"The Strategic Eye": Another Look

Review Exchange

S.M. KOSSLYN
Department of Psychology, Harvard University, 830 William James Hall, 33 Kirkland Street, Cambridge, MA 02138, USA; E-mail: smk@wjh.harvard.edu; fax: 617-496-3122

I was delighted to read Mark Rollins' review and reflection on my book, *Image and Brain* (1994), for two reasons: First, this is a wonderful example of ways in which philosophers can help laboratory scientists. They not only can sweep away cobwebs of "conceptual confusion," but – at least as important – can suggest testable hypotheses. Rollins has done a bit of both in his review. Second, Rollins has deeply understood the central thrust of my project; I only wish I could explain it so clearly and succinctly myself. I am flattered that he has taken the time to understand it so well, and grateful for his kind words.

But of course, Rollins doesn't agree with all that I wrote, and argues that there are two major problems with my theory and approach. The first is that the "particulars of perceptual strategy might be false." The second is that the variability that arises from my emphasis on strategy "poses a prima facie threat to explanation." In addition, Rollins suggests another perspective on the issue of format (i.e., the type of code) used in imagery. Let me consider each topic in turn.

1. The Particulars Might Be False

This discussion focuses primarily on the Chambers and Reisberg (1985, 1992; Reisberg and Chambers, 1991) research, which purported to show that patterns in visual mental images – unlike patterns in pictures – cannot be reinterpreted. Although Chambers and Reisberg's position evolved over time, an essential point has been that during imagery there is no "raw material" that can be consulted after the fact; rather, the interpretation associated with a stimulus dictates and shapes the nature of the subsequent mental image of the stimulus. In my book, I offered two kinds of accounts for their findings. The first was that if patterns are sufficiently complex, all portions cannot be retained in a mental image long enough to reorganize them. I think that Rollins has a point in dismissing this view; this notion is not sufficient to explain why people can visualize a map and scan it without difficulty but do have difficulty mentally rotating a shape and "seeing" that it is a map of Texas.

The second argument was that people verbally code patterns, and those codings in turn dictate the form of subsequent images (I suggested a mechanism whereby this could occur, but the details are not pertinent here). Rollins' has three difficulties with this account: First, "verbal coding does not always interfere with the reinterpretation of images." In response, I first note that verbal coding per se is not

Minds and Machines **11:** 287–291, 2001.

always effective. My view is that when verbal coding is not effective as a primary means of storing information, it is also least likely to interfere with reinterpretation. This is, of course, an empirical issue. Second, "if verbal coding is taken to imply conceptualization" it isn't clear how this account applies to reversible figures such as Necker cubes. I don't mean to limit verbal coding to conceptualization; rather, verbal coding "locks in" a particular perceptual organization by labeling it. This process can occur with any pattern, ambiguous or not. Third, "we need to know why verbal coding does not interfere with reinterpretation in perception." It's not clear to me that it doesn't. If we measured the time to reinterpret a visible ambiguous figure that previously was coded in terms of the opposite reading, my bet is that people would in fact take longer to "see" the new read. (Again, this is an empirical question.) But Rollins asks how people can see a new interpretation at all. A major difference between imagery and perception is that an image fades rapidly once it is formed; in perception, the world is relatively stable, providing a kind of external long-term memory. Thus, the "raw material" that allows one to reinterpret the figure is readily available in perception, even after the first percept is shaped by one's prior label. In imagery, things are not so easy – the image is not only fading, but also is often incomplete.

Having said all of this, I nevertheless find myself in general accord with Rollins' key suggestion: Attentional strategies lie at the root of an account for why people cannot reinterpret ambiguous patterns in mental images. Rollins' perceptive ideas suggested to me that verbal coding has a major role (although perhaps not the only role) in directing attention, not simply in dictating perceptual organization. For example, if one verbally codes a part of a shape as "the top," one later will need to over-ride this coding when trying to "look at it from a different perspective." Rollins points out that this kind of hypothesis shifts the issue: We are no longer directed to focus on whether the format of visual perceptual representations differs from that of visual mental images, butrather are led to consider how attentional strategies may differ in the two cases. I agree. I also agree that the key question becomes "What is it about some of thereinterpretation tasks that interferes with re-deploying attention?". However, although I find Rollins' suggestions intriguing, this is an empirically tractable question, and only the results from future experiments can answer it.

2. Variability May Undercut Explanation

Rollins is quite right to highlight my emphasis of the role of strategy. This emphasis necessarily introduces a range of possible accounts for any given phenomenon, which is inconvenient for the scientist. I've gone even further along these lines since my 1994 book was written. In my earliest work (e.g., Kosslyn, 1980), I assumed that processing was lock-step, determined by algorithms as rigid as any of those in a word processor. In the subsequent neuroscience-oriented work, I argued that strategies are pervasive in the waythat Rollins correctly characterizes. In my

most recent work, I've continued to borrow a page from Simon (1969) and have emphasized the importance of task demands on the nature of processing. However, my current view is that the nature of the processes themselves depends on what is required in a task. That is, there is no fixed "library of subroutines" that are drawn upon in different combinations, as I assumed in 1994. Rather, I've adopted a much more constructivist view, with the notion that whole swaths of the brain can be dynamically reconfigured in response to task demands.

Perhaps a metaphor might help. Consider a ping pong ball resting on apicnic blanket. The position of the ball represents the processing occurring at any point in time (the entire configuration of brain activity). Under the blanket, randomly placed on the lawn, are nozzles (rather like those used in sprinkler systems) that shoot up blasts of air that vary in strength and frequency. As eachjolt of air occurs, the blanket is billowed upwards, and the ball rolls accordingly. Depending on which combinations of jets are active, and to what degree, the ballwill be in different places. This metaphor illustrates two key points: First, theexact same pattern of jets (inputs to the system) will have different effects on the ball when it starts in different places. And in fact, in recent research we havefound that the exact same task, mental rotation of block-like Shepard-Metzler stimuli, activates different parts of the brain if subjects had just previouslyperformed a task requiring mental rotation of pictures of hands than if they hadjust previously performed a task requiring mental rotation of block-like forms. Indeed, primary motor cortex is activated when subjects imagine rotating Shepard-Metzler figures if they had just previously imagined rotating hands, but not otherwise.

Second, the position of the ball is not a simple additive function of thestrength of the jets; depending on their timing, the same jets can have radically different effects on the ball. Bly and Kosslyn (1997) reviewed all of the then-available neuroimaging data on visual object recognition. There was a shockingamount of variability; most of the brain was activated in this task, when one considered the sum total of activated areas in the different studies. Although the studies superficially appeared similar, small task differences apparently made large differences in which parts of the brain were used.

This metaphor applies not only to the implementation of processes in the brain, but also to processing itself. Just as the ball can be any number of locations, new "processes" can arise in different tasks. For example, motor processes are qualitatively different from other types of processes; depending on the context, they may or may not be used. But such processes are not like subroutines that are called or not called. Rather, the precise nature of the process is constructed, depending on task depends (consider the way your hand molds to the shape of whatever it is you are holding). For instance, in mental rotation, one could imagine rotating a figure by grasping it, by turning an imaginary crank extending from it, by putting one's finger (or thumb) on a part and spinning itaround, and so on. And each of these different movements involves different processes and will activate at least some different parts of the brain. The "strategy" and "set of processes" that can be used

to perform the task is open-ended, and new variations can be invented on the spot (how about spinning thatform with a foot?).

However, we must also consider history: As the blanket sits over the pummeling jets for weeks on end, portions become threadbare and worn; such wear-and-tear is uneven, varying for different portions. When a portion becomes worn, less air is required to billow it to a given height. Thus, the system develops tendencies to respond in certain ways. This is what I had in mind with the idea of "cognitivestyle." Thus, I am agreeing with Rollins that "the *brain's visual* system might exhibit a variety of subpersonal strategies, even though they are not very useful."

I agree that life would have been simpler if the brain had been constructed along the lines of large computer programs, with clear modularity. But it just doesn't look that way to me.

3. Depiction

My view of depiction has little to do with resemblance, other than as abyproduct of the nature of the format. Space considerations prevent a detailed treatment here, but the key idea is that in depictive representations space is used to represent space. The space can be purely functional (as is true for an array in a computer) or actual space laid out on cortex (as occurs in topographically organized areas of the visual system). Patterns in this space correspond to portions of (typically the planar projection of) parts of objects such that the "distances" among these patterns in the representational space reflects the actualdistances among the corresponding parts of the objects. Thus, geometric ideas are at the core of this characterization of the depictive format (for details, see my1994 book). Rollins favors what he calls the *Recognition Theory*, which relies on the idea that "a picture is defined as an object that activates the same perceptual processes as the item or scene which the picture represents, although the picture is not itself that item or scene." This sounds reasonable to me, but is orthogonal to the format issue. The same could be said if the internal representation of an object or scene were entirely descriptive. Defining depiction, as Rollins does, in terms of "perceptual equivalence" seems to miss the point. It is clear earlier in the review that Rollins understands this distinction (between the "content' – the information conveyed – and the "format" – the type of code), so I can only conclude that I've failed to grasp the thrust of his position here.

In conclusion, Rollins has offered a stimulating and insightful reflection on my book. I'm sure that any author would be delighted to have such treatment from a reviewer!

References

Bly, B.M. and Kosslyn, S.M. (1997), 'Functional anatomy of object recognition in humans: Evidence from PET and fMRI'. *Current Opinion in Neurology*, 10, pp. 5–9.

Chambers D. and Reisberg, D. (1995), 'Can mental images be ambiguous?' *Journal of Experimental Psychology: Human Perception and Performance*. 11, pp. 317–328.

Chambers, D. and Reisberg, D. (1992), 'What an image depicts depends on what an image means.' *Cognitive Psychology* 24, pp. 145–174.

Kosslyn, S.M. (1980), *Image and Mind*. Cambridge, MA: Harvard University Press.

Kosslyn, S.M. (1994), *Image and Brain: The Resolution of the Imagery Debate*. Cambridge, MA: MIT Press.

Reisberg, D. and Chambers, D. (1991), 'Neither pictures nor propositions: what can we learn from a mental image?' *Canadian Journal of Psychology* 45 pp. 336–352.

Simon, H.A. (1969), *The Sciences of the Artificial*. Cambridge, MA: MIT Press.