loftus et al., 2004; Melmoth & Grant, 2006; Melmoth et al., 2009; Servos & Goodale, 1994). Visual feedback was unavailable in our experiment, and so we cannot draw direct conclusions from our data regarding the contribution of binocular vision to online control. However, we would point out that as with grasp programming, this does not necessarily mean the system is constructed to rely selectively on binocular signals. It could instead reflect the fact that, under normal viewing conditions, binocular disparity provides relatively precise information about the separation-in-depth of the object and fingers (Anderson & Bingham, 2010, see Chapter 4), whereas the available monocular cues (principally occlusion) do not. To determine if the system relies selectively on binocular information during online control requires experiments similar to those we have conducted for grasp programming, in which the informativeness of the relevant signals is controlled (See Chapter 4). Consistent with the cue-integration account, during object placement and grasping tasks the orientation of the hand has been shown to change online in response to perturbations of either binocular or monocular cues to slant (Greenwald et al., 2005; Greenwald & Knill, 2009b; van Mierlo et al., 2009).

3.16. *The role of uncertainty in grasp programming*

It has been proposed that the grip aperture increases that occur when binocular cues are removed reflect a 'conservative strategy', or compensation in response to loss of information (Loftus et al., 2004; Watt & Bradshaw, 2000; see also Schlicht & Schrater, 2007). Cue-integration theory provides a formal framework in which to understand this, because it explicitly describes the uncertainty in overall estimates of object properties (see General Introduction: Equation 3), and how it changes in different conditions. In Experiment 3.2 we found a relationship between the uncertainty of size estimates, and the opening of the grasp, largely independent of the available depth cues. In Experiment 3.3, we found similar changes in grip aperture when instead of removing a cue we made the information from it (binocular disparity) less precise. Taken together, these results are consistent with previous findings that the degree of uncertainty in the visuo-motor system's estimates of object properties are taken into account, and that a margin-for-error is programmed into the movement (Christopoulos & Schrater, 2009; Schlicht and Schrater, 2007). Moreover, they suggest that previously observed increases in grip aperture with the loss of binocular information reflect an increase in uncertainty *per se*, and not the loss of specific depth cues.

It is worth noting that while the idea of an adaptive margin-for-error in grasping is intuitive, it is far from trivial computationally. In order to generate an appropriate motor programme the system must not only encode the degree of uncertainty, but the relationship between different possible movement parameters and the probability of an error must also be known. Also, in real-world situations, the costs of errors are unlikely to always be equal (reaching for a mug brimmed with hot coffee vs. one filled with cold water) and this should also be factored into the movement programme (Christopoulos & Schrater, 2009; Körding & Wolpert, 2004; Trommershäuser, Maloney & Landy, 2003). Considering all these factors, along with the natural variations in the informativeness of different depth cues, even within reaching space, it can be seen that there would be potentially significant costs to the system of relying on one signal (binocular vision) alone.

3.17. Conclusions

We performed three experiments to examine whether binocular and monocular (texture) cues are integrated in the programming of grasping movements, and whether perceptual uncertainty or the availability of different depth cues can account for changes in grip apertures. In Experiment 3.1, we measured the uncertainty in size estimates in binocular- and or monocular-only conditions, and when both cues were available, in order to predict effects of removing each cue on grip apertures. In Experiment 3.2 we measured grasping performance in the same stimulus conditions, and found that removing either cue resulted in increased grip apertures. This finding is broadly consistent with cue-integration models that predict that both classes of cue should be combined for grasp programming, and is inconsistent with the idea that there is a binocular specialism for grasping. In Experiment 3.3 we found that increased uncertainty *per se* in the initial estimate of object size resulted in larger grip apertures. We also found reasonable agreement between the pattern of increases in visual uncertainty in Experiment 3.1 and grip aperture increases in Experiment 3.2, independent of which cues were available. Taken together, these results confirm that previously observed increases in grip apertures when binocular information is removed do result from an increase in visual uncertainty, and not from the loss of a particular depth cue. Moreover, they demonstrate that the visuo-motor system encodes the certainty with which object size is estimated, and programmes an increased margin-for-error to compensate. Overall, our results suggested that the visuo-motor system integrates binocular and monocular depth information, allowing it to programme grasps with a smaller margin-for-error.

3.18. Control Experiments

To avoid the possibility that confounding factors contributed to the pattern of results in Chapter 3 we performed a number of control experiments.

3.19. Experiment C1: Does cue removal change perceived size?

In Experiment 3.2 we attributed changes to grip apertures across the different depth-cue conditions to changes in the reliability/precision of size estimates. It is also possible that grip apertures changed because size estimates changed across different depth-cue conditions. To examine this we measured the point of subjective equality (PSE) between objects defined by binocular-plus-monocular cues, and by each cue type alone, so we could determine the relative perceived size of objects in each depth-cue condition.

Methods

Seven of the eight participants from Experiment 3.1 took part (four female and three male, aged 21-42 years). The experimental setup and stimuli were identical to those used in Experiment 3.1. We measured PSEs for the two-cue versus single-cue stimuli using a 2-IFC procedure. The standard was the binocular-plus-monocular stimulus and the comparison either the monocular-only or binocular-only stimulus. The experiment was blocked by comparison type. All other procedures (including trial order, object distances, staircase procedures) were identical to Experiment 3.1. Because in the monocular-only condition we presented binocular and monocular (right eye only) stimuli within a single trial we could not use an eye patch. However, the left eye's 'anaglyph' filter blocked all visible light from the right eye's image. The PSE was defined as the 50% point of the fitted psychometric function.

Results and Discussion

For each participant, in each condition, we subtracted the PSE from the standard size (45 mm) to give the change in perceived size that resulted from removing binocular or monocular information. Figure C1 plots these changes, averaged across all participants. It can be seen that the objects appeared slightly smaller when binocular information was removed (monocular-only condition) and slightly larger when monocular information was removed (binocular-only condition).

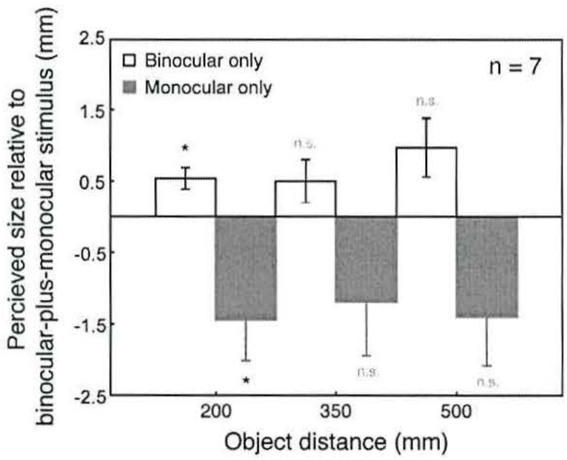


Figure C1. Perceived size relative to the binocular-plus-monocular stimulus in the binocular-only (open bars) and monocular-only (solid bars) conditions, at each object distance. Positive values indicate that the single-cue stimulus appeared larger. Error bars denote ±1 SEM. The asterisks denote size changes that were statistically significantly different to zero (one-sample *t*-test, $p < 0.05$).

We estimated the effect these small biases in object size estimates could have on grip apertures, in the following way. First, we computed perceived size changes for different object sizes, assuming that the change in perceived size when cues are removed is a constant proportion of object size. We then computed the change in grip aperture that would result from these changes separately for each participant in

Experiment 3.2. To do this, we calculated his or her grip aperture 'scaling function'—the relationship between grip aperture and object size—at each distance, in the binocular plus monocular condition. We then multiplied the change in perceived size by the slope of this function (which was typically in the range 0.6 to 0.7) to give the expected change in grip aperture due to changes in perceived size. Finally, we added this change to the grip aperture data in each case. Figure C2a plots the mean of these 'corrected' grip apertures in each condition as a function of object distance. As in Experiment 3.2, we plot grasps to the 30 mm object to minimise bias. It can be seen that the pattern of results is very similar to the uncorrected data (Figure 3.7). Moreover, the pattern of statistical effects of removing each cue is unchanged by compensating for changes in object size estimates (Figure C2b). This suggests that our overall pattern of effects results from changes in uncertainty in the estimates of object properties, not from changes in perceived size across viewing conditions (see also Experiment 3.3).

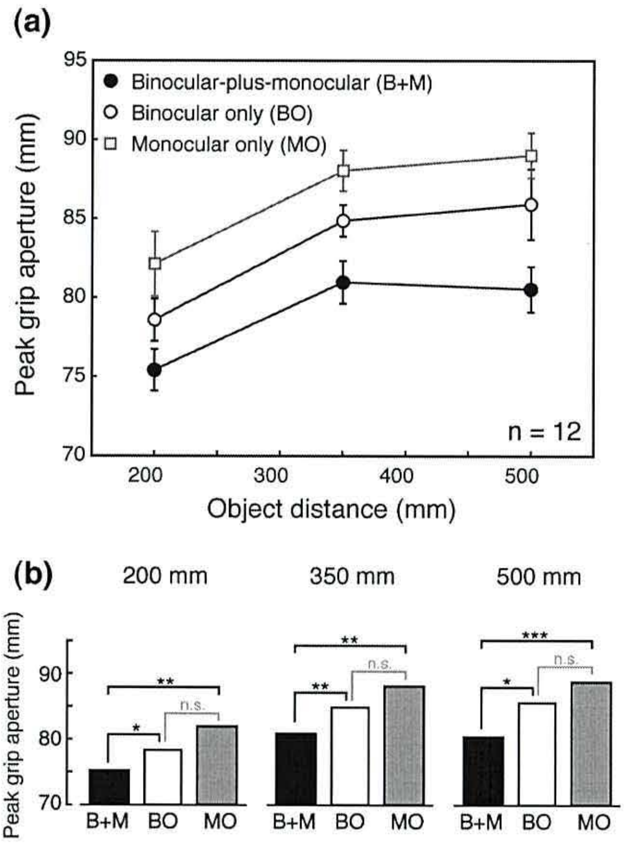


Figure C2. Mean PSE-corrected peak grip apertures for grasps to the 30 mm object in Experiment 3.2. (a) Mean peak grip aperture in each depth-cue condition as a function of object distance (collapsed across object size). Error bars denote ±1 SEM. (b) The results of planned comparisons (linear contrasts) examining the effect of depth-cue condition on peak grip aperture at each viewing distance. Grey lines show non-significant differences ($p > 0.05$). Black lines denote significant differences at the $p < 0.05$ (*), $p < 0.01$ (**), and $p < 0.001$ (***) levels.

3.20. Experiment C2. Effects of cue removal on distance estimates?

The data from our PSE experiment indicate that estimates of object size were slightly biased between the different depth-cue conditions. A possible explanation for these findings is that the estimate of distance varied between conditions. Emmert's law (Emmert, 1881) tells us that for a fixed retinal-image size, an overestimation of distance will result in the object being perceived as larger and an underestimation of distance will result in the object being perceived as smaller. Biases in the estimate of distance between the depth-cue conditions would therefore bias the perceived size of the object and the programmed grip aperture, independent of any changes in reliability. It is also possible that noise in the estimate of distance varied between the depth-cue conditions. Increased variance in the estimate of distance may have resulted in participants not knowing where to place their hand in space. If so, they may have compensated for this uncertainty by increasing their gip aperture. To examine these possibilities we ran an experiment in which participants pointed to the stimuli in order to test both the bias in the estimates of distance and variability between the different depth-cue conditions.

Methods

Five of the eight participants from Experiment 3.1 took part in the experiment (two female and three male, aged 22-42 years). The experimental setup and stimuli were identical to those used in Experiment 3.1 (binocular-plus-monocular, binocular-only, monocular-only). The procedure was similar to that used in Experiment 3.2, except that instead of participants grasping the stimuli, they were asked to point to the middle, of the front edge of the object. Participants were asked to point to the object as quickly and naturally as possible. Each trial began with the participant pressing the start button down with their right index finger. The movement was initiated on the participant hearing the start signal. Participants pointed under visually open-loop conditions to a worksurface that was clear of real objects and therefore provided no haptic feedback about the accuracy of the movements. The visual object had a fixed size of 45 mm and the experiment was blocked by depth-cue condition. Participants completed 2 blocks of 10 repetitions to each distance (60 trials) for each depth-cue condition (180 trials total). All other aspects of the procedure were exactly as in Experiment 3.2, including the timings and object distances used.

Results and Discussion

We inferred distance estimates from the radial distance each participant reached. The variability of distance estimates was inferred from the standard deviation of the radial distance. The results are plotted in Figure C3. Figure C3a shows the mean pointed radial distance and Figure C3b shows the variability (standard deviation) of the radial distance. The mean pointed radial distance increased with distance and did not alter between the different depth-cue conditions, indicating that estimates of distance were not significantly biased by the removal of binocular or monocular cues. This was confirmed by a 3 × 3 (depth-cue condition × distance) analysis of variance that showed no significant main effect of depth-cue condition ($F_{(2,8)} = .87$, $p > .05$) or depth-cue × distance interaction ($F_{(4,16)} = .28$, $p > .05$), but did show a significant main effect of object distance ($F_{(2,8)} = 1325.35$, $p < .001$).

The variability in the radial pointed distance increased with distance, consistent with the signal dependent noise associated with larger movements (Harris & Wolpert, 1998). Removing binocular or monocular cues from the binocular-plus-monocular stimulus did, generally increase the variance in the pointed distance. However, the changes in variance between the depth-cue conditions were not statistically significant. A 3 × 3 (depth-cue condition × distance) analysis of variance showed no significant main effect of depth-cue condition ($F_{(2,8)} = 3.21$, $p > .05$) or depth-cue × distance interaction ($F_{(4,16)} = .84$, $p > .05$) but did show a significant main effect of object distance ($F_{(1.14.3)} = 14.42$, $p < .01$).

The results suggest that distance estimates were similar across the depth-cue conditions used in Experiment 3.2. This is consistent with the finding that peak velocities were not affected by the removal of information in Experiment 3.2. The changes in perceived size observed in experiment C1 can therefore not be explained by changes in the estimate of viewing distance. Further, the results suggest that removing information did not increase the variability of distance estimates. However, it is likely that motor noise predominates over sensory noise in the pointing task used here, making such effects difficult to detect (see Knill & Kersten, 2004). Indeed cue-integration theory predicts that removing information should reduce the certainty in the estimate of distance. It is likely that the variance in the estimate of distance was reflected in the JND measurements from Experiment 3.1.

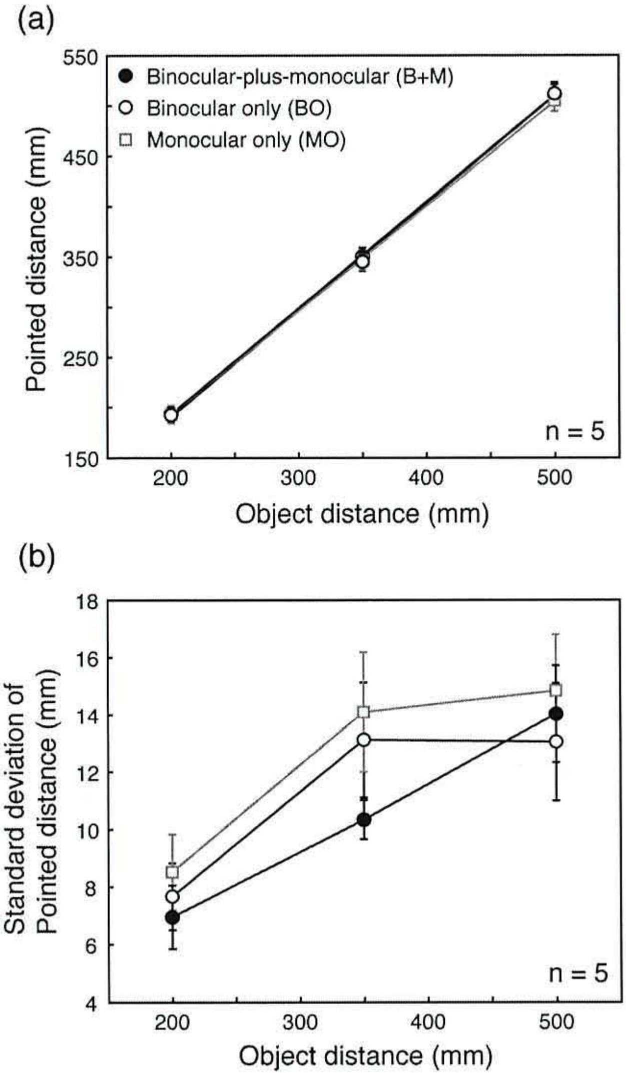


Figure C3. Results of the pointing experiment. (a) Mean pointed radial distance in each depth-cue condition plotted as a function of object distance. Solid lines show the best fitting linear regressions to the data in each case. (b) Standard deviation of the radial pointed distance in each depth-cue condition plotted as a function of object distance. The solid lines show point-to-point fits to the data. Error bars denote ±1 SEM.

3.21. Experiment C3: Was perceived size affected by the manipulation of uncertainty in Experiment 3.3?

It is possible that the increases in grip aperture in the transparent condition in Experiment 3.3 are due not to increased uncertainty in estimates of object properties, but to increased size estimates. To check this, we measured the difference in perceived size between the transparent and opaque stimuli, using a 2-IFC procedure (similar to Experiment C1, above) in which we measured the PSE between the two stimulus types.

Methods

Five of the participants from Experiment 3.3 took part in the experiment (one female and four male, aged 21-42 years). The apparatus, stimuli and procedure were identical to those used in the size-discrimination experiment in Experiment 3.3, with the exception that on each trial the standard stimulus was the transparent object and the comparison was the opaque object. The PSE was defined as the 50% point of the resulting psychometric function in each case.

Results and Discussion

Figure C4 plots the mean difference between the perceived size of the transparent object and the opaque object, at each object distance (positive values indicate the transparent object was perceived as larger). It can be seen that the transparent object was perceived as only very slightly (< 1 mm) larger than the opaque object. One-sample *t*-tests showed none of these differences were significantly different from 0 ($p > 0.05$). Moreover, it can be seen that at the far distance, where the large increase in grip apertures was observed in Experiment 3.3, there was no effect of the transparency manipulation on perceived size. This indicates that the increases in grip aperture in the transparent condition observed in Experiment 3.3 can be attributed to an increase in perceptual uncertainty *per se*.

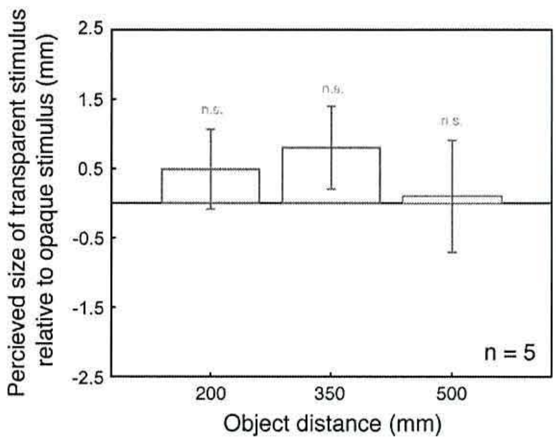


Figure C4. Mean perceived size of transparent stimulus relative to opaque stimulus, at each object distance. Positive values indicate the transparent object was perceived as larger. Error bars denote ±1 SEM.

Chapter 4: The relative precision of binocular and monocular cues to finger position determines their contribution to the online control of grasping

4.1. Introduction

The results of the previous chapter suggest that the contribution of binocular and monocular cues to grasp planning is likely to be determined by the reliability of information they convey. In this chapter we examine whether the same is true for the online control of the movement. Once initiated, a reach-to-grasp movement is updated with online visual feedback detailing the position of the digits, relative to contact surfaces on the object. This process allows the visuomotor system to correct for errors in the initial plan or changes in the world such as a moving target (Castiello, Bennett, & Stelmach, 1993; Gentilucci, Chieffi, Scarpa, & Castiello, 1992; Pauligan et al., 1991a 1991b; Soechting & Lacquiniti, 1983). The importance of online control is evident because removing vision of the hand or object typically results in larger grip apertures and longer time spent in the final approach of the reach (Berthier, Clifton, Gullapalli, McCall, & Robin, 1996; Connolly & Goodale, 1999; Churchill, Hopkins, Rönnqvist, & Vogt, 2000; Jackobson & Goodale, 1991; Jackson et al., 1997; Loftus et al., 2004; Wing et al., 1986). Again, we can use cue-integration theory as a framework to examine the online control of movements. Specifically, cue-integration theory predicts that the visuomotor system will use all available sources of information, independent of the particular depth cues by which it is specified, in order to minimise uncertainty. This contrasts with the view that binocular cues might be particularly important for online control (Anderson & Bingham, 2010; Bingham et al., 2001; Melmoth & Grant, 2006; Morgan, 1989; Watt & Bradshaw, 2000). Here we explored whether binocular signals are relied on exclusively for online control by comparing the effects of removing them in situations where they might be the most reliable source of information, and situations where monocular cues may be more informative.

4.2. *Evidence for a critical role of binocular cues in online control*

As reviewed earlier (see General Introduction: Evidence for a binocular specialism for grasping), one of the reasons binocular cues are considered of particular importance for grasp control is due to the quality of depth information they provide. For instance, binocular cues can specify the full metric properties of the scene (Rogers & Bradshaw, 1993), such as the absolute position of the digits relative to the object. This information could be used during online control to calibrate the pre-shaping of the grasp to that of the objects dimensions. Several researchers have also pointed out another way that disparities could aid online control. Relative disparities can, in principle, specify the relative distance between the hand and object with great precision (Anderson & Bingham, 2010; Bingham et al., 2001; Bradshaw & Elliot, 2003; Bradshaw, Parton & Glennerster, 2000; Melmoth & Grant, 2006; Morgan, 1989). Monitoring reductions in the relative disparity between the hand and object would allow the visuomotor system to effectively 'null' the relative disparity signal to achieve hand-object contact. It has been suggested that disparity nulling could be successfully implemented if the visuomotor system were to maintain a constant proportion between relative disparity and its rate of change (Anderson & Bingham, 2010). This strategy, termed 'proportional rate control' would produce a constant deceleration that reached zero velocity at object contact, proving an effective method for the online control of action, without requiring estimates of the absolute distance between the hand and object (Anderson & Bingham, 2010). The proportional rate control strategy has been observed in participants who performed pointing movements to a single target (Anderson & Bingham, 2010), suggesting that relative disparity could provide useful information for the online control of grasping. However, as we will discuss later, changes in viewing geometry frequently produce situations in which relative disparity is uninformative, suggesting that the system should not be hard wired to rely on these signals.

To be effective, online control has to supply accurate information concerning the position of a fast moving hand relative to its target. Binocular cues have been shown to provide no loss of depth sensitivity at high velocities faster than those of the moving hand (Morgan & Castet, 1995), indicating that they may be particularly useful for the online control of prehension. Another constraint is with regard to the speed of processing information. Online control requires the visuomotor system to both collate and use information. Given that reaching movements are typically on the order of 350-750 ms, the longer this process takes the more likely the information will be out of date or the movement will be over. Thus, information that is processed faster will presumably be of greater use for the online control of grasping. There is, at present, conflicting evidence concerning the time course of online processing for binocular and monocular cues for motor control (Greenwald et al., 2005; van Mierlo et al., 2009). In two separate studies, participants placed a virtual cylindrical object on a virtual surface slanted in depth. The slanted surface was composed of both binocular and monocular cues that could be perturbed during the course of the movement. Responses to these perturbations were used to determine the speed of processing for binocular and monocular cues. Greenwald et al. (2005) found that binocular cues were processed faster than monocular cues. However, van Mierlo et al. (2009) found that the time course of cue processing varied with methodological constraints. When the perturbation of the slanted surface was masked, as in the Greenwald, Knill & Saunders study, participants responded quicker to changes from binocular cues. However, when there was no mask, and participants saw the perturbation, they responded quicker to changes from monocular cues. Thus, it is unclear whether binocular cues provide an advantage of quicker processing times over monocular cues. Of course, the fact that variations in either signal produced changes in the movement is itself consistent with a cue-integration account of grasping. We explore further evidence for this in the next section

4.3. *Cue integration in reaching and grasping*

As discussed in the General Introduction (see *Cue integration in grasping?* section) there is a growing body of evidence to show that different visual depth cues are integrated for the control of action (Greenwald and Knill, 2009a, 2009b; Greenwald et al., 2005; Knill, 2005; Knill & Kersten, 2004; van Mierlo et al., 2009). Indeed, the results from Chapter 3 suggest that the visuomotor system integrates information from binocular and monocular depth cues for grasp planning, and several studies have also examined online control. During object placement and grasping tasks, the orientation of the hand has been shown to change online in response to perturbations of either binocular or monocular cues to slant (Greenwald et al., 2005; Greenwald & Knill, 2009b; van Mierlo et al., 2009). Further, Saunders and Knill (2004, 2005) selectively perturbed motion or position information of the hand during a pointing movement. They found that responses to these perturbations were well fit by an optimal integration model that sets the contribution of motion and position information according to their relative reliabilities. These results suggest that different sources of information are readily integrated, according to their reliability, for the online control of action.

4.4. *Effects of removing binocular cues*

Removing or degrading binocular cues during closed-loop grasping typically results in increased movement times with prolonged decelerations, larger grip apertures, and longer time spent in contact with the object (Melmoth et al., 2009; Melmoth & Grant, 2006; Watt & Bradshaw, 2000; Servos et al., 1992). However, these results do not allow us to separate the contribution of binocular cues to planning and online control, as the same information was available throughout both stages. Studies have therefore selectively introduced or removed binocular cues at movement onset in an effort to delineate their contribution to online control. Removing binocular cues only during online control results in longer movements and deceleration phases (Jackson et al., 1997; Servos & Goodale, 1994) and has sometimes, though not always, resulted in larger grip apertures (see Jackson et al., 1997 Experiment 3). Introducing binocular cues after an initial monocular view produces the opposite effects, as shown by shorter movements with shorter decelerations (Bradshaw & Elliott, 2003; Jackson et al., 1997; Servos & Goodale, 1994) and has sometimes, though not always, resulted in smaller grip apertures (see Jackson et al., 1997 Experiment 3). These results have led to the claim of a special role for binocular cues in the online control of grasping (Melmoth & Grant, 2006; Watt & Bradshaw, 2000). However, as noted earlier, while this experimental design tells us that binocular cues contribute to the online control of grasping in typical viewing situations, it tells us little about the nature of the contribution, because it confounds the removal of a particular *source* of information (binocular cues) with an overall reduction in the available information *per se*. The results are thus consistent both with a binocular specialism for the online control of grasping, and with the cue-integration account. The change in performance when an eye is covered could reflect either the system switching from exclusively using binocular cues to non-preferred monocular cues (Marotta et al., 1997), or degradation of a signal based on multiple cues, resulting in a less precise overall estimate of the hand and object properties.

It is well documented that online feedback is utilised throughout a movement (Saunders & Knill, 2003, 2005). However, its influence may be more evident late in the reach, due to the time it takes to process and use information from visual depth cues. Perturbation studies show this process typically takes the visuomotor system between 100–300ms (e.g. Castiello et al., 1993; Brenner & Smeets, 1997; Prablanc & Martin, 1992; Saunders & Knill, 2003, 2004, 2005; Soechting & Lacquiniti, 1983). Indeed, studies investigating the removal or introduction of binocular cues during online control typically find no effects on early reach kinematics such as peak velocity, but do on later kinematics such as deceleration. (Servos & Goodale, 1994; Bradshaw & Elliott, 2003, although see Jackson et al., 1997). Another reason for the lack of effect on early kinematics may be due to the fact that unreliable information would be expected to produce smaller online corrections than reliable information. In the final quarter of the reach the hand passes from peripheral to foveal vision, providing more reliable feedback (Brown, Halpert & Goodale, 2005; Sivak & MacKenzie, 1990; Schlicht & Schrater, 2007; Rizzo & Darling, 1997). Thus, it may be that online corrections that occur during the early part of the reach are smaller due to the poor reliability of information on which they are based, and are therefore harder to detect.

4.5. *Effects of viewing geometry on the precision of feedback*

A key problem with the binocular specialism account of feedback control is the idea that the visuomotor system would rely exclusively on relative disparity. This would be a sub-optimal solution, because the precision of binocular (and monocular) feedback varies with viewing geometry. Figure 4.1 illustrates this in cartoon form. When the hand and object approached separation in the depth plane (surface parallel to the line of sight), monocular cues become increasingly uninformative and performance becomes increasingly dependent upon an estimate of depth from binocular disparity (Figure 4.1a). This is because binocular feedback provides accurate information concerning the separation of the approaching hand and the object in the depth dimension (Anderson & Bingham, 2010; Bingham et al., 2001; Bradshaw & Elliot, 2003; Bradshaw et al., 2000; Melmoth & Grant, 2006; Morgan, 1989). The opposite is true in the frontoparallel plane (surface perpendicular to the line of sight), here monocular cues are informative and binocular cues are not (Figure 4.1b). This is because binocular feedback does not provide information concerning the separation of the approaching hand and object in the frontoparallel plane, whereas monocular feedback does (Westheimer, 1981). As discussed earlier, relative disparities would be relatively informative when the hand and object were separated primarily in the depth plane as they could be used to guide the fingers to contact points on the object. However, when the hand and object were separated primarily in the frontoparallel plane, relative disparities would be relatively uninformative, as they would only provide information concerning whether the hand was too high or too low relative to the object.

In natural situations, we grasp objects viewed from many orientations, but quite rarely are the hand and object separated only in the depth dimension (imagine grasping objects on your desk, for example). Thus it seems likely that in many common viewing situations binocular cues may be relatively uninformative, and a system designed to rely on them exclusively would perform poorly compared to one that integrated all sources of information according to the precision of information they convey. We therefore used changes in viewing geometry to manipulate the relative precision of binocular and monocular feedback in this experiment.

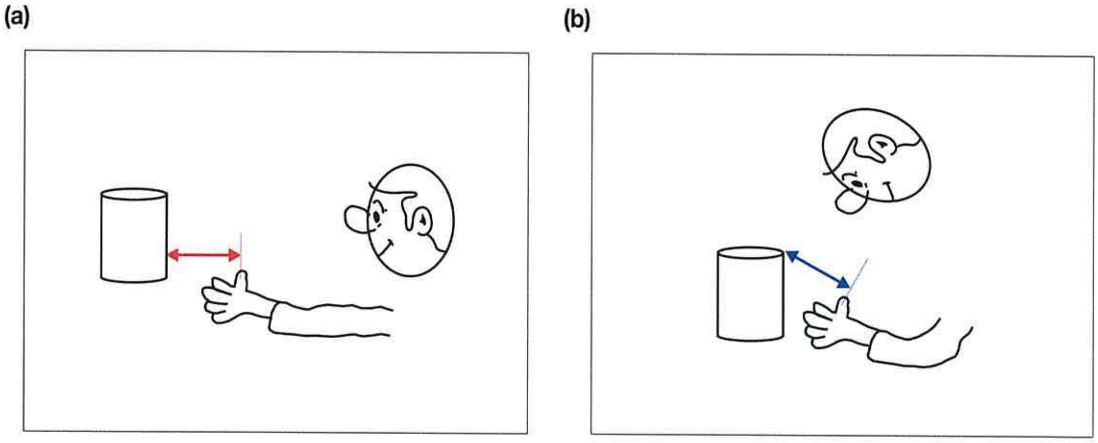


Figure 4.1. A cartoon to illustrate how precision of feedback changes with viewing geometry. (a) The fingers and object surfaces are separated primarily in the depth dimension (surface parallel to the line of sight). Here binocular cues are relatively informative and monocular cues are not. (b) The fingers and object surfaces are separated primarily in the frontoparallel plane (surface perpendicular to the line of sight). Here monocular cues are relatively informative and binocular cues are not.

4.6. *Aims of the current study*

In this study we explored whether changes in online control of movements are best explained by a system built to rely exclusively on binocular cues. We altered the relative precision of binocular and monocular feedback by changing the viewing geometry across conditions. When the hand and object are separated primarily in the depth plane, binocular cues should provide more precise feedback. However, when the hand and object are separated primarily in the frontoparallel plane, monocular cues should provide more precise feedback. The precision of information concerning the distance of the hand relative to the target object would be expected to change dynamically over the course of a movement and as a result, would be extremely difficult to measure. Rather than try to measure the dynamic reliabilities of the cues we directly manipulated the relative precision of feedback from monocular and binocular cues by using changes in viewing geometry. This allowed us to make general predictions based on the relative precision of binocular and monocular feedback without the need for direct measurement of this information during online control.

Participants reached to grasp objects both when binocular cues were available throughout the movement, or removed at movement onset. By examining changes in the kinematics of the movements, we were able to determine the contribution of binocular cues to online control across conditions. We manipulated viewing geometry across three conditions in order to change the precision of feedback from binocular and monocular cues. Each condition is referred to in terms of the slant of the table surface relative to observer at the middle of three viewing distances[4.1.]. In the first, '75-degree' condition, participants were positioned so that the hand and object were separated primarily in the depth dimension (Figure 4.2a). This set up is similar in design to previous studies of online control (Bradshaw & Elliott, 2003; Jackson et al., 1997; Servos & Goodale, 1994). Here, binocular cues should provide more precise feedback than monocular cues and their removal should result in a deterioration in performance, similar to that observed in previous studies (Jackson et al., 1997; Servos & Goodale, 1994). In another, '0-degree' condition, participants were positioned so that the hand and object were separated primarily in the frontoparallel plane (Figure

---

[4.1.] Slant is defined as the difference between the surface normal and the line of sight.

4.2c). Here monocular cues should provide more precise information than binocular cues. The binocular specialism view predicts that the removal of binocular cues will result in a loss of critical information and therefore a deterioration in performance. Cue-integration theory makes different predictions. According to Equation 3 (see General Introduction: *Integrating information from multiple depth cues*) removing a less precise source of information (0-degree condition) will result in a smaller overall reduction in precision than removing a more reliable source of information (75-degree condition). Therefore, we predict that the removal of binocular cues will lead to a smaller deterioration in performance in the 0-degree condition compared to the 75-degree condition. In the '37.5-degree' condition, participants viewed their hand and object at an orientation intermediate to the two orientation conditions described above (Figure 4.2b). In this condition, the separation of the hand and the object is closer to the frontoparallel plane then in the 75-degree condition. Binocular cues should therefore provide less precise feedback compared to the 75-degree condition but more precise feedback than in the 0-degree condition. The removal of binocular feedback in this condition should therefore result in a smaller deterioration in performance compared to the 75-degree condition, but a larger deterioration in performance compared to the 37.5-degree condition.

There are several kinematics that appear to index reduced performance during online control. Removing reliable information during online control results in increased decelerations and movement times (Bradshaw & Elliott, 2003; Jackson et al., 1997; Servos & Goodale, 1994) and has sometimes, though not always, resulted in larger grip apertures (see Jackson et al., 1997 exp 3). We therefore expect that movement time and time in the slow phase (time from peak deceleration to last contact, see

General Methods) will increase with the removal of precise feedback, in order to allow the visuomotor system more time to implement online corrections and to guide the fingers to the pre-selected contact points on the object. Increasing the time in the slow phase will also result in the end of the movement being slower, reducing the chance of knocking into the object when faced with less precise feedback. The removal of precise feedback may also be expected to result in larger, more conservative, grip apertures to avoid knocking into or missing the object. However, as peak grip aperture occurs quite early, and the reliability of feedback increases towards the end of the reach, changes in the precision of information available online may not affect this measure. Similarly, we do not expect to find any effects on the relatively early kinematic of peak velocity in this experiment. This is supported by previous findings that show that removing binocular feedback at movement onset does not significantly affect peak velocity (Jackson et al., 1997; Servos & Goodale, 1994). Removing reliable information during closed-loop grasping has also been shown to increase time spent in contact with the object (Melmoth & Grant, 2006; Servos & Goodale, 1994). We expect therefore, that the removal of precise feedback will result in an increase in time spent in contact with the object (the time taken between first contact and movement end, see General Methods) to allow longer for accurate guidance of the digits to their pre-selected contact points.

Current psychophysical data allow us to make predictions for grasping performance between the orientation conditions used in this experiment. When judging the distance between two planes that were slanted in depth and composed of both binocular and monocular cues, observers' thresholds decreased as the planes' separation approached the frontoparallel plane, (Gepshtein & Banks, 2003; Gepshtein, Burge, Ernst &

Banks, 2005; Takahashi, Diedrichsen & Watt, 2009). At slightly further viewing distances to this study, a three-to-fivefold decrease in thresholds was observed in the frontoparallel plane compared to the depth plane. Therefore, we predict that these substantial increases in the precision of feedback, when both binocular and monocular feedback is available, will lead to a steady improvement in performance as the separation between hand and object approaches the frontoparallel plane.

4.7. Methods

Participants

Sixteen right-handed participants took part in the experiment (10 male, 6 female, aged 21-45 years). All had normal or corrected to normal vision and stereoacuity better than 40 arcsec. Participants gave informed consent and were paid for their participation. All procedures were in accordance with the Declaration of Helsinki.

Apparatus and Stimuli

Participants reached to grasp objects in a lit environment. To vary the precision of information provided by binocular and monocular feedback, we varied the participant's view of their digits relative to the object. This was achieved by changing both the orientation of the object, and the supporting table, and the position of the participant's eyes in each orientation condition, using the experimental apparatus shown in Figures 4.2 and 4.3. Objects were placed on a ground plane that could rotate around a central axis. Setting the ground plane to different rotations enabled us to alter the position of the object and the hand relative to the observer. To change the position of participants' eyes between the orientation conditions, we repositioned their heads using an adjustable chin rest. To aid this process, eye position was calibrated using a sighting device that was aligned with the nodal point of the eye. Both lateral and vertical parallax errors were avoided when using the sighting device. Moving both the participant and the table was necessary to achieve the desired orientation conditions while maintaining an acceptable comfort level for the participant.

Three orientation conditions were used. In the 75-degree condition, the object and hand were viewed so that the relative positions of the digits and the object differed primarily in the depth dimension. Here binocular feedback should provide more precise information than monocular feedback. In the 0-degree condition participants looked directly down on their hand and the object, so that they were separated in the frontoparallel plane. In this case monocular feedback should provide more precise information than binocular feedback. In the 37.5-degree condition participants viewed their hand and object at a position intermediate to the two orientation conditions described above.

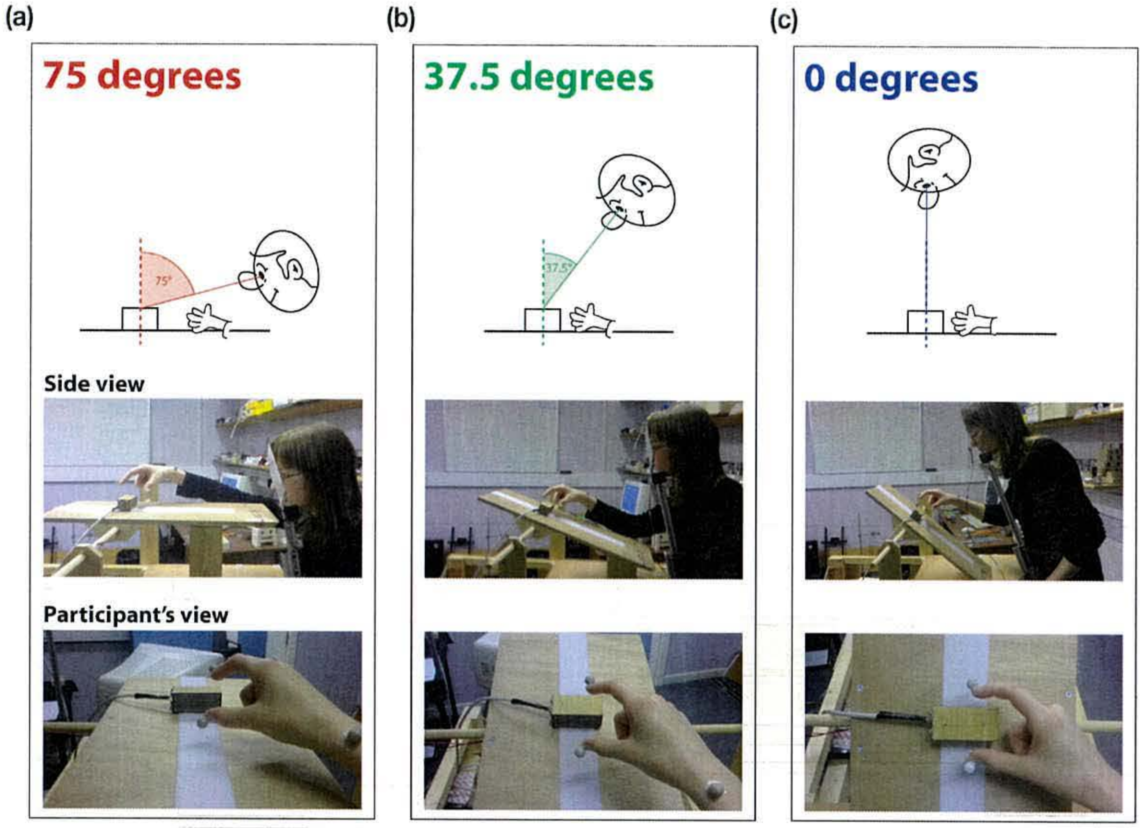


Figure 4.2. Cartoons and photographs to illustrating the position of the participant and his or her view of the hand relative to the object in each condition. (a) The 75-degrees condition in which the digits and object are separated primarily in the depth dimension. (b) The 37.5-degree condition in which the digits and object are separated intermediate between the depth plane and frontoparallel plane. (c) The 0-degree condition in which the hand and object are separated primarily in the frontoparallel plane. The top row shows a cartoon to illustrate the position of the participant relative to the object in each condition. The middle row shows how this was actually achieved, by varying the position of the participant and the orientation of the table in each condition. The bottom row shows photos taken from the participant's viewpoint in each case, showing the separation of the hand relative to the object.

Because the reliability of disparity falls off rapidly with increases in viewing distance (Hillis et al., 2004) we controlled for the average viewing distance. For objects placed 350 mm from the start button, the distance between the eye and the object was constant at 450 mm in all three orientation conditions. The average viewing distance was 450.33 mm in the 75-degree condition, 451 mm in the 37.5-degree condition, and 452.67 mm in the 0-degree condition. We wanted to isolate the effects of removing binocular information on the online control of grasping. We therefore tried to minimise differences in the quality of information during the planning phase, across conditions. To that end, we deliberately used a small number of object sizes and viewing distances, to enable participants to learn the stimulus set (Keefe & Watt, 2009). We hope that additional learned information, in addition to a binocular initial view, meant that a similar quality of information was available for the planning of the reach, so that any effects that were observed should be due to differences in the precision of feedback.

Figure 4.3 shows the participant's view of the ground plane, start switch, and object. The ground plane measured 600 mm in length and 255 mm in width, and its centre was aligned with the participant's midline. It could be rotated around a central axis located 400 mm from its front edge (nearest the observer). A start switch was located along the centre of the ground plane, 50 mm from its front edge (participant's view shown in Figure 4.3a). This allowed participants to comfortably rest their hand on the ground plane while pressing down on the start switch. The start switch was constructed from two drawing pins embedded in the ground plane. One pin was attached to a 5V connection and the other to a digital input connection on the LabJack input/output device. When the participant's index finger and thumb were pressed together, resting on the pins, the circuit was closed; at movement onset the circuit was broken, allowing for an accurate reading of movement onset.

To control feedback, we had participants wear glasses with no optical correction to which the privacy screen was attached over the left lens. The privacy screen completely occluded the left eye's view while opaque and could be made transparent to enable a binocular view. In conditions where monocular online feedback was provided, the privacy screen became opaque at movement onset, when the circuit on the start switch was broken. To ensure stable and accurate object placement during conditions where the ground plane was inclined, we attached Velcro to the bottom of the objects and along the centre of the ground plane (participants view shown in Figure 4.3b). This stabilised the position of the objects, while allowing them to be easily picked off the ground plane.

The stimuli were wooden blocks with depths of 25, 35 and 45 mm (depth in the grasping direction). The width and height of the objects were constant at 70 and 30 mm respectively. Both the front and rear contact surfaces of the objects were embossed with an open circuit (participants view shown in Figure 4.3c) constructed from a conductive fabric (woven nylon coated with silver; Schlegel Electronic Material, Inc., Belgium). Each circuit was created from two intersecting grids that covered the surface of the object without overlapping. One grid was attached to a 5V connection and the other to a digital input connection. The circuit was closed when the finger or thumb came into contact with, and covered the separate grids, allowing for an accurate reading of object contact time for each digit from the LabJack digital input/output device. This was possible due to the conductance of the skin allowing a small current to pass through the circuit. To aid skin conductance, we had participants moisturise their hands prior to the start of the experiment.

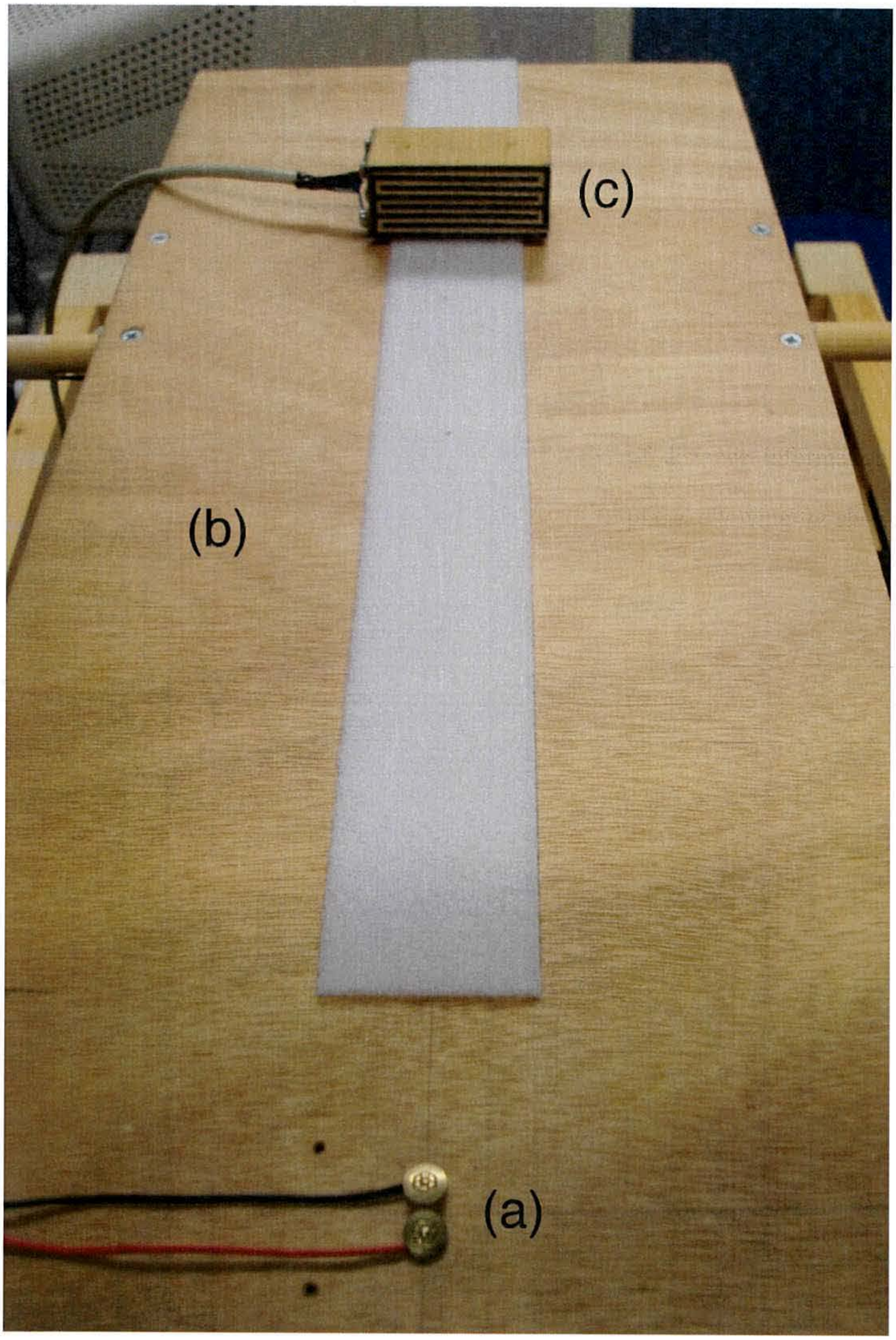


Figure 4.3. Photograph of the apparatus: (a) Start switch; (b) ground plane, and (c) object with open circuit.

Procedure

Each trial began with the participant pressing down on the start switch, with their thumb and index finger held together; the room was in complete darkness. Participants were asked to keep their eyes closed while the experimenter positioned the object on the ground plane using a torch. Once the block was positioned the participant opened their eyes and the trial was initiated by illuminating the work surface with a table lamp (controlled by computer). The object was viewed binocularly for 2000 ms, followed by an audible beep that was the participant's start signal to initiate their movement. On binocular feedback trials the privacy screen remained transparent throughout the reach. On monocular feedback trials the privacy screen became opaque at movement onset, occluding the left eye. The light remained on throughout the reach and was turned off at the end of each trial. Participants were instructed to fixate on the object throughout the experiment, and to pick it up as quickly and naturally as possible, front-to-back using their thumb and index finger. Movements that began before the start signal, or longer than 600 ms after it, were considered void and repeated at the end of the block. In these cases the light was turned off at movement onset providing the participants with feedback of the timing error. Objects were presented at distances of 300, 350 and 400 mm from the start switch along the participants' midline.

The experiment was blocked both by the type of feedback (binocular or monocular) and the orientation condition (75-degree, 37.5-degree and 0-degree), giving a total of 6 blocks. We did this because when online feedback of the hand is randomised on a trial-by-trial basis (closed-loop vs. open-loop), participants perform as if online feedback of the hand is unavailable even when it is present (Jacobson & Goodale, 1991). One block consisted of six repetitions of each object x distance combination, giving a total of 54 trials per block. Trial order was randomised within blocks and the block order was counterbalanced across participants. No participants reported any discomfort during the experiment, and they were encouraged to take breaks when needed in order to stay focused on the task. In the 75-degree and 37.5-degree conditions, the participant remained seated. In the 0-degree condition, they stood while resting their left hand on the fixed surface to provide support. In order to increase the comfort and ergonomics of the task we allowed the participants to rest their reaching arm on raised cardboard boxes that were attached to the fixed surface in the 75-degree and 37.5-degree orientation conditions. Across all conditions, participants could see their grasping hand in their peripheral vision when it was placed on the start switch and they fixated on the object. This ensured feedback concerning the hand's position was available throughout the reach.

Dependent measures

To analyse the effect of removing binocular feedback, we examined a number of kinematics measures from the three-dimensional coordinates. These were: (1) The peak velocity of the reach; (2) peak grip aperture (the maximum separation between the thumb and index finger); (3) the time in the slow phase (the time from peak deceleration to last contact); (4) time spent in contact with the object (the time between first contact and movement end), and (5) movement time (the time between contact with the start switch being broken and last contact). In all cases, last contact describes the time at which the second digit contacted the object.

4.8. Results

We first calculated the mean values of each of the dependent variables across all participants. To examine the effects of feedback in each orientation condition, we performed planned comparisons (these were 1-tailed paired samples *t*-tests, collapsed across distance and size). To compare the effects of feedback across different orientation conditions, we first computed the difference between binocular and monocular feedback trials, by subtracting the binocular value from the monocular value. The resulting difference score indexed the effect of removing binocular feedback and provided a value that could be directly compared across conditions. Planned comparisons (1-tailed paired samples *t*-tests) were performed on these difference scores to examine the effect of removing binocular feedback between the orientation conditions. Lastly, to test our predictions based on changes in the overall precision of feedback between the orientation conditions, we performed planned comparisons between the different orientation conditions (1-tailed paired samples *t*-tests) for the binocular feedback condition, in which both binocular and monocular feedback was available.

Movement velocity analysis

We first examined changes in peak velocity across the three orientation conditions. Figure 4.4 plots the average peak velocity across all participants in the (a) 75-degree, (b) 37.5-degree, and (c) 0-degree conditions, as a function of object distance. Movement velocities scaled in the normal manner with object distance in all conditions (Jeannerod, 1984, 1988) indicating that sufficient information was available to support reaching. Removing binocular feedback resulted in slower peak velocities in all conditions although the overall effect was smallest in the 37.5-degree condition. Planned comparisons showed that removing binocular feedback led to a significant decrease in peak velocity in the 75-degree and 0-degree conditions, but not in the 37.5-degree condition (Table 4.1). Despite these differences, planned comparisons on the difference scores showed that removing binocular feedback did not result in different effects on peak velocity between the different orientation conditions (Table 4.2).

Planned comparisons between the orientation conditions showed peak velocities to be significantly slower in the 75-degree condition, compared to both the 37.5-degree, and 0-degree conditions. No significant difference in peak velocity was found between the 37.5-degree condition and the 0-degree condition (Table 4.3). These results suggest that more precise overall feedback was available as the hand and object approached separation in the frontoparallel plane (Gepshtein & Banks, 2003; Gepshtein et al., 2005; Takahashi et al., 2009), allowing for faster, less conservative reaches. Another possibility is that the reliability of distance information increased as the orientation approached the frontoparallel plane, allowing for faster reaches.

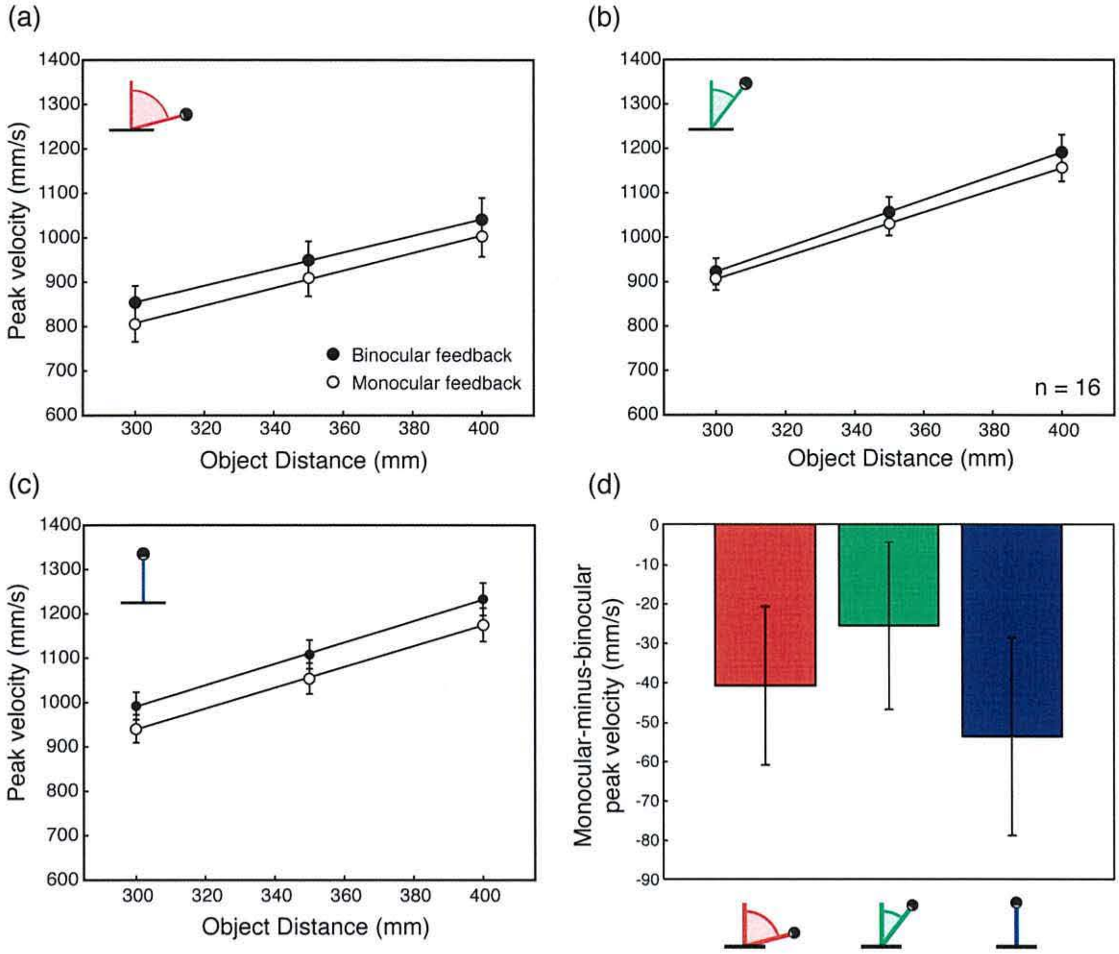


Figure 4.4. Peak velocity in the (a) 75-degree, (b) 37.5-degree, and (c) 0-degree condition as a function of object distance. The solid black circles denote the binocular feedback condition and the open black circles denote the monocular feedback condition. Solid lines show the best fitting linear regressions to the data in each case. (d) The mean difference in velocity between binocular and monocular feedback (monocular-minus-binocular) in each condition. Error bars denote ±1 SEM.

Grip aperture analysis

Figure 4.5 plots the normalised peak grip aperture across all participants in each orientation condition, as a function of object size. Grip apertures were found to scale reliably with object size in all conditions (Jeannerod, 1984, 1988) indicating that sufficient information was available to support grasping. Removing binocular feedback resulted in larger grip apertures in the 75-degree condition but not in the 37.5-degree or 0-degree conditions. Although the trend of changes in grip aperture was in the predicted direction, planned comparisons showed no statistically significant effects of feedback condition in any of the orientation conditions (Table

4.1). Consistent with these findings, planned comparisons on the difference scores found no significant difference in grip aperture between the different orientation conditions (Table 4.2).

When both binocular and monocular feedback was available, planned comparisons showed that peak grip apertures were significantly larger in the 75-degree condition compared to both the 37.5-degree and 0-degree condition. No significant difference in peak grip aperture was observed between the 37.5-degree and 0-degree conditions (Table 4.3). These findings suggest that more precise overall feedback was available as the hand and object approached separation in the frontoparallel plane (Gepshtein & Banks, 2003; Gepshtein et al., 2005; Takahashi et al., 2009), allowing for smaller, less conservative, grip apertures. It is also possible that reliability in the estimates of object properties increased as the orientation approached the frontoparallel plane, allowing for smaller grip apertures.

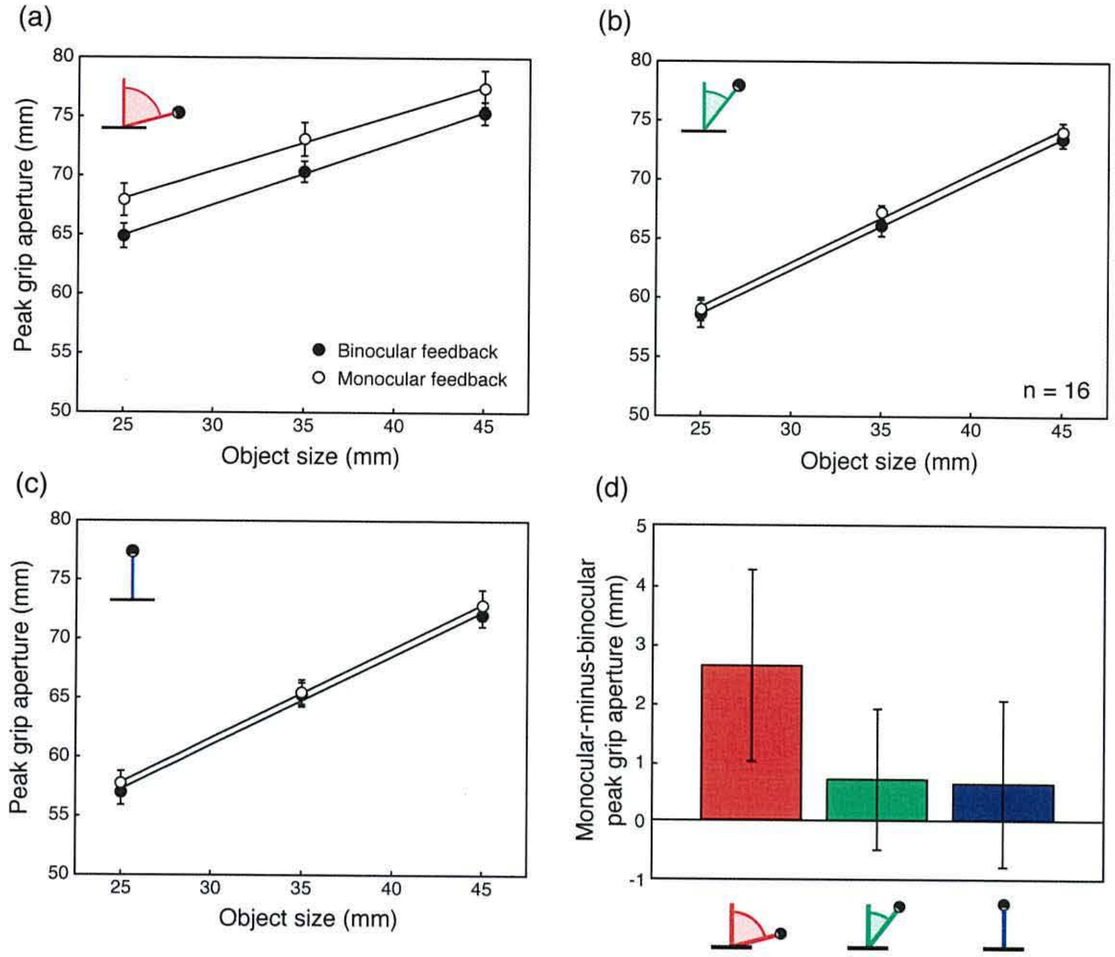


Figure 4.5. Normalised peak grip aperture in the (a) 75-degree, (b) 37.5-degree, and (c) 0-degree condition as a function of object size. The solid circles denote the binocular feedback condition and the open black circles denote the monocular feedback condition. Solid lines show the best fitting linear regressions to the data in each case. (d) The mean difference in grip aperture between binocular and monocular feedback (monocular-minus-binocular) in each condition. Error bars denote ±1 SEM.

Time in the slow phase analysis

Figure 4.6 plots the mean time in the slow phase across all participants in each condition as a function of object distance. Removing binocular feedback resulted in a significantly longer slow phase in the 75-degree and 37.5-degree conditions, but not in the 0-degree condition (Table 4.1). These results suggest that binocular feedback contributed to movements in the 75-degree and 37.5-degree condition but not in the 0-degree condition where the hand and object were separated primarily in the frontoparallel plane. Planned comparisons on the difference scores showed that removing binocular feedback resulted in significantly longer time spent in the slow phase in the 75-degree condition compared to the 0-degree condition. No significant differences were observed between the 75-degree and 37.5-degree conditions or the 37.5-degree and 0-degree conditions (Table 4.2).

In the binocular feedback condition, planned comparisons showed that times spent in the slow phase were significantly longer in the 75-degree condition compared to both the 37.5-degree and 0-degree condition. No significant difference in time spent in the slow phase was observed between the 37.5-degree and 0-degree conditions (Table 4.3). These results suggest that feedback became more precise overall, as the hand and object approached separation in the frontoparallel plane (Gepshtein & Banks, 2003; Gepshtein et al., 2005; Takahashi et al., 2009), allowing for shorter decelerations.

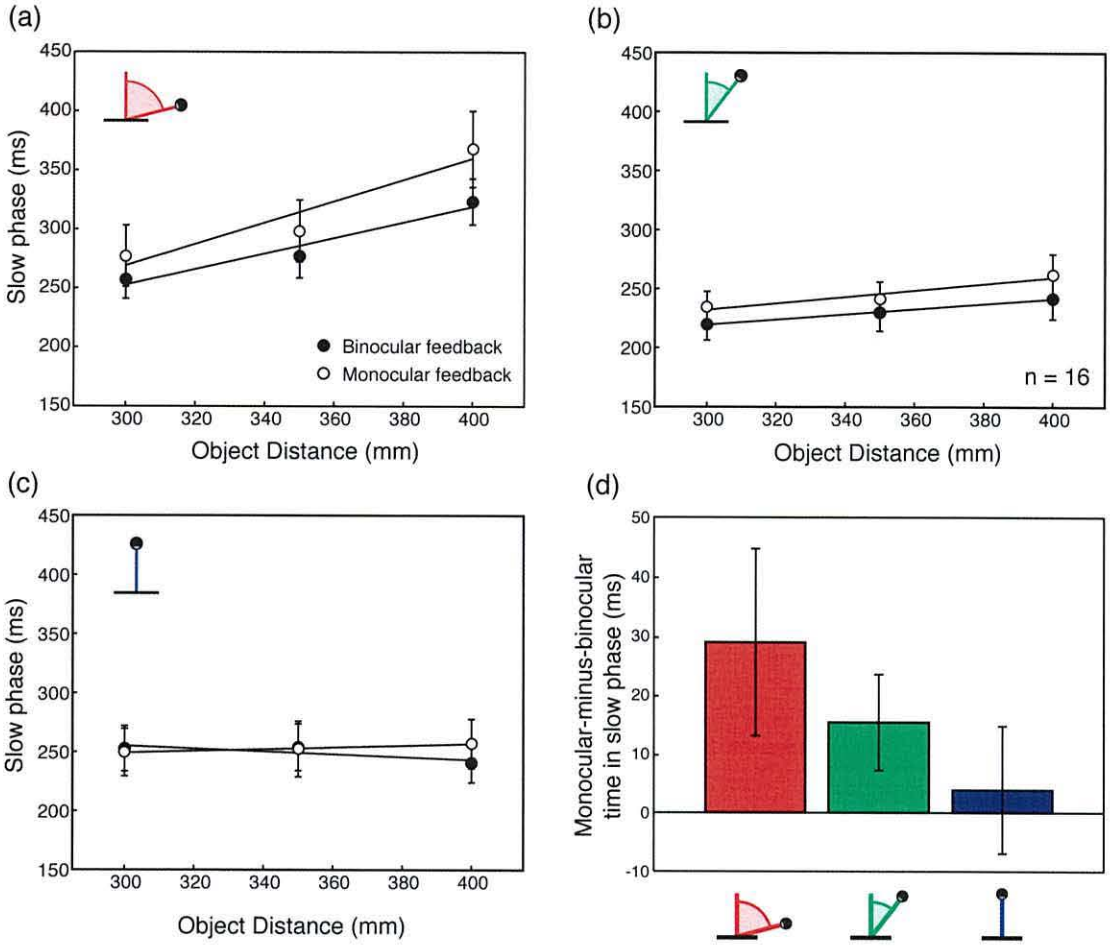


Figure 4.6. Time in the slow phase in the (a) 75-degree, (b) 37.5-degree, and (c) 0-degree condition as a function of distance. The solid black circles denote the binocular feedback condition and the open black circles denote the monocular feedback condition. Solid lines show the best fitting linear regressions to the data in each case. (d) The mean difference in time in the slow phase between binocular and monocular feedback (monocular-minus-binocular) in each condition. Error bars denote ±1 SEM.

Time spent in contact with the object analysis

Figure 4.7 plots the mean time spent in contact with the object across all participants in each condition. Removing binocular feedback resulted in more time spent in contact with the object in the 75-degree and 37.5-degree conditions but not in the 0-degree condition (Table 4.1). These results suggest that binocular feedback contributed to the movement in the 75-degree and 37.5-degree conditions but not in the 0-degree condition where the hand and object were separated primarily in the frontoparallel plane. Planned comparisons on the difference scores showed that removing binocular feedback resulted in significantly longer time spent in contact with the object in both the 75-degree and 37.5-degree conditions, compared to the 0-degree condition. No significant difference was observed between the 75-degree and 37.5-degree conditions (Table 4.2).

Time spent in contact with the object, when both binocular and monocular feedback was available, was significantly longer in the 75-degree condition, compared to both the 37.5-degree and 0-degree conditions. No significant difference in time spent in contact with the object was observed between the 37.5-degree and 0-degree conditions (Table 4.3). These findings suggest that more precise overall feedback became available as the hand and object approached separation in the frontoparallel plane (Gepshtein & Banks, 2003; Gepshtein et al., 2005; Takahashi et al., 2009), allowing for less time spent in contact with the object. Another possibility is that the reliability in the estimates of object properties increased as the orientation approached the frontoparallel plane, allowing for shorter times spent in contact with the object.

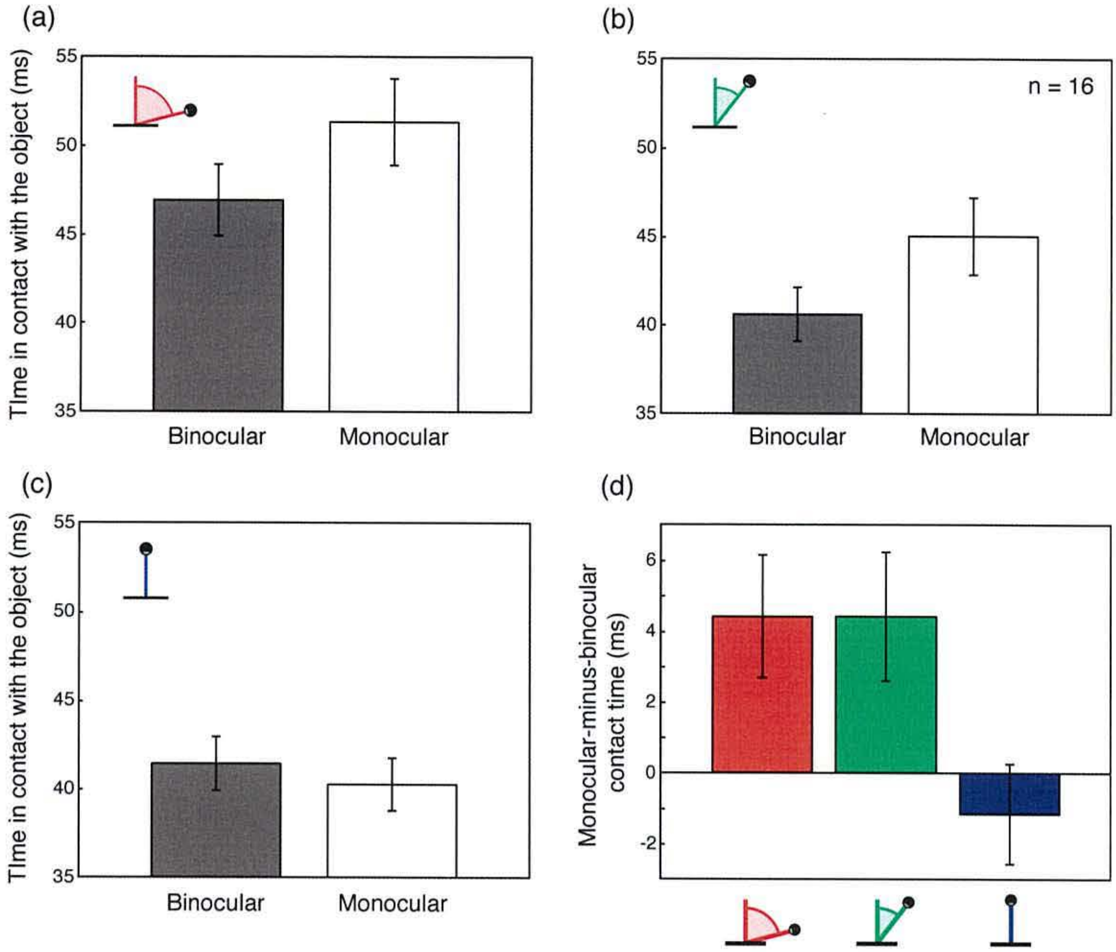


Figure 4.7. Time spent in contact with the object in the (a) 75-degree, (b) 37.5-degree, and (c) 0-degree condition. The blue bars denote the binocular feedback condition and the red bars denote the monocular feedback condition. (d) The mean difference in object contact time between binocular and monocular feedback (monocular-minus-binocular) in each condition. Error bars denote ±1 SEM.

Movement time analysis

Figure 4.8 plots the mean movement time across all participants in each condition. Removing binocular feedback resulted in longer movement times in the 75-degree and 37.5-degree conditions but not in the 0-degree condition (Table 4.1). These results indicate that binocular feedback contributed to movements in the 75-degree and 37.5-degree condition, but not in the 0-degree condition where the hand and object were separated primarily in the frontoparallel plane. However, planned comparisons of the difference scores found that removing binocular feedback did not result in different effects between the conditions (Table 4.2).

During binocular feedback, planned comparisons showed that movement times were significantly slower in the 37.5-degree and 0-degree conditions, compared to the 75-degree condition. No significant difference was observed between the 37.5-degree and 0-degree conditions (Table 4.3). These results suggest that more precise feedback was available as the hand and object approached separation primarily in the frontoparallel plane (Gepshtein & Banks, 2003; Gepshtein et al., 2005; Takahashi et al., 2009), allowing for shorter movement times. It is also possible that the reliability of distance information increased as the orientation approached the frontoparallel plane, allowing for shorter movement times.

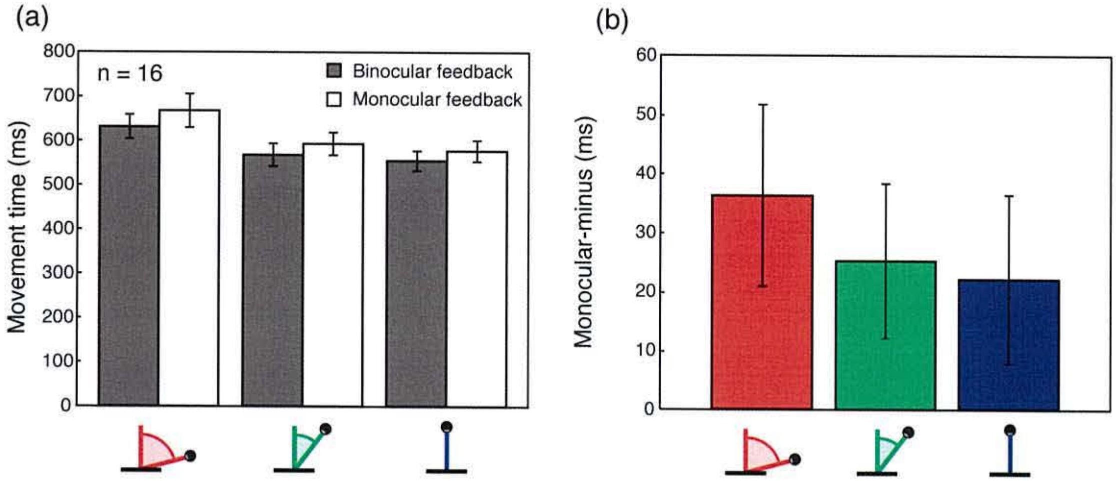


Figure 4.8. (a) Movement time in the 75-degree, 37.5-degree, and 0-degree condition collapsed across object size and distance. Grey bars denote the binocular feedback condition and white bars denote the monocular feedback condition. (b) The mean difference in movement time between binocular and monocular feedback (monocular-minus-binocular) in each condition. Error bars denote ±1 SEM.

Table 4.1. Planned comparisons (1-tailed) to examine the effects of feedback condition. Bin refers to the binocular feedback condition and Mon refers to the monocular feedback condition. Asterisks indicate a significant result.

Effects of feedback condition Planned comparisons (1-tailed).					
	PV	PGA	Slow phase	MT	Time spent in contact with the object
Bin Vs Mon 75-degree	$p < .05$*	$p > .05$	$p < .05$*	$p < .05$*	$p < .05$*
Bin Vs Mon 37.5-degree	$p > .05$	$p > .05$	$p < .05$*	$p < .05$*	$p < .05$*
Bin Vs Mon 0-degree	$p < .05$*	$p > .05$	$p > .05$	$p > .05$	$p > .05$

Table 4.2. Planned comparisons (1-tailed) conducted on difference scores computed between the binocular and monocular feedback conditions (monocular-minus-binocular). Asterisks indicate a significant result.

Difference scores for the feedback condition (monocular-minus-binocular) Planned comparisons (1-tailed).					
Orientation condition comparison	PV	PGA	Slow phase	MT	Time spent in contact with the object
75-degree - 37.5-degree	$p > .05$	$p > .05$	$p > .05$	$p > .05$	$p > .05$
75-degree - 0-degree	$p > .05$	$p > .05$	$p < .05$*	$p > .05$	$p < .05$*
37.5-degree - 0-degree	$p > .05$	$p > .05$	$p > .05$	$p > .05$	$p < .01$*

Table 4.3. Planned comparisons (1-tailed) to examine effects across orientation conditions during binocular feedback. Asterisks indicate a significant result.

	Effects across orientation conditions for binocular feedback Planned comparisons (1-tailed).				
Orientation condition comparison	PV	PGA	Slow phase	MT	Time spent in contact with the object
75-degree - 37.5-degree	$p < .001$*	$p < .01$*	$p < .001$*	$p < .001$*	$p < .001$*
75-degree - 0-degree	$p < .001$*	$p < .01$*	$p < .05$*	$p < .001$*	$p < .01$*
37.5-degree - 0-degree	$p > .05$	$p > .05$	$p > .05$	$p > .05$	$p > .05$

4.9. Discussion

4.10. *Summary of results*

In this experiment we examined the contribution of online binocular and monocular feedback to grasping performance. Changes in viewing geometry were used to manipulate the relative precision of each signal. We recorded visually closed-loop grasping movements and compared performance when binocular feedback was available throughout the reach with when it was removed at movement onset. When the hand and object were separated primarily in the depth dimension (75-degree condition) and binocular feedback was relatively informative, its removal resulted in longer decelerations and longer time spent in contact with the object. However, when the hand and object were separated primarily in the frontoparallel plane (0-degree condition) and binocular feedback was relatively uninformative, its removal did not result in longer decelerations or time spent in contact with the object. These results suggest that the precision of information from different visual depth cues determines their contribution to the online control of grasping. When binocular feedback was relatively uninformative, it was relied upon less for the online control of grasping.

Moreover, online performance was good when only monocular feedback was available. Overall, the findings are consistent with a visuomotor system that integrates information from multiple sources of information according the precision of information they convey (Greenwald and Knill, 2009a, 2009b; Greenwald et al., 2005; Knill, 2005; Knill & Kersten, 2004; Saunders & Knill, 2004, 2005, van Mierlo et al., 2009).

4.11. *No evidence for a binocular specialism in the online control of grasping*

Our findings do not support the claim that the online control of grasping depends on information from binocular cues (Melmoth & Grant, 2006, Watt & Bradshaw, 2000). As found previously, when binocular cues were relatively informative (75-degree condition), their removal resulted in movements with longer decelerations (Bradshaw & Elliott, 2003; Jackson et al., 1997; Servos & Goodale, 1994). However, the removal of binocular cues did not result in a decline in performance when they were relatively uninformative (0-degree condition), indicating that they did not contribute significantly to online control in this condition. Importantly, *overall* performance did not deteriorate under monocular viewing in the 0-degree condition. It is possible that degrading binocular feedback, while increasing the precision of monocular feedback, (0-degree condition) could have resulted in a significant deterioration in performance. This finding would be broadly consistent with the system being designed to rely on binocular cues meaning that it would not be possible to rule out a binocular specialism account for online control. However, when binocular feedback was relatively uninformative (0-degree condition), overall performance (both binocular and monocular feedback available) was improved compared to when binocular feedback was relatively informative (75-degree condition), as shown by improvements in all of the kinematics measured in this study. This finding indicates that in the presence of relatively unreliable binocular feedback, monocular feedback contributed significantly to the performance of the online control of grasping. This is not surprising as viewing geometry tells us that monocular feedback should be relatively informative in the 0-degree condition.

If the visuomotor system were designed to rely solely on relative disparities for online control, we should observe a significant deterioration in performance in conditions in which the relative disparity signal is made uninformative. We observed the opposite effect. In the 0-degree condition, in which relative disparities would have been relatively uninformative, overall performance (both binocular and monocular feedback available) improved compared to the 75-degree condition, in which relative disparities would have been relatively informative. These findings are inconsistent with claims that visuomotor system is specialised to rely on relative disparities (Anderson & Bingham, 2010; Bingham et al., 2001; Bradshaw & Elliot, 2003; Melmoth & Grant, 2006; Morgan, 1989).

Overall, these results, and those of previous studies, show monocular feedback contributes to the online control of grasping, even when binocular feedback is available simultaneously (Greenwald and Knill, 2009a, 2009b; Greenwald et al., 2005; Knill, 2005; Knill & Kersten, 2004; van Mierlo et al., 2009). The binocular-specialism account of grasping implies that different control processes are employed during 'normal' binocularly guided movements, and when the system must 'switch' to relying on monocular input (Marotta et al., 1997). In contrast, our data (and the results of the studies by Knill, Mierlo, and colleagues) suggest that the quality of information available, rather than specific cues, determines grasping performance. This is a fundamentally different account, because only the available information changes in different orientation conditions, and not the underlying control processes.

4.12. *Comparison of observed and expected effects of uncertainty on grasping*

In line with previous findings, we found that the removal of reliable binocular feedback resulted in an extended deceleration phase. This is consistent with the visuomotor system compensating for a reduction in the precision of feedback by allowing more time for online corrections and to gather information, while reducing motor noise (Fitts, 1954; Harris & Wolpert, 1998). Consistent with this idea, we predicted that movement times would increase with the removal of binocular feedback in the 75-degree condition relative to the 0-degree condition. Although there was a significant increase in movement time in the 75-degree condition, but not in the 0-degree condition, we found no significant difference between the conditions as indexed using the difference scores. Given that the time spent in the slow phase did increase significantly between the 75-degree and the 0-degree conditions, we expected to see this reflected by extended movement times. One possible explanation for this is that noise in the data is to account for the lack of a significant increase in movement time with the removal of binocular feedback between the 75-degree and 0-degree conditions. Indeed, the fact that a significant increase in movement time with the removal of binocular feedback is observed in the 75-degree condition, but not in the 0-degree condition suggests that this is the case.

We did not predict to see any significant effects of removing binocular feedback on the early kinematic index of peak velocity due to the time it takes for the visuomotor system to take account of new visual information and effect a change, and due to the fact that the precision of feedback would be poorer early in the reach while the hand was still in peripheral vision. Less precise feedback would be expected to lead to smaller changes upon online control that would be harder to detect. However, contrary to previous findings (Jackson et al., 1997; Servos & Goodale, 1994), we observed a significant reduction in peak velocity with the removal of binocular feedback at movement onset in both the 75-degree and 0-degree conditions. This result may have been due to the fact that trials were blocked by viewing condition, meaning that participants anticipated less precise feedback for online control in the monocular feedback blocks, and therefore programmed slower peak velocities. This is supported by the finding that removing relatively unreliable binocular feedback in the 0-degree condition also resulted in slower peak velocities. However, if this is the case, it is unclear why no effect was observed in the 37.5-degree condition. Another possibility is that binocular feedback made a significant contribution to the reach because its removal led to slower peak velocities. This is consistent with the finding that online feedback has been demonstrated to occur throughout the reach (Saunders & Knill, 2003, 2005) and not only late in the reach as has previously been suggested (Carlton, 1981; Rizzo & Darling, 1997).

In the 37.5-degree condition we predicted that the precision of binocular feedback would be intermediate between the 75-degree and 0-degree conditions, and that removing binocular feedback in this condition would therefore produce effects between those observed in the 75-degree and 0-degree conditions. For changes in the slow phase and movement time, this trend, although not significant, was in the predicted direction.

The overall increase in performance in the 0-degree and 37.5-degree conditions, compared to the 75-degree condition, matched predictions based on psychophysical data that shows that the precision of information increases significantly as the separation of two surfaces slanted in depth approaches the frontoparallel plane (Gepshtein & Banks, 2003; Gepshtein et al., 2005; Takahashi et al., 2009). The finding that overall performance increased as the precision of information increased provides further evidence for a visuomotor system that integrates all sources of information, according to the precision of information they convey, for the online control of grasping (Greenwald and Knill, 2009a, 2009b; Greenwald et al., 2005; Knill, 2005; Knill & Kersten, 2004; Saunders & Knill, 2004, 2005, van Mierlo et al., 2009).

4.13. Conclusions

Overall our results suggest that the contribution of different depth cues for the online control of grasping is determined by the precision of information they convey. The results clearly show that the visuomotor system does not preferentially rely on binocular cues for online control, and instead support the idea that different depth cues are combined, according to their reliability, for the online control of grasping.

Chapter 5: The role of binocular vision in grasping: a small stimulus-set distorts results

5.1. Introduction

The majority of studies examining the role of binocular vision in grasping have compared movement kinematics (maximum velocity, maximum grip aperture etc.) under normal binocular viewing and when binocular information is removed by covering one eye (Jackson et al., 1997; Keefe & Watt, 2009; Loftus et al., 2004; Melmoth & Grant, 2006; Melmoth et al., 2009; Servos et al., 1992; Servos & Goodale, 1994; Watt & Bradshaw, 2000; for a review, see Melmoth & Grant, 2006). The above experimental design has been directly adopted from "classical" perception experiments, which have examined the effects on depth perception either of subtracting specific depth cues from the visual scene, or of introducing them into an impoverished scene (e.g. Künnapas, 1968; see Sedgewick, 1986). The logic is straightforward: the extent to which adding or removing a specific cue causes a change in bias or sensitivity is thought to index the role that cue plays in recovering depth. In order to interpret results unambiguously, however, great care must be taken to eliminate sources of depth information that are not under experimental control, and stimulus artefacts that can 'artificially' improve performance. In this paper we highlight ways in which the design of typical grasping experiments differs from that of typical perception experiments, with a resulting reduction in experimental control. We examine the consequences of this for a specific example: investigating the role of binocular information in providing the initial estimate of object properties to programme a grasp.

Depth stimuli potentially contain many geometrical artefacts, which could allow an observer to complete the task using information other than the depth cue(s) under study. For example, consider a simple experiment in which an observer must indicate the perceived slant of a planar stimulus. The angular size of a plane of fixed dimensions varies directly as a function of its slant. Observers could therefore base their judgements of slant magnitude directly on the size at the retina (or size on the screen) of the stimulus, without recovering slant, since there is a one-to-one mapping between angular size and slant magnitude. To control for this, experimenters typically either match the angular size of the stimulus on each trial, or vary it randomly (within a certain range), so as to render this signal uninformative (Hillis et al., 2004; Knill & Saunders, 2003).

Psychophysical depth perception studies routinely employ such ‘controls’ to eliminate stimulus artefacts. However, such measures are extremely uncommon in studies of visually guided grasping. Almost all grasping studies use a small set of widely-spaced object distances and sizes—rarely more than three of each, and typically fewer—and the dimensions that are not manipulated experimentally (object height, for example) are of a fixed physical size. This means that low-level image properties vary directly with object properties in the manner described above. For example, since object width is normally held constant, an object’s width at the retina uniquely specifies its distance. Moreover, each object × distance combination is likely to be unique in terms of low-level image properties such as retinal size so, in principle, these simple cues uniquely specify which stimulus configuration is being presented on a given trial.

These stimulus artefacts may be particularly problematic in grasping experiments, because of another difference from classical perception experiments: the provision of feedback about performance. In depth perception experiments observers usually receive no feedback about the 'correctness' of their responses. In contrast, participants in reach-to-grasp experiments almost always successfully grasp the stimulus, thereby receiving kinaesthetic feedback about the true size and distance of the stimulus on every trial. Over repeated trials the consistent relationship between simple low-level image properties and stimulus parameters could therefore be learned.

Even if this lack of control over low-level stimulus properties is overcome (by matching retinal size etc.) learning of the stimulus set via kinaesthetic signals could still present problems. If the visuo-motor system has learned the sizes and distances used in the study, the intended task of recovering an object's properties from the available depth cues is reduced to *identifying* which object × distance combination is presented on a given trial. Stimuli in grasping studies are normally widely spaced along the dimensions of interest (object distances typically differ by at least 100-150 mm, and object sizes by 15 mm or more). So even rather imprecise visual information, which on its own would result in poor estimates of object properties, could be sufficient to identify whether the stimulus is the "small", "medium" or "large" object, for example, allowing its properties to be retrieved from memory[5.1.]. Essentially, the stimuli could become familiar objects (and distances), through the course of the experiment. Familiar size is known to be an effective depth cue (Holway & Boring, 1941; O'Leary & Wallach, 1980), and information from familiar size has previously been shown to contribute to the control of grasping movements (Marotta & Goodale, 2001; McIntosh & Lashley, 2008).

[5.1.] In Experiment 3.1, which uses the same stimuli as this experiment, we have found that just-noticeable differences in object depth under monocular viewing are typically 2-3 mm.

The above discussion illustrates ways in which learning the stimulus set could provide an additional, unintended source of information about stimulus properties. What are the likely effects of this? Since the intended (binocular and monocular) depth cues remain available on each trial, it is informative to consider how multiple sources of information are combined to estimate depth. Psychophysical studies have shown that the combined depth estimate from *n* independent information sources is well approximated by a weighted sum of depth estimates from each, where each signal is weighted according to its reliability (the reciprocal of its variance) (e.g. Ernst & Banks, 2002). A key feature of this cue combination 'strategy' is that the variance, and therefore the uncertainty, of the resulting depth estimate is always increased by removing signals. Assuming the noise in each signal is independent and Gaussian, the variance of the combined estimate ($\sigma^2_{\text{combined}}$) is given by Equation 4 (see Oruç et al. 2003 for a derivation):

$$\sigma^2_{\text{combined}} = \frac{\prod_n \sigma_n^2}{\sum_n \sigma_n^2} \tag{4}$$

As demonstrated in Chapter 3, this general principle has also been shown to account well for visuo-motor performance (Greenwald and Knill, 2009a, 2009b; Greenwald et al., 2005; Knill, 2005; Knill & Kersten, 2004; van Mierlo et al., 2009). In response to increases in perceptual uncertainty the visumotor system adds a margin of error to the grasp by programming larger grip apertures (Christopoulos & Schrater, 2009; Schlicht and Schrater, 2007, see Chapter 3). Using the reliability-based cue-weighting framework as a normative model, we can therefore make general predictions about the effects on grip apertures of removing binocular cues in the presence and absence of learned stimulus information, by considering the changes in variance (uncertainty) of depth estimates in each case.

Figure 5.1 plots changes in the variance of hypothetical depth estimates when binocular information is removed, both when learned information is available (left-most column) and when it is not (middle column, see caption for details). In the top row, we assume the simplest case in which all three sources of information have equal variance ($\sigma^2_{\text{monoc}} = \sigma^2_{\text{binoc}} = \sigma^2_{\text{learned}}$). Removing binocular information always increases the variance of the depth estimate (plot c; Equation 4), but this effect is smaller when learned information is also available. Therefore we would expect the removal of binocular information to result in a smaller increase in grip apertures when learned information is available than when it is not (Christopoulos & Schrater, 2009; Schlicht & Schrater, 2007, see Chapter 3). The bottom row in Figure 5.1 models the (arguably more realistic) case in which monocular cues are relatively unreliable ($\sigma^2_{\text{monoc}} = 2\sigma^2_{\text{binoc}} = 2\sigma^2_{\text{learned}}$). Here the confounding effect of learned stimulus information is greater. Removing binocular information results in a much larger increase in variance when learned stimulus information is unavailable, and so would be expected to have a much larger effect on grip apertures under such conditions. The availability of learned stimulus information would therefore be expected to cause misestimation of the effects of removing binocular information.

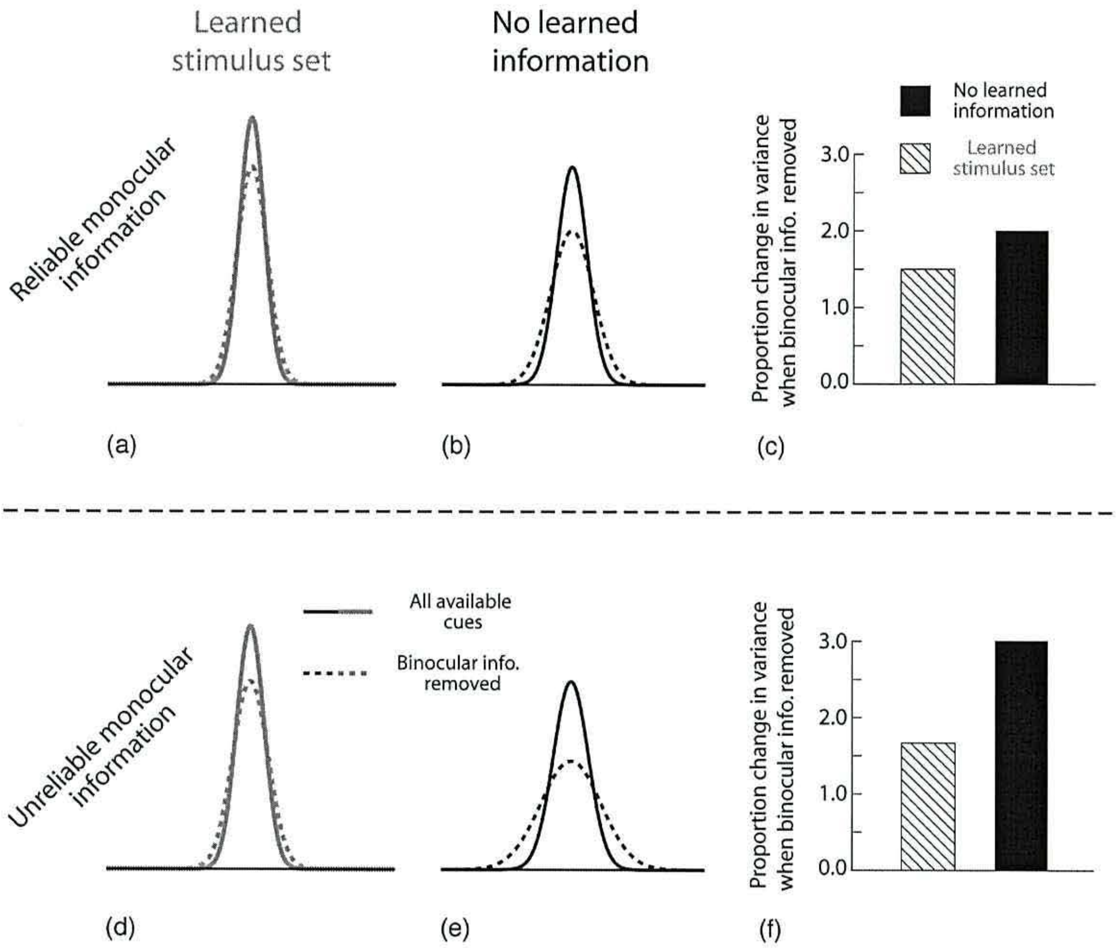


Figure 5.1. Changes in the variance of estimates of object properties when binocular information is removed. (a) Hypothetical probability density functions for an estimate of object size when learned information is available. The solid line shows the estimate when all cues are available, and the dashed line shows the estimate when binocular information is removed. (b) Similar probability density functions for an estimate of size from binocular and monocular information only, when learned stimulus information is unavailable. (c) The proportion increase in the variance of size estimates that results from removal of binocular information in each case. The top row models a case in which the reliability (inverse variance) of all three information sources is equal ($\sigma^2_{\text{monoc}} = \sigma^2_{\text{binoc}} = \sigma^2_{\text{learned}}$). The bottom row (d, e, f) plots the same information for a case in which monocular cues are relatively unreliable ($\sigma^2_{\text{monoc}} = 2\sigma^2_{\text{binoc}} = 2\sigma^2_{\text{learned}}$). Probability density functions were computed using Equation 4.

We examined this prediction directly by comparing the effects of removing binocular information on grasp kinematics in two conditions: (i) a conventional “small stimulus-set” condition, and (ii) a “randomised stimulus-set” condition, designed to prevent learning. In the small stimulus-set condition, participants reached to three fixed object sizes at three fixed distances, so learning could occur in the manner described above. In the randomised stimulus-set condition, each trial was different. We used a near-continuous range of stimulus properties, and there was no consistent relationship between low-level image features and object properties. Therefore participants could do the task only by using information available from the intended binocular and/or monocular depth cues presented on a given trial. We examined visually open-loop grasping to explore the contribution of binocular vision to the initial estimate of object properties, used to programme the movement. The removal of binocular cues during open-loop grasping often results in rather small effects (Melmoth & Grant, 2006). If this is caused, at least in part, by learning of the stimuli, we should see a clear effect of removing binocular information when stimulus learning is prevented (randomised stimulus set), allowing us to determine unambiguously the contribution of binocular information in this initial phase.

5.2. Methods

Participants

Separate groups of 14 right-handed participants completed the small (7 female, 7 male; aged 19-35) and randomised (10 female, 4 male; aged 19-43) stimulus-set conditions, in return for payment. All had normal or corrected to normal vision and stereoacuity < 40 arcsec. Participants gave informed consent prior to taking part, and all procedures were in accordance with the Declaration of Helsinki.

Apparatus and stimuli

The experimental setup and an example of the stimuli are shown in Figure 5.2. A chin rest was used to stabilise head position and to locate the participants' eyes 400 mm above a table surface, directly above a small "start button" fixed to the table on the body midline. Each participant was carefully positioned, and his or her inter-pupillary distance taken into account, to ensure that the visual stimuli were geometrically correct. The visual stimuli were virtual objects, presented on a TFT monitor, and drawn stereoscopically using red-green anaglyph glasses. The monitor was placed face down, 500 mm above the table, and participants viewed a reflection of the stimuli in a horizontal mirror, placed so that the monitor surface was optically coincident with the table surface. The mirror occluded the participants' hands.

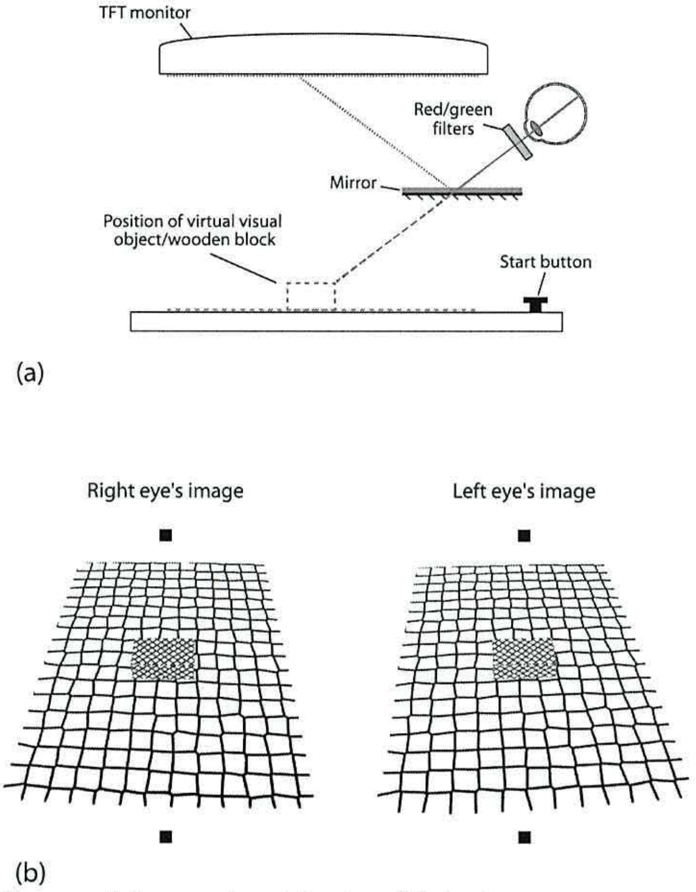


Figure 5.2. (a) A diagram of the experimental setup. (b) A stereogram representative of the visual stimuli. Cross-fuse to view. In monocular conditions only the right eye's image was presented.

The visual stimuli were rectangular blocks, lying on a textured ground plane coincident with the table surface. The ground plane and object surfaces were defined by the perspective projection of Voronoi diagrams (de Berg et al., 2000; see Figure 5.2b), which have been shown to provide effective binocular and monocular depth cues (Hillis et al., 2004; Knill & Saunders, 2003). We used anti-aliasing to position stimulus elements with subpixel accuracy. Participants wore anaglyph glasses throughout the experiment, but in the monocular conditions only the right eye's image was drawn, and a patch covered the left eye. Both experiments were completed in the dark, so only the stimulus was visible. Participants grasped real wooden blocks, the relevant dimensions of which exactly matched the depth and distance of the visual objects.

Procedure

On each trial participants held down the start button with their thumb and index finger. The visual stimulus was then displayed, and after 2 sec there was an audible beep, which was the signal to grasp the object. Participants were instructed to pick up the object front-to-back, quickly and naturally, using their thumb and index finger only. Releasing the start button extinguished the visual stimulus (the hand was never visible). Movements initiated before the start signal, or more than 600 ms afterwards, were considered void and were repeated at the end of the block.

The apparatus, stimuli, and procedure were identical in the small and randomised stimulus-set conditions with the exception of differences in the exact stimulus parameters used, and the numbers of repetitions. Different participants completed the small stimulus-set, and randomised stimulus-set conditions.

Small stimulus-set condition

The small stimulus set consisted of three object distances (200, 350, & 500 mm from the start button) and three object depths (30, 45, & 60 mm). The visual objects were always 60 mm wide and 25 mm high, and were presented along the body midline. Participants completed six repetitions of each object-distance combination, blocked by viewing condition (binocular or monocular) (6 x 3 x 3 x 2 = 108 trials). Trial order was randomised within blocks.

Randomised stimulus-set condition

The randomised stimulus set consisted of six 'base' object distances (200, 250, 300, 350, 400 & 450 mm), and on each trial a random distance (uniform distribution) between 0 and 30 mm was added to the base distance. It was not practical to fully randomise object depth because an equivalent real object was to be grasped. Instead we used a (large) set of seven object depths (30, 35, 40, 45, 50, 55 & 60 mm). We also randomised the width of the visual object on each trial in the range 60 ± 20 mm, to prevent the retinal image size of the object, or its aspect ratio, serving as a simple cue to distance and/or depth. In this condition, participants completed just one repetition of each combination, and trials were again blocked by viewing condition (1 x 7 x 6 x 2 = 84 trials). Therefore, in addition to the randomising of object parameters, they were never presented with the same object/distance combination twice. Trial order was randomised within blocks.

5.3. Results

Many kinematic indices can be computed to characterise grasping movements, and they are often highly correlated. Here we restrict our analysis primarily to peak grip aperture, because our predictions refer specifically to this measure. We also computed peak velocity of the movements (by measuring the three-dimensional velocity of the thumb marker), since this is a primary measure of the transport, or reach component of the movement (Jeannerod, 1984, 1988), and has sometimes, though not always, been found to reduce significantly when binocular information is removed (Jackson et al., 1997; Loftus et al., 2004; Melmoth & Grant, 2006; Servos et al., 1992; Watt & Bradshaw, 2000). Estimates of kinematic parameters for 2 of the 14 participants in the small-set condition were extremely unreliable and so these participants' data were excluded from the analyses.

Movement velocity analysis

The grasping system could compensate for increased perceptual uncertainty not only by increasing grip apertures, but also by reaching more slowly (Harris & Wolpert, 1998; Wing, Turton & Fraser, 1986). If reach velocity and grip aperture are 'traded-off' differently across different depth-cue conditions changes in grip aperture (the margin-for-error) could not be attributed unambiguously to perceptual uncertainty. Before analysing grip apertures, we tested for this by comparing peak velocity of the movements in each cue condition. Peak velocity data for each condition are plotted in Figure 5.3. Figure 5.3a plots the mean peak velocities for all participants in the small-set condition, as a function of object distance. The solid lines are linear regression fits to the data in each case. Figure 5.3b plots the data for the participants in the randomised condition. Because in this condition each participant reached to a different set of distances, we first summarised his or her data by putting the object distances into 50 mm 'bins'. We then computed the mean distance and peak velocity for each bin, before computing overall means across participants. Note that the data in Figure 5.3b therefore include variance due to different participants reaching to different distances, and so the error bars are not directly comparable to those in Figure 5.3a. Figure 5.3 shows that reach velocity scaled reliably with object distance, in the normal fashion, under all viewing conditions (Jeanerrod, 1984, 1988), indicating that the information available in each condition was sufficient to support reliable reaching. However, the figure clearly suggests that removing binocular information did not affect peak velocity in either the small-set, or randomized condition. For the small-set condition a 2 × 3 (depth-cue condition × distance) analysis of variance confirmed that there was a significant main effect of distance ($F_{(2,22)} = 766.89$, $p < 0.001$), but no main effect of depth-cue condition ($F_{(1,11)} = 3.79$, $p > 0.05$), and no significant depth cue × distance interaction ($F_{(2,33)} = .32$, $p > 0.05$). In the randomised condition a 2 × 6 (depth-cue condition × distance) analysis of variance showed a significant main effect of distance ($F_{(1.3,16.9)} = 207.89$, $p < 0.001$, Greenhouse-Geisser corrected), but no main effect of depth-cue condition ($F_{(1,13)} = .17$, $p > 0.05$), and no significant depth cue × distance interaction ($F_{(2.1,27.9)} = .83$, $p > 0.05$, Greenhouse-Geisser corrected). These findings confirm that it is meaningful to compare grip-aperture effects across depth-cue conditions, as the removal of binocular cues did not significantly affect peak velocity in either the small-set or randomised condition.

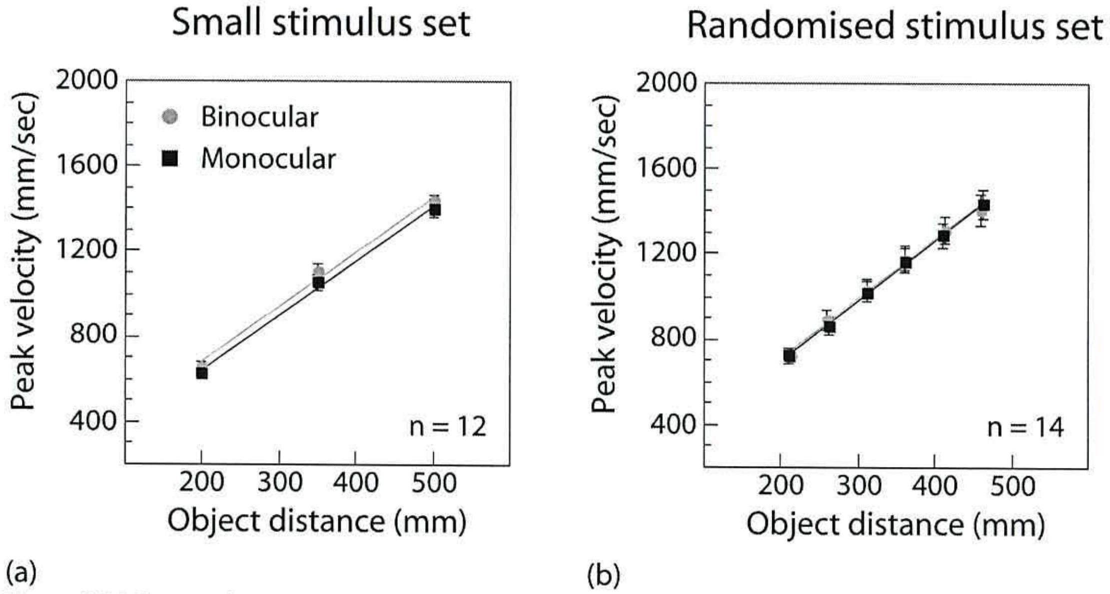
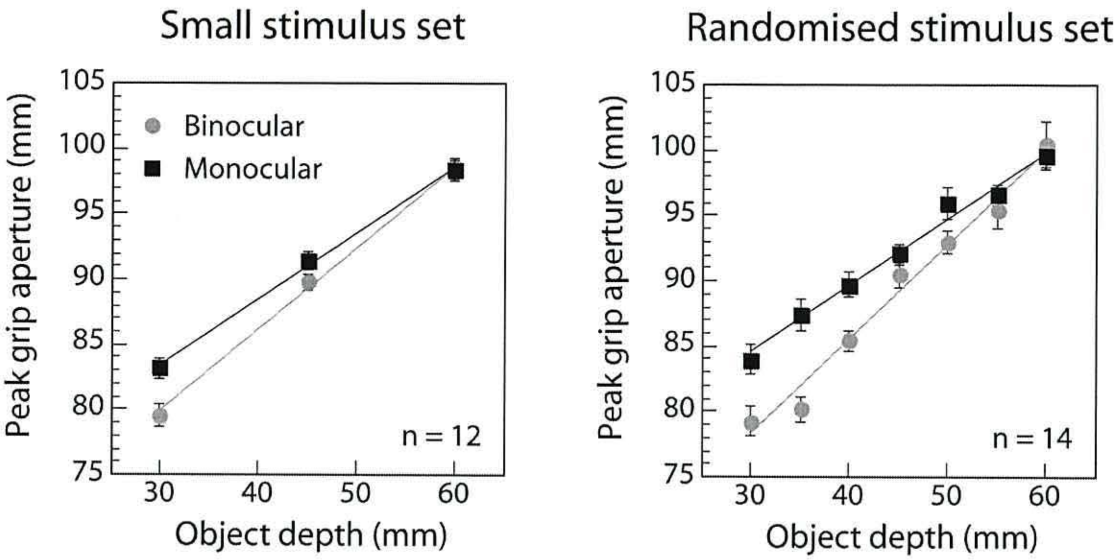


Figure 5.3 Mean peak movement velocity as a function of object distance in (a) the small stimulus-set condition, and (b) the randomised stimulus set condition. In the randomised condition, object distances were put in 50 mm bins, and the average distance and velocity were computed for each bin (see main text for details). The grey circles denote the binocular condition and the black squares denote the monocular condition. The solid lines are the best-fitting linear regression to the data in each case. The error bars indicate ± 1 SEM.

Peak grip aperture analysis

To account for differences in grip apertures across individuals (due to differences in hand sizes, for example) we normalised peak grip apertures by expressing them as a proportion of each participant's average peak grip aperture across all trials. For ease of interpretation, and to allow comparisons with previous work, we multiplied these proportion values by the average peak grip aperture across all participants, within the small-set, and randomised conditions, to give normalised grip apertures in millimetres.

Figures 5.4a and 5.4b plot the mean peak grip aperture for all participants in binocular and monocular viewing conditions for the small stimulus set, and randomised set, respectively, as a function of object depth. The solid lines denote the best-fitting linear regressions in each case. It can be seen that in all conditions peak grip apertures increased linearly with object depth in the normal fashion (Jeanerrod, 1984, 1988). This confirms that grasping movements remained reliable under all conditions. It can also be seen, however, that removing binocular information lead to a clear change in peak grip apertures in both the small stimulus-set, and randomised conditions, and that as predicted, this effect was larger in the randomised conditions.



(a) (b)

Figure 5.4. Mean peak grip apertures as a function of object depth in (a) the small stimulus-set condition, and (b) the randomised stimulus set condition. The grey circles denote the binocular condition and the black squares denote the monocular condition. The solid lines are the best-fitting linear regression to the data in each case. The error bars indicate ± 1 SEM.

To test our specific predictions statistically, we computed the best-fitting linear regressions to each participant's grip aperture data, and carried out statistical analysis on these fit parameters. Peak grip apertures are linearly related to object size (Jeannerod, 1984, 1988), and using these 'grip aperture scaling functions' we can reliably characterise the overall pattern of each individual's grip apertures, and directly compare grasps made to the different stimulus sets. Figure 5.4 shows that in both the small and the randomised conditions the effect of removing binocular information was largest for small object sizes, and progressively decreased to zero with increasing object size. This pattern of data has been observed previously (e.g. Watt & Bradshaw, 2000, see Chapter 3), and presumably reflects the fact that, as grip aperture increases overall, progressively increased effort is required to open it yet wider. This effect is also likely to be more pronounced during open-loop grasping, because grip apertures are larger overall when visual feedback is prevented (Jakobson & Goodale, 1991). Therefore although one might expect the principal effect of removing binocular information to be an increase in the intercept of the grip aperture scaling function (i.e. larger grip apertures), it should also be expected to lead to a reduction in its slope.

Figures 5.5a and 5.5b plot the mean slopes and intercepts, respectively, of the linear regression fits to each individual's data in the small-set and randomised-set conditions. Planned pair-wise *t*-tests confirmed that removing binocular information had a significant effect on grip aperture scaling functions in both stimulus-set conditions. In the small-set condition, the slope of the grip scaling function was significantly less under monocular viewing ($t(11) = 6.33$, $p < .001$, one-tailed), and the intercept was significantly higher ($t(11) = 5.50$, $p < .001$, one-tailed). The same pattern of results was observed for the randomised-set condition (slope: $t(13) = 4.54$, $p < .001$; intercept: $t(13) = 5.49$, $p < .001$; both one-tailed). Figure 5.5 also shows that, in line with our predictions, removing binocular information resulted in a greater *change* in grip aperture scaling functions in the randomised condition. To test this statistically we computed the difference in (i) slope, and (ii) intercept in the monocular and binocular conditions for each participant (i.e. the effect on each participant of removing binocular information). We then ran independent-measures *t*-tests comparing this effect in the small and randomised conditions, for each regression parameter. This analysis confirmed that removing binocular information resulted in a significantly larger change in both slopes ($t(24) = 1.74$, $p < .05$, one-tailed) and intercepts ($t(24) = 2.01$, $p < .05$, one-tailed) of the grip aperture scaling functions in the randomised stimulus-set condition.

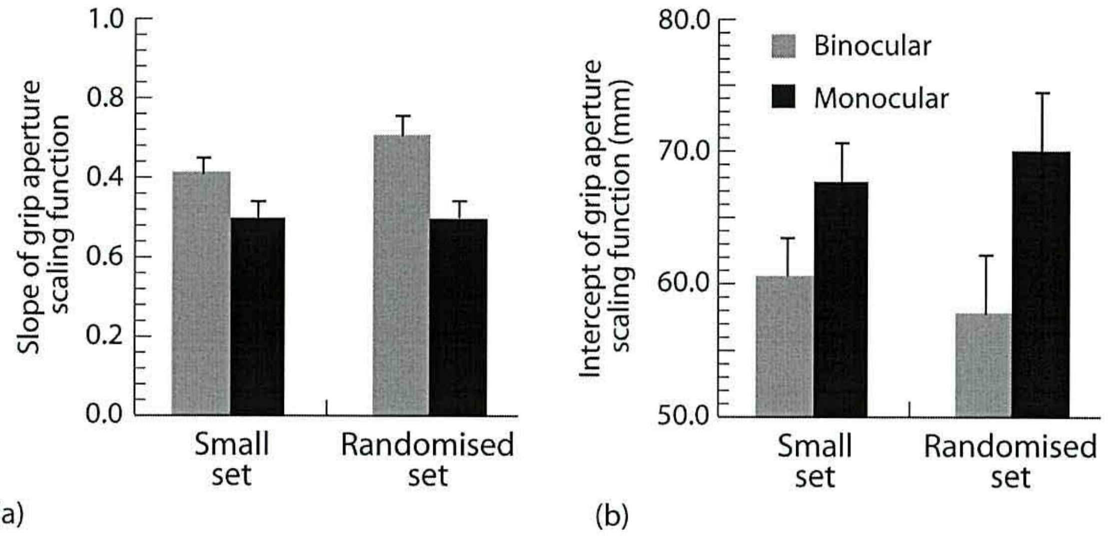


Figure 5.5. (a) The mean of the slopes for the best-fitting linear regressions to each participant's grip aperture scaling function, under each experimental condition. (b) The mean intercept of the same fitted curves. Grey bars denote binocular viewing and black bars denote monocular viewing. The error bars are ± 1 SEM.

The magnitude of the effects of removing binocular information

The analysis of the regression parameters, above, demonstrates that removing binocular information caused a significantly larger change in grip aperture scaling functions when the stimulus could not be learned. How does this relate to actual grip apertures? We used the fit parameters of each individual's grip aperture scaling function to reliably estimate their peak grip aperture for a 30 mm object depth. In the small-set condition the average increase in grip apertures with removal of binocular information was 3.58 mm, compared to 6.27 mm, in the randomised condition. That is, removing binocular information had a 75% greater effect on grip apertures when the stimulus set could not be learned.

Isolating the effect of learning the stimulus set

As described above, increasing object size, and so increasing grip apertures overall, leads to a reduction in the effects of removing information (see Figure 5.4). Similar effects are to be expected with increasing object distance, because grip apertures also generally increase with object distance (Loftus et al., 2004; Watt & Bradshaw, 2000), presumably to compensate for increased variability in movement endpoint with larger motor responses (Harris & Wolpert, 1998). An analysis of grasps made to the subset of near and small objects should therefore yield the least biased estimate of the effect of removing binocular information, allowing us to quantify the effects of learning *per se* more accurately.

Figure 5.6 plots mean peak grip apertures for grasps made to the "near/small" subset of trials. For the small-set condition, we included data from all grasps to the smallest object (30 mm depth) at the near distance (200 mm). For the randomized condition, in order that a sufficient number of trials contributed to the analysis, we included grasps to 30, 35 and 40 mm object depths at a distance of < 250 mm (mean object distance with binocular and monocular viewing was very similar: 217 and 215 mm, respectively). Note that the average object depth and distance in the randomised condition (35 and 216 mm, respectively) were therefore both slightly larger than in the small-set condition.

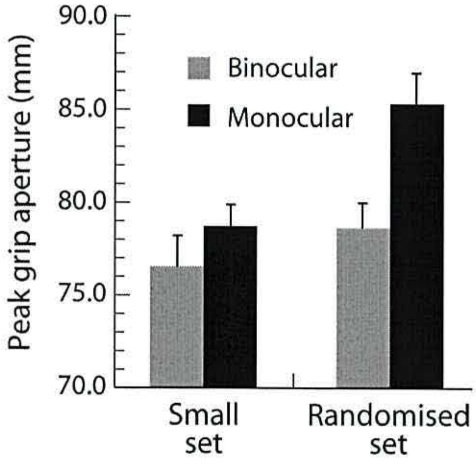


Figure 5.6. Mean peak grip apertures for grasps made to the subset of small/near objects. See main text for details. Error bars denote ± 1 SEM.

The overall pattern of data in Figure 5.6 is consistent with our predictions. In the small-set condition, removal of binocular information resulted in a small increase in grip apertures but this did not reach statistical significance ($t(11) = 1.68$, p = .06, one-tailed). In the randomised set, however, grip apertures were significantly larger under monocular viewing (*t*(13) = 3.13, p < .01, one-tailed). Moreover, an independent *t*-test comparing monocular-minus-binocular "difference scores" in each condition showed that the effect of removing binocular information was significantly larger in the randomised condition (t(24) = 1.80, p < .05, one-tailed). Therefore, when the confounding effects of large overall grip apertures were minimised, the effect of learned stimulus information was substantial. Removing binocular information resulted in a three times larger increase in grip apertures with the randomised set than with the small set (6.70 mm vs. 2.19 mm).

Testing the assumption that learning occurred

The above analysis assumes that the differences between conditions are attributable to learning the stimulus set. This seems reasonable, given that the experimental conditions differed only in terms of the small versus randomised stimulus sets used. Nonetheless, we also looked for direct evidence of learning, which we would expect to see in our small-set condition (but not in our randomised condition). Learning the stimulus set should reduce uncertainty, leading to reduced grip apertures. Under monocular viewing, when learning should have the largest effect, we would therefore expect to see a reduction in grip apertures during the experiment.

To examine this we compared peak grip apertures at the beginning and end of the experiment in both monocular conditions. Given the dependence of the experimental effects on object size we again analysed grasps to the small objects (defined as above). In order to include enough trials to reliably estimate grip apertures, we analysed the first six, and last six grasps to the small objects, across all distances. In the monocular/small-set condition peak grip apertures were significantly smaller at the end of the experiment than at the beginning (mean reduction of 1.8 mm; $t(11) = 2.40$, $p < .05$), suggesting that uncertainty in estimates of object properties reduced following repeated exposure to the same stimuli. In the monocular/randomised-set condition there was no significant change in grip apertures across the experiment (mean increase of 1.2 mm; $t(13) = 0.82$, $p = .43$). Taken together, these results suggest that learning of the stimulus set did occur in the small-set condition, but not in the randomised condition.

5.4. Discussion

According to the model of sensory integration set out in the Introduction, since closing an eye removes information, it should always result in increased uncertainty in estimates of object properties. This in turn would lead to increased grip apertures, in order to build an increased 'safety margin' into the movement (Christopoulos & Schrater, 2009; Schlicht and Schrater, 2007, see Chapter 3). All else being equal, however, removing one of two cues will have a greater effect than removing one of three. So we expected to see a larger effect of removing binocular information when learned information was unavailable (leaving only monocular cues) than when it was available (leaving monocular cues *plus* learned stimulus information). The pattern of data we observed was consistent with these predictions. In both small and randomised stimulus-set conditions, removing binocular information lead to significant changes in grip aperture scaling functions (larger grip apertures for grasps to small objects, and reduced grip aperture scaling with object size). However, this effect was significantly larger in the randomised condition, when the stimuli could not be learned, than in the conventional small-set condition. We also found more direct evidence that participants learned the small stimulus set through the course of the experiment: in the monocular/small-set condition only, grip apertures were significantly smaller at the end of the experiment than at the beginning, consistent with a reduction in uncertainty about object properties. Taken together, these results suggest that participants do learn the properties of small stimulus sets in typical grasping experiments, and that this (uncontrolled) information contributes to grasping performance. Similar to previous findings (e.g. Watt & Bradshaw, 2000, see Chapter 3, Chapter 6), effects of removing binocular information reduced with increasing object size (and distance), and were absent for grasps to the largest objects. We attribute this to the increased effort required to open the grasp yet wider, as overall grip aperture increases. The analysis of grasps to near and small objects, which should best isolate the effect of learning during the experiment (Figure 5.6), suggests that the contribution of learned stimulus information to grasping performance can be substantial.

One obvious implication of our findings is that the effects of removing binocular information may previously have been underestimated. Another way to state this is that 'monocular' grasping performance has been overestimated, because monocular cues were not presented in isolation. To draw this conclusion definitively it is useful to consider the magnitude of our effects in comparison to similar studies. In their Experiment 3, Jackson et al. (1997) measured grasps to one 22.5 mm object presented at three distances, and found a grip aperture increase of ~3.2 mm when binocular information was removed[5.2.]. In Loftus et al.'s (2004) Experiment 4, participants grasped a 25 mm object at three distances, and grip apertures increased by 3.0 mm

---

[5.2.] Jackson et al. (1997) reported peak grip apertures in terms of the angle formed by the finger, wrist and thumb markers. We estimated the conversion into millimetres using the average hand size of our observers.

when binocular information was removed. The effect we observed—an increase of a 3.58 mm for grasps to the 30 mm object (across all distances)—was therefore not atypical. By comparison, in our randomised condition, in which participants could not learn the stimulus set, grip apertures at 30 mm object size increased by 6.27 mm (a 75% greater increase), suggesting that previous studies have indeed underestimated the effect on grip aperture of removing binocular information.

The effect observed in our randomised condition clearly suggests that binocular information makes a significant contribution to the initial estimate of object properties used to programme a grasp (see Chapter 3). However, as we demonstrated in Chapter 3, this finding does not imply a binocular specialism for grasping. Reliability-based cue weighting predicts that removing *any* reliable source of information will affect grasp kinematics. So observing that removing a given cue leads to a decrement in performance does not, on its own, constitute evidence for that cue having a particular or unique role over any others (see Chapter 3). Furthermore, since we restricted the available depth cues—information was available only from texture/perspective, binocular disparity and convergence—we may have observed a greater effect of removing binocular information than would be expected when the normal range of cues is available. For example, information from motion parallax has been shown to contribute to the control of natural grasping (Marotta et al., 1998), but was unavailable in our experiment because head movements were prevented.

One potentially surprising aspect of our data was the finding that removing binocular information had no effect on peak velocities in either stimulus-set condition. This accords with previous findings (Jackson et al., 1997; Loftus et al., 2004; see Chapter

3, Chapter 6, but see Bingham & Pagano, 1998). In principle, however, one might expect increased uncertainty about object distance to lead to slower, more "cautious" reaches (i.e. reduced peak velocity). The fact that we did not find this in our randomised condition suggests that the lack of effects in previous studies cannot be attributed to learning of distance information. It should be noted, however, that various other aspects of the stimuli were invariant across trials, possibly introducing further uncontrolled information. The stimuli were always presented along the midline, on the same table surface, and eye-height was fixed above the table. If these parameters were learned, variations in the height-in-scene of the objects could have been a very reliable cue to distance. In which case monocularly available cues to location may have been sufficiently reliable that overall precision was not affected significantly by removing binocular cues.

The effects we have reported indicate that learned stimulus information, acquired during the experiment, contributes to grasp planning. This finding is consistent with the more general idea that grasping is not planned solely on the basis of visual "measurements" made at a given instant, but also on the basis of previously learned information about objects. Previous studies have examined this issue directly by examining the effects on grasping movements of presenting conflicting information from familiar size and other 'standard' depth cues, under binocular and monocular viewing. In a study by Marotta and Goodale (2001), participants first completed a training phase in which they grasped repeatedly for plain, self-illuminated spheres of a constant size. This continued in the experimental conditions, but they were also presented with smaller or larger 'probe' spheres on a small number of trials. The conflicting familiar-size information lead to an increased number of online corrections, indicating reaches to the wrong location, under monocular viewing but not under binocular viewing. More recently, McIntosh and Lashley (2008) ran a similar 'cue-conflict' experiment, but using common household objects (matchboxes), which participants were already familiar with. They used actual-size, as well as scaled-up and scaled-down replica matchboxes, to manipulate the conflict between familiar size and the other available depth cues. Conflicting familiar-size information resulted in biases in reach magnitude, and grip apertures, in the predicted direction. For example, participants over-reached to the scaled-down matchbox, suggesting that the smaller-than-normal retinal image caused an overestimate of object distance. These effects were present in both binocular and monocular conditions, but were larger when binocular cues were removed. These findings have important implications for understanding grasp control in real-world settings, because we very often grasp objects that are highly familiar to us (McInstosh & Lashley, 2008; Melmoth & Grant, 2006).

The results of this study are consistent with the theoretical idea (formalised within the reliability-based cue-weighting framework), as well as the results of Chapters 3 and 4 and a growing body of literature, suggesting that the visuo-motor system acts as a near-optimal 'integrator', combining all sources of information available to it (Greenwald and Knill, 2009a, 2009b; Greenwald et al., 2005; Knill, 2005; Knill & Kersten, 2004; van Mierlo et al., 2009). According to this view, the visuo-motor system does not 'switch' strategies depending on which cues are available, but rather weights *all* the available information according to the relative reliabilities of the signals involved. The familiar-size studies described above are also consistent with this idea. Reliability-based cue weighting predicts that familiar size will have a greater influence when binocular information is unavailable (see Introduction). Furthermore, it predicts that familiar size will have an increasing effect the more reliable it is. It seems very likely that the learned representation of size was much more reliable in McIntosh and Lashley's (2008) study, than in Marotta and Goodale's (2001). In the former case people interacted with everyday, functional objects they had commonly encountered for many years, whereas in the latter the objects were arbitrary shapes, and participants were only exposed to them during the experiment. Thus the model provides a good account of why McIntosh and Lashley (2008) found an effect of familiar size even when binocular cues were available, whereas Marotta and Goodale (2001) did not.

5.5. Conclusions

We have found that the effect of removing binocular information on grip aperture scaling is significantly larger when grasping to a randomised stimulus set than to a typical stimulus set comprised of a small number of fixed object sizes and distances. We suggest therefore that participants learn small stimulus sets, and that this introduces an additional, uncontrolled source of information, leading to misestimates of the effect of removing binocular depth cues. We have concentrated here on binocular vision but this argument applies, in principle, to manipulations of any source of depth information. Studies of visually guided grasping have almost all recognised the need to use several object/distance configurations to prevent 'stereotyped' movements. Our findings suggest, however, that even using three objects and three distances—a higher-than-typical number of stimulus configurations—is insufficient to prevent learning of the stimulus set, with resulting effects on grasp kinematics. Unfortunately, it seems likely that effects of stimulus learning are not limited to the manipulation used here, but may also be present in the data from a variety of studies of visuo-motor control. Fortunately, preventing them is relatively straightforward.

Chapter 6: ‘Strategic’ changes in grasping in response to visual uncertainty

6.1. Introduction

The precision with which grasping movements can be executed depends on a variety of factors, including the level of sensory and motor noise. For instance the quality of available information varies with viewing geometry and the available signals (Hillis et al., 2004; Knill & Saunders, 2003, see Chapters 3 & 4), meaning that precision in the estimates of object properties used to programme the grasp will vary across different situations. Also, moving faster results in greater variability in the reach and grasp, thereby reducing precision of the movement (Harris and Wolpert 1998). Empirical results show that removing or degrading information results in either slower movements or, more consistently, larger grip apertures (Berthier et al., 1996; Christopoulos & Schrater, 2009; Connolly & Goodale, 1999; Churchill et al., 2000; Jackobson & Goodale, 1991; Jackson et al., 1997; Keefe & Watt, 2009; Loftus et al., 2004; Melmoth et al., 2009; Melmoth & Grant, 2006; Schlicht & Schrater, 2007; Watt & Bradshaw, 2000; Wing et al., 1986, see Chapters 3 & 5). These results suggest that the system is strategic and responds adaptively to increases in perceptual uncertainty by adjusting the margin-of-error in the programmed movement. This behaviour has been described as a 'conservative strategy' because it reduces the probability of making a grasping error, such as knocking into or missing the object (Watt & Bradshaw, 2000). The execution of such behaviour is by no means a trivial task for the visuomotor system. Assuming that the cost of failing to grasp the object is constant, the visuomotor system needs to encode the relationship between different types of movements and the probability of making a grasping error alongside perceptual uncertainty. Here we explore whether this is the case by exploiting the fact that this relationship is different for different tasks.

For example, to pick up a tennis ball requires guidance of the finger and thumb to the opposing sides of the ball. For grasps to this object, the probability of error as a function of grip aperture is asymmetrical: increasing grip aperture is less likely to result in an error whereas the same decrease in grip aperture is more likely to result in an error (as grip aperture approaches and eventually becomes smaller than the physical depth of the object). Under different circumstances this relationship alters. Consider the case of grasping a U-shaped object that has to be picked up by guiding the finger and thumb to press against its opposing inner surfaces. For this movement the probability of an error with changes to grip aperture has the opposite asymmetry to the case of the tennis ball: increasing grip aperture is more likely to result in an error than the same decrease in grip aperture. Now consider the case of grasping a bowling ball, where the finger and thumb must be inserted into two small holes a fixed distance apart. Here, the probability of error as a function of grip aperture is symmetrical: increasing or decreasing grip aperture (from the appropriate grip aperture required to grasp the object) is equally likely to result in error. This is shown in cartoon form in Figure 6.1 which plots hypothetical probability distribution functions of grasping error for different movements to a fixed object size. The yellow line illustrates the case of picking up the tennis ball. Here the probability of making a grasping error decreases as the programmed grip aperture increases. The red line shows the case of picking up the U-shaped object. For movements to this object the probability of making a grasping error decreases as the programmed grip aperture is reduced. Lastly, the green line illustrates the case of picking up the bowling ball. Here, both increasing, or decreasing the programmed grip aperture (from the appropriate grip aperture), increases the probability of making a grasping error. We refer to these three conditions as 'normal grasping', 'reversed grasping' and 'equal grasping' respectively.

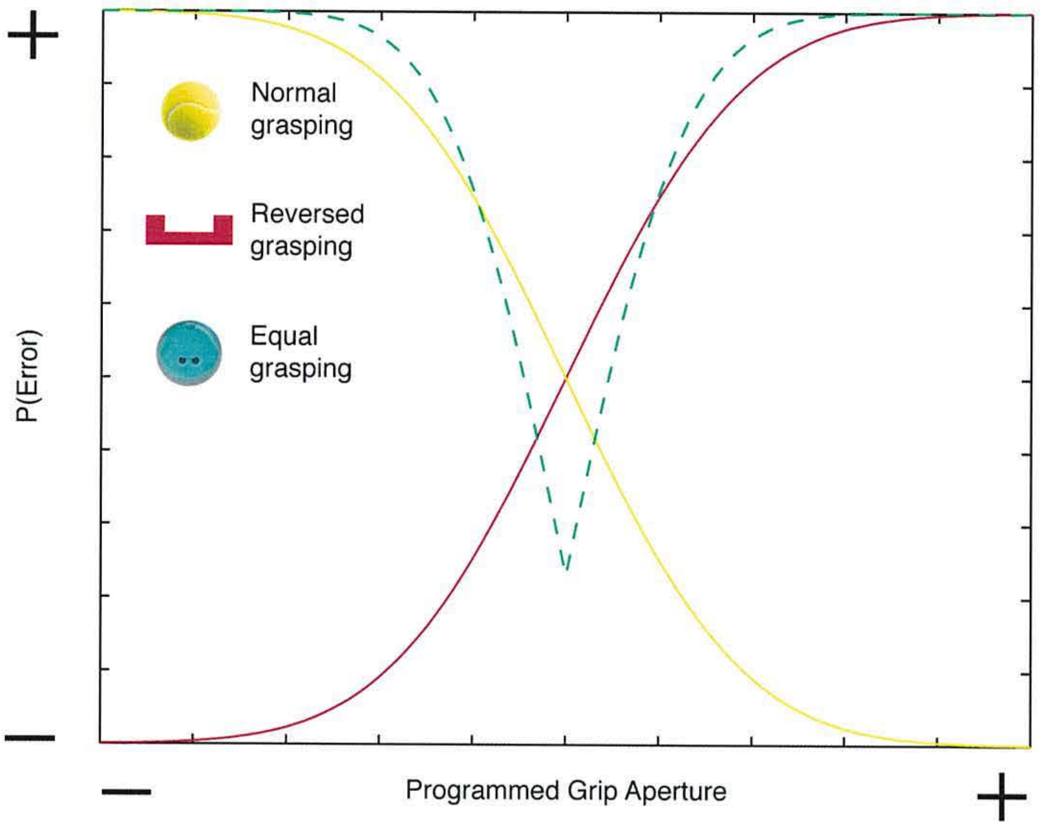


Figure 6.1. Hypothetical probability distribution functions of grasping error across a range of programmed grip apertures to a fixed object size. The different coloured lines plot the probability distribution for grasps under 'normal grasping' (yellow line), 'reversed grasping' (red line) and 'equal grasping' (dashed green line). We assume a level of motor noise across conditions which gives rise to the sigmoidal functions in the normal and reversed grasping conditions, and prevents the probability of error reaching 0 in the equal probability condition. In the absence of motor noise we could expect the distributions to more closely resemble step functions.

The examples illustrated in Figure 6.1 tell us that the visuomotor system should encode both the probability of making an error and perceptual uncertainty in order to programme successful movements that reduce the probability of making a grasping error. If such behaviour were demonstrated, grasping behaviour could indeed be described as 'strategic' as has been previously suggested (Watt & Bradshaw, 2000).

In a recent study, participants were found to programme reach trajectories that minimised the probability of a grasping error under conditions of uncertainty (Christopoulous & Schrater, 2009, see General Introduction: *Visual uncertainty and grasp programming* section). This suggests that the visuomotor system can represent the probability of making a grasping error given different movements, and modify the programmed trajectory of a reach accordingly. Here, we use the term 'probability landscape' to refer to the different probabilities of error associated with increasing or decreasing one's grip aperture for a given movement. We examined visually open-loop grasps made under conditions in which this 'probability landscape' and perceptual uncertainty were manipulated in order to determine if changes in grip aperture are truly strategic. Three viewing conditions were used, normal binocular viewing, blurred binocular viewing, and monocular viewing. Thus, perceptual uncertainty was increased in two ways: by removing binocular cues and by blurring the visual scene. According to cue-integration theory, increasing perceptual uncertainty using either method should have qualitatively similar affects on grip aperture kinematics. In the examples detailed in Figure 6.1, increases in perceptual uncertainty would be expected to result in larger grip apertures to the tennis ball (normal grasping), smaller grip apertures to the U-shaped object (reversed grasping) but for grip aperture to be held constant for grasps to the bowling ball (equal grasping), thereby reducing the probability of error in all cases.

6.2. Method

Participants

Seven participants took part in the experiment (4 male, 3 female, aged 18-23). All had normal or corrected to normal vision and stereoacuity better that 40 arcsec.

Participants gave informed consent and were paid for their participation. All procedures were in accordance with the Declaration of Helsinki.

Apparatus and stimuli

The experimental apparatus and examples of the stimuli are shown in Figures 6.2 and Figure 6.3. A chin rest was employed to reduce head movement and ensure stable presentation of stimuli. Participants were positioned so that their eyes were 400 mm above the table surface, directly above a start button aligned with their mid-sagittal plane. A square pad 5 mm high and 10 mm in width and breadth was attached to the table surface 120 mm in front of the start button for positioning of the index finger in the reversed grasping condition. This was to ensure that participants started the movement with their finger and thumb apart by a fixed amount. The stimuli were viewed on an illuminated matt black table surface during the experiment. To control stimulus presentation we used a 'privacy screen' (PolyVision^{TM}, Witham, Essex; see General Methods). The privacy screen was placed 5cm in front of the participant's eyes so that it fully occluded the apparatus and stimuli while opaque. Participants viewed and grasped wooden objects. Different types of objects were used in the 'normal grasping', 'reversed grasping' and 'equal grasping' conditions. In the normal grasping condition the objects were made from a single rectangular block and were picked up by pressing the thumb and forefinger against the objects opposing sides (Figure 6.3a). In the reversed grasping condition participants reached to grasp U-shaped blocks that were picked up by pressing the upper surfaces of the thumb and forefinger against the objects opposing inner sides (Figure 6.3b). In equal grasping condition, the objects were constructed from three rectangular blocks. The middle block, which participants grasped, was flanked by rectangular blocks placed either side of it. This provided a condition in which the finger and thumb had to be placed into two small gaps to pick up the object by pressing against the opposing sides of the middle block (Figure 6.3c). Throughout the experiment, participants wore ribbed rubber thimbles on their finger and thumb (Figure 6.3). These were used to increase friction in the reversed grasping condition so that the object could be successfully picked up.

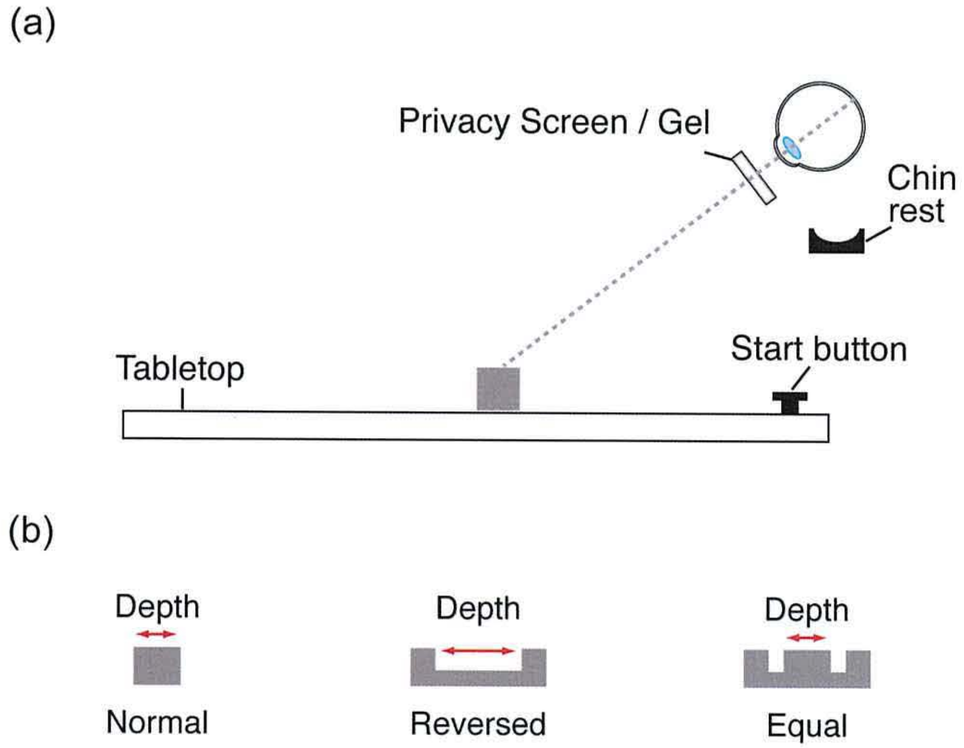


Figure 6.2. (a) Diagram of the experimental setup. (b) Sagittal views of the objects used in the experiment. Red arrows show the 'depth' of the object for the part that was grasped in each case.

There were seven object sizes in each object condition. In the normal grasping condition the width and height of the objects were 70 and 30 mm respectively. In the reversed, and equal grasping conditions the width of the objects was 75 mm and their height was 37 mm. In all cases the term depth refers to the part of the object that was grasped. This is shown in Figure 6.2b. In the normal grasping condition, objects measured 20, 25, 30, 35, 40, 45 and 50 mm in depth. In the reversed grasping condition the inner distance between the two outer blocks was the portion of the object that was grasped and had depths of 70, 75, 80, 85, 90, 95, and 100 mm (the outer blocks measured 25 mm in depth). In the equal grasping condition the middle block, which was grasped, had depths of 20, 25, 30, 35, 40, 45, and 50 mm. For these objects there were two blocks that flanked the middle block, each measuring 25 mm in depth and positioned so that the gaps in which participants had to place their thumb and finger had a constant depth of 22 mm. In both the reversed and equal grasping conditions, the gaps in which participants placed their fingers were 25 mm in height and the base of the object measured 12 mm in height.

To manipulate perceptual uncertainty we used three viewing conditions: binocular viewing, blurred binocular viewing, and monocular viewing. In the binocular condition the scene was viewed normally with both eyes. In the blurred binocular viewing condition the scene was viewed through a translucent plastic film attached to the privacy screen directly in front the eyes. The film was made of 'gel', as used in stage lighting. The film both blurred the visual scene and reduced contrast thereby increasing perceptual uncertainty. In the monocular viewing condition the participants wore an eye patch over their right eye, removing all binocular cues and so increasing perceptual uncertainty.

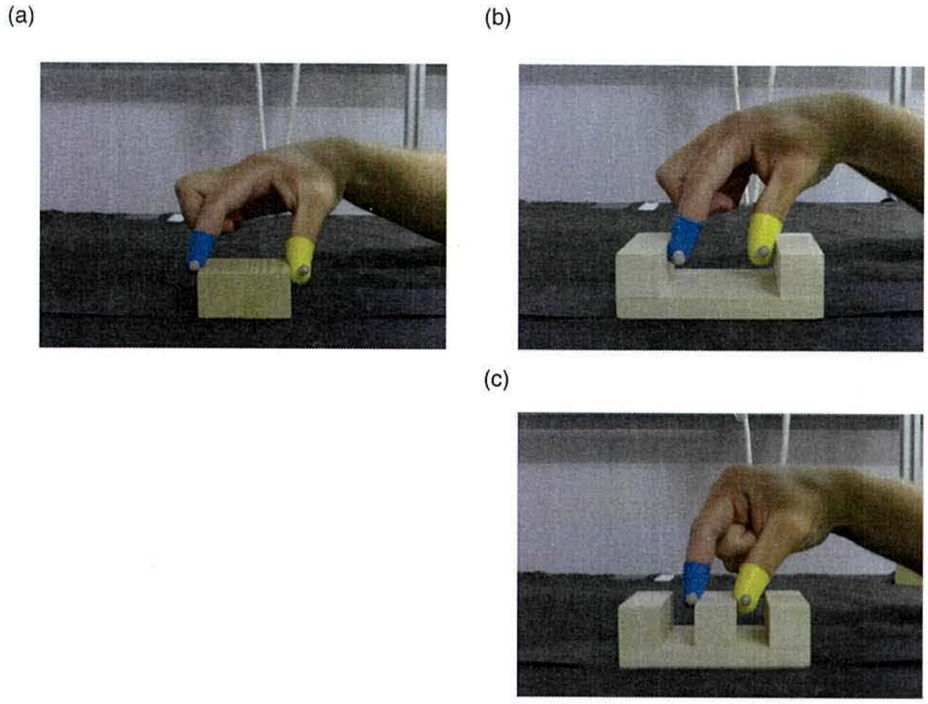


Figure 6.3. Photographs showing each type of grasping movement and object: (a) normal grasping, (b) reversed grasping, and (c) equal grasping.

Procedure

In the normal grasping and equal grasping conditions, participants began each trial by pressing down on the start button with their thumb and index finger pinched together. In the reversed grasping condition participants began each trial by pushing the start button down with their thumb while placing their index finger on a pad located 12 cm in front of the start button.

In all conditions the object was presented for 2 seconds followed by an audible beep which was the start signal to initiate the movement. In the normal and equal grasping conditions participants were instructed to pick up the object front to back using their thumb and index finger (grasping the central block in the case of the equal grasping condition; see Figure 6.3a,c). In the reversed grasping condition participants were instructed to pick up the objects by pushing against its opposing inner surfaces using their thumb and index finger (Figure 6.3b). Releasing the start button extinguished the participants' view so that grasps were performed visually open-loop. Reaches that were initiated before the start signal, or >600 ms after it, were considered void and repeated at the end of the block. In an attempt to minimise learning of the object properties (Keefe & Watt, 2009, see Chapter 5) we used 7 object sizes (see apparatus and stimuli) and randomised distance between 200 and 500 mm (in 5 mm increments). To do this we used 6 'base' distances of 200, 250, 300, 350, 400 and 450 mm. On each trial a random value between 0 and 50 mm (in 5 mm increments) was added to this 'base' distance in order to distribute the distances used throughout the range of 200 to 500 mm. The objects were presented on the table surface along the mid-sagittal plane. Trials were blocked by viewing condition (binocular, blurred binocular, and monocular) and grasp condition (normal grasping, reversed grasping and equal grasping). One block consisted of each object size presented at each object distance (6 x 7 = 42 trials per block) and each block was repeated twice (18 blocks total). Trial order was randomised within each block and the block order was counterbalanced across participants. In total, the experiment took around 5 hours to complete and was split across several sessions.

6.3. Results

Movement velocity analysis

Because perceptual uncertainty can affect both grip apertures and movement velocity we first examined velocities of the reach. Movement velocities were analysed from the 3-D coordinates of the thumb marker on each trial. Each participant's data was summarised by putting the object distances into 50 mm 'bins'. For example, for a given viewing/object condition, all reaches between 200 mm and less than or equal to

250 mm were averaged, creating one data point at 225 mm. This was repeated in 50 mm steps, allowing us to compute the mean peak velocity for each 'bin' (225, 275, 325, 375, 425, 475 mm) before computing overall means across participants. The results are plotted in Figure 6.4 which shows that velocity scaled reliably with object distance under all conditions (Jeannerod 1984, 1988), indicating that normal movements occurred in all of the grasping conditions. Increasing perceptual uncertainty by removing binocular cues or blurring the scene did not appear to affect peak velocity in any of the grasping conditions. These observations were confirmed by a 3 x 6 (viewing condition x object distance) repeated measures analysis of variance, performed separately for each grasping condition. A significant main effect of distance was found in all grasping conditions. However, there was no significant main effect of viewing condition or interaction between viewing condition and distance in any of the grasping conditions used in the experiment (results reported in Table 6.1). These results show that participants did not compensate for increased perceptual uncertainty by moving slower in any of the grasp conditions.

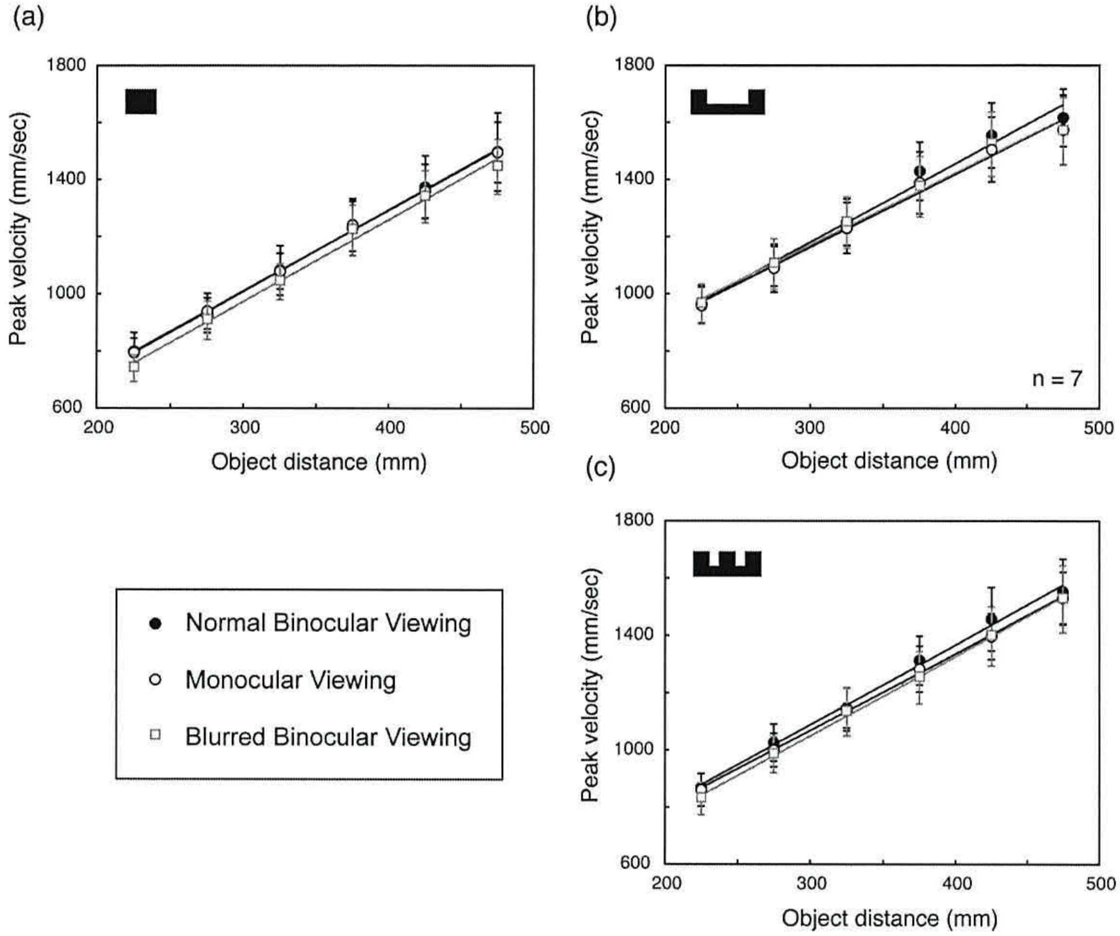


Figure 6.4. Mean peak velocity as a function of object distance in (a) the normal grasping condition, (b) the reversed grasping condition, and (c) the equal grasping condition. The black circles denote the normal binocular condition; the hollow black circles denote the monocular condition and the hollow grey squares denote the blurred binocular condition. The solid lines represent the best-fitting linear regressions to the data in each case. The error bars denote ±1 SEM.

Table 6.1. Effects of viewing condition and distance on peak velocity for each grasp condition.

	Summary of ANOVAs for peak velocity in each grasp condition		
Grasp condition	Viewing condition	Distance	Viewing condition x distance
Normal grasping	$F_{(2,12)} = 1.21$, $p > .05$	$F_{(1.1,6.5)} = 101.57$, $p < .001$	$F_{(10,60)} = .62$, $p > .05$
Reversed grasping	$F_{(2,12)} = 1.29$, $p > .05$	$F_{(1.3,7.9)} = 123.27$, $p < .001$	$F_{(10,60)} = 1.48$, $p > .05$
Equal grasping	$F_{(2,12)} = 2.38$, $p > .05$	$F_{(1.2,7.1)} = 123.75$, $p < .001$	$F_{(10,60)} = .82$, $p > .05$

Grip aperture analysis

Maximum grip aperture is a reliable kinematic of motor planning, because grasps made to normal objects produce maximum grip apertures that scale reliably with object size (Jeannerod, 1984, 1988). In this study we used changes in grip aperture to index motor planning in normal grasping as well as two unconventional cases: reversed and equal grasping. It has been shown previously that the temporal coupling of grip force and load force observed when grasping a normal object is maintained under reversed grasping (Flanagan & Tresilian, 1994). However, to our knowledge the relationship between grip aperture and object size has not been reported.

To check that measures based on grip apertures were informative in all grasp conditions, we examined average grip aperture profiles as a function of object size in all three grasp conditions. The results for one participant are shown in Figure 6.5 which plots grip aperture profiles (the distance between the index finger and thumb) as a function of time in each of the grasp conditions. As expected, there are clear maximum grip aperture peaks in the normal and equal grasping conditions. In both cases the first peak was measured as peak grip aperture - additional peaks may have been caused by online corrections. In the reversed grasping condition grasp profiles approximately resembled the inverse of normal grasping profiles. Here, there was a clear minimum grip aperture that could easily be identified. The first minimum 'trough' was measured as minimum grip aperture. Importantly, maximum and minimum grip apertures scaled reliably with object size in the equal and reversed conditions, respectively. This scaling was evident in all participants data, and can be seen in the main results (Figure 6.6) We are confident, therefore, that maximum grip aperture in the equal grasping condition, and minimum grip aperture in the reversed grasping condition provide a reliable kinematic indicator of motor planning. It is interesting to note that peak grip apertures in the equal grasping condition had shallower peaks than those in the normal grasping condition. This likely reflects the different constraints imposed by the two different tasks. Unlike normal grasping, equal grasping has little room for error, so the maximum and contact grip aperture will likely be very similar, leading to the flattened peaks that we see in Figure 6.5.

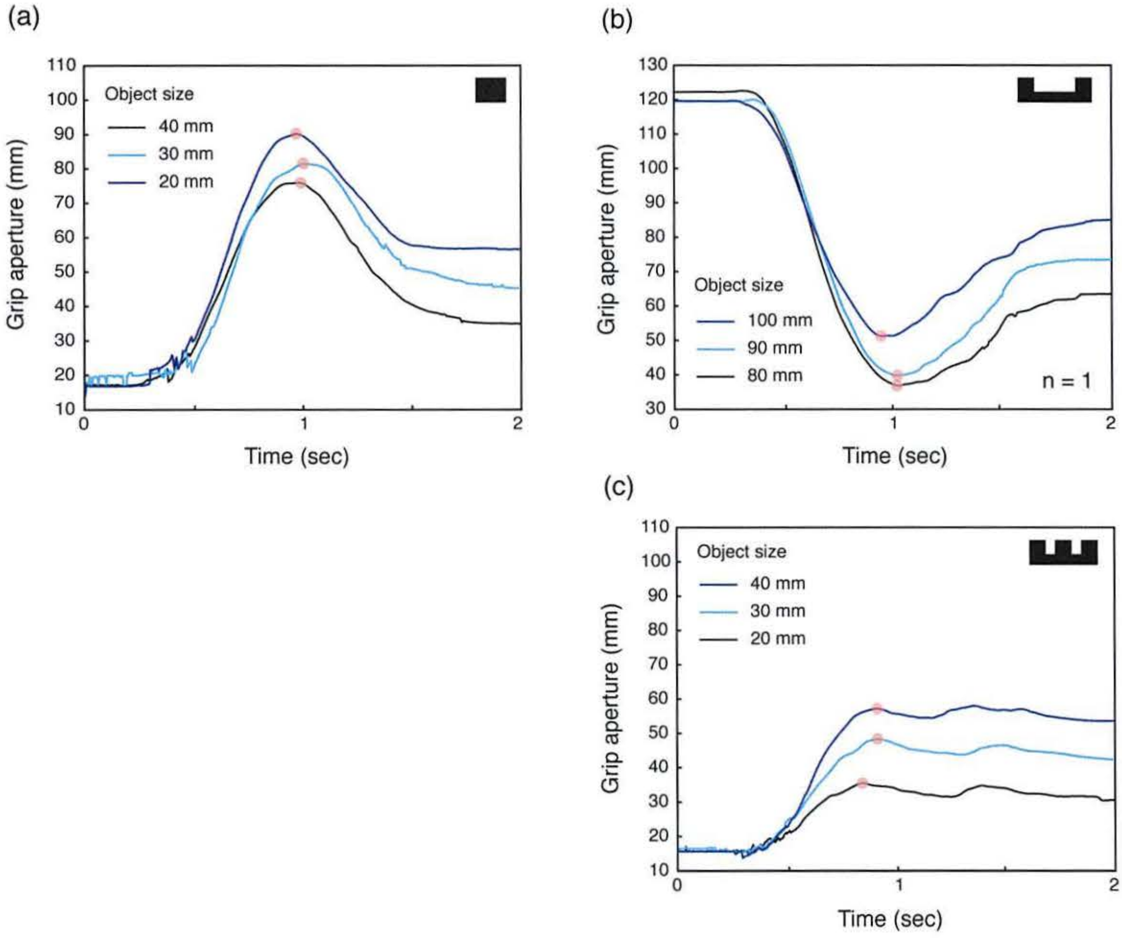


Figure 6.5. Average grip aperture profiles – the distance between the finger and the thumb - in each grasp condition. Profiles are plotted as a function of time in the: (a) normal grasping; (b) reversed grasping, and (c) equal grasping condition. Each profile is plotted from one participant's data - an average of 12 grasps to one object size – under normal binocular viewing. The transparent red circles denote the maximum/minimum grip aperture in each case.

Figure 6.6 shows maximum and minimum grip apertures as a function of object size for the normal, equal, and reversed grasping conditions. In all conditions, grip aperture scaled reliably with object size, indicating that participants were able to perform the task correctly in each case. To confirm these effects, in each grasp condition we performed a 3 x 7 (viewing condition x object size) repeated measures analysis of variance (results reported in Table 6.2). We predicted that increases in perceptual uncertainty would lead to larger grip apertures in the normal grasping condition, smaller grip apertures in the reversed grasping condition, and no effect on grip apertures in the equal grasping condition. Figure 6.6 shows that this was the pattern of results observed. Increasing perceptual uncertainty by removing binocular cues or blurring the visual scene lead to larger grip apertures in the normal grasping condition, smaller grip apertures in the reversed grasping condition and did not affect grip apertures in the equal grasping condition. To confirm these effects we performed planned pair-wise comparisons (linear contrasts, collapsed across distance and size) between normal binocular and monocular viewing, and normal binocular and blurred binocular viewing in each condition (Table 6.3). Consistent with our predictions, in the normal grasping condition, increasing perceptual uncertainty by removing binocular cues or by blurring the visual scene led to significant increases in grip aperture (Table 6.3). These effects were smallest at the largest object sizes, presumably reflecting the increased effort to open the hand yet wider as it approaches its maximum aperture (Keefe & Watt, 2009; Watt & Bradshaw, 2000). The pattern of results in the reversed grasping condition was opposite to that in the normal grasping condition. Increasing perceptual uncertainty by removing binocular cues, or blurring the scene, caused a significant decrease in grip aperture (Table 6.3). These effects were smallest at the smallest object sizes presumably because grip apertures approached the minimum (thumb and finger pressed together) and could not decrease further. For equal grasping, increasing perceptual uncertainty by removing binocular cues, or blurring the visual scene resulted in no significant changes in grip aperture (Table 6.3). Grip aperture scaling in the equal grasping condition was similar to that in the normal grasping condition, although maximum grip apertures were smaller overall. As mentioned earlier, this likely reflects that fact that there is less margin for error in the equal grasping case, leading to a flattening of maximum grip aperture.

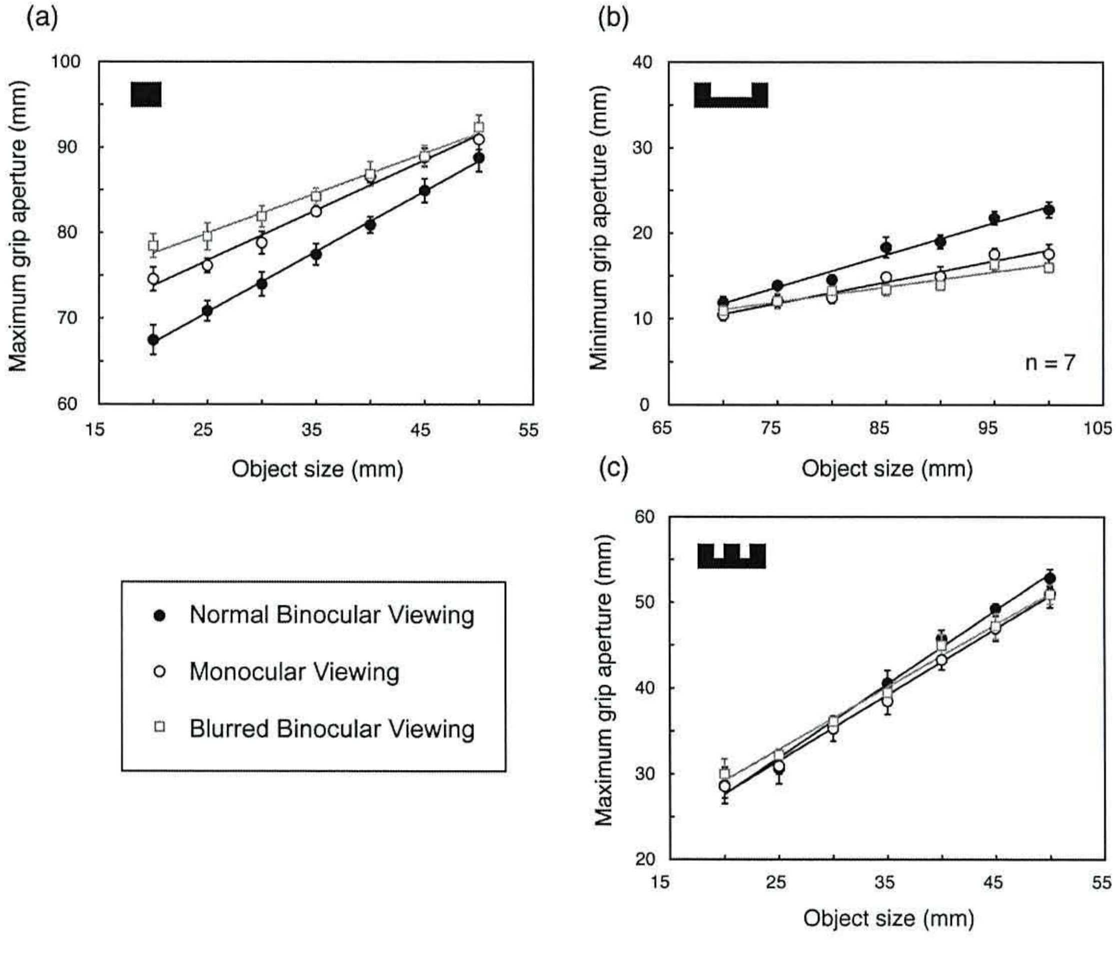


Figure 6.6. Mean maximum (normal, and equal grasping) and minimum (reversed grasping) grip apertures as a function of object size in (a) the normal grasping condition, (b) the reversed grasping condition and (c) the equal grasping condition. The same symbols are used to denote viewing condition as in Figure 6.5. The solid lines represent the best-fitting linear regressions to the data in each case. The error bars denote ±1 SEM.

In the normal grasping and reversed grasping conditions there was a significant interaction between object size and viewing condition (Table 6.2). These interactions occurred because increasing perceptual uncertainty by removing binocular cues or blurring the visual scene resulted in shallower grip aperture scaling functions in both the normal grasping and reversed grasping conditions (presumably due to ceiling/floor effects in either case). There was no significant interaction in the equal grasping condition (Table 6.2), demonstrating that grip aperture scaling functions were not affected by increasing perceptual uncertainty.

It is important to note that scaling functions in the reversed grasping condition exhibited shallower slopes than those in the normal or equal grasping conditions. Given that the same quality of information was available as in the other grasping conditions, we speculate that this reflects the novelty of the task. We rarely pick up objects using the outer surfaces of our thumb and forefinger, so it is possible that the mapping between object size and grip aperture may be poorly specified for reversed grasping, compared to normal and equal grasping.

Table 6.2. Effects of viewing condition and object size on maximum (normal and equal grasping) and minimum (reversed grasping) grip apertures.

	Summary of ANOVAs for maximum/minimum grip aperture		
Grasp condition	Viewing condition	Object size	Viewing condition x object size
Normal grasping	$F_{(2,12)} = 10.27, p < .01$	$F_{(1.3,7.9)} = 60.29, p < .001$	$F_{(12,72)} = 4.56, p < .001$
Reversed grasping	$F_{(2,12)} = 11.20, p < .01$	$F_{(1.9,11.8)} = 40.12, p < .001$	$F_{(12,72)} = 4.84, p < .001$
Equal grasping	$F_{(2,12)} = .32, p > .05$	$F_{(6,36)} = 117.53, p < .001$	$F_{(12,72)} = 1.14, p > .05$

Table 6.3. Effects of viewing condition on maximum (normal and equal grasping) and minimum (reversed grasping) grip apertures. Bin refers to normal binocular viewing. Mon refers to monocular viewing, and Blur refers blurred binocular viewing.

	Summary of planned comparisons for maximum/minimum grip aperture	
Grasp condition	Bin - Mon	Bin - Blur
Normal grasping	$F_{(1,6)} = 15.94, p < .01$	$F_{(1,6)} = 11.44, p < .05$
Reversed grasping	$F_{(1,6)} = 13.51, p < .05$	$F_{(1,6)} = 25.28, p < .01$
Equal grasping	$F_{(1,6)} = .43, p > .05$	$F_{(1,6)} = .06, p > .05$

6.4. Discussion

It is clear from the results that increases in perceptual uncertainty lead to changes to grip aperture in the direction predicted by the probability landscape of a given task. Figure 6.1 illustrates the different probability landscape associated with each task. In order to reduce the probability of a grasping error with increases in perceptual uncertainty, larger grasps should be programmed under normal grasping, smaller grasps should be programmed during reversed grasping, and, during equal grasping, grip aperture should be held constant. These probability landscapes qualitatively predicted performance in the experiment.

As observed previously (Christopoulous & Schrater, 2009; Schlicht & Schrater, 2007; see Experiment 3.3), grip apertures increased with increases in perceptual uncertainty during normal grasping. However, when the probability landscape was 'reversed' during reversed grasping, increasing perceptual uncertainty resulted in smaller grip apertures. No change in grip aperture was observed with increases in perceptual uncertainty for equal grasping. Moreover, this response to uncertainty results in changes in grip aperture in the direction that minimises the probability of error. It seems therefore, that the visuomotor system is able to encode both perceptual uncertainty and the probability of making an error given different movements, in order to programme grasps in a strategic fashion across different probability landscapes.

It is likely that the probability landscape for normal grasping may have already been learnt, given that movements that require normal grasping are experienced in everyday settings. The same is true, to an extent, for equal grasping movements: we often encounter objects in cluttered environments, such as picking the correct jar from a spice rack or the red pen from a pen pot (these examples would share similar probability landscapes to the equal grasping condition used here). The same is not true for reversed grasping, as we rarely pick up U-shaped objects using the outer surfaces of our thumb and finger. Based on this prior experience we imagine that the probability landscape will be known for both normal and equal grasping but not for reversed grasping. Thus, we might expect that early in the experiment, the probability landscape is already known in the normal and equal grasping conditions and so uncertainty should have the predicted effects. In contrast, in the reversed grasping condition, the probability landscape has to be learnt over the course of the experiment, so the effects of uncertainty might be evident only in later trials. We can gain insight into this from our data by examining grip apertures at the beginning and end of the experiment. To do this we computed means for the first 6 grasps and the last 6 grasps to each object size for normal binocular viewing and blurred binocular viewing in each grasping condition (participants completed a total of 12 grasps to each object size in each grasp x viewing condition). We compared results between normal binocular viewing and blurred binocular viewing because this was where we observed the largest effects of perceptual uncertainty and would therefore expect to find the largest effects of learning. Figure 6.7 shows the results for the normal grasping, reversed grasping and equal grasping conditions. To examine the effects of perceptual uncertainty in the first and second half of the experiment we conducted paired-samples *t*-tests.

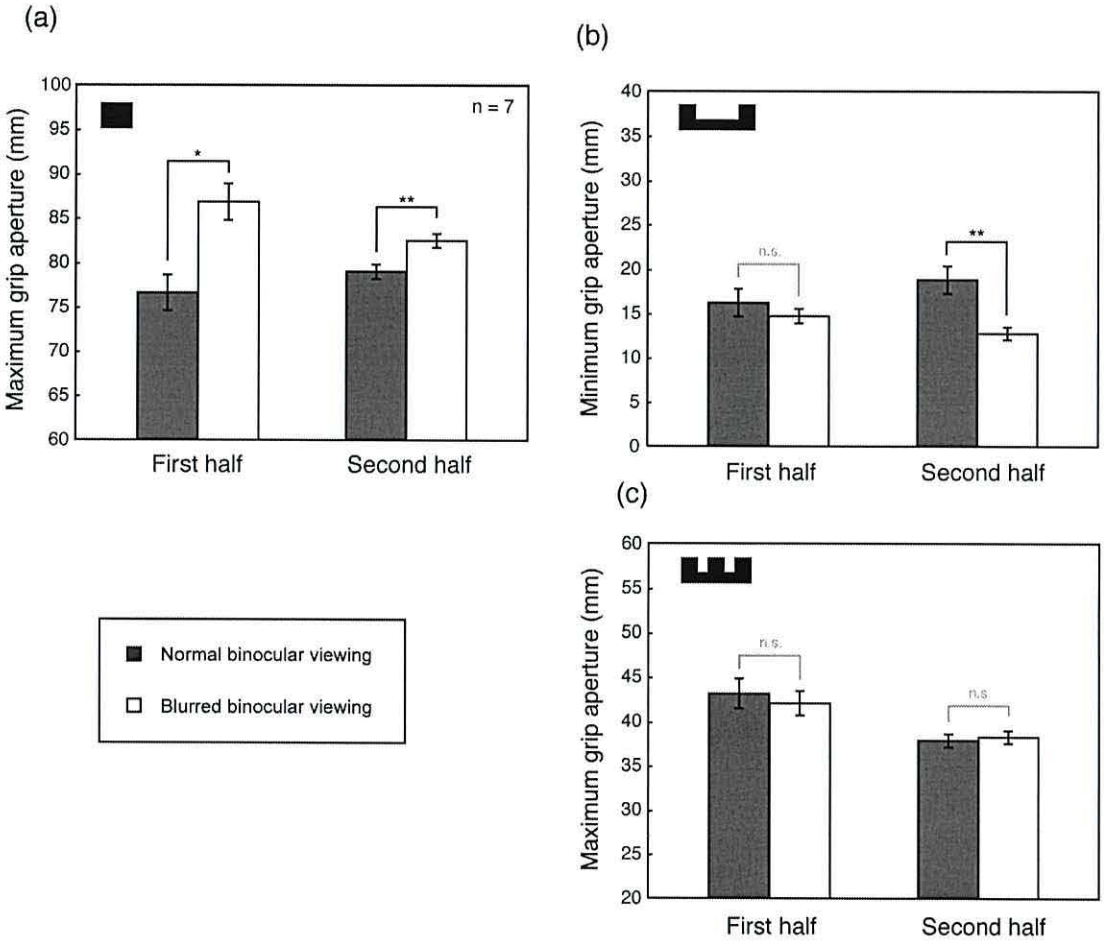


Figure 6.7. Grip apertures in the first and second half of the experiment for normal binocular viewing and blurred binocular viewing. (a) normal grasping, (b) reversed grasping, and (c) equal grasping. Each data point contains the data from 42 grasps (6 grasps x 7 object sizes). The error bars denote ±1 SEM. The connecting lines between bars denote statistical significance (paired-samples *t*-test, 1-tailed). Grey lines show non-significant differences ($p > 0.05$). Black lines denote significant differences at the $p < 0.05$ (*), and $p < 0.01$ (**) levels.

In the normal grasping condition grip apertures were significantly larger under increased levels of perceptual uncertainty in the first half of the experiment ($p < 0.05$). The same effects, although smaller, were observed in the second half of the experiment ($p < 0.01$). These results suggest that probability landscape was already known for the normal grasping condition, and is likely due to the fact we encounter movements that require normal grasping in everyday situations. Interestingly, in the normal grasping condition the effects of perceptual uncertainty resulted in smaller increases in grip apertures in the second half of the experiment compared to the first. One explanation for this finding is that participants were overly conservative in the first half of the experiment, increasing the programmed margin of error in conditions in which perceptual uncertainty was high. Over time, additional information may have become available (the probability landscape may become better known and/or properties of the objects may begin to be learnt, decreasing uncertainty; Keefe & Watt, 2009), leading to grasps that were less conservative). Decreasing the magnitude of grip apertures would reduce the energy expended by the visuomotor system, reflecting a sensible adaptation.

In the reversed grasping condition, minimum grip apertures were significantly smaller under increased levels of perceptual uncertainty in the second half of the experiment ($p < 0.01$) but not in the first ($p > 0.05$). These findings suggest that in the reversed grasping condition the probability landscape was not known at the start of the experiment but was learnt over time. In the second half of the reversed grasping condition minimum grip apertures were not as small under binocular viewing as in the first half of the experiment. This could be because once the probability landscape was learnt, the visuomotor system was less conservative under conditions in which perceptual uncertainty was low, reducing the margin of error and thus the magnitude of grip apertures and energy expenditure.

In the equal grasping condition there were no significant effects of perceptual uncertainty on grip apertures in either the first ($p > 0.05$) or second half of the experiment ($p > 0.05$). There are two explanations for these results. As suggested earlier, they may indicate that the probability landscape was known at the start of the experiment because of previous experience with objects in cluttered environments. Another possibility is that with increased levels of perceptual uncertainty, in the equal grasping condition, participants were unsure how to modify their grip apertures and therefore kept them constant

In the equal grasping condition, grip aperture cannot be changed to compensate for increases in perceptual uncertainty (Christopoulous & Schrater, 2009; Schlicht & Schrater, 2007, see Chapter 3), so it is reasonable to ask if the system compensates in a different way. One such way is to slow down (Wing et al., 1986). Although we demonstrated that peak velocity did not change, it is also possible that participants spent longer in the final portion of the movement with increases in perceptual uncertainty, allowing them greater time to execute grip closure. To examine this we computed the time each participant spent decelerating - time from peak velocity to object contact – in the equal grasping condition. Object contact was calculated as the time at which the thumb marker reached the distance to the near edge of the object (the edge of the object closest to the participant). We used the same method for sorting the data into 50 mm bins as was used for the velocity data (see Method). The method for extracting object contact produced high variability in the data. To overcome this, data greater than two standard deviations from the mean were removed. The results are shown in Figure 6.8, which plots the time spent decelerating as a function of distance. It is clear that the time spent decelerating increased as a function of distance and that perceptual uncertainty did not appear to affect deceleration time. To confirm these effects data were entered into a 3 x 6 (viewing condition x distance) repeated measures analysis of variance. A significant main effect of distance was found ($F_{(5,30)} = 3.05$, $p < .05$), however, no effect of viewing condition ($F_{(2,12)} = .22$, $p > .05$), or interaction between viewing condition x distance was observed ($F_{(10,60)} = .70$, $p > .05$). These results show that participants did not compensate for changes in perceptual uncertainty by extending the time they spent decelerating in the equal grasping condition.

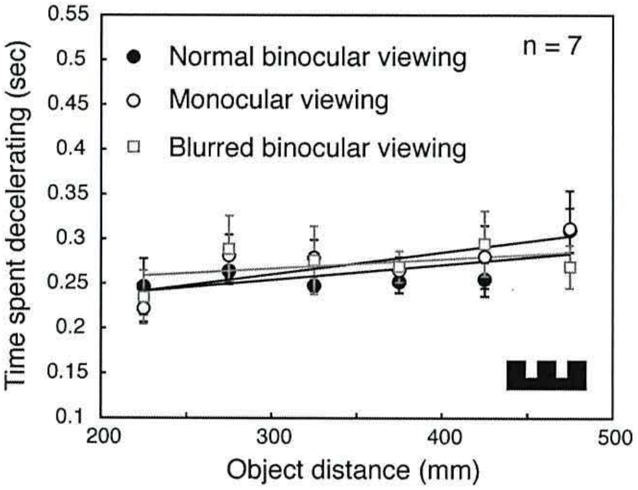


Figure 6.8. Time spent decelerating in the equal grasping condition. Time from peak velocity to object contact is plotted as a function of distance. The same symbols are used to denote viewing condition as in Figure 6.5. The solid lines represent the best-fitting linear regressions to the data in each case. The error bars denote ±1 SEM.

Participants did not compensate for changes in perceptual uncertainty in the equal grasping condition by changing grip aperture, moving slower, or extending the time they spent decelerating. It is likely therefore that increased perceptual uncertainty propagated into the movement programme. If so, we should see a particularly high number of online corrections in this condition with increases in perceptual uncertainty. To investigate this we counted the number of online corrections across all of the conditions used in the experiment. Online corrections were defined as a second maximum/minimum grip aperture that was greater than 5 mm in height from the first identifiable maximum/minimum grip aperture. This was to ensure that we identified online corrections rather than motor noise. We compared results between normal binocular viewing and blurred binocular viewing because this was where we observed the largest effects of perceptual uncertainty and would therefore expect to find the largest number of online corrections.

Figure 6.9 plots the proportion trials in which online corrections occurred in each grasping/viewing condition. It is clear that a greater overall proportion of online corrections were made in the equal grasping condition compared to the normal and reversed grasping conditions. Moreover, increases in perceptual uncertainty appeared to result in a greater proportion of online corrections in all three of the grasping conditions.

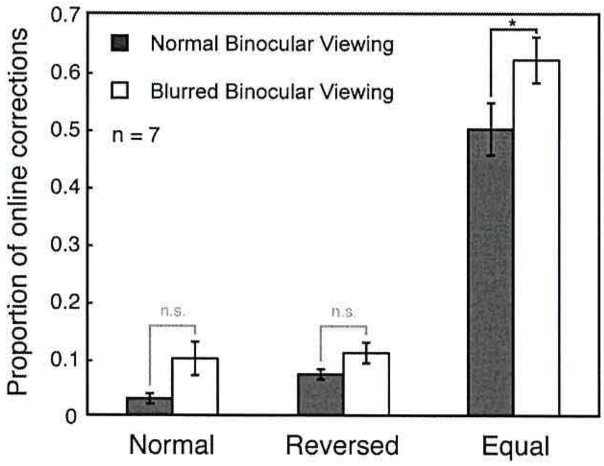


Figure 6.9. The proportion of online corrections across all conditions. The same symbols are used to denote viewing condition as in Figure 6.7. The error bars denote ±1 SEM. The connecting lines between bars denote statistical significance (post hoc Tukey's tests). Grey lines show non-significant differences ($p > 0.05$). Black lines denote significant differences at the $p < 0.05$ (*) level.

To specify the exact nature of these effects we conducted post-hoc pair-wise comparisons (Tukey's tests, collapsed across distance and size) between all three grasping conditions. We performed further pair-wise comparisons between normal binocular and blurred binocular viewing in each grasping condition. As expected there was a significantly higher proportion of online correction in the equal grasping condition compared to the normal ($p < 0.001$) and reversed grasping conditions ($p < 0.001$). Moreover, there was no significant difference in the proportion of online corrections between the normal and reversed grasping conditions ($p > 0.05$). The fact that more online corrections were made in the equal grasping condition likely reflects the fact that small errors in the initial motor plan would result in a grasping error and that participants were not able to reduce the probability of these errors by changing their programmed grip aperture as they were in the normal and reversed grasping conditions. Of particular interest is whether there were significantly more online corrections made in the equal grasping condition as perceptual uncertainty increased. We found that increasing perceptual uncertainty in the normal ($p > 0.05$) and reversed grasping ($p > 0.05$) conditions did not lead to a significant increase in the proportion of online corrections that were made. However, in the equal grasping condition, increasing perceptual uncertainty by blurring the visual scene did lead to a significant increase in the proportion of online corrections ($p < 0.05$). Participants in the equal grasping condition were not able to compensate for changes in perceptual uncertainty by changing their programmed grip aperture. Furthermore, they did not compensate for increased perceptual uncertainty by moving slower or by spending longer decelerating. Instead, they made more online corrections with increases in perceptual uncertainty. This finding helps to explain the effect of increasing perceptual uncertainty in this condition. Because participants were unable to/did not compensate for increases in perceptual uncertainty, they inevitably made more grasping errors as indexed by the significant increase in online corrections.

In this study, we made the assumption that the cost of an error – failing to pick up the object – was known to the visuomotor system and was constant across conditions. This is reasonable, given that the tasks were similar in each case. In the real world, however, the cost of failure may well differ across tasks, and so should also be factored into the movement. Research has shown that the visuomotor system can indeed represent the costs associated with a specific task and behave accordingly. In a speeded aiming task, Trommershäuser et al., (2003) made the cost of error explicit, and found that participant's pointing movements were in a direction that optimised gain across different levels of cost. This raises the question of how manipulating the cost of an error would affect performance in the tasks reported here. An obvious prediction is that increasing the cost of error should increase the adaptive margin-for-error.

In this study, we found that increasing perceptual uncertainty by removing binocular cues or blurring the visual scene lead to a qualitatively similar effect on the adaptive margin-for-error programmed in the grasp. These results suggest that removing binocular cues does not result in a loss of critical information (Marotta et al., 1997), but rather increase uncertainty in the estimates of object properties used to programme the grasp. This finding is consistent with our results from Chapters 3.

6.5. Conclusions

In conclusion, the results of this study show that the visuomotor system is able to encode both the probability of making an error and perceptual uncertainty in order to program grasps that are truly strategic.

Chapter 7: General discussion

7.1. *Overview*

This thesis addressed two questions, namely the role of binocular cues in grasping, and the response of the visuomotor system to changes in perceptual uncertainty. Cue-integration theory states that a combined estimate has less uncertainty than any single cue estimate because the redundancy in multiple cues is exploited. In this work we used the cue-integration framework as a normative model to study the role of binocular cues, both for the planning and online control of grasping. Further, we examined whether learnt information contributes to grasp planning alongside visual depth cues in a manner consistent with cue-integration theory. Lastly, we examined how uncertainty in estimates of object properties is factored into grasping movements by exploring movements, for which a different response to uncertainty might be expected. This was to determine whether perceptual uncertainty and the probability of error, given different movements, are encoded for grasp planning. In this general discussion we will re-state the principle hypotheses tested in this thesis, followed by a review of the empirical chapters in respect of their support for these hypotheses. Lastly we will discuss what these results tell us about the underlying processes that control grasping.

7.2. *The role of binocular cues in grasping*

In Chapter 3 we examined the competing hypotheses of whether cue type or visual uncertainty *per se* determines the margin-for-error in programmed grip apertures. In so doing, we directly tested whether the binocular specialism[7.1.], or cue-integration

---

[7.1.] It is important to note that most studies examining the role of binocular cues in grasping were performed before MLE cue-integration was demonstrated for perception (Ernst & Banks, 2002), and were therefore not considered alongside a cue-integration framework.

accounts of grasp programming best explains the empirical data. Previous studies that have examined the role of binocular cues in grasping by having participants close an eye are unable to distinguish between these two accounts of grasping because they confound the removal of a particular source of information (binocular cues) with an overall reduction in the available information (see Melmoth & Grant, 2006 for a review). In Experiment 3.1 uncertainty in object-size estimates was measured for binocular- and monocular-only conditions, and when both cue types were available together. Changes in viewing distance were used to manipulate the reliability of information, creating conditions in which removing either binocular or monocular information had similar effects on the uncertainty of size estimates. In Experiment 3.2 we measured visually open-loop grasps to the same stimuli, allowing us to examine the relationship between variations in perceptual uncertainty and grasp opening. According to the binocular specialism view, the selective removal of monocular cues should have no effect on grasping, because the binocular input on which the system depends remains intact (Marotta et al., 1997). According to the cue-integration model, in this situation, removing monocular *or* binocular depth cues will reduce the precision of the available depth information by the same amount (see General Introduction: Equation 3), and so should result in similar increases in the margin-for-error programmed into the grasp. Our results showed that removing binocular or monocular information resulted in similar increases in grip aperture, consistent with the similar overall reduction in the precision of depth information in each case. The finding that both binocular and monocular cues contributed to grasping performance is consistent with a cue-integration account of grasping where all sources are used (weighted according to their reliability), and inconsistent with the claim that binocular cues are critical for grasping performance (Melmoth et al., 1997).

In Chapter 4 we extended the findings of Chapter 3 to examine whether cue type, or the precision of information, determines the contribution of binocular and monocular information to the online control of grasping. It has been claimed that binocular cues may be of particular importance for the online control of grasping (Anderson & Bingham, 2010; Bingham et al., 2001; Melmoth & Grant, 2006; Morgan, 1989; Watt & Bradshaw, 2000) based partly on the consideration that the visuomotor system could 'null' the relative disparity signal to achieve hand-object contact (Anderson & Bingham, 2010; Bingham et al., 2001; Melmoth & Grant, 2006; Morgan, 1989, see Chapter 4). Studies examining the role of binocular cues for online control of grasping have typically examined movements made when the hand and object are separated in the depth plane and binocular cues provide relatively precise information (Bradshaw & Elliott, 2003; Jackson et al., 1997; Servos & Goodale, 1994). However, we often pick up objects in different viewing conditions. For example, to pick up an object from our desk we often look directly down at the hand and object so that they are separated in the frontoparallel plane. In this example binocular cues would be relatively uninformative, whereas monocular cues would be relatively informative. Thus a system built to rely exclusively on binocular information would perform poorly. However, a system that integrates all sources of information would not, as it would be able to use information from the more informative monocular signal. Therefore, in Chapter 4 we used changes in viewing geometry to manipulate the relative precision of binocular and monocular feedback. In so doing, we directly tested if binocular signals are relied upon exclusively for online control. In one condition the observers' head was positioned so that, as the hand approached the object, the relative positions of the digits and the object differed primarily in the depth dimension. Here binocular disparity should provide more precise feedback than monocular cues. In another condition, participants looked directly down on the hand/object, so they were separated in the frontoparallel plane. Here monocular cues are highly informative, and disparity is not. We predicted that removing informative binocular information would lead to a deterioration in performance whereas removing uninformative binocular information would not. Our results showed that this was the case. The removal of informative binocular feedback at movement onset led to a deterioration in the online control of grasping as indicated by more time spent in the slow phase of the reach and in contact with object. However, when binocular feedback was uninformative, its removal did not affect online grasping performance. These findings are consistent with a cue-integration account for the online control of grasping in which all signals are used (weighted according to their reliability), and inconsistent with the claim that online control relies exclusively on binocular signals (Anderson & Bingham, 2010; Bingham et al., 2001; Melmoth & Grant, 2006; Morgan, 1989; Watt & Bradshaw, 2000).

Taken together, our results (Experiments 3.1 and 3.2, and Chapter 4) and those of previous studies (Hillis et al., 2004; Knill & Saunders, 2003), demonstrate that changes in viewing geometry will create situations in which binocular cues provide relatively unreliable information both for the planning and online control of grasping (this contradicts the 'textbook' view that binocular cues are always reliable at near distances within reaching space). Under these conditions, a system that relied preferentially on binocular cues would perform worse than one that integrates all sources of information for grasping. For example, a system designed to rely exclusively on relative disparities for the online control of grasping would perform poorly when the hand and object were separated primarily in the frontoparallel plane and relative disparity is uninformative. Therefore, in order to understand the 'role' of visual depth cues in grasping it is necessary to consider the informativeness of the available signals across changes in the natural grasping environment. By using this approach in the experiments reported here (Experiments 3.1 and 3.2, and Chapter 4), we have shown that the effects of removing binocular cues on grasping are explained by a loss of information, not specific cues (binocular cues). This suggests that the visuomotor system does not rely preferentially on binocular cues for the control of grasping. Given these results, we expect that the removal of any reliable source of information should have similar effects on grasping performance.

Interestingly, single unit physiology work in the late '90s and early 2000s on the macaque intraparietal sulcus showed disparity sensitive neurons that are selective for 3-D properties of visual objects (surface orientation, and axis orientation), and are involved in the sensorimotor transformations required for hand manipulation (Murata et al., 2000; Sakata et al., 1995; Sakata et al., 1999; Taira, et al., 1990; Taira et al., 2000). These findings have been cited as evidence for a binocular specialism in grasping (Watt & Bradshaw, 2000). More recently, studies have found that these neurons are also sensitive to the monocular cues of texture and linear perspective (Tsutsui et al., 2001; Tsutsui et al., 2002). Therefore, overall these studies are consistent with the idea that different visual depth cues contribute to the control of grasping. Evidence for human homologues of macaque area AIP comes from both lesion and fMRI studies which have shown that human area AIP is strongly involved in grasping (Binkofski, Dohle, Posse, Stephan, Hefter, et al., 1998). Consistent with lesions to macaque AIP, human patients with AIP lesions also show grasping deficits (but, interestingly the lesion does not affect the reach) (Binkofski et al., 1998). Therefore, it seems likely that the human intraparietal sulcus also processes both binocular and monocular cues for use in the control of grasping. This is consistent with our own data and those of Knill and colleagues (Greenwald and Knill, 2009a, 2009b; Greenwald et al., 2005; Knill, 2005; Knill & Kersten, 2004; van Mierlo et al., 2009).

7.3. *Other depth cues*

The results of Experiments 3.1 and 3.2, and Chapter 4 suggest that different visual depth cues are integrated for grasping. Given this, we might expect other signals, not just disparity and texture, to be integrated by the visuomotor system for the control of grasping. Such a case is familiar size information that has been previously shown to affect grasping performance (Marotta & Goodale, 2001; McIntosh & Lashley, 2008). Familiar size information is often available in natural scenes (presumably you know the size of your coffee cup) and can in principle provide an estimate of the metric size of objects. It is also a particularly important case because it may represent a confound in cue-removal studies of grasping. Most grasping studies have used repeated movements to a small number of objects. Thus, the stimuli could themselves become familiar, adding another uncontrolled depth signal. In Chapter 5 we examined whether repeated grasps to a small stimulus-set (few object sizes and distances) results in learnt, familiar size information, that contributes to the programming of grasps. The effects of removing binocular cues on visually open-loop grasping were measured using (i) a conventional small stimulus-set, and (ii) a pseudo-randomised stimulus set that could presumably not be learned. We reasoned that removing binocular cues in the small stimulus-set condition would lead to smaller grip apertures than in the randomised condition because of the availability of an additional depth signal, which would reduce uncertainty in the estimates of object properties used to programme the grasp. As expected, removing binocular cues resulted in a significantly larger increase in grip aperture in the randomised stimulus set condition, suggesting that learned information contributed to grasping performance in the small stimulus-set condition. Overall the results, although not a quantitative test, are consistent with a cue-integration account of grasp planning in which uncertainty in the estimate of object properties is reduced as more signals become available.

Experiments 3.1 and 3.2, and Chapters 4 and 5 suggest that all sources of depth information are integrated by the visuomotor system in order to minimise uncertainty in estimates of object properties. Given this we would expect other sources of depth information not tested in this thesis to be readily integrated for the control of grasping. For example, accommodation has been demonstrated to provide distance information (Watt, Akeley, Ernst & Banks, 2005) that is used alongside other visual depth cues to help recover depth. Motion parallax, is also known to provide reliable depth information that could potentially be used for the control of grasping (Rogers & Graham, 1982; Bradshaw & Rogers, 1996). As discussed in Chapter 3, a previous study that examined the role of motion parallax in grasping found, surprisingly, that it does not support grasping (Watt & Bradshaw, 2003). However, the relative informativeness of motion parallax was not considered in this study and it is therefore not possible to draw conclusions regarding the 'role' of motion parallax in grasping. To determine if motion parallax contributes to the control of grasping one needs to consider its informativeness in a given situation. Given that there are many other monocular cues in natural scenes, it would be interesting to know if these cues, too, are factored into grasp control.

7.4. *Uncertainty and grasping*

Our findings from Chapters 3 and 5 suggest that perceptual uncertainty is encoded by the visuomotor system for the programming of grasping. This implies a 'strategic system that can appropriately control the probability of making a grasping error (knocking into or missing the object). Uncertainty in sensory estimates might be expected to have an important role in motor acts such as grasping because (i) there are often consequences (costs) for making errors, and (ii) the relationship between the probability of these errors and a movement can change for different tasks. Thus, the visuomotor system should ideally encode both perceptual uncertainty and the probability of error in order to optimise the movement (Christopoulos and Schrater, 2009; Schlicht & Schrater, 2007). In Chapter 6, we examined how uncertainty in estimates of object properties is factored into grasping movements by exploring different movements, for which a different response to uncertainty might be expected. Performance was compared for three different grasping movements for which increasing, decreasing, or keeping grip aperture constant as perceptual uncertainty increased reduced the probability of making a grasping error (knocking into or missing the object). This reflects the different task demands of real world behaviour and was designed to see if the visuomotor system encodes both perceptual uncertainty and the probability of error for grasp planning. If so, we hypothesised that the system would adapt the margin-for-error of the grasp in the direction that minimised the probability of error. Indeed, this was the pattern of grasping behaviour that was observed, suggesting that the visuomotor system encodes both perceptual uncertainty and the probability of making an error in order to programme grasps in a 'strategic' fashion. In the experiment reported in Chapter 6, it is likely that the cost of making a grasping error was constant. However, in the real world we will encounter objects for which the cost of making a grasping error changes. For example, the costs of grasping error associated with a hot cup of coffee would likely be different to the costs associated with an empty cup due to the risk of scolding oneself. This raises the question of how manipulating the cost of an error would affect performance. Understanding the costs and probability of errors in natural tasks is difficult because the statistics of these costs and probabilities are not known, and cannot be directly measured. One approach, therefore, could be to explicitly manipulate costs (for example, financial penalties/rewards) and the probability of success/failure. This approach has proved useful in understanding the programming of manual aiming movements (Trommershäuser et al., 2003) and could be extended to grasping. An obvious prediction is that increasing the cost of error should increase the adaptive margin-for-error.

In Experiment 3.3 we examined whether increases in perceptual uncertainty, independent of changes in the available depth cues, resulted in increases in grip apertures. The results of this experiment showed that increasing in perceptual uncertainty resulted in significant increases in the margin-of-error programmed into the movement. This finding suggests that it is meaningful to use changes in grip aperture as an indirect measure of the underlying precision of object properties (Christopoulos & Schrater, 2009; Schlicht and Schrater, 2007), and is further supported by the good qualitative agreement between the psychophysical (Experiment 3.1) and grasping performance (Experiment 3.2) in Chapter 3. It also suggests the effects of removing binocular cues on grip apertures is due to changes in the precision of information rather than the removal a particular *source* of information (binocular cues). This is also supported by the findings of Chapter 4, which showed that removing binocular cues or blurring the visual scene led to a qualitatively similar deterioration in the performance of movement.

In Chapter 4, published psychophysical data indicated that the combined precision of feedback from binocular and monocular cues would increase substantially when the hand and object were separated primarily in the frontoparallel plane compared to when they were separated primarily in the depth plane (Gepshtein & Banks, 2003; Gepshtein et al., 2005; Takahashi et al., 2009). Consistent with the overall increase in precision of information we found that online grasping performance improved significantly when the hand and object were separated primarily in the frontoparallel plane compared to when they were separated primarily in the depth plane. This finding adds further support to a cue-integration account of grasping in which different sources of information are weighted according to their reliability.

7.5. *Cue-integration in grasping*

Cue-integration theory predicts that multiple estimates of the same property are exploited by the system in order to estimate object properties more precisely than with any one cue alone (see General Introduction: Equation 3). To demonstrate cue-integration in grasping therefore, it is necessary to show that performance is improved by adding sources of information. In Experiment 3.2 we found that grip apertures were consistently smaller when both binocular and monocular cues were available compared to when either cue was available alone. This is the first demonstration to show that the grasping system exploits the redundancy from having multiple signals simultaneously available to improve performance beyond single-cue level (see Knill & Kersten, 2004). Consistent with this, in Chapter 5 we observed smaller grip apertures when both monocular cues and learnt familiar size information was available, compared to when monocular cues were available on their own. Based on the assumption that more precise estimates of object properties allow a smaller margin-for-error in grip apertures to be programmed (see Experiment 3.3 and General Introduction: *Visual uncertainty and grasp programming*), these findings suggest that the visuomotor system was able to estimate object properties with less uncertainty with both signals than with either one alone. Thus, our data suggest that information from binocular and monocular depth cues, and from familiar size, are integrated for grasp programming.

7.6. *Optimal cue-integration in grasping?*

Psychophysical procedures, such as those used in Chapter 3, provide us with a direct measure of the underlying variance in estimates of object properties. This provides a powerful tool for studying perception and has allowed researchers to draw strong conclusions concerning the manner in which different sources of information are integrated for perception (e.g. Ernst & Banks, 2002). The aim of this thesis was to examine how different visual depth cues are integrated for the control of grasping. To do this we used changes in kinematic measures, namely grip aperture, to infer the uncertainty in the estimates of object properties used by the visuomotor system. This is problematic because whereas psychophysical methods provide a relatively direct measure of the uncertainty in estimates used by the visual system, kinematic measures provide only a rather indirect measure of uncertainty in estimates used by the visuomotor system, making quantitative predictions difficult. Furthermore, grasping movements are inherently noisy which makes effects arising from changes in perceptual uncertainty harder to detect (see Knill & Kersten, 2004). Because of this, studies have yet to demonstrate if cue-integration in grasping is *statistically optimal*, as it is in perception, in the sense that it gives the minimum-variance unbiased estimate of object properties (Ghahramani et al., 1997).

7.7. *Broader implications of the current work*

As discussed in the general introduction, Milner and Goodale's (1995) 'two visual systems' hypothesis proposes a functional separation between the dorsal stream which processes absolute, egocentric information for the control of action, and the ventral stream which processes relative, allocentic information for perception[7.2.]. This idea is linked with the proposal of a binocular specialism in grasping partly because binocular cues, can in principle specify metric depth (Howard & Rogers, 2002), whereas monocular cues typically specify only relative depth and are thought to be processed by the ventral stream. Consistent with this, patient D.F. who has a lesion to her ventral stream has impaired object recognition but preserved grasping performance that appears to rely on binocular cues (Carey et al., 1998; Dijkerman et al., 1996; Marotta et al., 1997). If the processes for perception and for action are separate it is possible that different visual depth cues could operate under different principles for the control of grasping and for perception. Thus, a finding of cue-integration for perception and binocular specialism for the control of grasping would be consistent with the two visual-systems model. The results observed in Chapters 3

[7.2.] It is important to note that the strict distinctions between absolute and relative depth cues do not exist under a cue-integration framework because relative depth cues can be promoted by other sources of information to provide absolute depth (Landy et al., 1995).

and 4 of this thesis, however, are part of a growing body of work that suggests both monocular and binocular cues are integrated for the control of grasping in a similar manner as for perception (Greenwald and Knill, 2009a, 2009b; Greenwald et al., 2005; Knill, 2005; Knill & Kersten, 2004; van Mierlo et al., 2009). Overall, these results do not support the two visual-systems model, however it is also possible that separate functional processes exist for perception and the control of action, both of which integrate different sources of information to minimise uncertainty in their estimates of object properties.

7.8. Conclusions

Overall our results suggest that binocular information is not critical for the planning or online control of grasping, as previously suggested (Anderson & Bingham, 2010; Bingham et al., 2001; Goodale and Milner, 2004; Marotta et al., 1997; Melmoth et al., 2009; Melmoth & Grant, 2006; Morgan, 1989; Watt & Bradshaw, 2000). Rather, the findings reported here are consistent with the idea that all sources of information are integrated, according to the precision of information they convey. Moreover, our results suggest that the system exploits redundancy in multiple signals to reduce uncertainty, allowing for more efficient movements with smaller margin-of-error. The system therefore appears to encode the degree of uncertainty in visual information and use this appropriately in motor control.

References

Akerstrom, R. A., & Todd, J. T. (1988). The perception of stereoscopic transparency. *Perception & Psychophysics, 44,* 421-32.

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near optimal bimodal integration. *Current Biology, 14*, 257–262.

Anderson, J. & Bingham, G. P. (2010). A solution to the online guidance problem for targeted reaches: proportional rate control using relative disparity τ. *Experimental Brain Research, 205,* 291–306.

Backus, B. T., & Banks, M. S. (1999). Estimator reliability and distance scaling in stereoscopic slant perception. *Perception, 28,* 217-242.

Banks, M. S., Gepshtein, S., & Landy, M. S. (2004). Why is spatial stereoresolution so low? *Journal of Neuroscience, 24,* 2077–2089.

Baumann, M. A., Fluet, M.-C., & Scherberger, H. (2009). Context-Specific Grasp Movement Representation in the Macaque Anterior Intraparietal Area. *Journal of Neuroscience, 29*, 6436-6448.

Berthier, N. E., Clifton, R. K., Gullapalli, V., McCall, D. D., & Robin, D. (1996). Visual information and object size in the control of reaching. *Journal of Motor Behavior, 28,* 187-197.

Bhushan, N., & Shadmehr, R. (1999). Computational nature of human adaptive control during learning of reaching movements. Biological Cybernetics, *81*, 39-60.

Binkofski, F., Dohle, C., Posse, S., Stephan, K. M., Hefter, H., Seitz, R. J., & Freund H. J. (1998). Human anterior intraparietal area subserves prehension: a combined lesion and functional MRI activation study. *Neurology, 50*, 1253–1259.

Bingham, G. P., Bradley, A., Bailey, M., & Vinner, R. (2001). Accommodation, occlusion, and disparity matching are used to guide reaching: a comparison of actual versus virtual environments. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 1314–1334.

Bingham, G. P, & Pagano, C. C. (1998). The necessity of a perception action approach to definite distance perception: monocular distance perception for reaching. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 145–168.

Bingham, G. P., Zaal, F., Robin, D., & Shull, J. A. (2000). Distortions in definite distance and shape perception as measured by reaching without and with haptic feedback. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1436–1460.

Bishop, P. O. (1989). Vertical disparity, egocentric distance and stereoscopic depth constancy: a new interpretation. *Proceeding of the Royal Society London B: Biological Sciences, 237,* 445-469.

Bootsma, R. J., Matenuik, R. G., MacKenzie, C. L., & Zaal, F. T. (1994). The speed-accuracy trade-off in manual prehension: effects of movement amplitude, object size and object width on kinematic characteristics. *Experimental Brain Research, 93,* 531-541.

Bradshaw, M. F. & Elliott, K. M. (2003). The role of binocular information in the 'on-line' control of prehension. *Spatial Vision, 16,* 295–309.

Bradshaw, M. F., Glennerster, A., & Rogers, B. J. (1996). The effect of display size on disparity scaling from differential perspective and vergence cues. *Vision Research, 36,* 1255–1264.

Bradshaw, M. F., Parton, A. D., & Glennerster, A. (2000). The task-dependent use of binocular disparity and motion parallax information. *Vision Research, 40*, 3725-3734.

Bradshaw, M. F., & Rogers, B. J. (1996). The interaction of binocular disparity and motion parallax in the computation of depth. *Vision Research, 36,* 3457–3468.

Bradshaw, M. F., & Rogers B. J. (1999). Sensitivity to horizontal and vertical corrugations defined by binocular disparity. *Vision Research, 39,* 3049–3056.

Brenner, E., & Smeets, J. B. (1997). Fast responses of the human hand to changes in target position. *Journal of Motor Behavior, 29*, 297–310.

Brenner, E., & van Damme, W. J. M. (1999). Perceived distance, shape and size. *Vision Research 39,* 975–986.

Brown, L. E., Halpert, B. A., & Goodale, M. A. (2005). Peripheral vision for perception and action. *Experimental Brain Research 165*, 97–106

Bülthoff, H. H., & Mallot, H. A. (1988). Integration of depth modules: Stereo and shading. *Journal of the Optical Society of America A, Optics and Image Science, 5*, 1749–1758.

Castiello, U. (2005). The neuroscience of grasping. *Nature Reviews Neuroscience, 6,* 726–36.

Castiello, U., & Begliomini, C. (2008). The cortical control of visually guided grasping. *Neuroscientist, 14*, 157-70.

Castiello, U., Bennett, K. M. B., & Stelmach, G. E. (1993). Reach to grasp: the natural response to perturbation of object size. *Experimental Brain Research, 94,* 163-178.

Carey, D. P., Dijkerman, H. C., & Milner, A. D. (1998). Perception and action in depth. *Conscious Cognition, 7,* 438–453.

Christopoulos, V. N. & Schrater, P. R. (2009). Grasping Objects with Environmentally Induced Position Uncertainty. *PLOS Computational Biology, 5, (10)*, e1000538.

Churchill, A., Hopkins, B., Rönnqvist, L. & Vogt, S. (2000). Vision of the hand and environmental context in human prehension. *Experimental Brain Research, 134,* 81-89.

Clark, J. J., & Yuille, A. L. (1990). *Data Fusion for Sensory Information Processing Systems.* Boston: Kluwer Academic Publishers.

Connolly, J. D., & Goodale, M. A. (1999). The role of visual feedback of hand position in the control of manual prehension. *Experimental Brain Research, 125*, 281-286.

Cuijpers, R. H., Brenner, E., & Smeets, J. B. J. (2008). Consistent haptic feedback is required but it is not enough for natural reaching to virtual cylinders. *Human Movement Science, 27,* 857–872.

de Berg, M., van Kreveld, M., Overmars, M., & Schwarzkopf, O. (2000). *Computational Geometry: Algorithms and Applications* (2nd ed.). New York: Springer.

Desmurget, M., Pélisson, D., Rossetti, Y., & Prablanc, C. (1998). From eye to hand: planning goal-directed movements. *Neuroscience and Biobehavioral Reviews, 22,* 761–788.

Dijkerman, H. C., Milner, A. D., & Carey, D. P. (1996). The perception and prehension of objects oriented in the depth plane 1. Effects of visual form agnosia. *Experimental Brain Research, 112,* 442–451.

Emmert, E. (1881). 'Grossen verhalnisse der Nachbidder'. Klinische Monatsblätter für Augenheilkunde, 19.

Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *Journal of Vision, 7,* 1–14.

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

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47,* 381–391.

Flanagan, J. R., & Tresilian, J. R., (1994). Grip-load force coupling: a general control strategy for transporting objects. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 944-957.

Foley, J. M. (1980). Binocular distance perception. *Psychological Review, 87*, 411–434.

Gallese, V., Murata, A., Kaseda, M., Niki, N., & Sakata, H. (1994). Deficit of hand preshaping after muscimol injection in monkey parietal cortex. *Neuroreport, 5*, 1525–1529.

Gårding, J., Porrill, J., Mayhew, J. E. W., & Frisby, J. P. (1995). Stereopsis, vertical disparity and relief transformations. *Vision Research, 35,* 703–722.

Gentilucci, M., Chieffi, S., Scarpa, M., & Castiello, U. (1992). Temporal coupling between transport and grasp components during prehension movements: effects of visual perturbation. *Experimental Brain Research, 47,* 71-82.

Gepshtein, S., & Banks, M. S. (2003). Viewing geometry determines how vision and haptics combine in size perception. *Current Biology*, *13*, 483–488.

Gepshtein, S., Burge, J., Ernst, M. O., & Banks, M. S. (2005). The combination of vision and touch depends on spatial proximity. *Journal of Vision, 5*, 7, 1013–1023.

Ghahramani, Z., Wolpert, D. M., & Jordan, M. I. (1997). Computational models of sensorimotor integration. In P.G. Morasso & V. Sanguineti (Eds.), *Self-organization, Computational Maps and Motor Control*. Amsterdam: Elsevier Press.

Glover, S. (2004). Separate visual representation in the planning and control of action. *Behavioural and Brain Sciences, 27*, 3-24.

Goodale, M. A., Jakobson, L. S. & Keillor, J. M. (1994). Differences in the visual control of pantomimed and natural grasping movements. *Neuropsychologia, 32,* 1159–1178.

Goodale, M. A., & Milner, A. D. (2004). Sight *Unseen: An Exploration of Conscious and Unconscious Vision.* Oxford: Oxford University Press.

Green, D. M., & Swets, J. A. (1974). *Signal detection theory and psychophysics*. New York: Robert E. Krieger.

Greenwald, H. S., & Knill, D. C. (2009a). Cue integration outside central fixation: A study of grasping in depth. *Journal of Vision, 9,* 1–16.

Greenwald, H. S., & Knill, D. C. (2009b). A comparison of visuomotor cue integration strategies for object placement and prehension. *Visual Neuroscience, 26,* 63–72.

Greenwald, H. S., Knill, D. C. & Saunders, J. A. (2005). Integrating visual cues for motor control: A matter of time. *Vision Research, 45,* 1975–1989.

Grill-Spector, K. (2003). The neural basis of object perception. *Current Opinion in Neurobiology, 13,* 159–166.

Harris C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature, 394,* 780–784.

Hillis, J. M., & Banks, M. S. (2001). Are corresponding points fixed? *Vision Research, 41,* 2457– 2473.

Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: optimal cue combination. *Journal of Vision, 4,* 967–992.

Holway, A. H., & Boring, E. G. (1941). Determinants of apparent visual size with distance variant. *American Journal of Psychology, 54,* 21–37.

Howard, I. P., & Rogers, B. J. (2002). *Seeing in Depth: Depth Perception.* Toronto: I. Porteous.

Jackson, S. R., Jones, C. A., Newport, R., & Pritchard, C. (1997). A kinematic analysis of goal-directed prehension movements executed under binocular, monocular, and memory-guided viewing conditions. *Visual Cognition, 4,* 113–142.

Jacobs, R. A. (1999). Optimal integration of texture and motion cues to depth. *Vision Research, 39,* 3621–3629.

Jakobson, L. S., & Goodale, M. A. (1991). Factors affecting higher-order movement planning: a kinematic analysis of human prehension. *Experimental Brain Research, 86,* 199–208.

Jeannerod, M. (1984). The timing of natural prehension movements. *Journal of Motor Behavior, 16,* 235–254.

Jeannerod, M. (1988). *The Neural and Behavioural Organization of Goal-directed Movements.* Oxford: Clarendon Press.

Jeannerod, M., Decety, J., & Michel, F. (1994). Impairment of grasping movements following a bilateral posterior parietal lesion. *Neuropsychologia, 32*, 369-80.

Keefe, B. D., & Watt, S. J., (2009). The role of binocular vision in grasping: a small stimulus-set distorts results. *Experimental Brain Research, 194,* 435–44.

Knill, D. C. (1998). Discrimination of planar surface slant from texture: Human and ideal observers compared. *Vision Research, 38*, 1683–1711.

Knill, D. C. (2005). Reaching for visual cues to depth: the brain combines depth cues differently for motor control and perception. *Journal of Vision, 5,* 103–115.

Knill, D. C., & Kersten, D. (2004). Visuomotor sensitivity to visual information about surface orientation. *Journal of Neurophysiology, 91,* 1350–66.

Knill, D., & Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences, 27,* 712–719.

Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research, 43,* 2539–2558.

Koenderink, J. J., & van Doorn, A. J. (1991). Affine structure from motion. *Journal of the Optical Society of America A, 8,* 377-385.

Körding, K. P. and Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature, 427,* 244–247.

Künnapas, T. (1968). Distance perception as a function of available visual cues. *Journal of Experimental Psychology 77,* 523-529

Landy, M. S., & Kojima, H. (2001). Ideal cue combination for localizing texture-defined edges. *Journal of the Optical Society of America A, 18,* 2307–2320.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. J. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research, 35,* 389–412.

Loftus, A., Servos, P., Goodale, M. A., Mendarozqueta, N., & Mon-Williams, M. (2004). When two eyes are better than one in prehension: monocular viewing and end-point variance. *Experimental Brain Research, 158,* 317–327.

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. *Experimental Brain Research, 116,* 113–121.

Marotta, J. J., & Goodale, M. A. (2001). The role of familiar size in the control of grasping. *Journal of Cognitive Neuroscience, 13,* 8–17.

Marotta, J. J., Kruyer, A., & Goodale, M. A. (1998). The role of head movements in the control of manual prehension. *Experimental Brain Research, 120,* 134–138.

Marotta, J. J., Perrot, T. S., Nicolle, D., & Goodale, M. A. (1995). The development of adaptive head movements following enucleation. *Eye, 9,* 333–336.

Marotta, J. J., Perrot, T. S., Nicolle, D., Servos, P., & Goodale, M. A. (1995). Adapting to monocular vision: grasping with one eye. *Experimental Brain Research, 104,* 107–114.

Mason, C. R., Gomez, J. E., & Ebner, T. J. (2001). Hand synergies during reach-to-grasp. *Journal of Neurophysiology, 86*, 2896–2910.

Mason, C. R., Theverapperuma, L. S., Hendrix, C. M., & Ebner, T. J. (2004). Monkey hand postural synergies during reach-to-grasp in the absence of vision of the hand and object. *Journal of Neurophysiology, 91*, 2826–2837.

McIntosh R., & Lashley, G. (2008) Matching boxes: familiar size influences action programming. *Neuropsychologia, 46,* 2441–2444.

Melmoth, D. R., Finlay, A. L., Morgan, M. J., & Grant, S. (2009). Grasping deficits and adaptations in adults with stereo vision losses. *Investigative Ophthalmology & Visual Science*, *50,* 3711–3720.

Melmoth, D. R., & Grant, S. (2006). Advantages of binocular vision for the control of reaching and grasping. *Experimental Brain Research, 17,* 371–388.

Melmoth, D. R., Storoni, M., Todd, G., Finlay, A. L. & Grant, S. (2007). Dissociation between vergence and binocular disparity cues in the control of prehension. *Experimental Brain Research, 183,* 283–298.

Miall, R. C., Weir, D. J., Wolpert, D. M., & Stein J. F. (1993). Is the cerebellum a Smith predictor? *Journal of Motor Behaviour, 25*, 203–216.

Milner, A. D., & Goodale, M. A. (1995). *The Visual Brain in Action*. Oxford: Oxford University Press.

Milner, A. D., & Goodale, M. A. (2006). *The visual brain in action* (2nd ed). Oxford, UK: Oxford University Press.

Mon-Williams, M., & Dijkerman, H. C. (1999). The use of vergence information in the programming of prehension. *Experimental Brain Research, 128,* 578–582.

Morgan, M. J. (1989). Vision of solid objects. *Nature, 339,* 101–103.

Morgan, M. J., & Castet, E. (1995). Stereoscopic depth perception at high velocities. *Nature, 378*, 380-383.

Murata, A., Gallese, V., Luppino, G., Kaseda, M., & Sakata, H., (2000). Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. *Journal of Neurophysiology, 83,* 2580–2601.

Ogle, K. N. (1950). *Researches in Binocular Vision.* Philadelphia, PA: W. B. Saunders.

O'Leary, A., & Wallach, H. (1980). Familiar size and linear perspective as distance cues in stereoscopic depth constancy. *Perception & Psychophysics, 27*, 131–135

Oruç, I., Maloney, L. T., & Landy, M. S. (2003) Weighted linear cue combination with possibly correlated error. *Vision Research, 43,* 2451–2468.

Paulignan, Y., Jeannerod, M., MacKenzie, C., & Marteniuk, R. (1991a). Selective perturbation of visual input during prehension movements. 2. The effects of changing object size. *Experimental Brain Research, 87*, 407-420.

Pauligan, Y., Mackenzie, C., Marteniuk R., & Jeannerod, M. (1991b). Selective perturbation of visual input during prehension movements. 1. The effects of changing object position. *Experimental Brain Research, 83*, 502-512.

Prablanc C., & Martin, O. (1992). Automatic control during hand reaching at undetected two-dimensional target displacements. *Journal of Neurophysiology, 67,* 455–469.

Previc, F. H. (1990). Functional specialisation in the lower and upper visual fields in humans: Its ecological origins and neurophysiological implications. *Behavioral and Brain Sciences, 13,* 519–575.

Rizzo, M. & Darling, W. (1997). Reaching with cerebral tunnel vision. *Neuropsychologia, 35*, 53-65.

Rogers, B., & Bradshaw, M. F. (1993). Vertical disparities, differential perspective and binocular stereopsis. *Nature, 361,* 253–255.

Rogers, B., & Graham. M. (1979). Motion parallax as an independent cue for depth perception. *Perception, 8*, 125-134.

Rogers, B., & Graham, M. (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision Research, 22*, 261–270.

Roy, A. C., Paulignan, Y., Farnè, A., Jouffrais, C., & Boussaoud, D. (2000). Hand kinematics during reaching and grasping in the macaque monkey. *Behavioural Brain Research, 117*, 75-82.

Roy, A., Paulignan, Y., Meunier, M., & Boussaoud, D. (2006). Prehension movements in the macaque monkey: effects of perturbation of object size and location. *Experimental Brain Research, 169*, 182-193.

Sakata, H., Taira, M., Murata, A., & Mine, S., (1995). Neural mechanisms of visual guidance of hand action in the parietal cortex of the monkey. *Cerebral Cortex, 5,* 429–438.

Sakata, H., Taira, M., Kusunoki, M., Murata, A., Tsutsui, K., Tanaka, Y., Shein, Y., & Miyashita, Y. (1999). Neural representation of three-dimensional features of manipulation objects with stereopsis. *Experimental Brain Research, 128,* 160–169.

Saunders J. A., & Knill, D. C. (2005). Humans use continuous visual feedback from the hand to control both the direction and distance of pointing movements. *Experimental Brain Research, 162*, 458-473.

Saunders J. A, Knill D. C. (2003). Humans use continuous visual feedback from the hand to control fast reaching movements. *Experimental Brain Research, 152,* 341–352.

Saunders J. A., & Knill, D. C. (2004). Visual feedback control of hand movements. *Journal of Neuroscience, 24,* 3223–3234.

Savelsbergh, G. J. P., Steenbergen, B., & van der Kamp, J. (1996). The role of fragility information in the guidance of the precision grip. *Human Movement Science, 15,* 115–127.

Schenk, T., & McIntosh, R. D. (2010). Do we have independent visual streams for perception and action? *Cognitive Neuroscience, 1,* 52-78.

Schlicht, E. J., & Schrater, P. R. (2007). Effects of visual uncertainty on grasping movements. *Experimental Brain Research, 182,* 47–57.

Servos, P., & Goodale, M. A. (1994). Binocular vision and the on-line control of human prehension. *Experimental Brain Research, 98,* 119–127.

Servos, P., Goodale, M. A., & Jakobson, L. S. (1992). The role of binocular vision in prehension: a kinematic analysis. *Vision Research, 3,* 1513–1521.

Serwe, S., Drewing, K., & Trommershäuser, J. (2009). Combination of noisy directional visual and proprioceptive information. *Journal of Vision, 9,* 1–14.

Sheedy, J. E., Bailey, I. L., Buri, M., & Bass, E. (1986). Binocular vs. monocular task performance. *American Journal of Optometry and Physiological Optics, 63,* 839–846.

Sivak, B. & MacKenzie, C. L. Integration of visual information and motor output in reaching and grasping: the contributions of peripheral and central vision. *Neuropsychologia, 28*, 1095-1116.

Smeets, J. B. J., & Brenner, E. (1999). A new view on grasping. *Motor Control, 3,* 237–271.

Soechting, J. F., & Lacquiniti, F. (1983). Modification of trajectory of a pointing movement in response to a change in target location. *Journal of Neurophysiology, 49*, 584-564.

Taira, M., Mine, S., Georgopoulos, A. P., Murata, A., & Sakata, H. (1990). Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Experimental Brain Research, 83,* 29–36.

Taira, M., Tsutsui, K. J., Jiang, M., Yara, K., & Sakata, H. (2000). Parietal neurons represent surface orientation from the gradient of binocular disparity. *Journal of Neurophysiology, 83,* 3140–3146.

Takahashi, C., Diedrichsen, J., & Watt, J. (2009). Integration of vision and haptics during tool use. *Journal of Vision, 9*, 3, 1-13.

Trommershäuser, J., Maloney, L. T., & Landy, M.S. (2003). Statistical decision theory and the selection of rapid, goal-directed movements. *Journal of the Optical Society of America A, 20*, 1419–1433.

Tsirlin, I., Allison, R. S., & Wilcox, L. M. (2008) Stereoscopic transparency: constraints on the perception of multiple surfaces. *Journal of Vision, 8,* 1–10.

Tsutsui, K., Jiang, M., Yara, K., Sakata, H., & Taira, M. (2001). Integration of perspective and disparity cues in surface-orientation-selective neurons of area CIP. *Journal of Neurophysiology, 86,* 2856–2867.

Tsutsui, K., Sakata, H., Naganuma, T., & Taira, M. (2002). Neural correlates for perception of 3D surface orientation from texture gradient. *Science, 298,* 409–412.

Turner, J., Braunstein, M. L., & Anderson, G. J. (1997). Relationship between binocular disparity and motion parallax in surface detection. *Perception and Psychophysics, 59*, 370-380.

van Ee, R., & Anderson, B. L., (2001). Motion direction, speed and orientation in binocular matching. *Nature, 410,* 690–694.

van Mierlo, C. M., Louw, S. Smeets, J. B., & Brenner, E. (2009). Slant cues are processed with different latencies for the online control of movement. *Journal of Vision, 9*, 25, 1-8.

Wallace, J. M., & Mamassian, P. (2004). The efficiency of depth discrimination for non-transparent and transparent stereoscopic surfaces. *Vision Research, 44,* 2253–2267.

Watt, S. J., Akeley, K., Ernst, M. O., & Banks, M. S. (2005). Focus cues affect perceived depth. *Journal of Vision, 5*, 7, 834-862.

Watt, S. J., & Bradshaw, M. F., (2000). Binocular cues are important in controlling the grasp but not the reach in natural prehension movements. *Neuropsychologia, 38,* 1473–1481.

Watt, S. J., Bradshaw, M. F., (2003). The Visual Control of Reaching and Grasping: Binocular Disparity and Motion Parallax. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 404–415.

Weir, P. L., MacKenzie, C. L., Marteniuk, R. G., & Cargoe, S. L. (1991). Is object texture a constraint on human prehension?: kinematic evidence. *Journal of Motor Behaviour, 23*, 205–210.

Weir, P. L., MacKenzie, C. L., Marteniuk, R. G., Cargoe, S. L., & Fraser, M. B. (1991). The effects of object weight on the kinematics of prehension. *Journal of Motor Behaviour, 23*, 192–204

Welchman, A. E., Deubelius, A., Conrad, V., Bülthoff, H. H., & Kourtzi, Z. (2005). 3D shape perception from combined depth cues in human visual cortex. *Nature Neuroscience, 8*, 820-827.

Westheimer, G. (1981). Visual hyperacuity. *Progress in Sensory Physiology, 1*, 1-37.

Wickelgren, E. A., McConnell, D. S., & Bingham, G. P. (2000). Reaching measures of monocular distance perception: Forward versus side-to-side head movements and haptic feedback. *Perception and Psychophysics, 62,* 1051–1059.

Wing, A. M., Turton, A., & Fraser, C. (1986). Grasp size and accuracy of approach in reaching. *Journal of Motor Behavior, 18,* 245–260.

Yuille, A. L., & Bülthoff, H. H. (1996). Bayesian decision theory and psychophysics. In D. C. Knill & W. Richards (Eds.), *Perception as Bayesian Inference*. New York: Cambridge University Press.