

4 Grasping the Past and Present: When Does Visuomotor Priming Occur?

Melvyn A. Goodale, Jonathan S. Cant, and Grzegorz Króliczak

4.1 Introduction

Humans are able to use vision to generate and control an impressive array of skilled actions. Much of this control presumably depends on processing within the intricate network of visual areas (more than 30) that have been charted in the cerebral cortex (for recent reviews, see Kaas, 2000; Grill-Spector & Malach, 2004; see also Felleman & Van Essen, 1991). Despite the bewildering complexity of the interconnections between these cortical visual areas, two broad “streams” of anatomically distinct visual projections have been identified in the macaque monkey brain: a ventral stream, projecting from area V1 to the inferotemporal cortex, and a dorsal stream, projecting from area V1 to the posterior parietal cortex (Ungerleider & Mishkin, 1982). These regions also receive inputs from a number of other subcortical visual structures, particularly the superior colliculus, which sends prominent projections to the dorsal stream (via the thalamus). Although some caution must be exercised in generalizing from monkey to human, recent neuroimaging evidence suggests that the visual projections from primary visual cortex to the temporal and parietal lobes in the human brain involve a separation into ventral and dorsal streams similar to that seen in the monkey (for review, see Culham & Kanwisher, 2001; Tootell et al., 2003; Van Essen et al., 2001).

Although the anatomical distinction between the ventral and dorsal streams of visual processing has been recognized for some time, the different functions of the two streams have undergone various interpretations. In their seminal paper, Ungerleider and Mishkin (1982) argued that the ventral stream plays a critical role in the identification and recognition of objects, while the dorsal stream mediates the localization of those same objects. Some have referred to this distinction in visual processing as one between “what” versus “where.” Although the evidence for this distinction initially seemed quite compelling, recent findings from a broad range of studies in both humans and monkeys has forced a reinterpretation of the division of labor between the two streams.

According to Goodale and Milner's (1992) reinterpretation of the functions of the two streams, the ventral stream plays the major role in constructing the perceptual representation of the visual world and the objects within it, while the dorsal stream mediates the visual control of actions directed at those objects (figure 4.1). In contrast to the "what versus where" model put forward by Ungerleider and Mishkin (1982), Goodale and Milner argued that the structural and spatial attributes of an object are being processed by both streams, but for different purposes. In the case of the ventral stream, information about a broad range of object parameters is transformed for perceptual purposes; in the case of the dorsal stream, some of these same object parameters are transformed for the control of actions. This is not to say that the distribution of visual inputs does not differ between the two streams, but rather that the main difference lies in the nature of the transformations that each stream performs on those two sets of inputs.

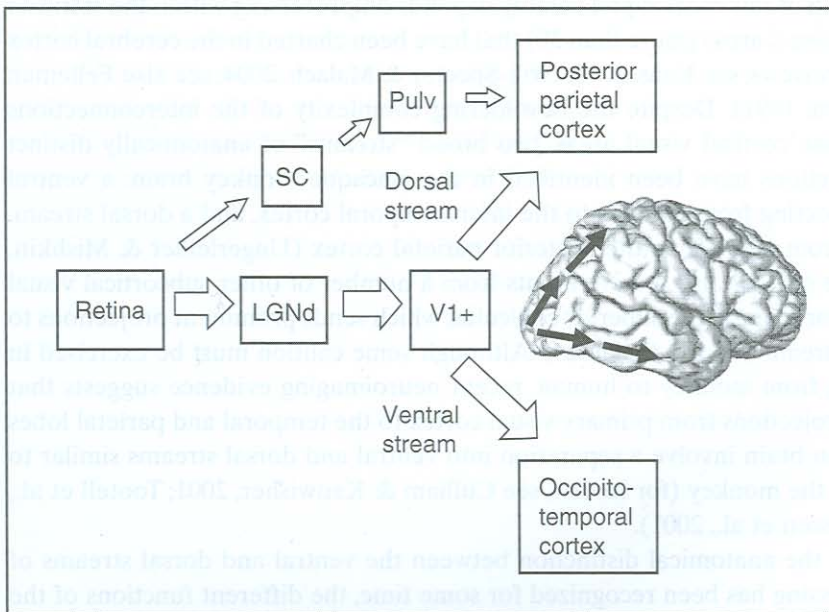


Figure 4.1

Schematic representation of the two streams of visual processing in human cerebral cortex. The retina sends projections to the lateral geniculate nucleus, pars dorsalis (LGNd), in the thalamus, which projects in turn to primary visual cortex (V1). Within the cerebral cortex, the ventral stream arises from early visual areas (V1+) and projects to regions in the occipitotemporal cortex. The dorsal stream also arises from early visual areas but projects instead to the posterior parietal cortex. The posterior parietal cortex also receives visual input from the superior colliculus (SC) via the pulvinar (Pulv). The approximate locations of the pathways are shown on the left side of the three-dimensional reconstruction of the pial surface of the brain made from an anatomical magnetic resonance image. The routes indicated by the arrows involve a series of complex interconnections.

4.2 Frames of Reference for Perception and Action

According to Goodale and Milner (1992), the two separate streams of visual processing evolved because of the different transformations required for vision for perception and vision for action. For one to be able to grasp an object successfully, for example, it is essential that the brain compute the actual size of the object—and its real distance and position with respect to the observer. In addition, these computations will take into account only the dimension of the goal object that is relevant for action, such as its width, for example, while ignoring other dimensions, such as its length, that may not be immediately useful in performing the task at hand (e.g., Ganel & Goodale, 2003). Moreover, the information about the orientation and position of the object must be computed in frames of reference that take into account the orientation and position of the object with respect to the effector that is to be used to perform the action (i.e., in *egocentric frames of reference*). In addition, because observers and goal objects often do not stay in a static relationship with one another, the required coordinates for action are most effectively computed immediately before the movements are initiated, that is, in *real time*. A corollary of real-time visuomotor transformation is that neither the coordinates for a particular action nor the resulting motor program need to be stored in memory. Indeed, such storage could create interference between competing action plans for multiple objects in the visual array, or between action plans to the same object following a change in the spatial relationship between target and actor. In fact, there is a good deal of evidence that object-directed actions, such as grasping, that are initiated *after* the goal object has been removed from view are qualitatively different from the actions that are programmed while the object is visible (for a review, see Goodale et al., 2003). As we shall see, these findings suggest that the control of actions to remembered objects may depend heavily on earlier perceptual processing—processing that does not typically intrude on the control of visually guided actions operating in real time.

In contrast to the vision-for-action system, the vision-for-perception system computes the size, location, shape, and orientation of an object primarily in relation to other objects and surfaces in the scene. In other words, the metrics of perception are inherently relative and the frames of reference are largely scene-based, which explains why we are so sensitive to size-contrast illusions and other visual illusions that depend on comparisons between different objects in the visual array (for reviews, see Goodale & Haffenden, 1998; Goodale & Milner, 2004; Goodale et al., 2003). Encoding an object in a scene-based frame of reference (sometimes called an *allocentric frame of reference*) permits a representation of the object that preserves the relations between the object parts and its surroundings without requiring precise information about the absolute size of the object or its exact position

with respect to the observer. Because of this reliance on allocentric computations, individual features of objects cannot be perceived in isolation. For example, explicit judgments of the width of an object cannot help but be affected by the length of that object (Ganel & Goodale, 2003) or by the apparent width of other objects in the scene (Hu & Goodale, 2000).

Vision for perception also operates over a much longer time scale than that used by vision for action. We can recognize objects we have seen minutes, hours, days—or even years—before. In fact, object recognition would not be possible unless perceptual information about previously encountered objects were stored in memory—and an allocentric representation system is ideal for storing this information. As we indicated earlier, there is evidence that the vision-for-perception system is recruited when a goal object is no longer visible at the moment an action is initiated. In other words, the programming and control of memory-driven actions make use of information that was processed and encoded earlier by vision for perception. As a consequence, memory-guided actions tend to be based on the relative rather than the absolute metrics of the target object. Indeed, there is evidence that grasping movements that are programmed and initiated after a goal object has been removed from view are qualitatively different from the actions programmed while the target remains visible (Goodale et al., 1994; Hu & Goodale, 2000; Milner et al., 1999; Milner & Goodale, 1995; Westwood & Goodale, 2003). This reliance on allocentric scene-based representations for guiding memory-driven or “offline” actions is presumably much more efficient than the use of a stored egocentric encoding of the target (or a stored motor program) that would have to be continuously updated as the observer’s position with respect to the target changed during the delay.

To reiterate: Vision for action and vision for perception require fundamentally different frames of reference and temporal scales—which helps to explain why two separate visual pathways have evolved in the primate cerebral cortex. Visually guided actions, it appears, depend on pathways that engage relatively encapsulated visuomotor mechanisms in the dorsal stream. These dedicated visuomotor mechanisms, together with related networks in the premotor cortex and brainstem, compute the absolute metrics of the target object and its position in the egocentric coordinates of the effector used to perform the action. Such real-time programming is essential for the production of accurate and efficient movements in a world where the location and disposition of a goal object with respect to the observer can change quickly and often unpredictably. In contrast, the perceptual mechanisms in the ventral stream (and associated cognitive structures) use relative metrics and scene-based or allocentric coding. Such computations make it possible not only to recognize objects but to plan and execute actions upon objects long after they have vanished from view.

4.3 Neurological Studies of Timing in Perception and Action

Studies with neurological patients have demonstrated clear temporal limits on the ability of the isolated dorsal system to guide manual prehension. Consider, for example, the case of D.F., a young woman who suffered damage to her ventral-stream pathway as a result of anoxia from carbon monoxide poisoning (Milner et al., 1991). Even though D.F. is unable to indicate the size, shape, and orientation of an object, either verbally or manually, she shows normal preshaping and rotation of her hand when reaching out to grasp it (Goodale et al., 1991). For example, even though D.F. was unable to report the width of objects that varied in their dimensions from trial to trial, she showed excellent size scaling of her grasp in flight as she reached out to pick up each object. On the basis of this dissociation between perceptual report and the visual control of action, Goodale and Milner (1992) argued that D.F.'s dorsal stream was still functioning relatively normally despite the apparent damage to her ventral stream, a conclusion that was recently confirmed in a neuroimaging study of D.F. (James et al., 2003). It appears that her ventral stream can no longer process form information for perception but her dorsal stream remains sensitive to form information about objects that are the targets of visually guided grasping movements. However, despite the fact that D.F. can pick up objects accurately in real time, she demonstrates extremely poor size scaling of her grip when she attempts to pick up objects that are no longer visible (Goodale et al., 1994).

Thus, when a delay of only 2 s was introduced between D.F.'s viewing a goal object and the instruction to initiate her grasping movement, the aperture of her grasp in flight showed no correlation at all with the size of the object. D.F.'s deficit contrasts sharply with the performance of normal participants in the same task, who showed only subtle differences in their visually guided and delayed grasping movements. The dissociation between D.F.'s ability to perform real-time actions and her inability to perform memory-driven actions suggests that D.F., unlike normal subjects, cannot use her memory of the goal object to program her delayed grasping movements. Since visual memory for object features depends on the perceptual mechanisms that reside in the ventral stream, which we now know is damaged in D.F. (James et al., 2003), she was precluded from setting up such memories. In other words, she had no memory of the object's dimensions because she failed to perceive those dimensions in the first place. Her damaged ventral stream meant that she did not form a percept of the object's form when she was exposed to it initially. In the same experiment, however, D.F. was able to pantomime grasping movements to familiar objects such as a pea or a tennis ball, on verbal instruction, presumably because the size information in this case could be retrieved from her long-term memory of these objects. Taken together, these results suggest that (1) the dorsal

system normally operates in real time, accessing transient visual signals about the target object at the time of movement programming, and (2) the ventral stream is required for creating the object representations that are maintained in memory for the control of later actions. These conclusions are supported by evidence from other neurological patients who have damage to the dorsal rather than the ventral stream. These patients show the opposite pattern of deficits and spared visual abilities from those seen in D.F. (a double dissociation). The optic ataxia patient A.T., for example, is quite unable to scale her grasp when reaching to unfamiliar objects—even though the objects remain visible both before and during the action. Presumably, this deficit arises because such actions are normally generated exclusively on the basis of online visual processing of the objects' features—processing that is thought to take place in the dorsal stream, which is damaged in this patient. A.T. shows much better scaling of her grasp when she reaches to familiar objects, however, where the appropriate actions can be programmed on the basis of stored semantic information about the object that is derived from long-term memory (Jeannerod et al., 1994). The memory is presumably triggered by her visual recognition of the familiar object using the intact perceptual mechanisms in her ventral stream—which she can then use to shape her grasp. But what about actions directed to unfamiliar objects that are no longer visible, objects whose features cannot be accessed from long-term memory but must instead be stored in short-term memory? If, as suggested earlier, the control of such memory-driven movements depends not on the visuomotor mechanisms of the dorsal pathway but rather on the perceptual mechanisms of the ventral pathway, then one might expect a patient with optic ataxia to show a paradoxical improvement in performance if the grasping movement is delayed until the goal object has been removed from view. In this situation the patient would be forced to rely on a stored representation of the target object, laid down moments earlier by the perceptual mechanisms of the ventral stream, rather than an online computation of the object's features, which would normally engage the visuomotor mechanisms in the (damaged) dorsal stream.

This is exactly what happened in a recent experiment with the optic ataxia patient I.G. (Milner et al., 2001). Like other patients with optic ataxia, I.G. is quite unable to open her hand and fingers appropriately when reaching out to pick up objects of different sizes. Yet despite this deficit in real-time grasping, I.G. showed good grip scaling when “pantomiming” a grasp for an object that she had seen earlier but that was no longer present. In fact, after some practice, I.G. was able to scale her grip when grasping a real target object that she had previewed 5 s earlier—even though the object was visible during the movement. In other words, despite the presence of real-time vision, her performance improved by virtue of the preview. By interposing catch trials in which a different object was covertly substituted for the original object during the delay between the preview and the grasp, the experimenters

were able to show that I.G. was using memorized visual information to calibrate her grasping movements. In other words, on these catch trials, her grip scaling reflected the size of the object she had previewed earlier rather than the size of the object that was now in front of her, in contrast to what happened with normal participants. These findings again provide compelling support for the idea that the control of grasping movements made after a delay depends on information derived from earlier perceptual processing of the object by mechanisms in the ventral stream. As will become evident in the next section, this ventral-stream perceptual processing plays a significant role in the apparent priming of some kinds of visuomotor responses by earlier exposure to the target stimulus.

4.4 Can the Visuomotor System be Primed?

The more recently we have seen an object, the easier it is to recognize when it is presented again, even when we cannot recall having seen the object earlier. The improvement in recognition from earlier presentation of the target stimulus, which is often called *priming*, has been the subject of extensive behavioral research over the last 2 decades. Although it is clear that the perceptual processing of a visual stimulus and/or access to various semantic associations, such as its name, show evidence of priming (e.g., Thompson-Schill & Gabrieli, 1999), it is not clear that this would be the case for visuomotor processing for actions directed at that same object. Indeed, the perception–action dissociation put forward by Goodale and Milner (1992) would suggest that such priming should not occur. As was discussed earlier, according to their account, the dorsal “action” system works in real time and does not make use of memory for the computations that are carried out in programming an object-directed action. Memory-guided actions, it seems, are mediated instead by the ventral “perception” mechanisms and thus make use of visual information that is quite distinct from that used to control real-time actions.

Despite these arguments, one recent study claims that visuomotor priming can and does occur. Craighero et al. (1996) demonstrated that the time to initiate a grasping response was reduced when participants viewed a priming stimulus (presented 100ms earlier) that was congruent in orientation with the target object (as compared to their response on trials in which an incongruent or neutral priming stimulus was presented). They interpreted this improvement in reaction time as evidence for a visuomotor priming effect, suggesting that earlier exposure to the congruent prime reduced the amount of processing required to program the goal-directed action. On the face of it, this idea is completely inconsistent with Goodale and Milner’s view that visuomotor computations are made immediately before the movement is made, with no computations being stored in memory that

can be used to execute subsequent goal-directed movements. How, then, can the results of Craighero et al. be explained in light of this real-time view of visuomotor transformation?

A critical aspect of the Craighero et al. (1996) study was that the target object was *never visible* to the participants. Participants were told by verbal instruction at the beginning of the trial what the orientation of the target object would be—and this instruction was always correct. Only later was the visual “prime” presented (just before the grasping movement was initiated). This meant, of course, that participants would have had to plan their action from memory rather than from direct visual input. As discussed earlier, memory-guided actions appear to engage quite different visuomotor mechanisms from those used by visually guided actions (and are, in fact, quite different from our everyday interactions with objects—in that we usually see the object that we wish to pick up). Thus, the fact that Craighero et al. used a memory-driven grasping movement—a movement that presumably is programmed using information derived from the perceptual processing in the ventral stream—might explain why they found evidence of priming in their task. However, what is needed to test this directly is a direct comparison of priming in memory-driven and visually driven grasping.

4.4.1 Visuomotor Priming with Novel Objects

Cant et al. (2005) recently conducted a series of experiments to address this issue. In their first experiment, they compared priming of memory-guided and visually guided grasping using a paradigm similar to that of Craighero et al. (1996). The prime and target objects were presented on a turntable (figure 4.2), which could be rapidly rotated after participants had viewed the prime, and viewing time was controlled by LCD goggles that could quickly be made transparent or opaque. In both the memory-guided and the visually guided tasks, participants were instructed to grasp the target object (a long, rectangular block) as quickly as possible upon hearing an auditory cue. Each trial began with an auditory verbal cue that indicated the orientation of the target object (i.e., the word *left* when the far end of the block was oriented 45° leftward from the participant’s midline, or *right* when it was oriented 45° rightward). After a 1,500-ms interval, the goggles became transparent and revealed the prime object for a duration of 500 ms. The goggles then were occluded for a randomly varied interstimulus interval (ISI) of 1,250, 1,500, or 1,750 ms. During the ISI, the turntable was rotated in order to position the target object in the work space. After the ISI, an auditory signal cued participants to respond. In the visually guided task, the goggles were made transparent coincidentally with the auditory cue and then vision was occluded either when the response was initiated or after 500 ms, whichever came first. Thus, participants received direct visual information about the target object during the movement-programming phase but not during



Figure 4.2

Turntable apparatus used to present the prime and target objects to the participants. (Reproduced with permission from Cant et al., 2005)

the online control phase. In the memory-guided task, the goggles remained occluded after the auditory cue, so participants received no direct vision of the target object at any time.

In both tasks, grasping movements were made under four different priming conditions (congruent, incongruent, no-prime, and neutral), which were randomly interleaved. In the congruent and incongruent conditions, a rectangular prime object, identical in shape to the target, was presented in either the same or different orientation with respect to the target object. In the no-prime condition, no object was presented during the prime-viewing period; this condition was included to determine a baseline reaction time for visually guided and memory-guided grasping. In the neutral condition, a circular disk was presented during the prime-viewing period; this condition was included to determine whether or not viewing any object at all—even one that bears no visual similarity to the upcoming target object—could influence the speed with which a grasping movement is initiated. A similar condition was also used by Craighero et al. (1996).

As figure 4.3 indicates, the results from this first experiment were clear: Memory-guided, but not visually guided, grasping movements were primed. In the memory-guided grasping task, participants initiated their grasp to the target significantly faster after viewing a congruent prime than after an incongruent prime. The fact that visually guided grasping movements were completely unaffected by the earlier presentation of the priming condition is consistent with the idea that dedicated visuomotor networks that mediate object-directed movements operate in real time, and the necessary computations are based on information derived directly from the retina rather than from memory of earlier presentations. When the object is not visible, however, the grasping movement has to be based on stored information—and this is presumably the reason that presenting congruent or incongruent stimuli just before the memory-driven movement was initiated had an effect on response latency. This latter finding, of course, replicates the Craighero et al. (1996) study—and indicates that “visuomotor” priming may in fact be limited to memory-guided actions.

Even so, it has to be acknowledged that the paradigm that Cant et al. (2005) adapted from Craighero et al. (1996) is somewhat unconventional. In most perceptual priming studies, no advance information is given about the features of the upcoming target stimulus. Yet in the Craighero et al. experiment (and in Cant et al.’s replication), participants were given accurate verbal information about the

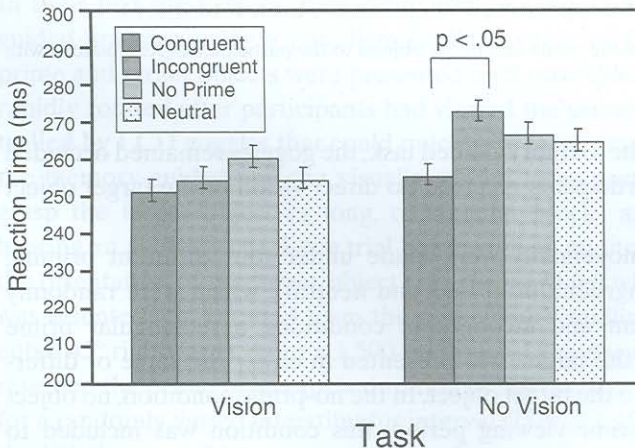


Figure 4.3

Reaction times for each experimental condition in the visually guided and memory-guided tasks used in the first experiment by Cant et al. (2005). Reaction times for visually guided grasping were not significantly different for the four conditions. For memory-guided grasping, however, reaction times for the congruent prime–target trials were significantly faster than those for the incongruent trials. In all cases, Bonferroni pairwise comparisons were used. (Reproduced with permission from Cant et al., 2005)

target's orientation well before the so-called priming stimulus was presented. Under these conditions, Cant et al. found no evidence for visuomotor priming (for visible targets). It is possible, however, that a priming paradigm of the sort known to produce robust priming in a naming task with visible objects (e.g., Wiggs & Martin, 1998) might result in visuomotor priming—even in a visually guided grasping task. Cant et al. explored this possibility in an additional experiment that made use of a more conventional priming paradigm.

In this experiment, each trial consisted of a prime stimulus followed a short time later by the target object, and the effects of presenting the prime were explored in both a visually guided grasping task and a naming task. As figure 4.4 illustrates, four novel objects that varied in shape were used; each object afforded the same whole hand grasp in which the thumb opposed the finger pads of three or four of the remaining digits. Participants were first taught to associate a name with each of the four shapes (*kiff*, *fid*, *tam*, and *sup*). Following this training period, the participants were tested in both the naming and the grasping tasks. In each task, the prime stimulus was presented for 500ms and then the target was presented following a variable ISI (1,250, 1,500, or 1,750ms). On some trials, the prime and the target were identical in shape and orientation. On others, they had the same shape but different orientations, or vice versa. On still other trials, they differed in both orientation

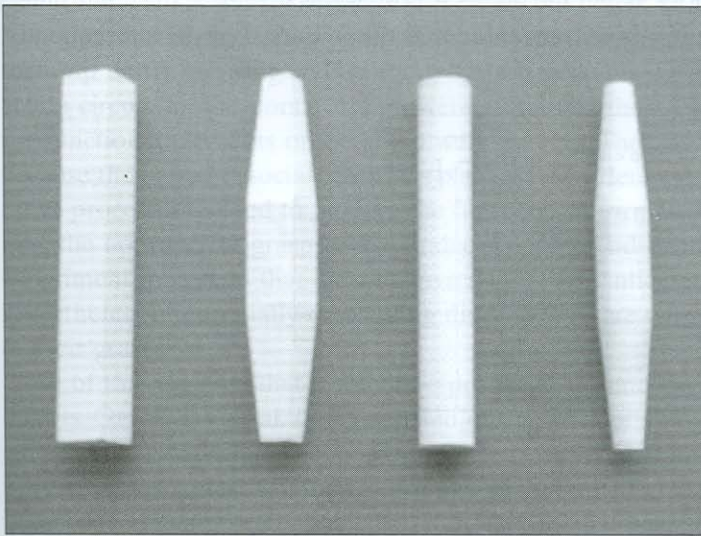


Figure 4.4

Stimuli used in the second experiment by Cant et al. (2005). Each object was given a nonsense name (i.e., *kiff*, *fid*, *tam*, and *sup*). The shape of the blocks varied on two dimensions: cross-section (circle vs. square) and tapering (tapered vs. nontapered). (Reproduced with permission from Cant et al., 2005)

and shape. Finally, there were “baseline” trials in which no priming stimulus was presented at all. All these different trials were randomly interleaved.

In the naming task, participants were faster to name the target when its shape was the same as the prime—but there was no effect of orientation (figure 4.5). This demonstrated that the paradigm could reproduce the standard priming effects on naming that have been reported many times in the literature. However, despite the robust effect on naming, the reaction time for grasping was completely unaffected by the orientation or the shape of the prime (see figure 4.5). This clear difference between naming and grasping confirmed that the null effect of priming on visually guided grasping in the first experiment was not due to the peculiarities of that particular paradigm.

It could be argued, however, that priming was not found in the grasping tasks because the visuomotor circuitry underlying visually guided grasping was not sufficiently engaged during the priming phase (whereas some sort of covert naming could have occurred). Thus, in a final experiment, Cant et al. (2005) required participants to grasp the prime before grasping the target object. Again, reaction time remained unaffected by the prime stimulus. This inherent lack of priming (for all visually guided grasping tasks across all three experiments) is entirely consistent with the real-time view of visuomotor control. To reiterate, the initial programming of visually guided grasping appears to be determined more by what is on the retina than by what is in memory. When the object is not visible, however, the “real-time” dorsal visuomotor circuitry is circumvented and the initiation of the subsequently

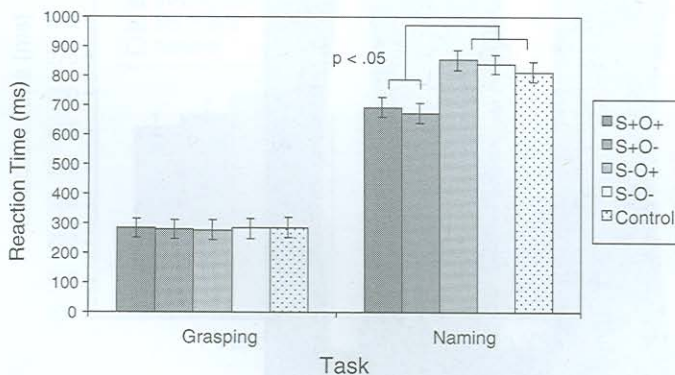


Figure 4.5

Reaction times for each experimental condition in the grasping and naming tasks used in the second experiment by Cant et al. (2005). Reaction times for grasping were not different for the five experimental conditions. For naming, reaction times were reduced for conditions in which prime and target objects were the same shape. S+, same shape; S-, different shape; O+, same orientation; O-, different orientation. (Reproduced with permission from Cant et al., 2005)

“blind” grasping movement is executed by the perceptual circuitry in the ventral stream.

4.4.2 Visuomotor Priming with Familiar Objects

But can priming ever affect the visuomotor system? Behavioral studies supporting the idea of the online nature of visuomotor control (including the experiments by Cant et al., 2005) have used simple wooden blocks and/or nonsense objects. Such objects have no particular function—and thus there is no stored representation of function that could be “primed” and later used by the vision-for-action system. Perhaps the situation would be very different for familiar objects, such as tools, that invoke functional hand postures in anticipation of their use.

There is certainly evidence that functional objects engage the visuomotor system in a rather different way than meaningless objects do. In a series of elegant experiments, for example, Creem and Proffitt (2001) showed that the processing of functional components of the grasp operates independent of the more metrical aspects of grasp scaling. In their studies, participants were presented with a series of tools and implements, such as a hammer, a toothbrush, and a screwdriver, with the handle turned away from them. When the participants reached out to pick up these objects, they grabbed the handle even though this meant adopting an uncomfortable posture. If, however, they were required to carry out a paired-associate memory task at the same time, the participants picked up the object as if blind to its function. Nonetheless, although they grasped the objects inappropriately, they still picked them up deftly, showing well-calibrated grasp postures. In other words, vision-for-action circuits in the dorsal stream were still doing their job as well as ever. Only the functional elements of the grasp were missing. They were missing presumably because the paired-associate task was placing heavy demands on the high-level cognitive processes needed to retrieve the functional information about the object and thus the appropriate grasp to use. Indeed, Creem and Proffitt showed in another experiment that tasks that did not have a large semantic component but that were nevertheless attentionally demanding did not interfere with the functional aspects of their grasps.

All of this suggests that it might be possible to prime the grasping of functional objects. Garofeanu et al. (2004) carried out a series of studies designed to investigate just this possibility. The motivation was twofold. First and foremost, the studies were designed to test whether or not priming of grasping would occur when familiar objects are used as goal objects. (Rare studies that have investigated facilitation of visuomotor responses with such objects—e.g., Tucker & Ellis, 1998; see also Tucker & Ellis, 2004—have used mainly depicted objects as stimuli and/or a button-press response as a measure of reaction time, a response that may depend more on initial ventral rather than dorsal-stream processing.) Second, the Garofeanu et al.

studies explored whether or not grasping and naming of familiar objects would be differentially affected by the nature of the priming event. (Behavioral studies of visual priming with familiar objects have concentrated almost entirely on recognition tests limited to line drawings or photographs, e.g., Williams & Tarr, 1999; Liu & Cooper, 2001; see also Biederman & Gerhardstein, 1993, and/or the stimuli often consisted of verbal [semantic] material, e.g., Thompson-Schill & Gabrieli, 1999; for a review, see Schacter & Buckner, 1998.)

Garofeanu et al. (2004) measured reaction time when participants grasped and/or named objects in a repetition-priming paradigm. In such a paradigm, an object is first presented during a study phase and then presented again during a test phase, along with other objects that were not presented before (see Schacter, 1987; Tulving & Schacter, 1990; Toth, 2000). The effects of earlier exposure on later performance—for example, decreased latency and/or better object identification for old as compared to new objects—are measured. If any effects are found, they are assumed to depend on implicit memory of that previous encounter. It is thought that implicit memory must be involved because the experimenter never makes any reference to the prior occurrence of the stimulus, and participants are never explicitly required to remember the studied objects for later recall or recognition tests.

In their first experiment, Garofeanu et al. (2004) looked at the effects of prior grasping on later grasping or naming of the same objects and at the effects of prior naming on later naming or grasping. In other words, there were two within-task conditions in which priming could occur (grasping–grasping and naming–naming) and two cross-task conditions (naming–grasping and grasping–naming). Forty common objects (from several superordinate categories, such as household objects, personal care objects, and tools) were presented. All the objects had a prominent axis of elongation and required a specific hand posture during grasping. These objects were presented one at a time at a distance of 60 cm in front of the participant, 20 of them at an orientation of 0° , where the handle pointed away from the participant, and the other 20 at an orientation of 135° , where the handle pointed to the right and toward the participant (for examples, see figure 4.6). In the study phase, participants were shown only 20 of the objects (half at an orientation of 0° and the other half at an orientation of 135°) and were asked either to grasp or to name them. In the later test phase, the same 20 objects were presented randomly interleaved with the 20 new objects—and again the participants were asked to either grasp or name all the objects. The orientation of the old objects did not change from study to test. In a second experiment, Garofeanu et al. manipulated the orientation of the object; on some blocks of test trials the orientation of the old objects was changed from study to test, and on others the orientation remained the same. In this second experiment, the two orientations were mirror symmetrical (135° vs. 225°) and demanded quite different hand postures. From the point of the view of the participants in both these

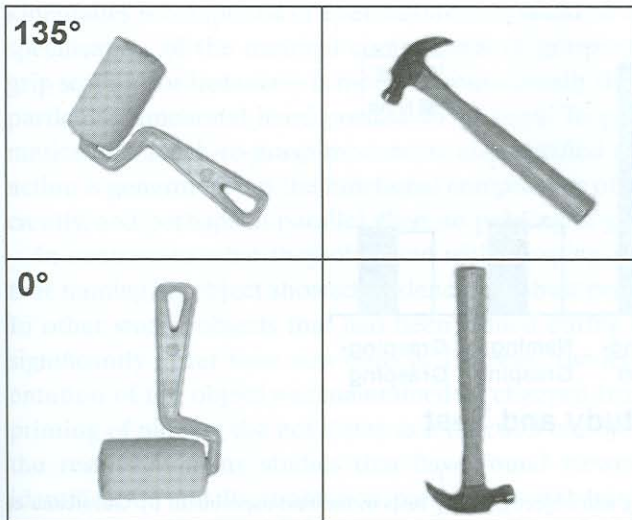


Figure 4.6

Examples of the objects used in the first experiment by Garofeanu et al. (2004). The objects were presented in one of two different orientations during the study and test phase of the experiment. (Reproduced with permission from Garofeanu et al., 2004)

experiments, however, there was no real difference between the study phase and the test phase, except that the latter lasted longer.

No priming of grasping was found in either experiment. In other words, the latency to grasp studied objects did not differ from the latency to grasp new objects (figures 4.7 and 4.8). Of course, in those cases in which the orientation of the object changed from study to test, it was perhaps not surprising that there was no reduction in latency, since the required hand movement would be quite different and an earlier encounter with the object would not be that helpful. However, even when the orientation of the object was maintained from study to test, there was no latency advantage for the studied over the new objects.

The complete absence of priming in this study is surprising given that the objects to be grasped were tools and implements—objects that are associated with particular functions and thus particular hand postures. Even though one might not expect the metrical aspects of the grasping movement to be subject to priming, one might still think it possible that the generation of function-appropriate hand postures might be subject to priming (by virtue of the fact that the identification of the object—and thus its function—was almost certainly primed). The absence of priming suggests that the visuomotor system has quick and automatic access to functional components of the grasp well before the programming of the reach-to-grasp

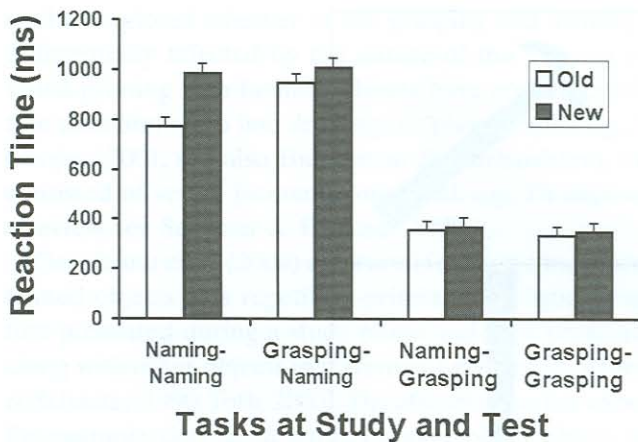


Figure 4.7

Reaction times in the object-naming and object-grasping tests in the first experiment by Garofeanu et al. (2004). There were four conditions: naming-naming, grasping-naming, naming-grasping, and grasping-grasping (at study and test, respectively) under which the old (studied) objects were encountered. The new (nonstudied) objects were named or grasped only in the test phase. Participants showed significantly faster naming of the objects they had named and grasped earlier. Participants did not grasp the studied objects any faster than they did the new objects, irrespective of whether the old objects had been named or grasped earlier. (Reproduced with permission from Garofeanu et al., 2004)

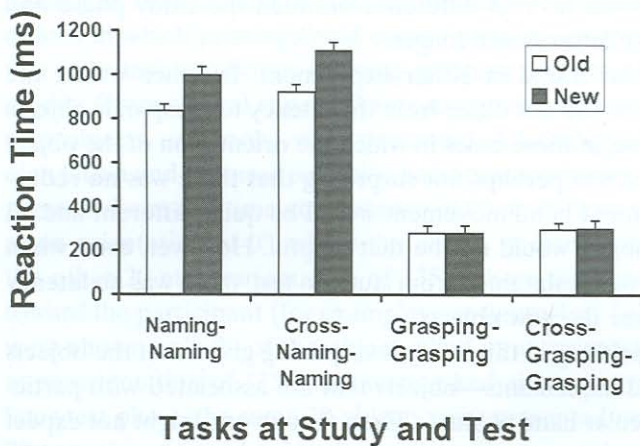


Figure 4.8

Reaction times in the object-naming and object-grasping tasks, with and without the change in object orientation. Irrespective of whether the orientation was maintained or changed (crossed) from study to test, the objects previously named were named significantly faster than new objects. The objects previously grasped were not grasped faster. (Reproduced with permission from Garofeanu et al., 2004)

kinematics is completed or even initiated. It could be the case, for example, that the specification of the metrical components of grasping—movement amplitude and grip scaling, for instance—is more computationally demanding than selection of the particular functional hand posture to be used. To put it another way, if the kinematics of a reach-to-grasp movement are specified *de novo* on every occasion an action is generated, but the functional components of the grasp can be accessed efficiently, and perhaps in parallel, then no priming of grasping might be evident.

In contrast to what they observed with grasping, Garofeanu et al. (2004) found that naming an object showed evidence of robust priming (see figures 4.7 and 4.8). In other words, objects that had been named during the study phase were named significantly faster than new objects. This was true irrespective of whether the orientation of the object was maintained or changed from study to test. The fact that priming of naming did not differ as a function of object orientation converges with the results of many studies that have found viewpoint independence in object identification and discrimination, particularly with mirror symmetrical presentations (e.g., Biederman & Gerhardstein, 1993; Cant et al., 2005; James et al., 2002). However, naming was also primed by previous grasping of the same objects. That is, even if the object was only grasped at study, its naming was then significantly faster at test. (As was already discussed, the opposite was not true: Grasping was not faster when preceded by naming.) Further tests also showed that naming was faster regardless of the method of examination used in the study phase. It did not really matter whether, during study, participants named the object, grasped the object, matched the orientation of their arm to the object, or simply observed the object; subsequent naming of these objects at test was always faster as compared to the naming of objects that had not been seen before (although explicit naming during the study phase did produce the best priming).

Even though Garofeanu et al. (2004) found no evidence for visuomotor priming, they did find some improvement of performance over the course of the experiment. Participants were faster at initiating their grasping movements to the target objects during the test phase than during the study phase—independent of whether or not the objects were old or new. In other words, participants got better as they became more experienced with the task. Practice helps with visuomotor performance, as anyone who has learned to ride a bicycle can attest to. Nevertheless, there is little evidence that the storage of specific visual information has much of a role to play in the programming of movements directed at visible targets. In contrast, visual memory appears to play a much bigger role in perceptual processes, such as object recognition and identification.

4.5 The Role of Visual Memory in the Control of Action

The results of Cant et al. (2005) and Garofeanu et al. (2004) demonstrate that prior perceptual/semantic processing of objects (including their orientation) does not prime later grasping of those objects. Even when a “transfer-appropriate” paradigm is used, in which the objects are grasped during both the study phase and the test phase, there is no indication of priming. Yet, at the same time, any kind of prior interaction with the objects, cognitive or visuomotor, appears to prime later naming of those objects. These findings converge nicely with a broad range of findings in both neurological patients and normal observers discussed earlier showing that the dedicated visuomotor mechanisms in the dorsal stream do not make use of visual information that is presented *prior* to the programming of a target-directed action. Indeed, recent work by Westwood and Goodale (2003) has shown that for information about target size, for example, to be reflected in the scaling of a grip aperture during a grasping movement, that information must be present on the retina during the programming of the grasp.

Nevertheless, in setting out this real-time account of how visuomotor processing occurs, we do not mean to imply that visual memory does not have a role to play in some aspects of visuomotor control. Visual memory contributes quite directly to certain aspects of motor programming, notably those that depend on information that cannot be derived directly from the retina. One example is the calibration of the initial grip and lift forces that are applied to an object when it is picked up. When people pick up familiar objects, the force they apply to the object is scaled appropriately for the object’s weight (and other characteristics) from the moment their hand makes contact, well before any somatosensory feedback is available (Gordon et al., 1993). This means that they must have retrieved the information about the object from memory, information that could be activated only when they recognized the object. Such recognition is presumably mediated by visual mechanisms in the ventral, not the dorsal, stream. At present it is not clear how this information is communicated to the motor system, although a number of plausible accounts have been put forward (for a discussion of this issue, see Goodale & Haffenden, 2003).

When people reach out and pick up objects they have not encountered before, they have to make an estimate of the object’s weight—an estimate that is presumably based on previous experience with similar objects. Thus, when people reach out to pick up objects that “look” heavy, such as objects that appear to be made of stone or metal, they apply more force than they do to objects that look as though they are made of lighter materials, such as polystyrene. Moreover, when participants are presented with objects of different sizes that are made of the same material, they typically apply more force to the larger object than they do to the smaller objects—even when the objects are actually the same weight (Gordon et al., 1991). When

participants heft such objects, they typically experience what has been called the “size–weight illusion” and conclude that the smaller object is heavier than the larger one. Even though it has been often assumed that the size–weight illusion is a consequence of the mismatch between applied force and the actual weight of the object, it turns out that this is not the case (Flanagan & Beltzner, 2000). When participants make repeated lifts of a set of objects that have the same weight but different sizes, they eventually apply the same initial lift force to all the objects. Nevertheless, the size–weight illusion is not at all diminished by this recalibration of lift force. In other words, participants continue to believe that the smaller object is heavier than the larger one even though they apply the same initial lift force to both.

Despite this apparent dissociation, there are good reasons to believe that both the calibration of the lift forces and the maintenance of the size–weight illusion depend on object recognition systems in the ventral stream that activate specific memories and/or related information about those particular objects. First of all, if the objects to which the observers have recalibrated their lift forces are replaced with new objects (with different surface features), which like the original objects have different sizes but the same weight, the observers now revert to applying less force to the smaller object in the new set (J. R. Flanagan, personal communication). Second, the patient D.F., who has visual form agnosia from ventral-stream lesions, does not show any evidence of a size–weight illusion when she picks up cylindrical objects of different sizes (but the same weight) using a handle attached to the top of the cylinders. In other words, when the only information that D.F. has about the size of the object is derived from vision, she correctly reports that the cylinders are the same weight. Moreover, unlike normal participants, D.F.’s grip and lift forces on the initial trials are not correlated with the size of the cylinder; that is, she does not apply greater force to the larger object (McIntosh, 2000). In contrast, when she picks up the cylinders across their diameter with either her eyes open or her eyes closed, she does show the size–weight illusion. In other words, when she has kinesthetic information about the size of the cylinders, she incorrectly perceives the smaller cylinders as weighing more than the larger ones. A similar dissociation between visual and kinesthetic size–weight illusions has recently been observed in patient S.B., who also has ventral-stream lesions (Dijkerman et al., 2004). Taken together, these findings suggest that stored information about objects that is activated by ventral-stream mechanisms plays an essential role both in computing the required forces for a successful grasp and in generating and maintaining the size–weight illusion.

As we saw earlier, the perceived function of an object can also determine the nature of the hand posture that we select to use when picking it up. Thus, we pick up screwdrivers and hammers by the handle rather than the “business end”—often adopting an awkward posture to accomplish this. And as we also reviewed earlier,

there is evidence that these functional components of grasping depend on high-level cognitive processing, which presumably makes use of object recognition circuitry in the ventral stream (Creem & Proffitt, 2001).

There are other aspects of our interactions with familiar objects in which visually driven memories can play an important role. For example, we typically pick up the same toothbrush and the same coffee cup every morning. As a consequence, non-metrical features of these objects, such as their color, can become associated with metrical properties such as their size and shape—and these reliable associations could then be used to program grip aperture. To put it another way, just as we apply the appropriate forces to pick up familiar objects, we can also use learned associations to scale our grasp to familiar objects. Using perceptual information to recover a motor routine could theoretically reduce the need for bottom-up computations of the metrics of the target object, thereby increasing the overall efficiency of the programming. That is, the incorporation of stored parameters such as grip aperture could reduce considerably the computational load on the visuomotor system. Thus, even though the grasp would have to be fine-tuned and adjusted to the particular situation, information retrieved from memory could provide the initial parameters for selecting the posture and scaling the grasp. In short, perceived object features such as color, which have no inherent link to the control of goal-directed actions, could through association provide a cue for the programming of such control. Nevertheless, experiments that have explored these learned associations have shown that the conditions under which such “shortcuts” are used are heavily constrained. For example, if the position of the familiar object is changed randomly from trial to trial, participants are less likely to make use of such learned motor routines (Haffenden & Goodale, 2002b). Moreover, the learned associations are more likely to be based on global cues derived from the entire surface of the object than on specific cues provided by isolated symbols or shapes on the object’s surface (Haffenden & Goodale, 2000a, 2002a). Nevertheless, there is clearly a role for visual memory in some aspects of visuomotor programming and control—particularly with respect to those features of the goal object that remain relatively constant over time, such as size, shape, weight, compliance, and surface friction.

However, despite the fact that many aspects of our everyday actions can be influenced by prior experience, the evidence from the priming studies reviewed earlier suggests that many of the kinematic parameters of object-directed reaching and grasping movements are programmed in real time on the basis of current sensory information—with little influence from earlier encounters with the goal object. This reliance on bottom-up input rather than stored information makes good sense, particularly with respect to features such as the orientation and location of the goal object, since it is unlikely that the disposition of the object with respect to the observer would remain constant over time. But even apparently immutable features

such as size or shape might not be expected to show evidence of priming either. After all, when the egocentric position of an object is changed, the affordance offered by that object can also change. In other words, the size of the grip and the configuration of the grasping fingers required to pick up an object are functions of the disposition of that object with respect to the hand. For this reason, then, prior information about the size or shape of a goal object in the real world should not be expected to affect the efficiency of movement programming—if that object is actually present and visible when the movement is programmed.

Although repetition priming has little effect on grasping, there is a long history of work showing that manipulation of attention to particular spatial locations can indeed reduce the response latency of orienting movements. In other words, precuing a particular location can speed up orienting responses to that location when the target is presented there. At present, it is not clear what the neural substrates are for this advantage, although there has been much speculation about what the circuitry might be (for a review, see Findlay & Gilchrist, 2003). The temporal parameters of this kind of spatial attention deployment are quite different from those involved in typical repetition-priming studies.

In summary, the visuomotor modules in the dorsal stream appear to work largely in real time—and earlier exposure to goal objects appears to have no apparent effect on the programming of the visually guided actions. The reliance of the dorsal stream on visible rather than stored information is clearly adaptive. In a world where the disposition and affordances of goal objects (with respect to the actor) can change from one moment to the next, it makes little sense to rely on visual memory. Only when vision of the goal object is not available at the time the movement is programmed does information from memory become useful. When this happens, offline information that was originally processed by the ventral stream is brought to bear on the programming of the memory-guided movement. In short, the two streams play complementary roles in the control of behavior.