CHAPTER 30

Eye movements both reveal and influence problem solving

Michael J. Spivey and Rick Dale

Abstract

Within the context of the theory of embodied cognition, our most frequent motor movements—eye movements—are sure to play an important role in our cognitive processes. Not only do eye movements provide the experimenter with a special window into these cognitive processes, they provide the individual with a way to modify their cognitive processes. This chapter examines this dual role of eye movements (both revealing and modifying cognition) in a variety of problem solving tasks. Experimenters can better understand the underlying cognitive processes that are involved in problem solving by recording eye movement patterns, and individuals can better perform their problem solving when they produce the right eye movement patterns.

Introduction

A well known puzzle, known as the river-crossing problem, places missionaries in boats with cannibals and asks you to transit them across a river in a manner that ensures that no missionary is eaten. Consider one such dire scenario: 'Three missionaries and three cannibals want to get to the other side of a river. There is a small boat, which can fit only two. To prevent a tragedy, there can never be more cannibals than missionaries together.' If the reader is like this chapter's authors, this puzzle induces a flurry of dynamic imagery, including a visualization of the nervous missionaries, salivating cannibals, a boat, a river, and perhaps even irrelevant factors like the protagonists' appearance, or a sun in the sky. The reader, like us, may manipulate the visualization and begin moving its 'pieces' in seeking a solution.

General solutions to versions of this problem have been proposed since the 19th century (Pressman and Singmaster, 1989), and into the mid-20th century in a burgeoning field of artificial intelligence (Amarel, 1968). These solutions have never attempted to integrate the rich 'perceptual simulation' (Barsalou, 1999) noted in the previous paragraph. In general, they have had highly formal logical characteristics, translating relevant parts of these problems into variables and operators often dubbed 'amodal' because they lack the characteristics of sensorimotor modalities involved in imagery. This has been a general feature of 'classical' approaches to our cognitive system: treating our internal thought processes as governed by abstract symbols not much different from a digital computer's states and processes (e.g. Newell, 1990; Newell and Simon, 1961). To the uninitiated, this may seem like a very complex, mathematical way of approaching these issues as potential explanations for how we ourselves solve them. But in fact this strategy has been an influential simplification, allowing cognitive scientists to develop computer programs that perform human-like problem solving in a diverse range of problem spaces (see Poole et al., 1997, for many examples).

There has been considerable debate about the usefulness of these formalistic approaches to understanding human cognition in general (e.g. Dreyfus, 1979; Marcus, 2001; Searle, 1980; Spivey, 2007). One charge against the classical approach has been its lack of perceptual and motor detail that we seem richly to experience in day-to-day cognitive functioning, as the first paragraph of this chapter attempts to demonstrate (see, e.g. Clark, 1997, for review). In the past two decades, the view that cognitive scientists should integrate sensorimotor, or 'modal,' representations in our theories of cognition has been gaining traction. Often dubbed the 'embodiment of cognition,' the approach sees cognitive processes as directly involving modal sensorimotor systems, e.g. visual, somatosensory, and auditory perception, as well as oculomotor and skeletomotor planning. In other words, even during 'high-level' cognitive processes such as problem solving, the brain recruits 'low-level' sensorimotor neural subsystems to assist in the cognitive computations.

AQ: please supply Marcus 2001 reference

In this chapter, we review research showing that eye movements support the embodiment of problem solving. When facing many kinds of puzzles and problems, our cognitive system makes use of implicit perceptuomotor activities to search the problem space for solutions. In what follows, we present a selective review of evidence for embodiment especially relevant to eye movements and problem solving. This review motivates two key predictions: Eye movements should: 1) reveal problem solving processes as they unfold, and 2) potentially aid problem solving directly. We then showcase recent empirical support for both predictions. The chapter ends with theoretical discussion, and future directions in which eye movement and other dynamic, temporal methodologies could contribute to our understanding of high-level cognition.

The embodiment of cognition

Consider again the river-crossing problem. This type of problem has some basic properties that many puzzles and problems have in common. First, it is conveyed through linguistic means. Problem solvers are provided statements that lay out a set of premises, assumptions, arrangements, etc. that are intended to establish basic rules or constraints on a scenario. Second, a given scenario may have some spatiotemporal conditions that approximate arrangements of things in the natural world. The river-crossing problem has folks travelling in boats, while other problem spaces involve moving blocks about pegs (e.g. Tower of Hanoi problem: Simon, 1975), or connecting dots on a two-dimensional surface (e.g. nine-dot problem: MacGregor et al., 2001). Both of these qualities—linguistic transmission and spatiotemporal arrangement—invoke cognitive processes that have embodied sensorimotor qualities. Evidence for this has come directly from eye tracking experiments.

In an obvious way, concrete spatial language that explicitly mentions positions or directions of movement may invoke sensorimotor representations. But even abstract figurative language can generate embodied effects, evidenced in eye movement patterns. For example, Matlock and colleagues have investigated the processing of sentences that figuratively invoke motion, known as 'fictive motion' descriptions (Matlock, 2004). Matlock and Richardson (2004) conducted a passive viewing experiment in which participants saw a scene of objects and listened to sentences. Some scene/sentence pairs contained fictive motion. For example, a view of a desert with a road down the centre of the scene could be described with 'The road *runs through* the desert' (fictive motion) or simply 'The road is in the desert'. Participants spent more time fixating the path of motion when they heard fictive motion sentences than when they heard the literal control sentences. Richardson and Matlock (2007) further demonstrated that when the sentences described movement difficulty (e.g. '. . . the desert is rough'), participants' eye movement patterns covaried with this movement difficulty. Scanpaths exhibited enhanced focus on the region of the fictive motion (e.g. along the road) under conditions of more difficult movement. These and numerous other studies have shown that a diverse range of language processes make use of sensorimotor representations (for reviews, see Anderson and Spivey, 2009; Barsalou, 2008; Glenberg and Kaschak, 2003).

As for spatiotemporal arrangement, river-crossing and other problems may induce visual imagery of manipulating this arrangement, during which modal visual areas of cortex are activated (see Kosslyn and Thompson, 2003; Kosslyn et al., 2003 for review). If this visual system activation is functionally important to visual imagery, then oculomotor output is likely to correspond to the structure of imagery. Indeed, Spivey and Geng (2001, expt. 1) have shown that visual imagery elicits eye movements (while viewing a blank screen) that correspond to the locations and movement directions in the spoken descriptions (Altmann, 2004; Brandt and Stark, 1997; Laeng and Teodorescu, 2002; see also Altmann, Chapter 54, this volume).

There may be good reasons for this oculomotor enactment of cognition: memory systems make use of external spatial locations as anchors for encoded information. The eyes should thus reflect these external 'indexes' (Pylyshyn, 1989) for recalled information. Richardson and Spivey (2000) have shown that when participants learn new facts in association with spatial regions on a screen, participants look to those now-blank regions when recalling the facts. In a series of experiments, participants saw a 2×2 grid, and in different cells viewed a video of a face presenting various statements (e.g. 'Australia's capital is Canberra'). When participants were queried to verify a version of one of those statements, they frequently looked to the corresponding empty cell in the grid while they answered the query—even though they could easily tell in their peripheral vision that it was empty, and the statement has been delivered auditorily anyway!

A natural question that follows is whether these eye movements to those locations can improve memory for that object or event. That is, can recreating during recall the eye movement pattern that took place during encoding instigate at kind of pattern completion process? It has recently been hypothesized that if eye movements are systematically manipulated by the experimental task, e.g. turned into an independent variable instead of a dependent variable, then one might observe different memory performance for trials in which the eyes are fixating the correct location versus an incorrect location (Ferreira et al., 2008). Although this is certainly a logical possibility, the existing data are not consistent with it (Richardson et al., 2009). In every test for such an improvement in memory accuracy conducted so far, none has been found (Hoover and Richardson, 2008; Richardson and Kirkham, 2004; Richardson and Spivey, 2000; Spivey and Geng, 2001, expt. 2). Recently, there is evidence for a slight decrease in reaction time to memory probes when the eyes are fixating the correct location compared to when they are fixating the wrong location, but no evidence for improvements in memory accuracy (Theeuwes et al., in press; Vankov, 2009). For example, Vankov (2009) presented participants with four line drawings of objects and then presented a memory probe consisting of a single word that might refer to one of the objects. In some conditions, the word showed up in the same location as the now-absent object, and in other conditions, it showed up in a different location. Vankov observed some subtle reaction time differences between these conditions, but the accuracy of memory performance was not affected by the location of the memory cue. Thus, rather than these memory-induced eye movements being evidence for an embodiment of cognition solely inside the sensorimotor system (where pattern completion of neural patterns can take place), they may be better evidence for an embeddedness of cognition in the environment (e.g. O'Regan and Noë, 2002; Spivey et al., 2004), whereby a small bit of internal semantic information regarding an object or event is linked to a location in the visual field where 'the rest' of the information is expected to be accessible (O'Regan, 1992).

The selective review above supports the involvement of embodied representations in language processes and visual imagery, both likely components of naturalistic problem solving. Memory processes accompanying both of these can be anchored in spatial locations in the world, and eye-movement signatures reveal this. In the next section, we briefly provide a more mechanistic account of the deep connection between eye movements and cognition, which leads to two broad predictions about how eye movements should be involved in problem solving.

Eye movements are coextensive with cognition

The key reason why eye movement patterns are so informative about mental activity is that they are part and parcel of it. Rather than being a slave output system that patiently awaits discrete finalized commands from the cognitive system, oculomotor processing is coextensive with cognitive processing. The ongoing processes of cognition continuously 'spill over' (before a decision is final) into the oculomotor system, causing it to prepare partially-active movement plans that are consistent with the gradually accumulating perceptual evidence (Gold and Shadlen, 2007). This is due in part to the fact that the eyes tend to make saccadic movements about every 200–300 ms, whereas it takes about 400–500 ms for a neural population code that embodies the representation of a visual object to achieve its maximal activation (e.g. Rolls and Tovee, 1995). As a result, the eyes are often moving to fixate a new object before the previous object has been fully and completely recognized. This continuously flowing perception-action cycle (Neisser, 1976) becomes an autocatalytic causal loop, in which cognition emerges (Spivey, 2007). Before a cognitive decision to respond to the stimulus environment has reached completion, the eyes will often move to fixate a new object or location in the visual field. This newly foveated perceptual information then changes the cognitive processes which led to that eye movement—and which were on their way to one decision but now may be on their way to a different one! Cognition somehow manages to simultaneously be a major cause of eye-movement patterns and partly a result of them.

A concrete demonstration of this continuity between a cognitive process (such as a decision) and an eye movement comes from a study by Gold and Shadlen (2000). As perceptual input accumulates over time to produce an overt decision and motor response, the continuous evolution of that decision can be seen in eye movement data. That is, the network of brain systems in which a perceptual decision is generated includes oculomotor nuclei, such as the frontal eye fields. Gold and Shadlen trained a monkey to respond to a pseudorandom dot-motion stimulus by looking at an upper response target when the central motion stimulus had upward motion in it, and at a lower response target when the central motion stimulus had downward motion in it. Then, just before a voluntary eye movement response was made, they microstimulated a region of the monkey's frontal eye fields (via mild electric current delivered through an electrode) to generate an involuntary rightward saccade. When the monkey was allowed only a couple hundred milliseconds to see the motion stimulus, the involuntary evoked saccade was a nearly pure rightward movement. However, with 300 or 400 or 500 ms to see the motion stimulus, the involuntary evoked saccade showed more and more vertical component in its direction. Essentially, the perceptual decision (of upward vs. downward motion) that was evolving over several hundred milliseconds was simultaneously generating a pattern of neural activity in the frontal eye fields that formed the beginnings of a voluntary upward or downward eye movement command. When the microstimulation caused the involuntary rightward saccade, the partly evolved voluntary eye movement command influenced the direction that the eyes actually went. Thus, rather than the oculomotor system waiting until a single confident perceptual decision was achieved, it was continuously 'listening to' and 'involved in' the gradual formation of that perceptual decision.

One might, in fact, wish to draw a distinction between the notions of the eye movement system continuously *listening* to cognition versus being continuously *involved in* cognition. This kind of distinction is often raised when discussing embodied cognition in general (Mahon and Caramazza, 2008), where some embodied cognition results may be best interpreted as a spreading of activation from cognitive to motor subsystems (e.g. Glenberg and Kaschak, 2002), and other embodied cognition results clearly demonstrate a functional role for motor patterns feeding back into cognitive processes (Pulvermüller et al., 2005). Accordingly, the next two sections in this chapter will address evidence for eye movements revealing certain cognitive processes related to problem solving tasks because those cognitive processes continuously spread their activation patterns into the oculomotor system, and also evidence for eye movements influencing the cognitive processes related to problem solving tasks because those eye movement patterns change the way the cognitive processes function.

Eye movements reveal problem solving

Mathematical problem solving

A number of studies have shown a relationship between the activity of the eyes and the activity of the mind during mathematical problem solving. Hess and Polt (1964) were the first to demonstrate that a person's pupils dilate when he or she is solving a difficult mathematical problem. As their multiplication problems got harder, pupil dilation steadily increased. It was a few years later that saccadic movements of the eyes were linked to problem-solving tasks. Yarbus (1967) showed that it is not merely visual input that determines eye movement patterns, but task constraints as well. In his experiment, the same painting (Ilya Repin's *An Unexpected Visitor*) elicited very different eye movement scanning patterns (scanpaths) when the viewer was probed with different queries, such as 'What are the ages of the people in the painting?,' 'What had the family been doing before the visitor arrived?', and 'How long had the visitor been away?' (For similar results in individuals with Asperger's syndrome, see Benson and Fletcher-Watson, Chapter 39, this volume.)

Using the record of saccadic eye movements to uncover the cognitive processes involved in mathematical problem solving was first systematically explored by Suppes et al. (1982). In testing their procedural theory of eye movements during addition and subtraction problems, they found that although arithmetic problems that require 'carrying' or 'borrowing' are more difficult and take longer than those that do not, actually 'carrying' or 'borrowing' values while solving the problem increases eye fixation durations only very slightly. When solving algebra problems, it is noteworthy that students do not simply fixate each number and symbol left-to-right. Instead, they will sometimes skip values (presumably perceiving them parafoveally) and other times they will refixate values (Salvucci and Anderson, 2001).

Where we can clearly observe the eyes spending substantially more time fixating difficult parts of a math problem is with word problems. Word problems can be written with inconsistent relational terms, such as 'Johnny has five cookies, which is three fewer than Mary has. How many cookies does Mary have?' This word problem is referred to as 'inconsistent' because the relational term 'fewer than' can trick students into using subtraction (instead of addition) to solve the problem. In fact, eye movement patterns (such as more regressive saccades back to the relational terms) suggest that this is exactly why students are sometimes unsuccessful with inconsistently worded problems (Hegarty et al., 1995; Verschaffel et al., 1992).

Eye movements during the linguistic delivery of math and logic problems have also been explored when there were no visual stimuli at all. For example, Demarais and Cohen (1998) presented participants with spoken recordings of transitive inference problems, such as 'a jar of pickles is below a box of tea bags; the jar of pickles is above a can of coffee; where's the can of coffee?' While staring at a completely blank display, participants tended to make more vertical (than horizontal) eye movements while hearing those above/below inference problems. When hearing inference problems where objects were aligned to the left and right of one another, participants made more horizontal (than vertical) eye movements. These results suggest that when people are solving transitive inference problems, with no external visual aids, they nonetheless generate some form of spatial mental model to assist their logical induction (Byrne and Johnson-Laird, 1989), and they perform eye movements consistent with the spatial characteristics of that mental model.

Eye movement patterns can also reveal the cognitive processes involved in solving geometry problems. Epelboim and Suppes (2001) tracked participants' eye movements while they looked at diagrams that had letters indexing the endpoints of lines and centroids of circles, and then some angles provided in degrees, and then a particular angle whose size (in degrees) was to be solved. The eye movement patterns exhibited a variety of interesting properties. For example, participants tended to look away from the diagram when they were performing mental arithmetic (as indicated by their verbal protocols). More importantly, the scanpaths revealed a great many redundant sequences consisting of many rescans of previously fixated diagram components throughout the solving of the problem. This observation is consistent with Epelboim and Suppes's model of geometry reasoning in which fixated components of a diagram tend to overwrite existing diagram components in visual working memory—suggesting that visual working memory may tend to store a surprisingly small amount of information at any one time (see also Ballard et al., 1997).

Mechanical problem solving

Carpenter and Just (1978; see also Just and Carpenter, 1985) were the first to show a relationship between eye movements and mental rotation. When participants compare a target shape to a set of alternative rotated shapes, to find the matching one, the eye movement patterns suggest that participants go through a few different procedures in sequence. They encode the target shape and *search* for a matching component among the shapes, then *transform* or rotate that component in one of the shapes, and then *confirm* whether the other components match those in the target shape after this mental transformation (Just and Carpenter, 1985). Interestingly, the eye movement data suggested that the increased latencies resulting from more complex shape comparisons appeared to have been due to longer encoding and confirmation phases, rather than longer transformation or rotation phases (Carpenter and Just, 1978).

While eyetracking a simplified version of the Tower of London problem, Hodgson et al. (2000) also found evidence for a search phase, followed by a transformation phase, and then a confirmation phase. Participants started each trial with multiple fixations of the target stack of balls, then multiple fixations of the stack of balls that they were to imagine manipulating to match the target, and then finishing with fixations of the target stack again. Interestingly, on trials where participants produced incorrect answers, their scanpaths were surprisingly conspicuously similar to those of the previous trial. Hodgson et al. suggested that participants sometimes get stuck from one trial to the next in an ocolumotor routine that may not be optimal for the new problem diagram.

Mental animation during mechanical reasoning can also be revealed in eye movement data. For example, Hegarty (1992; see also Hegarty and Just, 1993) recorded participants' eye movements while they looked at diagrams of pulley systems, and found evidence that when people mentally animate these diagrams, in order to verify/deny statements about them, their animations are piecemeal. Rather than imagining all the pulleys moving at once, as would actually happen when the rope threads its way through them, the eye movement patterns suggest that participants mentally animate small portions of the pulley system in sequential order, reflecting the causal chain of forces.

The same observation was made for eye movement patterns with more complex gear-and-belt diagrams. Rozenblit et al. (2002) tracked participants' eye movements while they looked at gear-and-belt diagrams in which the leftmost gear was specified to rotate in a particular direction, and it connected to several intermediate gears (which alternate the previous direction of rotation) and straight belts (which maintain the previous direction of rotation), culminating in a pendulum. The participant's task was to guess which direction the pendulum would swing. Practised eye movement coders could tell which directions a participant was mentally animating a set of connected gears and belts by watching the slow-motion video of the eye movement trace overlaid on the diagram. For example, a few fixations near the conjunction of two gears can reveal that the participant was mentally animating the left gear clockwise and the right gear counter-clockwise simply by the fact that the sequence of fixations went left-to-right on the *upper* portion of the left gear and then left-to-right on the *lower* portion of the right gear. The proof for this observation comes especially clear when that direction of animation is incorrect and thus allows the coder to accurately predict an incorrect guess from the participant on that trial. In fact, Rozenblit et al. showed that even unpractised coders—naïve experimental participants watching the slow-motion video of multiple trials (with their final portions edited out)—can use the eye movement traces to perform well above chance in predicting incorrect guesses. Thus, even when it is difficult to identity the specific properties of the eye-movement trace that reveal cognition, there is clearly information in those scanpaths that can nonetheless be used (even by non-experts) to infer cognitive processes.

Insight problem solving

So far, we have been discussing problem-solving tasks where the formulation of the solution plods along relatively linearly. With problems like that, participants tend to have accurate intuitions about how close they are to reaching the solution. With insight problems, by contrast, that is not the case at all. With insight problems—and there is a wide variety of them—participants usually reach a point partway through where they feel like they've run out of ideas, known as an *impasse*. When encouraged to keep trying anyway, some proportion of the participants will eventually achieve *insight*, in some cases even widening the eyes and declaring 'Aha!' Notably, on the way toward that 'Aha!'

moment, if they are asked to report how close to the solution they think they might be, their ratings do not correlate with their actual remaining time-to-solution. Insight problems tend to be difficult puzzles, so a substantial proportion of participants never find the solution at all (or they require guiding hints from the experimenter to get there). Participants typically describe the solution as having come to them 'suddenly out of nowhere.' Nonetheless, more subtle measures of performance, such as two-alternative forced-choice (Bowers et al., 1990), reaction-time priming (Bowden and Jung-Beeman, 1998), and eye-tracking (Knoblich et al., 2001), can often provide inklings of how the solution gradually emerges over time. This opacity to subjective reports makes insight problem solving particularly intriguing for theories of cognition, and a prime target of eye movement methodologies.

Before discussing the more standard insight problems, however, it might be useful to examine a kind of problem that falls somewhere in between non-insight problems and insight problems. This next eyetracking study that opens a window onto a mechanical problem-solving task is not typically considered an insight problem, per se, because it does not involve an impasse phase. Participants view a sequence of connected gears, are told that the leftmost gear will turn clockwise (or counter-clockwise), and then asked which direction the rightmost gear will turn. After several trials of solving these simple linear-gear-sequence problems, participants rather suddenly discover a new easy strategy for solving them. In fact, sometimes they even shout 'Aha!' when they suddenly realize that one need not mentally animate each gear to figure out which direction the last one rotates; one can just count whether there is an even or odd number of gears (Dixon and Bangert, 2004). If there is an even number of linearly-connected gears, then the last gear will rotate in the*opposite* direction of the first gear. If there is an odd number of gears, then the last gear will rotate in the*same* direction as the first gear. What is interesting about this discovery process is that the statistical structure of the eye-movement patterns presages that moment of discovery. That is, over the course of the two to five trials that precede the 'Aha' moment of figuring out the even-odd strategy, the eye-movement trace shows a statistically reliable increase in entropy, a measure of disorder in the data (Stephen et al., 2009). Stephen et al. interpret this as evidence that the mental animation strategy is becoming unstable, and is on the brink of yielding to the emerging even-odd strategy. On the one trial immediately preceding the trial on which they first use their new even-odd strategy, the eye-movement trace shows a statistically reliable*decrease* in entropy. This suggests that right before the new strategy is implemented, participants' scanning behaviour reveals a renewed stability and order reflecting the completion of the cognitive restructuring from the old mental animation strategy to the newly-adopted even–odd strategy. Thus, eye-movement patterns can provide a window into the gradual emergence of a new strategy during problem solving—even predicting when a new strategy is on its way.

This particular view on cognitive processes provided by eye-tracking can even be seen in genuine insight problems, which involve both an impasse and then an 'Aha' moment (for a more extensive review, see Knoblich et al., 2005). Knoblich et al. (2001) tracked participants' eye movements while they attempted to solve/repair Roman numeral arithmetic problems written with matchsticks. For example, reposition a single matchstick to make the following equation true: IV = III + III. (Note that all numerals and operators are formed with straight matchsticks.) In this non-insight example, one simply needs to change the 'IV' to a 'VI', and the equation is true. Participants tend to solve that one quickly and easily. However, more difficult problems, such as III = III + III, will often induce an impasse; and when the participant discovers the solution, they experience an 'Aha' moment. The solution for converting that false equation into a true one, involves relaxing an implicitly assumed constraint (that only matchsticks forming numerals can be modified).

Knoblich et al. (2001) found that, during the impasse phase with the difficult problems, mean fixation duration increased significantly. Basically, when participants had run out of ideas, they stared at the display without exploring new solution strategies. While that finding shows the impasse stage of insight problem solving, this next finding shows how eye movements can reveal the representational change that takes place right before achieving insight. As would be expected from an implicit (and incorrect) assumption that only matchsticks forming numerals can be modified, participants typically began the difficult problems fixating the numerals almost exclusively, rarely fixating the operators.

Over the course of attempting to solve these difficult 'constraint relaxation' problems, the participants who would soon discover the solution (e.g. changing the plus sign to an equal sign with a single movement of one matchstick) showed a steady gradual increase in proportion of fixations on operators prior to their 'Aha!' moment.

Similar evidence for eye movements revealing the emergent temporal dynamics of a seemingly-sudden insight comes from work by Jones (2003). Jones used a diagram-based car park problem, where the tricky insight is that, in addition to moving other cars out of the way to form an exit path, one's own car must move to make room for some of those other cars' movements. Jones operationalized the impasse as a looking time for a particular move that exceeded that participant's mean looking time by two standard deviations. He found that all 30 (out of 37) participants who solved the problem experienced at least one such impasse. Moreover, the bulk of that impasse was focused on the part of the problem that required realizing that one's own car needed to be moved. Mean looking time for that move was five times greater than the mean looking time for other moves. However, about one-third of the participants exhibited impasses (long gazes without making any moves) quite early on, suggesting that they may have been strategizing multiple moves in advance. In fact, those participants with early impasses solved the car park problem faster and with fewer moves than participants who did not exhibit early impasses.

Eye movements influence problem solving

In addition to providing cognitive scientists with a peek into the mental activity involved in problem solving, eye movements may themselves be able to instigate insight for a person attempting to solve a problem. For example, Grant and Spivey (2003, expt. 1) showed that fixating certain regions of a tumour-and-stomach diagram, for Duncker's classic radiation problem, was correlated with achieving insight. The schematic diagram consisted of a solid oval, representing the tumour, with a circumscribing oval representing the stomach lining. While viewing that display, participants were told, 'Given a human being with an inoperable stomach tumour, and lasers which destroy organic tissue at sufficient intensity, how can one cure the person with these lasers and, at the same time, avoid harming the healthy tissue that surrounds the tumour?'

All participants reported experiencing a period of impasse, where they had run out of ideas. Only about a third of the participants produced the correct solution (i.e. to use multiple low-intensity lasers, which are too weak to damage the skin, and converge their incident rays onto the tumour to combine their intensities and burn it away). These successful participants routinely blurted out an 'Aha!' (or some equivalent) when they discovered this solution. Although the eye movement patterns were somewhat similar for successful and unsuccessful participants, one key difference stood out. Successful solvers tended to spend more time looking at the stomach-lining portion of the diagram than unsuccessful solvers. In fact, the pattern of eye movements that was most closely associated with achieving insight with this problem was one in which one or two saccades were made from the external portion of the diagram inward to the tumour (often stopping at the stomach-lining) and then back out to another external region. This triangular sequence of saccades, which almost 'paints a sketch' of multiple incident rays converging on the tumour, was especially common prior to the participant achieving insight. Could it be that people who *happened to* produce eye movement patterns like that were steered toward the cognitive insight by their sensorimotor patterns priming a 'perceptual simulation' (Barsalou, 1999) of multiple incident rays? Or is it merely that their cognitive processes were gradually forming the correct solution anyway, and this caused their eye movements to reflect that nascent solution?

In a first step toward addressing this question, Grant and Spivey (2003) conducted a second experiment in which a computer display animated either the tumour or the stomach-lining, with a one-pixel increase in diameter flashing at about 3 Hz. With attention and eye movements drawn toward the animated tumour, only a third of those participants solved the radiation problem—the same as with the static diagram in their first experiment. However, of the participants whose diagrams had the stomach-lining flashing, two-thirds discovered the insight of converging multiple weak lasers onto the tumour. Thus, in this second experiment, attracting participants' attention toward the stomach-lining region of the diagram (which is where successful solvers were spontaneously looking in the first experiment) *doubled* the solution rate for this difficult insight problem. Note, however, that because Grant and Spivey did not track eye movements in this second experiment, it is not clear whether it is eye movements, per se, that jump-started this cognitive insight, or whether it might be covert spatial attention that did so.

Following up on this issue, Thomas and Lleras (2007) used a similar schematic diagram with Duncker's radiation problem and added to it a secondary visual-tracking task that, for some subjects, *just happened to* make their eyes move in the same triangular fashion that Grant and Spivey (2003) first noticed in their eye-movement data. Eye movements were recorded to ensure that participants adhered to the visual-tacking task. Importantly, participants whose eye movements were guided (by the visual-tracking task) to reproduce this triangular pattern were more than twice as likely to discover the solution than participants whose eye movements were driven in neutral or irrelevant patterns. Since participants reported not detecting any relationship between the visual-tracking task and the radiation problem, Thomas and Lleras (2007) concluded that the eye movements were influencing the formation of insight in this problem.

Interestingly, since covert allocation of spatial attention is often likened to programmed eye movements that simply aren't executed (e.g. Sheliga et al., 1994), one might expect that patterns of covert spatial attention—independent of eye movements—might be equally capable of influencing cognition in a diagram-based problem like this. Indeed, Thomas and Lleras (2009) found that when eye position was fixed on the tumour, but visual cues directed covert spatial attention to the external portion of the diagram and then back to the tumour again and again, the solution rate almost tripled—compared to central fixation without attentional cues. Essentially, preparing eye movements (with or without executing them) that are spatially compatible with the convergence solution to the radiation problem is sufficient to prime the perceptual simulation that embodies the cognitive insight. Thus, eye movement processes do more than simply *reveal* (to the experimenter) the temporal dynamics of cognition during problem solving. Eye movement processes can also *influence* (for oneself) the cognitive processes that go into an attempt to solve a problem.

Conclusion

Classical approaches to explaining problem solving, described in the introduction, neglect the perceptuomotor processes that are likely involved in problem solving. The eye-tracking work described above shows that embodied representations both reflect unfolding problem solving, and may directly aid in solving certain kinds of problems. Eye-tracking methodologies are extremely well suited to identifying these representations as they occur. For one, the eye-movement record is a semi-continuous real-time source of information about cognitive processes, in many ways out of conscious strategic control of participants (Liversedge and Findlay, 2000). In addition, as a motor system, it maps out spatial coding of the sensorimotor processes themselves (see Spivey et al., 2009, for further review).

AQ: please supply Liversedge and Findlay, 2000 reference

These characteristics guarantee that eye movements will remain, well into the future, an important methodological window on high-level cognitive processes, such as problem solving. In conjunction with other near-continuous measures of cognition used recently, such as the dynamics of reaching and pointing movements (e.g. Spivey and Dale, 2006; Spivey et al., 2005; Stephen et al., 2009), it offers investigation into the time-course and fine-grained representational structure of problem solving. Spatially and temporally coarser techniques, such as coded video streams or think aloud protocols (Ericsson and Simon, 1993) are still applied extensively to complex cognitive processes, like knowledge acquisition, problem solving, metacognition, and so on. Advances in the analysis of semi-continuous methods continue to offer novel insights that those coarser methods cannot access, especially in these high-level contexts (e.g. Graesser et al., 2005).

Finally, the theoretical question regarding whether there can be a preferred account of problem solving, classical or embodied, is not at present resolvable. In many regards, classically formal approaches enjoy more currency in this domain (Poole et al., 1997). Recently, some artificial intelligence research has been integrating symbolic and spatial representations, and solutions to some problems come handily with both in operation, compared to either alone (e.g. Wintermute and Laird, 2008). This perhaps encourages one to consider that both theoretical schemes may be helpful for understanding what guides a problem solver. Such integration will depend on a number of important factors, such as the particular problem at hand, problem-solving context, and any potential expertise. Eye tracking and other semicontinuous methods will inevitably provide fine-grained explorations of these and other factors in the future, contributing to ongoing theoretical debate.

Acknowledgements

Work on this chapter was supported by a collaborative research grant from the NSF to both authors.

References

Altmann, G., and Kamide, Y.(2004). Now you see it, now you don't: Mediating the mapping between language and the visual world. In J. Henderson and F. Ferreira (eds.) *The interface of language, vision, and action: Eye movements and the visual world*. New York: Psychology Press.

Amarel, S. (1968). On representations of problems of reasoning about actions. In D. Michie (ed.) *Machine Intelligence 3* (Vol. 3) (pp. 131–171). Amsterdam: Elsevier/North-Holland.

Anderson, S.E. and Spivey, M.J. (2009). The enactment of language: Decades of interactions between linguistic and motor processes. *Language and Cognition*, **1**, 87–111.

Ballard, D.H., Hayhoe, M.M., Pook, P.K., and Rao, R.P.N. (1997). Deictic codesfor the embodiment of cognition. *Behavioral and Brain Sciences*, **20**(4), 723–742.

Barsalou, L.W. (1999). Perceptions of perceptual symbols. *Behavioral and brain sciences*, **22**(4), 637–660.

Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology*, **59**, 617–45.

Bowden, E.M. and Beeman, M.J. (1998). Getting the right idea: Semantic activation in the right hemisphere may help solve insight problems. *Psychological Science*, **9**(6), 435–440.

Bowers, K.S., Regehr, G., Balthazard, C., and Parker, K. (1990). Intuition in the context of discovery. *Cognitive Psychology*, **22**(1), 72–110.

Brandt, S. and Stark, L. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience*, **9**, 27–38.

Byrne, R.M.J. and Johnson-Laird, P.N. (1989). Spatial reasoning. *Journal of Memory and Language*, **28**(5), 564–575.

Carpenter, P.A., and Just, M.A. (1978). Eye fixations during mental rotation. In J. Senders, R. Monty, and D. Fisher (eds.) *Eye movements and the higher psychological functions* (pp. 115–133). Hillsdale, NJ: Erlbaum.

Clark, A. (1997). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT press.

Demarais, A.M. and Cohen, B.H. (1998). Evidence for image-scanning eye movements during transitive inference. *Biological Psychology*, **49**(3), 229–247.

Dixon, J.A. and Bangert, A.S. (2004). On the spontaneous discovery of a mathematical relation during problem solving. *Cognitive Science*, **28**(3), 433–449.

Dreyfus, H.L. (1979). *What computers can't do*: *The limits of artificial intelligence*. London: HarperCollins.

Epelboim, J. and Suppes, P. (2001). A model of eye movements and visual working memory during problem solving in geometry. *Vision Research*, **41**(12), 1561–1574.

Ericsson, K.A. and Simon, H.A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Ferreira, F., Apel, J., and Henderson, J.M. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Sciences*, **12**(11), 405–410.

Glenberg, A.M. and Kaschak, M.P. (2002). Grounding language in action. *Psychonomic Bulletin and Review*, **9**(3), 558–565.

Glenberg, A.M. and Kaschak, M.P. (2003). The body's contribution to language. *Psychology of Learning and Motivation*, **43**, 93–126.

Gold, J.I. and Shadlen, M.N. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature*, **404**(6776), 390–394.

Gold, J.I. and Shadlen, M.N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, **30**, 535–574.

Graesser, A.C., Lu, S., Olde, B.A., Cooper-Pye, E., and Whitten, S. (2005). Question asking and eye tracking during cognitive disequilibrium: Comprehending illustrated texts on devices when the devices break down. *Memory and Cognition*, **33**(7), 1235–1247.

Grant, E.R. and Spivey, M.J. (2003). Guiding attention guides thought. *Psychological Science*, **14**(5), 462–466.

Hegarty, M. (1992). Mental animation: inferring motion from static displays of mechanical systems. *Journal of experimental psychology. Learning, Memory, and Cognition*, **18**(5), 1084–1102.

Hegarty, M., and Just, M.A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, **32**(6), 717–742.

Hegarty, M., Mayer, R.E., and Monk, C.A. (1995). Comprehension of arithmetic word problems: A comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology*, **87**, 18–32.

Hess, E.H. and Polt, J.M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, **143**, 1190–1192.

Hodgson, T.L., Bajwa, A., Owen, A.M., and Kennard, C. (2000). The strategic control of gaze direction in the Tower of London task. *Journal of Cognitive Neuroscience*, **12**(5), 894–907.

Hoover, M.A., and Richardson, D.C. (2008). When facts go down the rabbit hole: Contrasting features and objecthood as indexes to memory. *Cognition*, **108**, 533–542.

Jones, G. (2003). Testing two theories of cognitive insight. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **29**(5), 1017–1027.

Just, M.A. and Carpenter, P.A. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review*, **92**, 137–172.

Knoblich, G., Ohlsson, S., and Raney, G.E. (2001). An eye movement study of insight problem solving. *Memory and Cognition*, **29**(7), 1000–1009.

Knoblich, G., Öllinger, M., and Spivey, M. (2005). Eye movements providing insight into insight problem solving. In G. Underwood (ed.) *Eye Guidance in Cognition* (pp. 355–375). Oxford: Oxford University Press.

Kosslyn, S.M. and Thompson, W.L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological Bulletin*, **129**(5), 723–746.

Kosslyn, S.M. Ganis, G., and Thompson, W.L. (2003). Mental imagery: Against the nihilistic hypothesis. *Trends in Cognitive Sciences*, **7**(3), 109–111.

Laeng, B. and Teodorescu, D.S. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science*, **26**(2), 207–231.

MacGregor, J.N., Ormerod, T.C., and Chronicle, E.P. (2001). Information processing and insight: A process model of performance on the nine-dot and related problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **27**(1), 176–201.

Mahon, B.Z. and Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology (Paris)*, **102**(1–3), 59–70.

Matlock, T. (2004). Fictive motion as cognitive simulation. *Memory and Cognition*, **32**(8), 1389–1400.

Matlock, T. and Richardson, D.C. (2004). Do eye movements go with fictive motion? *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.

Neisser, U. (1976). *Cognition and reality*. San Francisco, CA: WH Freeman.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Newell, A. and Simon, H.A. (1961). Computer simulation of human thinking. *Science*, **134**(3495), 2011–2017.

O'Regan, J.K. (1992). Solving the 'real' mysteries of visual perception: the world as an outside memory. *Canadian Journal of Psychology*, **46**, 461–488.

O'Regan, J.K. and Noë, A. (2002). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, **24**(5), 939–973.

Poole, D., Mackworth, A., and Goebel, R. (1997).*Computational intelligence: a logical approach* Oxford: Oxford University Press.

Pressman, I. and Singmaster, D. (1989). ''The jealous husbands'' and ''the missionaries and cannibals''. *The Mathematical Gazette*, **73**(464), 73–81.

Pulvermüller, F., Hauk, O., Nikulin, V.V., and Ilmoniemi, R.J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience*, **21**(3), 793–797.

Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, **32**, 65–97.

Richardson, D.C. and Kirkham, N.Z. (2004). Multimodal events and moving locations: Eye movements of adults and 6-month-olds reveal dynamic spatial indexing. *Journal of Experimental Psychology: General*, **133**(1), 46–62.

Richardson, D.C. and Matlock, T. (2007). The integration of figurative language and static depictions: An eye movement study of fictive motion. *Cognition*, **102**(1), 129–138.

Richardson, D.C. and Spivey, M.J. (2000). Representation, space and Hollywood Squares: looking at things that aren't there anymore. *Cognition*, **76**(3), 269–295.

Richardson, D.C., Altmann, G., Spivey, M.J., and Hoover, M. (2009). Much ado about eye movements to nothing. *Trends in Cognitive Sciences*, **13**, 235–236.

Rolls, E.T. and Tovee, M.J. (1995). Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *Journal of Neurophysiology*, **73**(2), 713–726.

Rozenblit, L., Spivey, M.J., and Wojslawowicz, J. (2002). Mechanical reasoning about gear-and-belt systems: Do eye movements predict performance? In M. Anderson, B. Meyer, and P. Olivier (eds.) *Diagrammatic representation and reasoning* (pp. 223–240). Berlin: Springer-Verlag.

Salvucci, D.D. and Anderson, J.R. (2001). Automated eye-movement protocol analysis. *Human-Computer Interaction*, **16**(1), 39–86.

Searle, J. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, **3**(3), 417–457.

Sheliga, B.M., Riggio, L., and Rizzolatti, G. (1994). Orienting of attention and eye movements. *Experimental Brain Research*, **98**(3), 507–522.

Simon, H.A. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology*, **7**(2), 268–288.

Spivey, M.J. (2007). *The continuity of mind*. Oxford: Oxford University Press.

Spivey, M.J. and Dale, R. (2006). Continuous dynamics in real-time cognition. *Current Directions in Psychological Science*, **15**(5), 207–211.

Spivey, M.J. and Geng, J.J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, **65**(4), 235–241.

Spivey, M.J., Richardson, D.C., and Fitneva, S.A. (2004). Thinking outside the brain: spatial indices to visual and linguistic information. In J.M. Henderson and F. Ferreira (eds.) *The interface of language, vision, and action: Eye movements and the visual world.* (pp. 161–189). New York, NY: Psychology Press.

Spivey, M., Grosjean, M., and Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences*, **102**, 10393–10398.

Spivey, M.J., Richardson, D., and Dale, R. (2009). The movement of eye and hand as a window into language and cognition. In E. Morsella, J. Bargh, and P.M. Gollwitzer (eds.) *The Psychology of Action* (Vol. 2) (pp. 225–249). New York: Oxford University Press.

Stephen, D.G., Boncoddo, R.A., Magnuson, J.S., and Dixon, J.A. (2009). The dynamics of insight: Phase transitions and entropy in mathematical discovery. *Memory and Cognition,* **37** *(8),* 1132–1149.

Suppes, P., Cohen, M., Laddaga, R., Anliker, J., and Floyd, H. (1982).Research on eye movements in arithmetic performance. In R. Groner and P. Fraisse (eds.) *Cognition and eye movements* (pp. 57–73). North-Holland.

Theeuwes, J., Kramer, A.F., and Irwin, D.E. (in press). Attention on our mind: The role of spatial attention in visual working memory. *Acta Psychologica*.

Thomas, L.E. and Lleras, A. (2007). Moving eyes and moving thought: On the spatial compatibility between eye movements. *Psychonomic Bulletin and Review*, **14**, 663–668.

Thomas, L.E. and Lleras, A. (2009). Covert shifts of attention function as an implicit aid to insight. *Cognition,* **111**(2), 168–174.

Vankov, I. (2009). Mind the gap: The cost of looking at nothing, or the performance implications of memory-induced attention shifts. *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 1318–1323). Austin, TX: Cognitive Science Society.

Verschaffel, L., De Corte, E., and Pauwels, A. (1992). Solving compare problems: An eye movement test of Lewis and Mayer's consistency hypothesis. *Journal of Educational Psychology*, **84**(1), 85–94.

Wintermute, S. and Laird, J.E. (2008). Bimodal spatial reasoning with continuous motion. *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence (AAAI-08)* (pp. 1331–1337). Chicago, IL: AAAI-08.

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.

AQ: please update Theeuwes if now published