Visual Neuroscience (2009), **26**, 51–62. Printed in the USA.
Copyright © 2009 Cambridge University Press 0952-5238/09 $25.00
doi:10.1017/S0952523808080899

Vision, eye movements, and natural behavior

MICHAEL F. LAND

Department of Biology and Environmental Science, University of Sussex, Brighton, UK

(Received June 27, 2008; Accepted October 31, 2008)

Abstract

Historically, the principal function of vision has been to provide the information needed to support action. Visually mediated actions rely on three systems: the gaze system responsible for locating and fixating task-relevant objects, the motor system of the limbs to carry out the task, and the visual system to supply information to the other two. All three systems are under the control of a fourth system, the schema system, which specifies the current task and plans the overall sequence of actions. These four systems have separate but interconnected cortical representations. The way these systems interact in time and space is discussed here in relation to two studies of the gaze changes and manipulations made during two ordinary food preparation tasks. The main conclusions are that complex action sequences consist of a succession of individual object-related actions, each of which typically involve a turn toward the object (if needed), followed by fixation and finally manipulation monitored by vision. Gaze often moves on to the next object just before manipulation is complete. Task-irrelevant objects are hardly ever fixated, implying that the control of fixation comes principally from top-down instructions from the schema system, not bottom-up salience. Single fixations have identifiable functions (locating, directing, guiding, and checking) related to the action to be taken. Several variants of the basic object-related action scheme are discussed, including single-action events in ball sports involving only one anticipatory gaze shift, continuous production loops in text and music reading, and storage–action alternation in copying tasks such as portrait sketching.

Keywords: Gaze control, Object-related actions, Extended tasks, Schemas, Anticipatory saccades

Introduction: How gaze assists action

For our neolithic ancestors, in the time before vision had become domesticated by writing and computers, the primary role of gaze would have been to provide the information needed to carry out the tasks of everyday life—hunting, foraging, constructing shelters, cooking, and so on. Perception, in the sense of discerning the relations of sensations to the external world, has a part to play in this, but of greater importance would have been the guidance of action: steering the individual through the environment and guiding the limbs in the execution of the activities required for survival.

This review is in three parts. First, the different subsystems involved in the visual control of action are briefly introduced. This is followed by a discussion of how these systems work together in practice, based on two well-studied accounts of the actions involved in the preparation of food. These were chosen mainly because they involve many different actions and reveal something of the way the brain controls and schedules complex tasks. Finally, three important variants of the basic control pattern are described. Many of the examples here have been presented in other reviews (Land, 2003, 2006). What is new here is a systematic attempt to discuss this material in terms to the interactions between the component systems involved, and the relations of these systems to known brain structures and pathways.

Address correspondence and reprint requests to: Michael F. Land, Department of Biology and Environmental Science, University of Sussex, Brighton BN1 9QG, UK. E-mail: m.f.land@sussex.ac.uk

The components of the gaze–action system

Several distinct but interacting brain systems are involved in the visual control of action, and what I hope to do in this article is to define their relative roles in different types of activity. Consider first a simple task such as picking up a mug. It is clear that three different systems are involved. First, the mug must be found. Assuming that the mug is not hidden, it is the job of the gaze system to initiate movements of the eyes, and usually of the head and trunk as well, to bring the image of the mug onto the foveas of the eyes. This done, the motor system that controls the arms and hands can move to make contact with the mug. Information from the visual system is needed for several purposes. It will first need to check that the object fixated is indeed a mug, then provide the motor system with appropriate directional guidance, and finally check details such as the orientation of the handle to allow the hand to prepare the correct grip. Thus, each of the three systems—gaze, motor, and visual—has a distinct role, namely locating, acting, and supplying information as needed. In addition, there has to be a fourth system that has overall control of the other three systems. The next task to be performed, perhaps chopping vegetables with a knife, also requires the gaze, motor, and visual systems, but now their specific tasks are quite different. There has to be a controller that not only contains the job list of activities to be

https://doi.org/10.1017/S0952523808080899 Published online by Cambridge University Press

performed but has the power to reprogram the three ''executive'' systems so that they produce a new coherent action. We refer to this controller here as the ''schema'' system, where schema is used to mean the internal representation of a task (following Norman & Shallice, 1986). The relations of these four systems are outlined in Fig. 1A.

The locations of these systems in the brain—at least their cortical components—are well known (Fig. 1B; review in Jeannerod, 1997). The gaze system consists of the frontal eye fields (FEFs) and the lateral intraparietal area of the parietal lobe. These are reciprocally connected and are also reciprocally connected to the superior colliculus and ultimately the oculomotor nuclei of the brain stem. The motor areas consist of the primary motor cortex, the premotor cortex, and various parietal areas concerned with reach and grasp. The visual system occupies the occipital lobe and much of the temporal lobe. Again there are extensive connexions to the parietal and frontal lobes. The region most associated with the schema system, and with purposeful action in general, is the dorsolateral prefrontal cortex (DLPFC). Damage to this region does not affect the execution of individual actions but compromises their organization into coherent sequences. The arrows linking the schema system to the other systems in Fig. 1A can be thought of as the routes of descending attention. Such influences on the gaze and visual systems are well documented, although the link to the motor system is rarely discussed. It is logically necessary, however, since the motor system must be linked appropriately to the activation of gaze and vision in the execution of each action. The fact that the functions outlined in Fig. 1A appear to have specific brain regions associated with them emphasizes their distinctness: actions do not happen as emergent consequences of global brain activity but through the coordinated contributions of separate component systems.

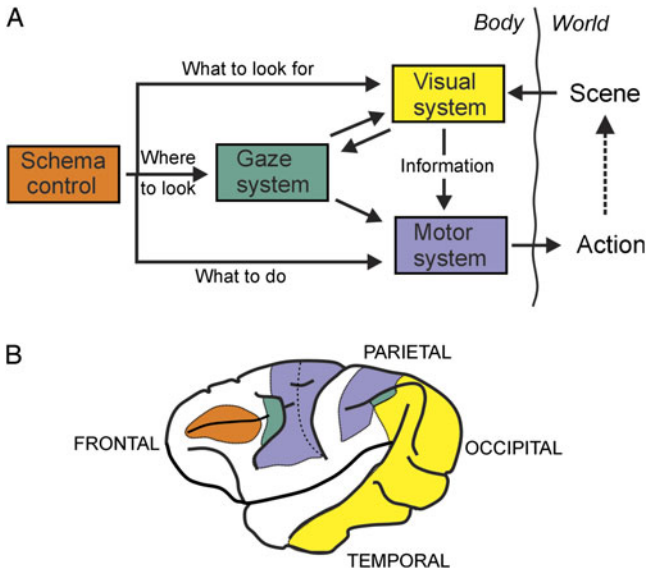


Fig. 1. (Color online) (**A**) Relations of the schema, gaze, visual, and motor systems during the performance of a visually controlled action. (**B**) The regions of the left macaque cortex known to be involved in the control of the systems outlined in **A**. The schema system is mainly associated with the DLPFC, the gaze system with the FEFs and the lateral intraparietal cortex, the motor system with the frontal premotor and motor cortices and parts of the parietal cortex, and the visual system with occipital and much of the temporal lobes.

The structure of extended tasks

To see how the scheme in Fig. 1 works in practice, I will use the results of two earlier studies on food preparation: making a cup of tea (Land et al., 1999) and making peanut butter and jelly sandwiches (Hayhoe, 2000; Hayhoe et al., 2003). The results of the two investigations were similar in most important respects (Land & Hayhoe, 2001). The reason for choosing these tasks was that they both involve many separate actions, and the transitions from one action to another provide important clues as to how the schema system manages the operation of the three executive systems. In both studies, eye movements and gaze direction were monitored using a head-mounted eye camera, and body and hand movements were filmed by a separate scene camera in the corner of the room. The position of relevant objects in the kitchen or on the counter was not known to the subjects before they began, although the overall room layout was. No particular instructions were given, other than to make a cup of tea or the sandwiches. The tasks each took several minutes. As much of the material has been reviewed elsewhere (Land & Hayhoe, 2001; Land, 2003, 2006), I will concentrate here on those aspects of the tasks that shed light on the way that the four subsystems in Fig. 1 interact.

Fig. 2A shows a scan path of the first 26 fixations of tea making—filling the kettle—and Fig. 2B the first 15 fixations during the making of a sandwich. Details of the actions accompanying the gaze movements are given in the figure caption. The principal conclusions that can be drawn from these scan paths are the following:

(1) Saccades are made almost exclusively to objects involved in the task, even though there were plenty of other objects around to grab the eye. This is important because it implies that ''bottom-up salience,'' in the sense used by Itti and Koch (2000), is of little if any relevance in these kinds of tasks. Of particular interest here was an observation by Hayhoe et al. (2003) in which the implements used for sandwich making were mixed with an equal number of irrelevant implements (pliers, forks, and tape). Before the instruction to begin was given, 48% of the subjects' fixations went to the irrelevant objects, but as soon as the task began, this fell to 16%. Clearly, the task constrains what the gaze system will target.
(2) The eyes deal with one object at a time. This corresponds roughly to the duration of the manipulation of that object and may involve a number of fixations on different parts of the object.

These findings emphasize the point that gaze and actions are temporarily bound to single objects, or sometimes object combinations such as jar and lid, and that the ''script'' of the overall action consists of a list of objects to be located and the acts to be performed on them. We have referred to these units as ''object-related actions'' (Land et al., 1999). In the terms of Fig. 1, each such unit represents the working out of a single schema.

The temporal and spatial relationships between gaze fixations and the actions they facilitate are of particular interest because they indicate the sequence in which the schema instructions are obeyed and the precision required for their execution. Does gaze always precede action? By what interval? Do we always look at the place where action takes place? If so, how close does foveal gaze need to be? We already have some answers to these questions from the pioneering studies of Dana Ballard and his colleagues in the early 1990s. They used a task in which a copy of a model made of colored blocks was assembled from a pool of dispersed

https://doi.org/10.1017/S0952523808080899 Published online by Cambridge University Press

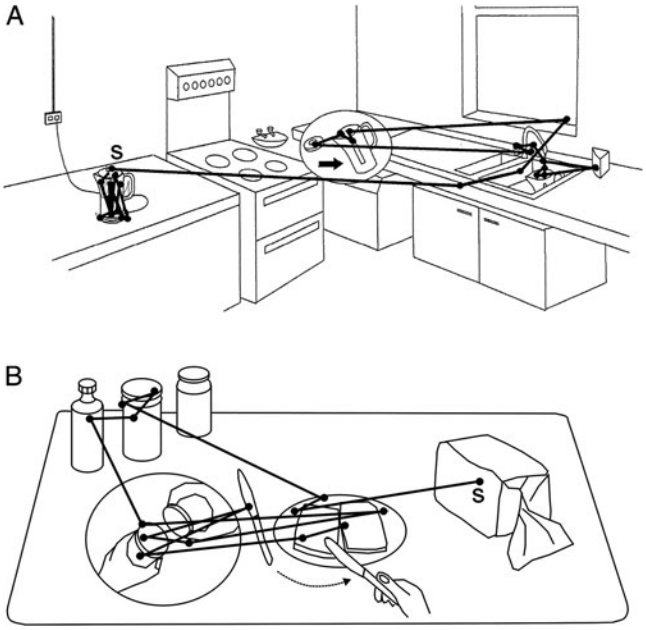


Fig. 2. (**A**) Scan path showing the first 26 fixations made by the author while filling a kettle prior to making a cup of tea. The record lasts about 13 s. The kettle is first inspected; three saccades to the sink before walking to it with the kettle, removing lid at the same time; selecting tap; one (irrelevant) fixation to sink tidy then watching water stream into kettle. From Land et al. (1999). (**B**) Scan path showing 15 fixations during part of a sandwich-making session. Bread viewed on plate; peanut butter jar selected (one irrelevant fixation on drink bottle), lifted and lid removed with one view back to plate; knife selected and used to scoop out butter, which is conveyed to bread and spread. From a video by Mary Hayhoe: excerpts in Hayhoe et al. (2003).

blocks, and they tracked the movements of eye and hand during the assembly process. They came up with two maxims: ''Do it where I'm looking'' (Ballard et al., 1992) and ''Look just in time'' (Ballard et al., 1995). In other words, fixation is close to the site of the action, and fixation precedes action by a short interval (usually less than a second). Fixation accuracy was hard to assess, because the objects involved were large—several degrees across in angular terms. In a more controlled study of object manipulation, Johansson et al. (2001) estimated the diameter of the ''effective fovea'' for fixations, based on the accuracy with which they were targeted, to be about 3 deg. In the tea and sandwich making studies, our aim was to see how the Ballard maxims applied in realistic unconstrained contexts.

In tea making, it was initially quite difficult to work out how the records of body movements, eye movements, and manipulations fitted together in time, and to begin with we had to resort to drawing out each 4-min sequence on a 40-ft roll of paper as a time line, recording the beginnings and ends of all acts and all the visual fixations that accompanied them. A tidied-up version showing the first 70 s of a tea-making episode is shown in Fig. 3A. What emerged from this exercise was that for most component acts, there was a definite temporal pattern. This involved first a movement of the whole body toward the next object in the overall sequence (this was not always required); second, a gaze movement to that object; and finally, the initiation of manipulation of the object by the hands. Good examples of this staggered progression can be seen in Fig. 3A at 5 and 35 s. We found that the best temporal marker to use to define the beginning of each act was the saccade that was made when the eyes left one object and moved on to the next, typically with a combined head and eye movement. These ''next object'' saccades could be used to ''chunk'' the task as a whole into separate object-related actions (which could consist of more than one sequential act on the same object). They could also act as time markers to relate the eye movements to movements of the body and manipulations by the hands. In this way, the different acts in the task could be pooled to get an idea of the sequence of events in a ''typical'' act. The results of this are shown in Fig. 3B. Not all acts could be included in these histograms because they involved waiting for something to happen—the kettle to fill, for example—which meant that the eyes tended to explore other parts of the room not relevant to the current action. On a few occasions, the two hands were engaged in different acts at the same time, and these too were excluded.

At first, it seemed surprising that it was the body as a whole that made the first movement in an object-related action, rather than the eyes. In fact, the reason was usually obvious: when the next object in the overall sequence was on a different work surface, a turn or a few steps were required before it could be reached with an eye movement. As Fig. 3B shows, the first saccade was made about 0.6 s after the start of the trunk movement. The first movements of the hands toward the object occurred about half a second after the first saccade. Thus, the eyes nearly always led the hands, presumably because manipulation required information from vision before it could begin. Interestingly, at the end of each object-related action, the eyes moved on to the next object about half a second before manipulation was complete. This means that the information that the eyes had supplied was still available in a store until the motor system required it.

Part of the reason why it had been so difficult to see the temporal pattern of the trunk, eye, and hand movements in the first place was that these movements were not simultaneous but staggered over a second or more. Equally important, however, was that these lags were quite variable. For the three subjects, the standard deviations of the lags were in the range 0.6–1.6 s, rather larger than the mean lag times, which only ranged between 0.4 and 0.7 s. Evidently, the brain is tolerant of asynchronies of up to 2 s or even 3 s in the timing of the components of an action. In sandwich making, the average lag times between eye and hand were much shorter, close to zero, possibly because the objects involved were closer together (Hayhoe et al., 2003). However, their standard deviations were just as large as those in tea making. From other studies, particularly of portrait drawing (see below), we observed similarly large variations in eye–hand lead times, so this apparent temporal sloppiness (better, flexibility) may be a feature of all self-paced activities. It probably has value in allowing the different component systems to complete their subtasks in an optimal way.

The sections that follow discuss the separate roles of the component systems of Fig. 1—the schema, gaze, visual, and motor systems—in the production of object-related actions. A summary diagram (Fig. 4) reflecting many of these comments is given at the end of this discussion.

Schema system

Our findings from the object-related action structure of tea and sandwich making confirm the reality of the schema as a basic unit of top-down control of action. The schema is the provider of the set of instructions for the performance of the next action in the task sequence, a set of instructions that determines where gaze

https://doi.org/10.1017/S0952523808080899 Published online by Cambridge University Press

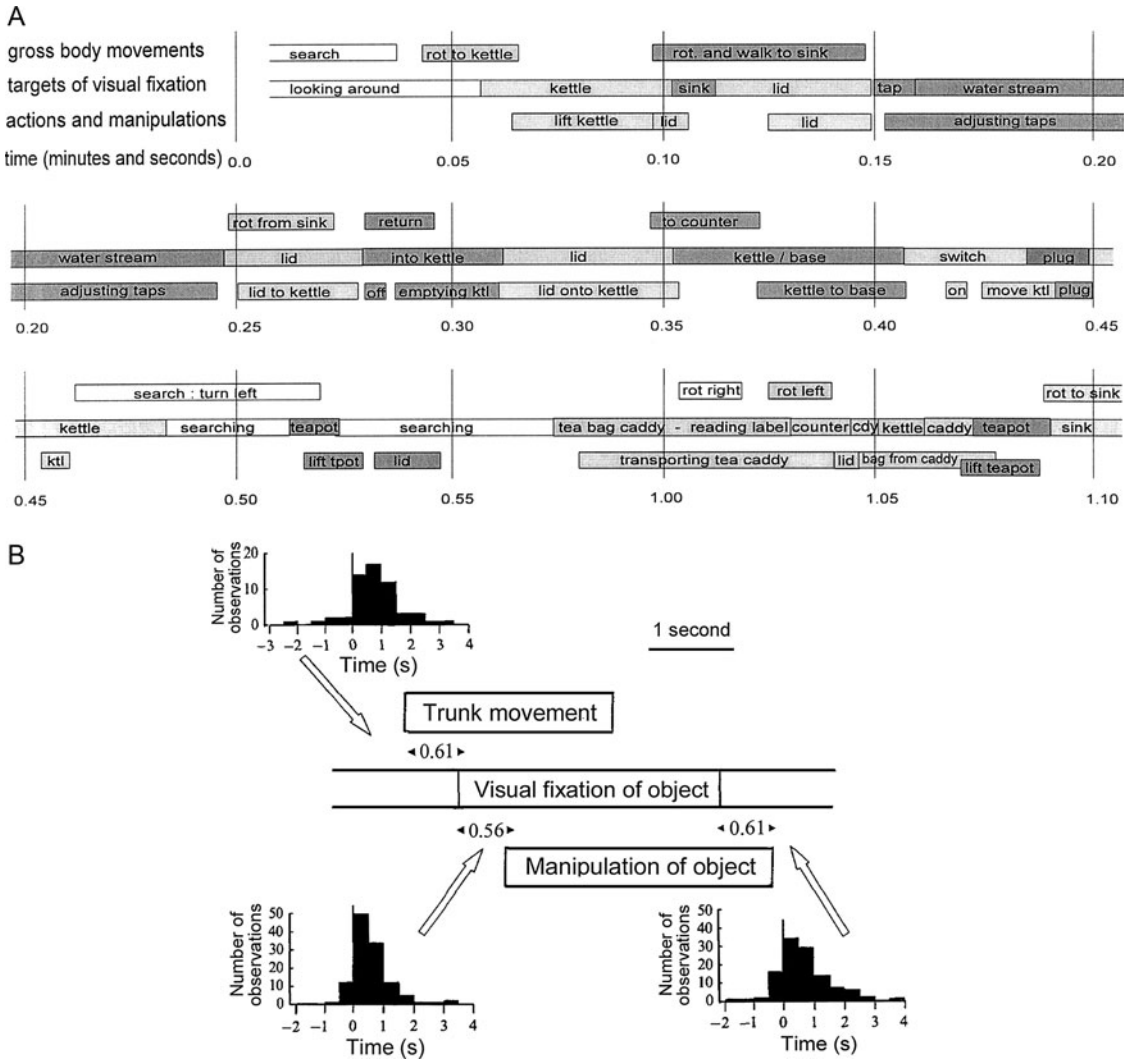


Fig. 3. (**A**) The relative timings of body movements, fixations on particular objects, and manipulation of those objects by the hands during the first 70 s of a tea-making session. There is a strong tendency for body movements (when made) to precede initial fixations, which precede manipulations. This is most clearly seen 0.05 and 0.35 s into the record. (**B**) Average relative timings of trunk movements, fixations on particular objects (which may involve several individual fixations), and manipulation of those objects. The data represent 94 object-related actions from three subjects. Typical actions last about 3 s (extended waiting periods, such as watching for the kettle to boil, have been excluded). Note that trunk movements (when present) precede fixations, which precede manipulation, by about 0.6 s in each case. Similarly, fixation moves on to the next object about 0.6 s before manipulation is complete. There are wide variations in timing (histograms), with manipulation sometimes preceding gaze fixation. From Land et al. (1999).

will be directed, what information the visual system will be called upon to provide, and what action will be taken. These instructions may be delivered simultaneously, but the start of their execution is spread—in both tea and sandwich making—over a period of about a second (Fig. 3B). It is generally accepted that a change of gaze requires the activation of covert and then overt attention (Wright & Ward, 2008), so it is presumably *via* attentional pathways that these instructions are delivered.

The role of the schema system as a whole is twofold. At the highest level, it must set the goal or goals of current behavior, and following from this, it must determine the sequence of actions that will achieve this—the script or schema list. Clearly, there is no one-to-one mapping of goal to schema sequence. For example, making tea, coffee, and soup share many of the same actions. The schema system must thus be able to draw on a pool of common behaviors, tailored to the circumstances of the current task. The ordering of the schemas also has some flexibility: it does not matter at what stage the mugs are brought to the counter, and indeed, this varied between participants. However, it does matter that the kettle has water in it before it is switched on. Some parts of the script are more negotiable than others.

One of the most influential models of this process of schema selection has been the supervisory attention system (SAS) of Norman and Shallice (1986). It distinguishes between routine operations that look after themselves and nonroutine operations that require supervisory intervention. Routine actions operate *via* an automatic system of schema selection they called ‘‘contention scheduling.’’ This is a process in which different schemas compete with each other to achieve an activation level that exceeds some selection threshold. A key feature of the contention scheduler is that it involves a ‘‘winner takes all’’ process, so that two incompatible actions are never performed at the same time. Such a mechanism

https://doi.org/10.1017/S0952523808080899 Published online by Cambridge University Press

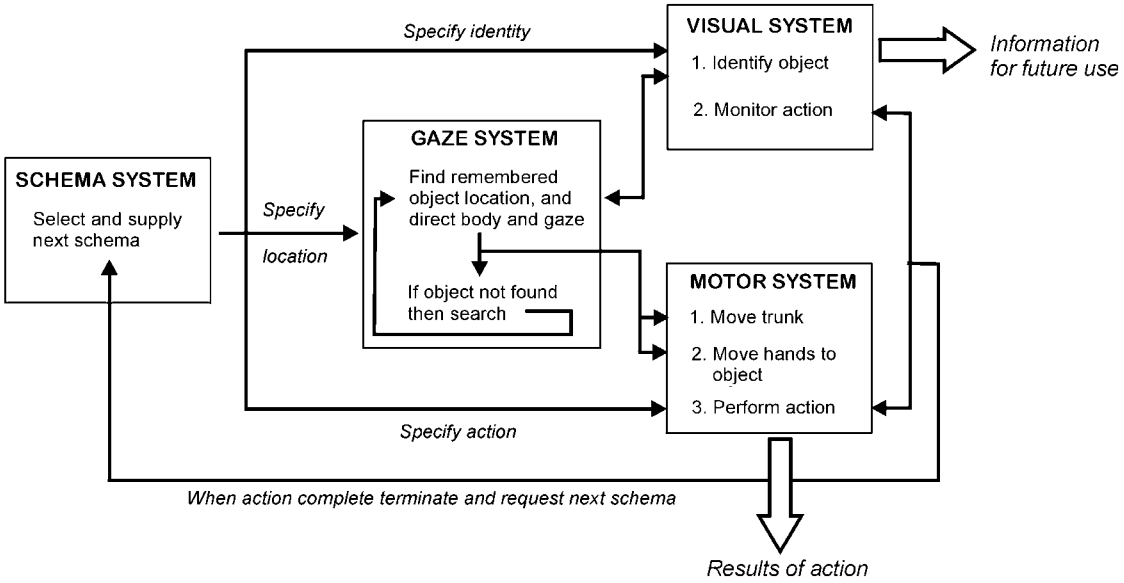


Fig. 4. Flow diagram, based on Fig. 1A, illustrating the interactions between the schema, gaze, visual, and motor systems during an object-related action. Note that in some cases, no action is involved, but information is stored for future use (as in a locating or look-ahead fixation). Details in the text.

for preventing motor conflict is clearly vital, and indeed, such conflicts are hardly ever seen. Even "multitasking" rarely means doing two things at exactly the same time. For nonroutine operations, Norman and Shallice introduced an additional SAS, more closely linked to conscious awareness, that could override the automatic system when that was likely to produce an undesirable result. It could also produce new schemas to deal with novel situations. Although we all recognize the sudden intrusion of consciousness into more routine action, as when a pedestrian steps into the road ahead of the car, it is not clear that this rather rigid distinction between automatic and attention-requiring actions is useful in more ordinary activities. For one thing, attention, in the sense that gaze direction is being under top-down control, seems to be present throughout action sequences. During food preparation, each new action requires attention to find the appropriate objects and also to monitor each action as it progresses. All this typically happens while one listening to the news on the radio at the same time, so although the attentional commitment to gaze control is certainly present, it is not at a high level. All that can really be said is that in some situations, attention is more closely coupled to conscious awareness than in others. Bargh and Ferguson (2000) take a similar view: that automatic and controlled processes are not distinct and also that even nonroutine processes have automatic control structures at a higher level.

In more recent models of supervisory attention (Shallice & Burgess, 1996; Shallice, 2002; Gilbert & Burgess, 2008), further "higher order" components were incorporated into the SAS. These included goal setting, problem solving, and strategy generation. The results of these operations are ultimately fed to the contention scheduler, which executes the schemas in conjunction with a special-purpose working memory process. A variety of evidence from patients with frontal lobe damage and from scanning studies indicates that the SAS itself operates from the dorsal prefrontal cortex, whereas at least some of the more routine contention scheduling may be based on selection processes operating in the premotor cortex and basal ganglia. Miller and Cohen (2001) take a similar view. The role of the prefrontal cortex is to maintain patterns of activity that represent goals and the means to achieve them, and that these provide "bias signals to other brain structures whose net effect is to guide the flow of activity along neural pathways that establish the proper mappings between inputs, internal states, and outputs needed to perform a given task." This formulation is quite consistent with the role of the schema system in Fig. 1.

Gaze system

Mechanisms of gaze change are sometimes thought of as being a part of the visual system and sometimes as components of the motor system. In this review, I have chosen to regard the gaze system as separate (although connected to both) because of its distinct role in locating and targeting the objects and places that are important for ongoing purposeful activity. Locating an object may only require it to be detected in peripheral vision and then fixated with the eyes and head. Often, however, the next object is initially out of sight, for example, on the work surface behind the subject, and its location has to be retrieved from memory. This then involves a trunk movement before the object can be brought into the visual field. As we have seen (Fig. 3B), such trunk movements typically precede the saccade that fixates the object by about half a second. Thus, gaze control is commonly a two-stage process, with the legs and trunk operating from memory, followed by eye and head movements working from a combination of memory and vision. The first (body movement) stage is likely to involve a translation from world-centered coordinates (knowing where the object is in the room) to trunk-centered coordinates (knowing what rotations and translations to make to reach it). It is difficult to determine how far in the periphery an object can be recognized and targeted by vision alone. All we know is that objects as far from the visual axis as 50 deg are often targeted with a single saccade, but beyond this there is always a secondary saccade of a few degrees before the final fixation. It seems likely that the second saccade is a visual correction of an initial memory-based estimate. The limits of visual detection are no doubt also influenced by object size. It is well known that the minimum recognizable

https://doi.org/10.1017/S0952523808080899 Published online by Cambridge University Press

letter size increases with increasing eccentricity (Anstis, 1974), and the same presumably also applies to domestic objects.

The eye movement component of the gaze system is based in the FEFs (see Fig. 1) which receive visual input and can initiate directed saccades. They are reciprocally connected to the parietal lobes (area lateral intraparietal cortex) and also to the prefrontal cortex. The head and trunk movements involved in gaze redirection presumably involve the premotor and motor cortices, but at present it is not clear how or where these different components of a gaze change are coordinated.

Motor system

The motor system of the hands and arms has a succession of different manipulations to perform, linked to the particular objects in use at the time. Schwartz et al. (1991) defined a minimal act as: ''the movement of an object from one place to another or a change in the state of an object,'' and that definition fits well with our object-related actions. Clearly, each movement of an object is different from any other, and although many of the ''changes of state'' have similar motor components—lifting a mug, jug, or teapot are similar actions—they differ in detail. The positions of the handles and the weights of the objects are different, for example. Other actions such as fitting a lid to a kettle or turning on a tap are complex activities that are quite different from a simple ''grasp and lift.'' The motor system must thus have a wide repertoire of learned actions at its disposal, and each of these actions is refined as it progresses by visual, proprioceptive, and haptic information.

In the scheme of Rizzolatti and Luppino (2001), the production of an action involves the cooperation of the premotor and parietal cortical regions, which are reciprocally connected. The premotor cortex (area F5) contains a variety of ''motor prototypes'' for reaching, grasping, and manipulative actions, which if activated can be converted to motor actions *via* the motor cortex. These prototypes are selected and shaped by the anterior intraparietal area (AIP), which has access to visual context information, including three-dimensional representations of the objects to be dealt with and information about their meaning *via* connexions from the temporal lobe. AIP also receives direct input from the DLPFC, and it is presumably *via* this route that the instructions from the schema system, about which action is to be performed, are conveyed.

Visual system

The visual system plays several roles in each action. First, the object being targeted by the gaze system must be identified. Either this occurs in the periphery and triggers the gaze change or, if the gaze change has been made from memory, when the object is close to the fovea. Sometimes an object is fixated, but no immediate action occurs. The location of the object may, however, be stored for future use. Land et al. (1999) referred to such fixations as *locating* fixations and Pelz and Canoza (2001) as *look ahead* fixations. Mennie et al. (2007) found that such fixations had a significant effect on subsequent eye–hand latency, showing that information was preserved over many intervening gaze changes. Other attributes of objects, such as their color, may also be acquired for future use. More commonly, fixations on objects precede manipulations that involve grasping and manipulation, and this involves passing visually acquired goal position information to the arm and hand control systems. This, of course, may be the same information that was involved in targeting gaze in the first place. Interestingly, such *directing* fixations are typically terminated before the hand has reached the target object, and so the act is completed without feedback—at least, not from the fovea. Fixations are also directed to places where an object is to be set down, again providing a spatial goal for the motor system to a location where, as yet, there is no target object.

Vision may be called upon to supervise an action more directly. *Guiding* fixations are concerned with manipulations involving more than one object, for example, a kettle and its lid. Here both objects have to be guided relative to each other so that they dock in an appropriate way. Most tool use is of this nature (e.g., spanner and nut, hammer and nail). It is more complicated than simple directing and usually involves a number of fixations that alternate between the two objects, and the action is normally completed under visual control. Some guided actions may involve more than two objects, for example, knife, bread, and butter. Finally, *checking* fixations determine when some condition is met, for example, the kettle is full, the water boiling, the top is off the bottle. These checking operations may require the eye to dwell on some appropriate region of an object, either in one long fixation or in a series of repeated fixations. When the specified condition is met, the present action is terminated and a new action is initiated; for example, when the kettle is full, the tap is turned off. Presumably, the conditions for terminating an action are part of the instructions supplied by the schema system at the beginning of the action, and when these are fulfilled, the schema system is informed and new instructions are supplied.

This functional classification of fixations, although appropriate here, does not fit all types of information gathering for which the eyes are employed. In particular, a fixation may simply establish a particular attribute of an object. Thus, in the block-copying study of Ballard et al. (1992), discussed later, the logic of the task dictates that the first fixation on the model establishes the color of a block, which is then used to locate a similar block for use in the copy. The next fixation on the model is used for another purpose, to check the block's location. In sketching and other copying tasks, one or more fixations acquire the form of a feature so that it can be reproduced (Fig. 8). Continuous tasks such as reading text or music are also rather different, as each fixation acquires a small part of an information stream which is then compounded internally with those preceding and succeeding it to produce a sequence that has meaning.

Interestingly, there are some things that are rarely, if ever, fixated during sequences of this kind. The hands themselves are never fixated, and once objects have been acquired by the hands, they are not fixated either. The implication of this somewhat surprising finding is that vision is a scarce and valuable resource and is disengaged from action as soon as other senses—touch and proprioception—can take over from it.

An important question concerns the restriction of the information provided by vision during the execution of an action. Is it confined to the immediate demands of the task, or are other more general features of the local environment picked up at the same time? Put another way, do the schema instructions constrain what can be seen? Mary Hayhoe and her colleagues have studied this question using a virtual reality setup in which subjects perform realistic tasks but in which changes can be made to the objects in those tasks either during saccades, when the subject is temporary blind, or when the object is in plain view. In one study (Triesch et al., 2003; Droll et al., 2005), the task was to sort bricks of two different sizes onto virtual conveyor belts. In 10% of the moves, the size of the brick was changed between pickup and put-down. These moves were invariably accompanied by a saccade, and eye movements were monitored to determine whether or not the

https://doi.org/10.1017/S0952523808080899 Published online by Cambridge University Press

change coincided with a saccade. The question was ''Under what circumstances would the change be noticed?'' The subjects were not told specifically to look for size changes but were instructed to report if anything odd happened. There were three conditions: (1) ''Pick up the bricks in front to back order and place them on the closer conveyor belt.'' Brick size is irrelevant here for both pickup and put-down decisions. (2) ''Pick up the tall bricks first and put them on the closer conveyor belt. Then, pick up the smaller bricks and also put them onto the closer conveyer belt.'' For this condition, size is important during the pickup phase but not the put-down. (3) ''Pick up the large bricks first and put them on the closer conveyer belt. Then pick up the small bricks and put them on the distant conveyer belt.'' Here brick size is relevant for both pickup and put-down decisions. The first conclusion was that the number of times a size change was reported at the time increased from about 2% in condition 1 to 20% in condition 2 and to 45% in condition 3. This suggests that changes are noticed in proportion to their importance in the performance of the task. When they are not relevant, the changes are simply not noticed. Remarkably, there were even occasions when the size change was made while the subjects were looking directly at the brick, but nevertheless they failed to notice or report the change. The second conclusion was that when a change was noticed, fixation times on the brick increased by about 300 ms, irrespective of the condition being tested. Unnoticed changes produced no increase in fixation times. These extended fixations must have been the result of the engagement of attention, not the changes in the retinal image. The overall conclusion of the Triesch et al. (2003) article is summed up by its title: ''What you see is what you need.'' We can probably infer from this that during the performance of a task, very little visual information is taken in that it is not of immediate use and that the schema system flags up in advance what the visual system has to look out for.

A summary of the events in an object-related action

Fig. 4 is an attempt to summarize the processes that occur in an object-related action, in the form of a flow diagram. The diagram is an expansion of Fig. 1A, incorporating many of the results of the discussion in the preceding sections; only a few extra comments will be added here. The schema system selects a single action to be performed and activates all three subsystems (gaze, visual, and motor) more or less simultaneously. Gaze control operates on the basis of position memory and/or on the visual identification of the specified object. It may require action by the motor system of the legs and trunk before the head and eyes can fixate the object. Either the gaze system or the visual system or both then provide the motor system of the hands with the location of the object. The visual system then either may guide the hands in the execution of the required action (and the trunk if this requires movement to a new place) or may simply store information about the object for future use (arrow top left), with no further immediate action taking place. This happens in *locating* fixations and when it is necessary to find out some attribute of an object, such as its color. If an action does take place, this is supervised by the visual system, often with minor readjustments of gaze. When the action is complete, the schema system is informed, the present instructions are canceled, and the next schema is prepared.

Variations

The scheme in Fig. 4 was intended to describe the control of a single action during an extended task made up of a sequence of different actions. This does not by any means cover all the types of action that involve eye–effector cooperation, and in this section, I briefly outline three of the more important variants. Fig. 5 illustrates these examples, in addition to the multiple-action sequence already discussed.

Single-action events (Fig. 5A)

In these cases, only one action occurs, and in the most pared-down instances, only a single eye movement may be involved. Ball sports provide many such examples in which the task is to intercept a moving ball and catch it or hit it with a controlled stroke. The key feature of these tasks is that they all require the observer to have an internal dynamic model of the ball's physical behavior, which makes it possible to anticipate its future trajectory (Hayhoe et al., 2005).

In cricket (or baseball), the time from delivery of a fast ball to the time it is struck can be as short as 400 ms, which means that only a single saccade is possible. In cricket, the ball usually bounces, and it is very important for the batter to see the bounce point clearly, partly because this provides timing cues for formulating the strike and partly because the ball may change trajectory after the bounce (Fig. 6A). The bowler will vary the bounce point, so the batsman must track the ball from the bowler's hand onto the pitch. Peter McLeod and I studied the eye and head movements of three batsmen striking balls from a bowling machine that delivered ''medium-pace'' balls at 25–30 m/s (Land & McLeod, 2000). The principal findings were (i) that the batsmen always watched the ball as it emerged from the machine; (ii) after a latency as short as 130 ms, they made a single saccade down to the estimated bounce point, where the eyes waited a few tens of milliseconds for the bounce itself; and (iii) after the bounce, the ball was tracked for 100–200 ms and then visual contact with the ball was usually lost just before the strike (Fig. 6B). Our heretical finding was that the batters did not ''keep their eye on the ball'' as they are usually exhorted to do. Instead, they made an estimate of the bounce point, based on the initial rate of decent of the ball from the machine, made a saccade that brought gaze to 5 deg or more below the ball, and lay in wait for it to bounce. They did this rather than track the ball (although they did not know this) because it takes about 100 ms or more for smooth tracking to be initiated, by which means that they would have missed many of the bounces. A saccade is much more efficient but is made without feedback and so requires a prior calculation (or in practice a reference to internal lookup tables acquired through years of practice).

The batsmen learn two things from this procedure. First, the time from the delivery to the bounce, and second the location of the bounce, measured (by them) as the angle of declination from the batsman's eye to the bounce point—the larger this angle is, the shorter the distance of the bounce to the bat. It turns out that these two variables are sufficient (apart from constants such as the length of the pitch) to determine when and where the ball will reach the bat and thus make it possible to formulate an attacking stroke (Land & McLeod, 2000). The precision required is impressive. Regan (1992) calculated that to make an effective stroke with the ''sweet spot'' of the bat requires timing to better than 5 ms and spatial accuracy of 5 cm.

Much the same anticipation of the bounce is seen when catching a bouncing ball (Hayhoe et al., 2005). Subjects initially fixate the hands of the thrower and then make a saccade to a point a few degrees above the future bounce point, arriving an average of 53 ms before the bounce. After the bounce, the ball is tracked

https://doi.org/10.1017/S0952523808080899 Published online by Cambridge University Press

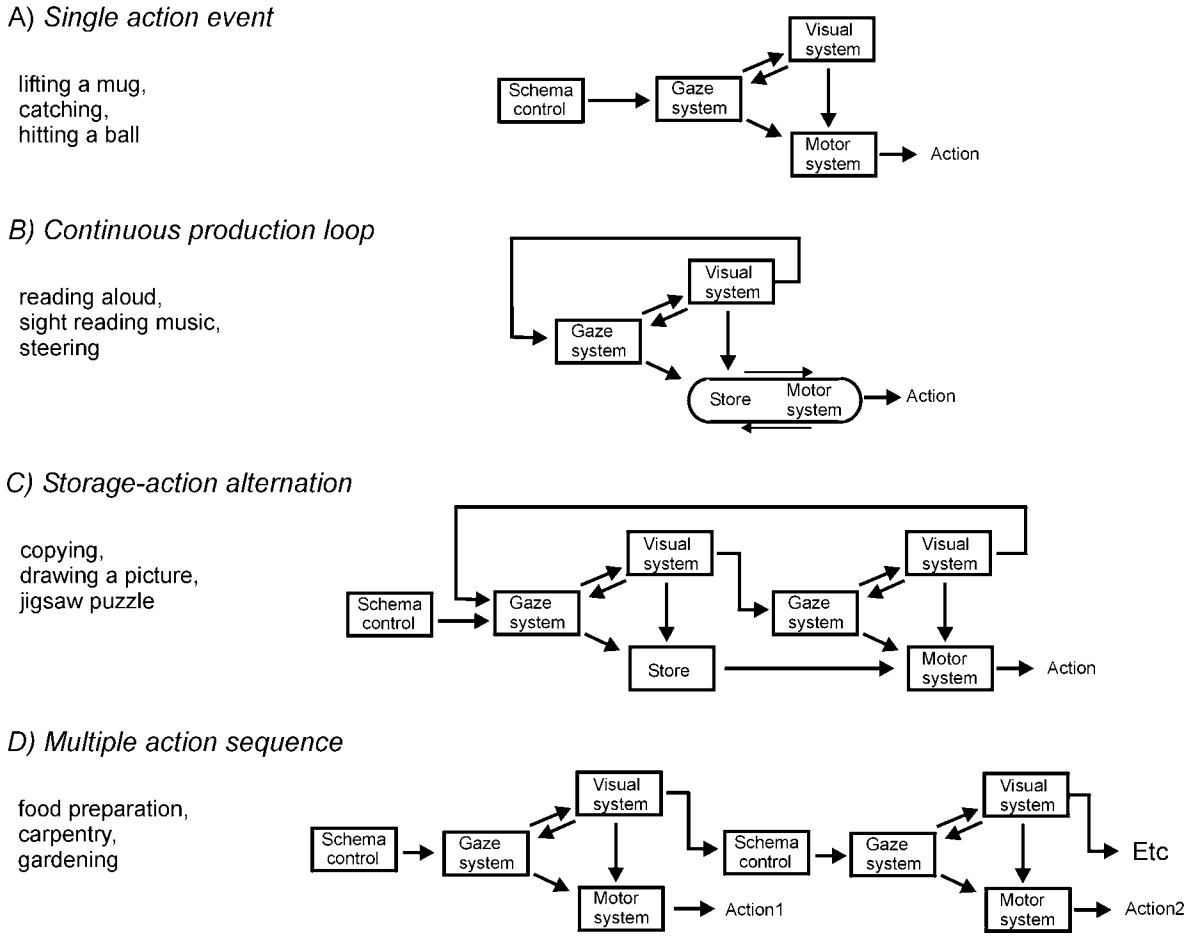


Fig. 5. Four variants of the basic control pattern of Fig. 1A, represented here as a *single-action event*. In a *continuous production loop*, information is fed semicontinuously into a loop that produces a continuous output. *Storage–action alternation* is a two-stage process where acquisition of information and its use are separated in time, and this is typical of copying tasks. *Multiple-action sequences* are found in extended tasks such as making tea or sandwiches where many different actions follow each other.

smoothly into the hands. Table tennis is similar (Ripoll et al., 1987; Land & Furneaux, 1997). The receiver watches the ball leave the opponent's bat and makes a saccade to a point about 5 deg above the anticipated bounce point, arriving 100–200 ms before the bounce itself. The timing and magnitude of the saccade are based on trajectory of the ball in the brief period after the stroke. Squash is even more impressive. McKinney et al. (2008) studied the eye movements of squash players receiving balls that had bounced off two walls of the court. They anticipated the first bounce off the service wall by about 150 ms, tracked it smoothly for about half the distance to the second wall, and then made a second anticipatory saccade, *not* to the second bounce point but to a point on the path of the ball *after* it had bounced, anticipating its position by an average of 186 ms. Thereafter, the ball was tracked until just before contact with the racket. Clearly, this second anticipatory saccade requires a sophisticated and flexible model of the behavior of bouncing balls. In their catching study, Hayhoe et al. (2005) demonstrated the updating of such models, by covertly replacing the normal ball with a bouncier one. Initially, catchers were unable to track the ball smoothly after it bounced and resorted to saccadic tracking, but after three repeats, their performance was back to normal. Subjects also adjusted the timing of the anticipatory saccade to the new ball dynamics.

These examples make the point that saccades can be made to locations that do not yet contain a visible "stimulus" (and so can hardly be "salient"), and that very accurate and specific information can be obtained from a single fixation, provided the player is adequately trained.

Continuous production loop (Fig. 5B)

In activities such as reading aloud, sight-reading music or touch-typing the sensory-motor pattern is not one of discrete acts (as in tea or sandwich making) but is a more or less continuous process. It is true that, as in other activities, the eye proceeds across the page in a series of saccades and fixations, but that is presumably to avoid the blur that continuous motion would cause, rather than there being any other procedural need to chunk the acquisition process; in speech processing, there are no equivalent discontinuities. Thus, in these activities, and some others such as walking on rough terrain or steering on a winding road, visual information is acquired as a semicontinuous stream and deposited onto the equivalent of a conveyor belt where it is transformed and from which it emerges after a processing delay as a motor output.

In reading English, about seven characters are taken in at each fixation (for a review, see Rayner, 1998). In reading aloud, the

https://doi.org/10.1017/S0952523808080899 Published online by Cambridge University Press

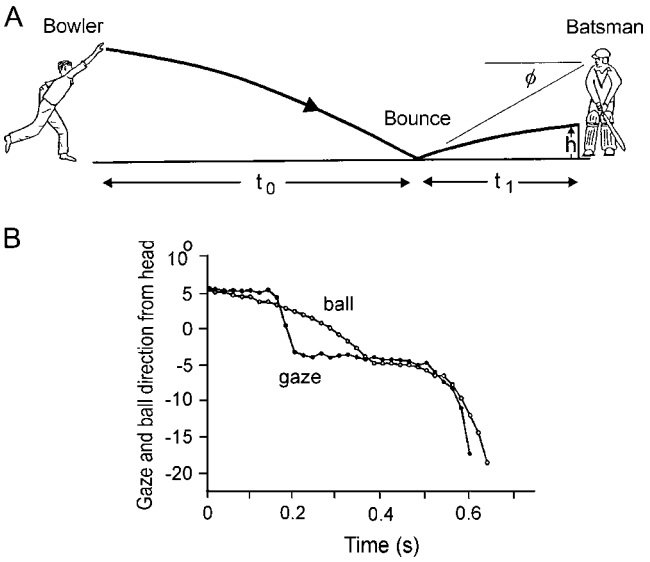


Fig. 6. (**A**) In cricket, the ball bounces once, and information from the bounce point is crucial if the batsman is to make a scoring stroke. From the time from delivery to bounce (t_0) and the location of the bounce (measured from the declination angle φ as seen by him), he can determine the time from bounce to bat (t_1) and the height at which it will reach the bat (h). These two variables (t_1 and h) are sufficient for the production of a well-timed shot. (**B**) Location of the vertical direction of gaze and of ball position relative to the batsman's head during the progress of the ball down the pitch. About 150 ms after the delivery (from a bowling machine in this case), the batsman takes his eyes off the ball and makes a downward saccade to the estimated bounce point and waits for the bounce (in this case for about 180 ms, but for a fast ball that bounces early, this delay will be much shorter). A good batsman will estimate the bounce point from the vertical decent speed of the ball before he makes the saccade. Note that by observing the delivery and the bounce, the batsman can obtain t_0 and φ. Modified from Land and McLeod (2000).

conversion of seen letters into words and then into speech takes a little less than a second. In a classic paper, Buswell (1920) showed that for high school readers, the eyes stay about 13 letters, or 0.79 s, ahead of the sound being uttered. This eye–voice span varies slightly with reading ability but is rarely longer than a second. Thus, visual information is taken in, processed, emitted, and discarded in a continuous stream. A trace may remain in the form of meaning, but the visual image of the letters and words is not retained. Continuous touch-typing involves a similar 1-s eye–hand delay (Butsch, 1932), although typists may also take in longer "chunks" of text and type them from memory (Fleischer, 1986).

Sight-reading a double stave of music is more complicated (Fig. 7). As Weaver (1943), the first person to study eye movements in music reading, remarked: "Notes on the treble and bass parts of the great staff are usually so far apart that both vertical and horizontal movements of the eyes must be used in preparing two parallel lines of material for a unified performance." Furneaux and Land (1999) confirmed Weaver's finding that the eyes do indeed alternate between the upper and lower staves, taking in about three note-heads per fixation. They also measured the time between looking at a note and playing it—the eye–hand span. As in reading aloud, this was close to 1 s, and interestingly, this time did not vary between players of different abilities. What did vary was the throughput—the number of notes that could be processed during that second. For beginners, this was about two; for professional pianists, it could be six or more. The other difference was in the variation in the eye–hand span. For a professional accompanist, playing a fast piece, the range was about 0.4 s; for a novice playing a simple piece, it varied chaotically over more than 2 s. The performance of the processor in music reading is even more remarkable than in reading, since the notes from the upper and lower staves are sampled serially, and yet, after decoding and conversion to finger movements, they are emitted simultaneously. On top of all this, a good pianist will also be able to inject style and emotion into the playing.

Walking over difficult or unpredictable terrain also requires vision to provide continuous information about the state of the ground ahead. Patla and Vickers (2003) got their subjects to step on a series of irregularly spaced "footprints" over a 10-m walkway. They found that subjects fixated the footprints on average two steps ahead, or between 0.8 and 1 s in time, an observation subsequently confirmed by Land (2006) for walkers avoiding cracks in pavement. Like walking, driving on a winding road requires a continuous translation of visual information about road curvature into steering wheel rotation. There is general agreement that to stay in lane and anticipate future bends, drivers require both feedback information from the lane markers a few meters ahead of the car (picked up in peripheral vision) and also feed-forward information from more distant road regions. Land and Lee (1994) found that on winding roads, drivers spend much of the time fixating the tangent point of the upcoming bend. This is the point on the inside of the bend that projects furthest into the roadway. On a bend of constant curvature, this point is stationary in the field of view, even though its physical position moves round the bend with the car. However, when road curvature changes, it moves further from or nearer to the car's heading direction, so providing the driver with a potential control signal to steer by. Land and Lee found that the delay between movements of the tangent point in the driver's field of view and corresponding changes in steering wheel angle was about 0.8 s, similar to the "processing time" in the systems we have already encountered. The tangent point is, however, not the only distant road region that can supply steering information. The flow field around points on the car's intended future path can also be used to guide steering (Wann & Wilkie, 2004). Salvucci and Gray (2004) have proposed a scheme in which the angular position of distant road features is not important, only their angular velocity, which means that any distant point can be used, including a vehicle in front. It is quite likely that drivers use different strategies depending on road conditions, but it will not be easy to settle this question.

The activities mentioned above all require additions to the basic arrangement of Fig. 1. The gaze system must be capable of advancing itself automatically, without referring back to the schema system for explicit instructions. The other requirement is a "conveyor belt module" able to receive one type of visual information semicontinuously and process it into a semicontinuous motor output. Remarkably, the processing time for all these varied transformations seems to be similar (~1 s). One would think that "working memory" would describe this process perfectly, but working memory as currently understood deals principally with thought and comprehension, rather than gaze control and action (Baddeley, 2007).

Storage–action alternation (Fig. 5C)

In some tasks, the obtaining of information and the control of action are disconnected in time, rather than being linked in the same vision–action cycle, as was the case during tea and sandwich making. This is particularly true in copying tasks such as drawing

https://doi.org/10.1017/S0952523808080899 Published online by Cambridge University Press

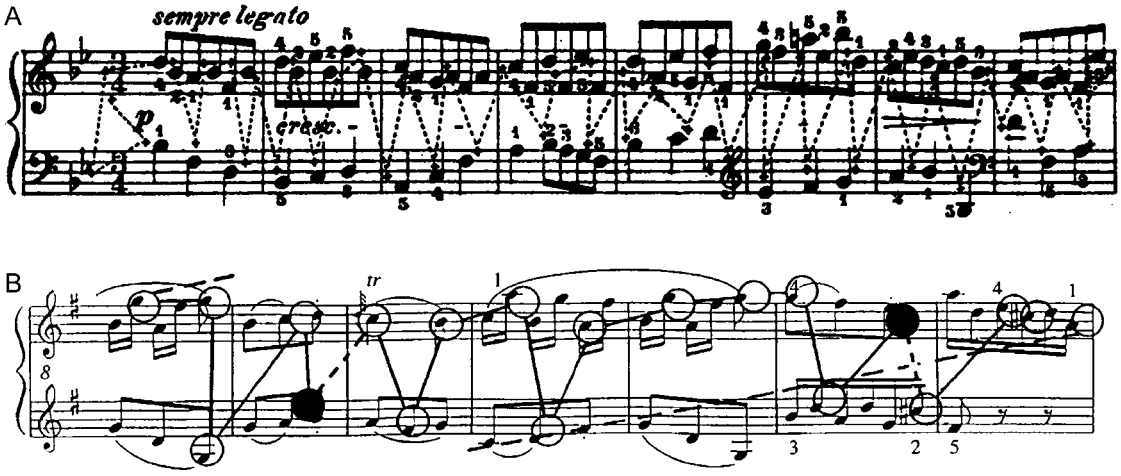


Fig. 7. (**A**) A pianist playing a Bach minuet, showing the alternation of gaze between the upper and lower halves of the stave. From Weaver (1943). (**B**) A professional pianist playing a Scarlatti sonata, showing a similar alternation. In this case, the time from observing a note to playing it was about 0.8 s. The two filled circles are fixations after which the pianist looked briefly down to the keys (about 30 deg below) and then back to the score with no disruption of the alternation pattern. From Furneaux and Land (1999).

from life or completing a jigsaw puzzle. One of the first studies of eye–action coordination was the block-copying task of Ballard et al. (1992, 1995), mentioned earlier. Here a copy of a model structure, consisting of a pattern of colored blocks, was made by selecting blocks from a random pool, transporting them to the copy site, and setting them down in the appropriate positions. In a typical cycle, gaze would go first to the model and choose a color, then move to the pool and pick up a block of that color. Next, gaze would return to the model to determine the location of the block and finally return to the selected block to supervise its move and drop into the correct position. Thus, the task of the gaze system alternated between obtaining information (color and position) and supervising action (lift selected block, and move and drop into position). Information has to be retained across fixations, but the use of memory is very parsimonious: the system seems to prefer to make an extra fixation on the model rather than remember both color and fixation at the same time, although this can be done if required.

Drawing a portrait also involves obtaining information from a sitter and then using this to shape the subsequent pen stroke. To find out what an artist learns while looking at the sitter, we studied the gaze movements of a painter friend (Nick Bodimeade) while he made quick portrait sketches (Land, 2006). In a sketch, as opposed to a detailed drawing in which there is much redrafting, each line is a single unrepeated action, so there is usually a clear relationship between the artist's intention, his eye movements, and the line he draws. A typical sketching cycle consists of a saccade from sketch to sitter, one to four fixations on the sitter, then a saccade back to the sketch to a location corresponding closely to the last fixation on the sitter (Fig. 8A). There were then several fixations on the sketch loosely linked to pen position, but interestingly none of them predicted where the next fixation on the sitter would occur. Each new line was, it seems, unrelated to the last. Each whole cycle took, on average, 1.7 s, divided between 0.8 s on the sitter and 0.9 s on the sketch (Fig. 8B). The drawing of each line corresponded roughly with the time that gaze was on the sketch, but often drawing continued after gaze had returned to the sitter. As in tea and sandwich making, the lengths of these half-cycles and relative timings of the beginning and end of each drawn line had standard deviations of half a second or more. The most intriguing aspect of each cycle was that the location of gaze, when it returned from sitter to sketch, did not mark the *beginning* of the new line but its *end point* (Fig. 8A). It seems that the artist, having arbitrarily decided where to begin his next line, used his time on the sitter to determine the line's trajectory and in particular where it would terminate. As in the block-copying task, only a very small

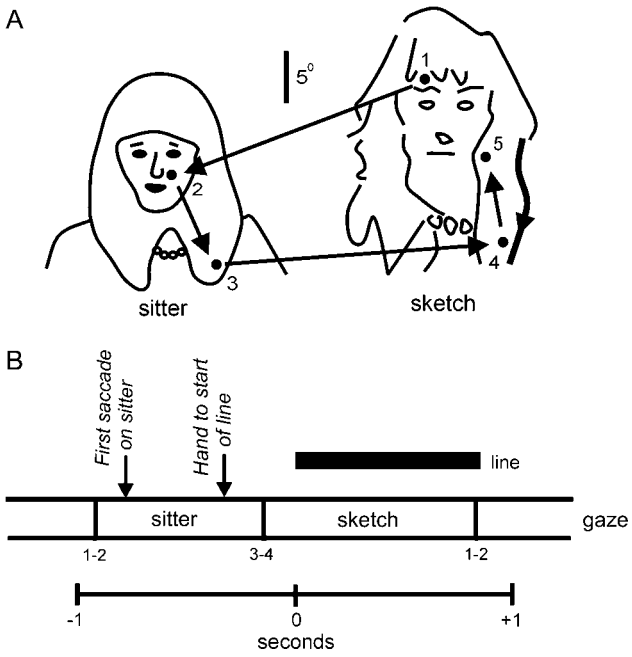


Fig. 8. (**A**) The cycle of looks from model to drawing when making a portrait sketch. The artist makes a saccade from sketch to model (1–2), makes one to four fixations on the model, finishing at (3), before transferring the final gaze point to the corresponding point on the sketch (4). This point usually corresponds to the end point of the next line he draws. Thereafter, gaze moves around the sketch (5). It appears that nothing predicts the start point of the line, which must therefore come from the artist's internal plan. (**B**) Timing of events in the cycle, derived by averaging all 25 cycles. The next line is typically started about 0.1 s after gaze has moved back from the sitter, but there is much variation in the timing. Modified from Land (2006).

https://doi.org/10.1017/S0952523808080899 Published online by Cambridge University Press

amount of information is held in memory during the move from observation to action: just enough to draw the one line.

When on the sketch, the eye tended to follow the pen in an irregular saccadic course, but when the eyes moved from sitter to sketch, gaze anticipated the future position of the pen. In a study of simple point-to-point line drawing, Tchalenko (2007) distinguished between two of these modes of eye–hand behavior which he referred to as Close Pursuit and Target Locking. In Close Pursuit, gaze followed the pen saccadically as the line was drawn, but in Target Locking, gaze went directly to the end point of the line, and the line was drawn to this new gaze location. A similar distinction was made by Sailer et al. (2005). They studied the acquisition of a skill in which a novel mouse-like device was used to move a cursor from one target to another on a screen. In the early stages of learning how to use this rather user-unfriendly device, gaze followed the cursor, providing feedback about the accuracy of the hand movements. Later, however, when the motor skill had become proficient, gaze went straight to the next target, with the cursor following behind. Pursuit and anticipation are also features of the ball sports discussed earlier, and it seems that these two forms of visuomotor control are likely to be common features of many skilled eye–hand tasks.

Conclusions

I hope to have demonstrated in this chapter that, in real-life tasks, the roles of vision are rich and varied and intimately bound up with the control of purposeful action. The visual, gaze, and motor systems have separate but interlocking roles in most activities, and if one is trying to understand how behavior is organized, it is helpful to try to study these contributions together, rather than separately as is usual in most laboratory situations. In general, eye movements to informative sources of visual information precede motor actions, which are then guided, monitored, and ultimately terminated under vision supervision.

It is also clear that these three systems (visual, gaze, and motor) are subject to joint programming by a supervisory system (referred to here as the schema system) that initiates their respective roles in the execution of particular tasks. It is assumed that the information for carrying out a variety of tasks already resides separately in the three systems and that the role of the schema system is to specify, *via* attentional pathways, which task is to be performed at any one time. A number of variations of the basic system outlined in Fig. 1 were discussed in the last section and illustrated in Fig. 5. Whether or not this list is complete is not clear at this stage, but those outlined, or combinations of them, seem to cover most types of visually controlled action.

In purposeful action, the role of bottom-up salience, in the sense of gaze being attracted to objects by their conspicuity, is of very limited relevance. Indeed, in a task that requires fixation on a sequence of specific objects, the capture of gaze by irrelevant salient objects would, in general, be an obtrusive nuisance. There now seems to be general agreement that top-down instructions dominate gaze behavior both during passive visual tasks involving visual search (e.g., Yarbus, 1967; Einhäuser et al., 2008) and in the performance of visually guided actions (Hayhoe, 2000; Land, 2003; Rothkopf et al., 2007). The challenge is now to find out how descending attentional mechanisms control gaze during purposeful action.

Acknowledgment

I am grateful to the Gatsby Foundation for a grant toward some of the work described in this article.

References

Anstis, S.M. (1974). A chart demonstrating variations in acuity with retinal position. *Vision Research* **14**, 589–592.

Baddeley, A. (2007). *Working Memory, Thought and Action*. Oxford: Oxford University Press.

Ballard, D.H., Hayhoe, M.M., Li, F. & Whitehead, S.D. (1992). Hand-eye coordination during sequential tasks. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **337**, 331–339.

Ballard, D.H., Hayhoe, M.M. & Pelz, J.B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience* **7**, 66–80.

Bargh, J.A. & Ferguson, M.J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin* **126**, 925–945.

Buswell, G.T. (1920). An experimental study of the eye-voice span in reading. *Supplementary Educational Monographs* **17**. University of Chicago Press.

Butsch, R.L.C. (1932). Eye movements and the eye-hand span in typewriting. *Journal of Educational Psychology* **23**, 104–121.

Droll, J.A., Hayhoe, M.M., Triesch, J. & Sullivan, B.T. (2005). Task demands control acquisition and storage of visual information. *Journal of Experimental Psychology. Human Perception and Performance* **31**, 1416–1438.

Einhäuser, W., Rutishauser, U. & Koch, C. (2008). Task-demands can immediately reverse the effects of sensory-driven saliency in complex visual stimuli. *Journal of Vision* **8**(2): article 2, 1–19.

Fleischer, A.G. (1986). Control of eye movements by working memory load. *Biological Cybernetics* **55**, 227–238.

Furneaux, S. & Land, M.F. (1999). The effects of skill on the eye-hand span during musical sight-reading. *Proceedings of the Royal Society of London. Series B, Biological Sciences* **266**, 2453–2440.

Gilbert, S.J. & Burgess, P.W. (2008). Executive function. *Current Biology* **18**, R110–R114.

Hayhoe, M. (2000). Vision using routines: A functional account of vision. *Visual Cognition* **7**, 43–64.

Hayhoe, M., Mennie, N., Sullivan, B. & Gorgos, K. (2005). The role of internal models and prediction in catching balls. *Proceedings of the American Association Artificial Intelligence* (Fall).

Hayhoe, M.M., Shrivastava, A., Mruczek, R. & Pelz, J.B. (2003). Visual memory and motor planning in a natural task. *Journal of Vision* **3**, 49–63.

Itti, L. & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research* **40**, 1489–1506.

Jeannerod, M. (1997). *The Cognitive Neuroscience of Action*. Oxford: Blackwell.

Johansson, R.S., Westling, G., Bäckström, A. & Flanagan, J.R. (2001). Eye-hand coordination in object manipulation. *Journal of Neuroscience* **21**, 6917–6932.

Land, M.F. (2003). Eye movements in daily life. In *The Visual Neurosciences*, ed. Chalupa, L.M. & Werner, J.S., pp. 1357–1368. Cambridge, MA: MIT Press.

Land, M.F. (2006). Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research* **25**, 296–324.

Land, M.F. & Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **352**, 1231–1239.

Land, M.F. & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities. *Vision Research* **41**, 3559–3565.

Land, M.F. & Lee, D.N. (1994). Where we look when we steer. *Nature* **369**, 742–744.

Land, M.F. & McLeod, P. (2000). From eye movements to actions: How batsmen hit the ball. *Nature Neuroscience* **3**, 1340–1345.

Land, M.F., Mennie, N. & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception* **28**, 1311–1328.

McKinney, T., Chajka, K. & Hayhoe, M. (2008). Pro-active gaze control in squash. *Journal of Vision* **8**(6): article 111, 111a.

Mennie, N., Hayhoe, M. & Sullivan, B. (2007). Look-ahead fixations: Anticipatory eye movements in natural tasks. *Experimental Brain Research* **179**, 427–442.

Miller, E.K. & Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience* **24**, 167–202.

Norman, D.A. & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In *Consciousness and Self Regulation.*

https://doi.org/10.1017/S0952523808080899 Published online by Cambridge University Press

Advances in Research and Theory, **Vol. 4**, ed. Davidson, R.J., Schwarts, G.E. & Shapiro, D., pp. 1–18. New York, NY: Plenum.

Patla, A.E. & Vickers, J.N. (2003). OHow far ahead do we look when required to step on specific locations in the travel path during locomotion. *Exp. Brain Res.* **148**, 133–138.

Pelz, J.B. & Canoza, R. (2001). Oculomotor behavior and perceptual categories in complex tasks. *Vision Research* **41**, 3587–3596.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* **124**, 372–422.

Regan, D. (1992). Visual judgments and misjudgments in cricket, and the art of flight. *Perception* **21**, 91–115.

Ripoll, H., Fleurance, P. & Caseneuve, D. (1987). Analysis of visual patterns of table tennis players. In *Eye Movements: From Physiology to Cognition*, ed. O'Regan, J.K. & Levy-Schoen, A., pp. 616–617. Amsterdam, The Netherlands: North-Holland.

Rizzolatti, G. & Luppino, G. (2001). The cortical motor system. *Neuron* **31**, 889–901.

Rothkopf, C., Ballard, D. & Hayhoe, M. (2007). Task and scene context determine where you look. *Journal of Vision* **7**(14): article 16, 1–20.

Sailer, U., Flanagan, J.R. & Johansson, R.S. (2005). Eye-hand coordination during learning of a novel visuomotor task. *Journal of Neuroscience* **25**, 8833–8842.

Salvucci, D.D. & Gray, R. (2004). A two-point control model of steering. *Perception* **33**, 1233–1248.

Schwartz, M.F., Reed, E.S., Montgomery, M.W., Palmer, C. & Mayer, N.H. (1991). The quantitative description of action disorganisation after brain damage: A case study. *Cognitive Neuropsychology* **8**, 381–414.

Shallice, T. & Burgess, P.W. (1996). Domains of supervisory control and the temporal organisation of behaviour. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **351**, 1405–1412.

Tchalenko, J.S. (2007). Eye movements in drawing simple lines. *Perception* **36**, 1152–1167.

Triesch, J., Ballard, D.H., Hayhoe, M.M. & Sullivan, B.T. (2003). What you see is what you need. *Journal of Vision* **3**, 86–94.

Wann, J.P. & Wilkie, R.M. (2004). How do we control high speed steering. In *Optic Flow and Beyond*, ed. Vaina, L.M., Beardsley, S.A. & Rushton, S.K., pp. 371–389. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Weaver, H.E. (1943). A study of visual processes in reading differently constructed musical selections. *Psychological Monographs* **55**, 1–30.

Wright, R.D. & Ward, L.M. (2008). *Orienting of Attention*. Oxford: Oxford University Press.

Yarbus, A. (1967). *Eye Movements and Vision*. New York: Plenum Press.

https://doi.org/10.1017/S0952523808080899 Published online by Cambridge University Press