Mental Imagery. The idea that mental imagery may have a role in reasoning goes back a long way – one of the earliest attempts to study empirically the role that such imagery played in thinking was carried out by Sir Francis Galton in 1883, when he surveyed a large number of people for their use of mental imagery, using his "breakfast table" visualization test. He was surprised to find that many people disavowed having any visual imagery and that scientists generally played down the role of imagery in their thinking. But the study of such mentalistic concepts soon fell on hard times under behaviorism. When behaviorist ideology began to lose its grip on psychology, about two thirds of the way through the twentieth century, mental imagery was one of the first mentalistic concepts to emerge into prominence. Initially this emergence showed up in the use of "imagery ratings," which overtook the venerable frequency-of-occurrence index as an "intervening variable" in traditional studies of associative learning and memory. Soon after the rehabilitation of mental imagery began, a large number of experiments were performed to explore the form and function of mental images in reasoning. These experiments showed that images could be "scanned," "rotated," and examined in the "mind's eye" in order to judge their size, shape or other visual properties. Other studies also showed that examining mental images "projected" onto perceived scenes exhibited many properties associated with vision. Such images could lead to interference with perceptual tasks, to visual-motor adaptation, and even to visual illusions of the sort that might be expected if the imagined forms were actually part of the stimulus. This persuaded many people that mental images themselves had many of the properties of visual stimuli—and in particular that they possessed metrical spatial properties, such as size and inter-object distances.

The consensus that soon emerged was that mental images were very similar to real visual stimuli, except that they were generated by the mind instead of by stimulation of the retina. One influential view took the form of a "Cathode Ray Tube" (or CRT) metaphor, in which an image was assumed to be projected onto a mental "display" that was then perceived by the visual system (Kosslyn, Pinker, Smith, & Shwartz, 1979). Another view suggested that imaginal memory could "feed into" the visual system at various levels – from the earliest stages of vision to later conceptual stages (Finke, 1980). Yet a third suggested that because imagery and vision use some of the same mechanisms, imaging bears a "second order isomorphism" to the similarity of patterns arising from vision (Shepard and Chipman, 1970). A common recurring theme is that cognition can draw a "picture" that it can then "reperceive" – a notion that clearly leads to a regress, especially since it is assumed that imagery and vision use the same inner display. That this obvious regress is not generally recognized is a tribute to how strong a grip our subjective impression of imagery, as something that an inner observer examines, has on us.

As claims of the similarity between vision and imagery multiplied, it became clear to some writers that something was amiss. It was beginning to look more and more like the peripheral visual apparatus as well as the visible world itself were being moved inside the mind. For example, experiments showed that "small" mental images were harder to see (and took longer to report details from) than "large" images, that the mind's eye could "squint," that it exhibited the "oblique effect" (in which oblique lines were harder to distinguish than similarly spaced horizontal or vertical lines), that it showed an acuity profile similar to that of the real eye, and so on. Since many of these properties arise in vision because of the neuroanatomy of the eye and its connections to the visual cortex, it appeared that the "mind's eye" might share all these properties (presumably it had a blind spot and in some cases might even need corrective glasses)! What was happening was that the pull of our subjective experience of "seeing with our mind's eye" was blinding us to the fact that our experience of imagery is an experience of *seeing a possible world*, not of examining causal information-processing mechanisms within the brain. Even more importantly, it was blinding us to alternative (and in most cases rather obvious) explanations of many of the empirical phenomena.

The alternative that was being overlooked is that when one is asked to "imagine" something, the natural interpretation of this task is that one should try to recreate as many aspects as one can, or as one believes to be relevant, of a situation in which one is actually viewing the imagined situation unfold. Thus if asked to imagine scanning attention across a mental map, one attempts to create (or simulate) a sequence of mental states in which one is attending to a sequence of places one knows to be along the route, (or, more likely, one computes time-to-contact durations to landmarks that we can see while imagining Pylyshyn, 2003b): The entire route need not be represented in detail anywhere, just the thought that one is looking "here" and then one is looking "there" and so on. The demonstratives *this* and *that* in this case can refer to objects that are actually perceived, perhaps visually if the eyes are open, or through audition or proprioception if the eyes are closed. Similarly, if asked to imagine a "small" thing one takes into account that small things are harder to see, have fewer visible details, and so on, and then one simulates, in terms of whatever responses one is required to make, what one believes would happen in such a situation. In the case of "mental scanning" experiments, we have demonstrated that if the experimental demands are removed (so, for example, when subjects are not invited to imagine that they are scanning their attention across the scene) no "scanning effect" relating time to distance ensues; Pylyshyn (1981). This "simulation based on tacit knowledge" explanation seems to fit the great majority of mental imagery findings reported in the literature. It also fits the following obvious fact about images: They are *our* images and we can make them have very nearly any property we wish– and we generally make them have the properties we believe would actually obtain if we were to see the real situation. (Cases not covered by this simple explanation are discussed in detail elsewhere Pylyshyn, 2002, 2003b).

If this were the entire story of research on the picture theory of mental imagery, cognitive scientists might well have lost interest in the topic. What has kept it alive is that the intuitively appealing picture theory has been given new life by recent research in neuroscience. Although the facts are not uncontested, even within neuroscience, the research has been taken to show that even the most peripheral part of the visual cortex (area V1) is active during episodes of mental imagery. Combined with the fact that activity in the visual cortex in primates is known to be retinotopic (i.e., it maps patterns of retinal activity in a continuously transformed, or locally affine, manner), these findings suggests a possible neural implementation of the elusive internal display long favored by picture theorists. Many writers jumped to the conclusion that these results showed the existence of spatial patterns in the brain that underwrite a pictorial form of mental images (indeed some have referred to this evidence as finally providing "the resolution of the imagery debate" Kosslyn, 1994). But this is far from being the case. Even if the neuroscience evidence were not problematic, the arguments against a picture theory that had been discussed over the past 30 years (not to mention Locke's argument against Berkeley's claim that ideas are images) remain unanswered. The interpretation placed on the neuroscience evidence by picture theorists is *highly* problematic and when the evidence examined even cursorily it is found to provide no support at all for the picture theory of mental imagery.

While the finding that some part of the visual system is active in mental imagery is (if sustained) itself quite interesting, it tells us nothing about the *nature* and *form* of the representation underlying mental images; representations in vision and imagery could have exactly the same form without either being pictorial (they could, for example, both take the form of symbol structures). But more importantly, the argument from activity in the visual cortex during imagery ignores the very significant (indeed, decisive) differences between retinal/cortical "images" and mental images, a few of which are summarized below.

(1) The pattern of activity in visual cortex is in retinal coordinates and, like vision, is limited to a field of view not much more than a few degrees of visual angle. In contrast, mental images (as shown by both phenomenological and experimental observations) are in allocentric or environmental coordinates and are panoramic or even cycloramic (360 degrees) in breadth. Consequently, the retinotopically-mapped pattern of activity in visual cortex is totally inappropriate for underwriting mental imagery.

(2) The only topographical mappings of perceived space found in the visual cortex are two-dimensional. Yet mental images are clearly three-dimensional, and indeed all imagery phenomena (e.g., mental scanning, mental rotation) occur as readily in three dimensions as in two. This means that explaining the three-dimensional versions of these phenomena would require postulating a different mechanism and a different form of representation – one that itself could not take the form of a neural display since there are no known 3D neural displays that map space.

(3) Information in cortical patterns and in mental images is accessed and interpreted very differently. If you construct a mental image, the result, however much it might feel like a picture, does not have the signature properties of visual perception. For example, imagined patterns of lines in 2D do not automatically lead to a 3D interpretation or to reversals between ambiguous interpretations. Try the following example. Imagine a parallelogram. Now imagine an identical parallelogram directly below it. Connect each vertex of the top figure to the corresponding vertex of the bottom figure. What does it look like? The striking difference between an image and a display can be made clear if you now draw the figure and look at it.

Accessing information from a mental image is different in many other ways from accessing information from a visual scene. If you were to write a word on the board you could easily read the letters in any order. But you can't do that with an image of the word. Mental images are also not like retinal images in that they are not subject to Emmert's law. If you have an image on your retina (e.g., an after-image), and you look at some surface in the distance, the apparent size of the image varies with the distance of the surface: the further away it is the larger the apparent size of the retinal image. This is not true of a mental image, as it should be if the image were actually a retinotopic pattern of cortical activity (in fact our informal observations suggest that an image projected onto a surface appears to get smaller, rather than larger, as the surface is moved away).

(4) If the same cortical display provides input to both vision and imagery, both should connect with the motor system in the same way. Yet they do not: Reaching for an imagined object does not exhibit the signature properties that characterize reaching for a perceived object (Milner & Goodale, 1995), and many visuomotor phenomena, such as smooth pursuit, do not occur with imagined motion.

(5) Even more important is the fact that retinotopic patterns in visual cortex have yet to be interpreted, while mental images *are* the interpretation; there is every reason to believe that images cannot be further reinterpreted *visually*. Of course one can think about them and figure out what would happen if we did things to them, like rotate them or combine them with other patterns, but we can only do so when the combinations are easy to infer (e.g., from fragmental cues), not when they involve a clearly visual (re)perception.

(6) Clinical neurological findings, often cited by adherents of picture theory, provide little support for the claim that the pattern of cortical activation corresponds to the mental image. If mental imagery and vision used the same cortical display, it is hard to see why vision and imagery capacities are so radically dissoc iated: There are many reports of normal imagery in people who have a variety of visual deficits; and there are many reports of normal vision in people with little or no mental imagery. Indeed, virtually all the experimental results cited in support of the picture theory have been obtained with blind people, though they may not be accompanied by the experience of "seeing," except in the recently blinded.

The recent interest in mental imagery has not clarified the puzzles that have been around since the time when Locke and Berkeley argued about them 300 years ago. Moreover, the new picture theories have not dealt with the recent empirical arguments (as outlined in Pylyshyn, 2003b). Yet despite the problematic state of theoretical understanding of mental imagery, most psychologists continue to assume that representations underlying mental imagery are very different from ones underlying other forms of thought, and in particular that such representations are in some important sense "spatial" (for an alternative interpretation of the apparent "spatial" nature of mental images that does not attribute this character to properties of an inner display, see the discussion in Chapter 7 of Pylyshyn, 2003a). This refractoriness of theorizing about mental imagery to counterarguments is no doubt attributable to the almost irresistible grip that our subjective impressions have on our inclination to accept certain kinds of theories, a grip that in earlier years had burdened theorizing about the physical world as well.

References cited

Finke, R. A. (1980). Levels of Equivalence in Imagery and Perception. *Psychological Review, 87*, 113-132.

Galton, Francis. (1883/1973). *Inquiries into human faculty and its development*. New York: Dutton.

Kosslyn, S. M. (1994). *Image and Brain: The resolution of the imagery debate*. Cambridge. MA: MIT Press.

Kosslyn, S. M., Pinker, S., Smith, G., & Shwartz, S. P. (1979). On the demystification of mental imagery. *Behavioral and Brain Science, 2*, 535-581.

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

Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review, 88*, 16-45.

Pylyshyn, Z. W. (2002). Mental Imagery: In search of a theory. *Behavioral and Brain Sciences, 25*(2), 157-237.

Pylyshyn, Z. W. (2003a). Return of the Mental Image: Are there really pictures in the brain? *Trends in Cognitive Sciences, 7*(3), 113-118.

Pylyshyn, Z. W. (2003b). *Seeing and visualizing: It's not what you think*. Cambridge, MA: MIT Press/Bradford Books.

Shepard, R. N., & Chipman, S. (1970). Second-Order Isomorphism of Internal Representations: Shapes of States. *Cognitive Psychology, 1*, 1-17.