

Cause and Effect Theories of Attention: The Role of Conceptual Metaphors

Diego Fernandez-Duque
University of Toronto

Mark L. Johnson
University of Oregon

Scientific concepts are defined by metaphors. These metaphors determine what attention is and what count as adequate explanations of the phenomenon. The authors analyze these metaphors within 3 types of attention theories: (a) “cause” theories, in which attention is presumed to modulate information processing (e.g., attention as a spotlight; attention as a limited resource); (b) “effect” theories, in which attention is considered to be a by-product of information processing (e.g., the competition metaphor); and (c) hybrid theories that combine cause and effect aspects (e.g., biased-competition models). The present analysis reveals the crucial role of metaphors in cognitive psychology, neuroscience, and the efforts of scientists to find a resolution to the classic problem of cause versus effect interpretations.

Everyone knows what attention is. (William James, 1890/1950, p. 261)

No one knows what attention is, and . . . there may even not be an “it” there to be known about (although of course there might be). (Pashler, 1998, p. 1)

The history of attention research is not just an ongoing debate about how to explain the phenomena of attention. It is equally a debate about what attention is. It is widely recognized in contemporary philosophy of science (Hanson, 1958; Hesse, 1966; Kuhn, 1962) that our scientific theories determine what we regard as data, how we individuate phenomena, and what the criteria for an adequate explanation should be. The phenomena are not just “given” in some theory-independent fashion. Rather, our theories and concepts partly determine what we will

take the relevant phenomena to be, and they thus determine what a good theory must account for.

This fact is quite evident in the field of attention research, in which even a cursory survey reveals that there is no general agreement about what a theory of attention ought to explain. Different theories have different views of what counts as attention. Some theories, for example, assume that there is a specific mechanism of attention, and they then ask how it works. Is it a cognitive system made of interacting subcomponents discretely localized in the brain? Or is attention a pool of resources we allocate to effortful tasks? Other theories are skeptical that there is such a “thing” as attention. They view attention instead as an epiphenomenon of the workings of multiple independent cognitive systems. Thus, whereas some theories conceptualize attention as a real “cause” of various cognitive events, others view it as a mere “effect” of multiple cognitive operations.

In this article, we argue that there is no way to identify attention independent of some theory of attention, and we argue that theories of attention are structured largely by conceptual metaphors. These metaphors provide the logic for our thinking and reasoning about the nature, structure, and processes of attention, and we cannot do without some set of metaphors, either in commonsense or in scientific models of the mind.

Diego Fernandez-Duque, Sunnybrook and Women’s College Health Sciences Centre, University of Toronto, Toronto, Ontario, Canada; Mark L. Johnson, Department of Philosophy, University of Oregon.

Jodie Baird, Patricia Carpenter, Jennifer Freyd, Morris Goldsmith, Asher Koriati, Mike Posner, and Albert Smith contributed helpful criticisms to the development of this article. During the preparation of this article, Diego Fernandez-Duque was supported by a postdoctoral fellowship from the Rotman Research Institute, by a grant from the Center for Consciousness Studies of the University of Arizona, and by a grant from the Heart and Stroke Foundation of Ontario.

Correspondence concerning this article should be addressed to Diego Fernandez-Duque, Sunnybrook and Women’s College Health Sciences Centre, University of Toronto, 2075 Bayview Avenue A421, Toronto, Ontario M4N 3M5, Canada. E-mail: diego@rotman-baycrest.on.ca

A conceptual metaphor is a mapping of entities, structures, and relations from one conceptual domain (the “source”) onto a different domain (the “target”; Johnson, 1987; Lakoff, 1987; Lakoff & Johnson, 1980). The entities and structures mapped from the source domain give rise to a parallel conceptual structure in the target domain. The cross-domain mapping actually constitutes a new conceptual structure in the target domain, rather than merely highlighting preexisting similarities between the source and target domains. For example, as shown subsequently, when we understand attention as a metaphorical spotlight that shines on various mental objects (i.e., representations), we use the entities and relations within the source domain of physical spotlights and the objects they illuminate to establish the nature of the mental operations that occur when we are attending to some stimulus or mental representation. Although the metaphors do not “create” the phenomena of attention, in the sense of causing processes to exist that did not exist before, they do constitute our conceptual understanding of phenomena associated with attention, and the metaphors give us the means for making sense of these phenomena.

Our strategy in this article is to show precisely how such metaphors define concepts of attention and guide scientific research. We do so by contrasting the very different metaphorical concepts of attention that underlie what are known in the scientific literature as “cause” versus “effect” theories.

Cause Versus Effect Theories of Attention

In cognitive psychology, *attention* typically is used to refer to a selective process by which attended information is processed more efficiently than nonattended information. This capacity to assign priority to one sensory stimulus over other equally salient stimuli is central to virtually all concepts of attention. But this raises a key question that has plagued all theories: What, precisely, is it that assigns this priority? What does the selecting of one stimulus over another?

In reaction to early behaviorist attempts to eliminate all “internal” systems, many cognitive psychologists since the 1950s have placed the locus of stimulus selection back into the “mind.” In place of stimulus-driven interpreta-

tions of behavior, selection began to be understood as an internal mechanism of attention. Whereas behaviorism and Gestalt psychology emphasized the properties of the external world as the site of stimulus control, cognitive psychologists argued that it was not possible to predict behavior merely on the basis of knowledge of the stimulus alone. This view is illustrated in Kahneman’s (1973) definition of *attention* as “a useful label for *internal* mechanisms that determine the significance of a stimulus” (p. 2). From this perspective, “the *mind* [italics added] is continually assigning priority to some sensory information over others, and this selection process makes a profound difference for both conscious experience and behavior” (Pashler, 1998, p. 2).

Most theories within this cognitive psychology tradition thus assume that attention “causes” changes over perception and other cognitive operations. According to cognitive theorists such as Broadbent (1958) and Kahneman (1973), selective attention to one stimulus over others frees people to orient their energies in certain directions, even in the face of stimuli that are more salient than the selected ones.

Causal theories can be contrasted with “effect” theories, which deny the existence of any central causal mechanism of attention. Whereas causal theories claim that attention “modulates” perception, effect theories conceptualize attention as an effect—a by-product—of the normal operation of various sensory and cognitive systems (James, 1890/1950; Johnston & Dark, 1986). For example, representations that rise to a certain level of activation are given temporary priority in the functioning of the organism, but there is thought to be no entity or substance bringing about this processing priority, and there is no central system that monitors the competition for processing. In the sections that follow, we examine the metaphorical bases of prototypical cause versus effect theories of attention, each with its own definite logic and corresponding knowledge structures.

Cause Theories of Attention

The Spotlight Metaphor

One of the best examples of “causal” theories is the attentional spotlight metaphor, in which an internal attentional system modulates the

processing of information carried out by sensory and cognitive systems. Scientists such as Hernández-Peón (cited in Watchel, 1967) have argued that

attention may be compared to a beam of light in which the central brilliant part represents the focus surrounded by a less intense fringe. Only the items located in the focus of attention are distinctly perceived whereas we are less aware of the objects located in the fringe of attention. (p. 418)

The spotlight metaphor consists of the conceptual mapping illustrated in Table 1.

Source-to-target mapping allows researchers to use their knowledge of the source domain (visual perception and devices of illumination) to construct a parallel knowledge of the target domain (attention). Such spotlight-based inferences have guided many years of attention research in cognitive psychology (Cave & Bichot, 1999; Fernandez-Duque & Johnson, 1999). In recent years, the internal structure and logic of the spotlight metaphor have also influenced research in cognitive neuroscience. Consider the following examples, in which neuroscientists explored the “physiological correlates of the ‘spotlight’ of visual attention” (Brefczynski & DeYoe, 1999, p. 370) by measuring the neuronal and hemodynamic response in areas of the visual cortex:

1. Several areas of the visual cortex have retinotopic maps of regions of the external world so that objects close or adjacent to each other in the world activate brain areas close or adjacent to each other in the visual cortex. If attention “sheds light” over sensory areas, then cuing attention to more central areas of the visual field should activate brain regions that map central locations, whereas cuing attention to a peripheral part of the visual field should increase blood flow in areas that map such a peripheral part of the field. Research testing

these metaphorical entailments has provided empirical support for such inferences (Brefczynski & DeYoe, 1999).

2. Because a spotlight moves in analog fashion, the target domain inference is that attention moves in an analog fashion. Therefore, the electrophysiological enhancement associated with the processing of attended stimuli should also move in an analog fashion, a prediction that has been tested and confirmed in the laboratory (Woodman & Luck, 1999).

3. The delay between the onset of a cue and the enhancement of the electrophysiological response at the cued location has been taken to be a measure of the amount of time necessary for the attentional spotlight to move to the cued location (Müller, Teder-Sälejärvi, & Hillyard, 1998). This prediction is based on models that conceptualize attention as a serial high-speed scanning mechanism (i.e., a spotlight) that moves from one location to another.

4. Another entailment of the spotlight metaphor is that the controlling agent is spatially distinct from the spotlight and from the field on which light is shed. In the target domain, this entails that the executive system is physically separate from the orienting system and from the sensory areas where attention is expressed. The concept of the executive system as defined by the spotlight metaphor led researchers to discover a network of cortical areas that participate in attentional control, moving attention from one location to another. Whether a stimulus is displayed at the attended location has no impact on the activation of these controlling areas. In other words, the perceptual systems that benefit from the attentional modulation appear to be separate from the neural system that controls the attentional spotlight and from the spotlight itself (Corbetta, Kincade, Ollinger, McAvoy, & Shul-

Table 1
The Attention Spotlight Metaphor

Source domain (spotlight)	Target domain (attention)
Spotlight mechanism	Orienting system (mechanism of attention)
Light	Attention
Agent who controls the spotlight	Executive system
Agent who sees	Awareness system
Potential field of vision	Representational space
Area illuminated by the spotlight	Attended representations

man, 2000; Hopfinger, Buonocuo, & Mangun, 2000; Martinez et al., 1999).

As these examples reveal, the spotlight metaphor has a highly articulated internal structure for the source domain (spotlight) that shapes, in very determinate ways, how scientists conceptualize the target domain (attention). The spotlight metaphor is thus a prime exemplar of causal conceptions of attention. The attentional “spotlight” is conceived as actually causing effects in cognitive processing.

The Limited Resource Metaphor

An equally important causal model is one built on a metaphor of attention as a limited resource that can be allocated by a general-purpose central processor in a graded fashion for the performance of different tasks. Resource models were most eloquently described in Kahneman’s influential book *Attention and Effort* (1973) and were further developed by researchers such as Norman and Bobrow (1975), Navon and Gopher (1979), and Hasher and Zacks (1979; for precursors of these ideas, see Moray, 1967). According to limited resource models such as that of Kahneman (1973), dual-task interference can be understood in terms of graded sharing of a single pool of limited mental resources (or capacity). This is a “causal” theory of attention, because the attentional resource has a modulatory effect on information processing.

Limited resource models have been, and continue to be, used to explain many psychological phenomena, such as dual-task interference (Christie & Klein, 1996), automaticity (Norman & Bobrow, 1975; Schneider & Shiffrin, 1977), priming (Posner & Tudela, 1997), and mental rotation (Carpenter, Just, Keller, Eddy, & Thulborn, 1999). They underlie theories of cognitive development (Case, 1985; Craik & Byrd, 1982; Harnishfeger, 1995), neuropsychological deficits (Schwartz et al., 1999), brain activity (Posner & Tudela, 1997), awareness (Baars, 1997; Dennet & Kinsbourne, 1992; Farah, 1994; Posner, 1994), and emotion (Posner & Rothbart, 1986).

The conceptual mapping of the limited resource metaphor is shown in Table 2. The concept of a limited resource plays a crucial role in scientific research only because psychologists hold certain metaphor-based assumptions about

Table 2
The Attention as a Limited Resource Metaphor

Source domain (resources)	Target domain (attention)
Valuable resource	Attention (capacity–arousal)
Quantity of resource	Amount of available attention
Allocation of resource	Allocation of attention to tasks
Budget	Task strategy for allocating attention
Controller of the resource	Executive system

how the mind works: assumptions about the existence of information processing, about a cognitive machinery that needs a source of energy for its proper functioning, about energy that can be manipulated and allocated, about the existence of a separate executive system that allocates the resource, and so forth.

It is important to emphasize that attention is not literally a substance (like water, soil, or gasoline) that can be quantified and parceled out. Neither is there a homunculus whose job it is to monitor this putative “scarce” and “valuable” resource. In other words, researchers derive implications for testing within attention experiments by using the knowledge of the source domain to construct a corresponding understanding of the target domain of attention. To get an idea of how this works, consider some of our shared basic knowledge of the source domain of allocation of limited, valuable commodities. In the source domain, there is (a) a *limited commodity* (e.g., some substance or object, or its surrogate, money) that is (b) *controlled* by someone who allocates it where he or she deems it most necessary, and (c) does so in a flexible way, *applying resources* to different goals in a *graded fashion*. (d) The *amount of resource needed varies* for each task or product, and (e) *affects its quality*.

When basic source domain knowledge of this sort is then applied, through cross-domain mapping, to our conceptualization of the target domain, we get a series of corresponding knowledge claims (or entailments) about attention. Therefore, for each piece of shared knowledge about the source domain (a through e above), there is a corresponding piece of knowledge about aspects of the target domain (a’ through e’ below). Thus, in the target domain:

(a') "attention has been used to refer to . . . all aspects of cognition having to do with *limited resources* [italics added] or capacity" (Shiffrin, 1988, p. 739).

(b') Such attentional resources are *controlled by an executive system*, or general-purpose central processor, at the very top of the decision-making tree, in charge of distributing the attentional resources: "[Attention] means a laborious process, whereby processing resources are *voluntarily allocated* [italics added] to a particular task, or activity at the expense of other tasks and activities" (Umiltà, 1988, p. 175).

(c') "The *allocation* of attention is a matter of *degree* Attention is . . . *controllable*. It can be *allocated* to facilitate the processing of selected perceptual units or the execution of selected units of performance" [italics added] (Kahneman, 1973, p. 201).

(d') "The *amount* of attention or effort exerted at any time *depends primarily on the demands* of current activities" [italics added] (Kahneman, 1973, p. 201).

(e') Finally, "allocating more attention to a given task enhances performance" (Pashler, 1998, p. 3).

The limited resource metaphor has provided a rich description of attention, and the knowledge it entails has influenced theory and guided research. For example, the models predict that, as the system becomes overloaded, performance will degrade steadily (Norman & Bobrow, 1975). When a task requires more resources than are available, decreased performance should be expected: "Once the capacity limit is exceeded, selection of the information to be processed will be required" (Lavie, 1995, p. 452). Task interference would be expected to be nonspecific, because all cognitive processes tap into the same unitary resource. Even if the tasks do not share any mechanisms of perception or response, interference will occur whenever the system's capacity is overloaded.

One of the chief advantages of the limited resource metaphor is that it allows us to model the graded allocation of attention to different tasks. In contrast, the spotlight metaphor has difficulty explaining such phenomena, because a spotlight does not allow the person who controls it to split the beam or to reduce its intensity in a way that would save the attentional light so that it could be used to illuminate another location. Whereas the spotlight metaphor entails

that the intensity of the light is invariant, the limited resource metaphor allows a reduction of intensity at the attended location when a secondary task is added.

The limited resource metaphor has motivated many studies with results that eventually raised serious challenges to models assuming a unitary resource (i.e., a single pool of attentional resources). For example, "effortful" tasks sometimes do not interfere with each other (perfect time sharing; Navon, 1984), so in some cases increasing the difficulty of one task is not followed by a cost in the performance of the secondary effortful task (difficulty insensitivity). Moreover, sometimes changes in the modality of one of the tasks affect performance of the other task, even in the absence of difficulty changes (structural alteration effect; Wickens, 1984). These challenges from empirical studies have led to a revision of the original single capacity (or single "pool") models in favor of a multiple resource model (Navon & Gopher, 1979; Pashler, 1998; Wickens, 1984). These models propose that interference between two tasks occurs only when the same reservoir of limited resources is accessed, but not otherwise.

Causal versions of the limited resource metaphor have also been criticized for attributing agency to the central conscious control mechanism. In the words of Pashler (1998), "whether or not attention is allocated to a stimulus is usually thought to depend on a voluntary act of will; in the metaphysics of folk psychology, this ultimately depends on the choice made by the self" (p. 2). Without an account of how the executive system works to decide the attentional resource allocation, the limited resource metaphor leads to a regress of homunculus-like control mechanisms.

Effect Theories of Attention: Competition Metaphors

To avoid the recurring homunculus problems that plague most causal models, many researchers, spurred by recent developments in cognitive neuroscience, have adopted "effect" theories that explain attention phenomena as by-products of information processing among multiple systems. It might be said that effect theories attempt to explain attention away, because they deny the existence of any causal force or mechanism lying behind the alleged

phenomena of attention: "Attention is not a high-speed mental spotlight that scans each item in the visual field. Rather, attention is an emergent property of slow, competitive interactions that work in parallel across the visual field" (Desimone & Duncan, 1995, p. 217). Such models claim that no executive system is required. Instead, perceptual objects are seen as vying for limited processing resources; that is, they compete for activation in areas of perceptual processing: "Objects in the visual field *compete* for processing in several cortical areas" (Rees, Frackowiak, & Frith, 1997, p. 835). If effect theories could successfully account for the full range of phenomena typically associated with attention, they would eliminate the need for a central executive and thus avoid homunculus problems.

The principal effect theories are based on what we label "competition" models, defined by the metaphorical mapping illustrated in Table 3. The competition metaphor is very different from any metaphor of the causal sort discussed earlier. Although the mapping does involve the idea of competition for scarce resources, the nature of these "resources" is markedly different from the concept articulated in the limited resource metaphor, wherein attention is a causally efficacious substance-like reality that modulates cognitive processes. By contrast, according to the competition metaphor, what we call "attention" is an emergent property or epiphenomenon of the fact that when various stimulus representations "compete" for processing "resources," one of them will "win." It is this "winning" that we validate when we say that we are attending to that particular stimulus or representation, but there is no "attention" that is being "won."

Effect theories employing this particular metaphor-based logic have prompted scientists to

ask a very different set of research questions than those generated by causal theory metaphors. For example, the competition metaphor naturally leads scientists to ask the following:

1. *What is it that "competes"?* Effect theories that use the competition metaphor have answered this question in various ways. One can say that *stimuli* are what compete for processing resources. Because stimuli are stimuli only relative to the perceptual and motor capacities of organisms, however, it might be more accurate to say that it is *stimulus representations* that are actually in competition (Làdavias, Pretronio, & Umiltà, 1990). But such representations exist for the organism in neuronal patterns of activation, which suggests that the proper subject of competition would be *neuronal units* (Cohen, Romero, Servan-Schreiber, & Farah, 1994). Finally, by extension, one can speak of entire *hemispheres* competing, insofar as these hemispheres are collections of systems of neuronal units (Kinsbourne, 1977).

2. *What is being competed for?* Most theories assume that what is competed for is neuronal activation, namely, a representation in a neuronal cluster. Thus, to say that a person is attending to *x* is to say that the neuronal unit or units involved in representing *x* are activated to a level "above" other competing units.

3. *What "decides" who wins the competition?* This is, of course, the big question for any competition theory, because the principal motive of such effect theories is to avoid the need to posit and explain the workings of an executive system. But if there is no control mechanism to determine which competing unit "wins," then what sense can one possibly make of winning a competition? Within this framework, the answer must be that "nobody" decides who wins; rather, winning is merely a consequence of rising above a thresh-

Table 3
The Competition Metaphor

Source domain (resources)	Target domain (attention)
Competing individuals	Stimuli-mental representations
Valuable resource	Mental resources-neuronal receptive fields
Goal: securing limited resources necessary for survival	Goal: securing resources for mental processing and aware perception
Competition for resources	Competition for neuronal activation
Survival of individual	Activation above a specified threshold

old that typically involves mutual inhibitory connections.

4. *What is the result of "winning"?* The answer would seem to be "awareness." In this metaphor, "being attended to" is simply the name we give to whatever "wins" the competition for processing. Stimuli that are not strong enough are eliminated, whereas strong stimuli succeed in gathering enough processing resources to become conscious (Dennet & Kinsbourne, 1992). In this view, the competitive process gives rise to awareness, which is stable but also evolving (i.e., a stream of consciousness), and attention and awareness are not supramodal systems but rather emergent processes of brain activity. Awareness is not a natural kind but a collection of successful inputs.

An example of this reasoning comes from the literature on binocular rivalry. When different images are presented to the two eyes, perception alternates such that each image is visible for a few seconds at a time. This binocular rivalry phenomenon is usually understood as a competition between neuronal responses, so at a given time one representation is dominant and the other is suppressed. Current views of binocular rivalry argue that "the neural representations of the two stimuli *compete* [italics added] for visual awareness" (Logothetis, Leopold, & Sheinberg, 1996, p. 621). Binocular rivalry is a process in which "each region of the retinal image in one eye seems to fight with the corresponding region in the other eye" (Wolfe, 1996, p. 588).

The entailments generated by the competition metaphor give rise to specific predictions and theoretical explanations that are in sharp contrast to those of other metaphors. For example, patients with right parietal lesions and hemispatial neglect respond very slowly when a target appears in an uncued location. Spotlight models explain this finding as a deficit in disengaging attention and "moving" it to the new location. In contrast, competition models argue for competitive weights that favor stimuli represented in the undamaged areas over stimuli represented in the lesioned areas. According to this view, lesions to the parietal cortex lead to a deficit in the processing of spatial information (which can be redescribed as an attention problem), but the parietal cortex is not a specialized attentional system.

Along the same lines, connectionist models of attention posit that "units" compete for resources (Cohen et al., 1994). These models assume the existence of "perception" units that detect the spatial location of the stimulus and send their output to a second layer of units. Units in the second layer are grouped in two assemblies, one for each visual field, projecting inhibitory connections toward the other assembly. These inhibitory connections put units of one side in competition with units of the other side. Unilateral "lesions" lead to a disengagement deficit, whereas bilateral lesions do not, because the interhemispheric balance is left intact. This pattern of predictions contrasts sharply with the prediction of the spotlight metaphor that bilateral lesions to the attentional network should produce a disengagement deficit at least as severe as that produced by unilateral lesions.

Some theorists have further argued that rival hemispheres have opposing orienting vectors that compete with each other (Kinsbourne, 1977). Lesions to one hemisphere lead to an imbalance in the opposite hemisphere. Thus, presenting a competing stimulus to the unaffected side exaggerates the bias, leading to an "extinction" of the stimulus represented by the impaired side. Because of the lack of competition from the lesioned hemisphere, perception in the unaffected side is better than normal (Lādavas et al., 1990; Seyal, Ro, & Rafal, 1995).

These examples illustrate the fact that effect theories raise a set of questions about the nature and mechanisms of attention that are strikingly different from the questions raised by causal theories. If attention is a by-product of a competition for limited information-processing resources, scientists should emphasize the description of stimuli and their processing rather than focusing on the properties of a central manager that effects change.

Critiques of Effect Theories and the Emergence of the Biased Competition Model

The move to stay away from top-down influences is welcomed by researchers eager to avoid the homunculus problems associated with causal theories of attention. However, the radical conclusion that attention is something we

can do without raises serious challenges to any effect theory. The chief problem is to explain how, in the absence of causal top-down control, an organism can free itself from the control of the most salient stimuli. How can a system in which attention is merely a by-product of bottom-up processing in independent, interacting units assign priority to less salient stimuli, which humans seem quite capable of doing? In response to this type of criticism, there has arisen a variation on the competition metaphor known as the *biased competition model*, which, although emphasizing bottom-up competition, also acknowledges the existence of top-down biases.

The biased competition model proposes that inputs compete for neuronal receptive fields, which represent a limited resource: "Receptive fields can be viewed as a critical visual processing resource, for which objects in the visual field must compete" (Desimone & Duncan, 1995, p. 197). When the target is highly discriminable, it has no problem securing its receptive fields. In contrast, when the target is surrounded by distractors that are similar to it, there is more competition, and the neuronal response to the target is reduced. Other factors, such as novelty and general significance, also influence which stimuli become most salient. Novel stimuli have "a larger neural signal in the visual cortex, giving them a *competitive advantage* in gaining control over attentional and orienting systems" (Desimone & Duncan, 1995, p. 202). Stimuli that fail to secure representation within a sufficient number of receptive fields fail to exist at the conscious level. In this sense, the biased competition model is an effect model that regards attention as a by-product of information processing.

However, the model is closer to cause theories insofar as it allows bottom-up processes to be modulated by top-down factors such as attentional templates or working memory representations. By holding active the feature or spatial location of the object of interest, an individual can bias competition in favor of one stimulus over another. Holding a target in mind is thought to activate target-sensitive neurons and to secure those neuronal receptive fields for when the target is finally presented. In other words, "multiple objects within a cell's receptive field compete for control over the cell's response, and attentional inputs favor relevant

objects. In this view, 'top-down' attentional influences can overrule 'bottom-up' stimulus-driven competition among stimuli in ventral [lobe] areas" (DeWeerd, Peralta, Desimone, & Ungerleider, 1999, p. 753).

By incorporating cause and effect aspects of attentional theories, and by grounding attention in biological constructs, the biased competition model accounts for a multitude of findings (Behrmann & Haimson, 1999; O'Craven, Downing, & Kanwisher, 1999; Rees et al., 1997; Treue & Martinez, 1999).¹ For example, when a good stimulus (i.e., one for which a neuron is very responsive) is presented in close proximity to a poor stimulus (i.e., one for which a neuron is unresponsive), the presence of the poor stimulus leads to a reduction in the neuronal response. According to the biased competition model, when the good and bad stimuli appear together they activate neurons that compete with each other, and attending to a stimulus biases this competition in the direction of the attended item (Reynolds, Chelazzi, & Desimone, 1999). If attention is directed to the poor stimulus, the response is further suppressed, but if attention is directed to the good stimulus, the response is enhanced. Consistent with this view, preparatory attention increases the neuronal firing of cells that respond preferentially to the stimulus, and functional magnetic resonance imaging studies reveal activation of brain areas that code for the expected features (Chawla, Rees, & Friston, 1999; Kastner, DeWeerd, Desimone, & Ungerleider, 1998).

Proponents of the biased competition model claim that the theory is safe from the homunculus problem, arguing that attentional processes are an emergent property of the stimulus information processing carried out by multiple cognitive systems (Desimone & Duncan, 1995). Attention is not localized in a single part of the brain, and neither are the top-down modulatory mechanisms. Instead, the concept of the central

¹ Although biological grounding has most frequently been explored by proponents of the competition models, it has also been successfully applied to limited resource models, in which it has been proposed that "just as *physical energy* systems require resources, so do neural energy systems. It is the consumption of different types of resources that may be indexed by various metabolically based neuroimaging methods, including [functional magnetic resonance imaging]. Specifically, greater task demand translates into greater resource demand" (Carpenter et al., 1999, p. 18).

executive is replaced by domain-specific feedback loops that bias the information processing of upcoming stimuli.

However, the partition of top-down effects into domain-specific feedback loops raises the question of how those loops are bound again into a coherent unit in a way that does not merely reinstate some form of homunculus. The question of who or what decides the allocation of top-down modulations remains unsolved, even if the homunculus is replaced by a multitude of homunculi. This problem was already evident to Broadbent (1958) and the other cognitive psychologists who, in their reaction to behaviorism, brought “mental” processes to the forefront of psychology research. The problem is that explaining attention away by linking it to other cognitive systems explains attention as a by-product (attention as an effect) but still does not account for attention as a top-down modulation (attention as a cause). In other words, “attention as a cause” cannot be reduced to “attention as an effect,” because the top-down modulation, by definition, should be somewhat independent from the bottom-up factors. Otherwise, behavior would be entirely stimulus driven.

Once the limited resource metaphor and the competition metaphor have been spelled out, the tension between attention as a cause and attention as an effect becomes evident. Causal theories of attention, such as Kahneman’s, emphasize the existence of an internal agent in charge of distributing attentional resources. In contrast, effect theories emphasize the existence of competing stimuli that aim to secure a place in representational space. The biased competition model combines the two features, explicitly acknowledging the competition among stimuli but also implicitly assuming the existence of multiple homunculi. The homunculus problem remains.

The Metaphorical Structure of Scientific Reasoning

We have been pursuing the hypothesis that conceptual metaphors, along with the cognitive models they support, reside at the heart of scientific reasoning. We illustrated the pervasiveness of such constitutive metaphors in the field of attention research, focusing specifically on versions of the spotlight, limited resource, and

competition metaphors. The core of our argument is that scientists’ ability to conceptualize, reason about, and study experimentally the phenomena of attention depends on the structure and logic of a specific set of metaphors. We showed how inferences made about source domains (e.g., spotlights or limited resources) constrain the possible inferences made about attention.

Because the ontologies and inference patterns are different for each metaphor, it is fair to ask whether there is any unified, comprehensive concept of attention or, rather, whether *attention* is a vague label that scientists from different orientations use to refer to distinct, disparate domains of cognitive activity. On the one hand, our analysis provides evidence that the scientific concepts of attention are irreducibly metaphorical and that there are multiple metaphors, each with its own structure and set of inferences. As has been shown, in cases in which the ontologies (i.e., the entities and properties specified by the distinct source domains) of different metaphors are inconsistent, each metaphor defines attention differently.

On the other hand, the different metaphors do not simply specify radically different concepts, as though they were each defining a completely different set of unrelated phenomena. Rather, the metaphors do seem to gravitate around a recurring and relatively stable set of phenomena that researchers tend to think of as unified in some as yet unknown way. Indeed, it is possible to list features that cognitive psychologists have repeatedly attributed to attention, such as the following: (a) Attention involves some form of stimulus selection; (b) attention enhances processes in the area that is the focus of attention; and (c) attention facilitates access to awareness (see Fernandez-Duque & Johnson, 1999, for a more complete list of typical features attributed to attention). One might think of these shared features as forming a literal notion of attention that any attention theory has to account for. But note that such a literal notion is far too sparse in structure to guide research by itself, even though it constrains the space of adequate source domains. It is the different metaphors that actually shape the research programs. However, brought together by overlapping research paradigms and a history of shared questions, scientists have tended to see themselves as all studying “attention.”

So do the phenomena of attention exist independent of the metaphors? The answer generated by our analysis is both yes and no. Yes, there are indeed observable behaviors that occur regardless of scientists' conceptualization of them. However, the answer is no once we acknowledge the constitutive role of metaphor in defining what counts as attention and how attentional phenomena are individuated and described. Thus, there are no theory-independent, metaphor-independent phenomena of attention as such. This can be seen clearly in the way the spotlight metaphor establishes a new set of structural relations that become part of the target domain. In the source domain—the domain of vision—there is a *spotlight* that gives off *light*, permitting *someone* to *see* something. The first thing to note is that none of these entities or acts exist independently in the target domain. Although there are metaphorical counterparts in the target domain for each of these aspects, there are no literal similarities between the entities in the source domain and those in the target. There is not, literally, light in the target domain. There is not, literally, a person who sees in the target domain. Instead, scientists use their knowledge of the source domain entities and operations to develop a parallel knowledge structure for the target domain (attention).

The metaphor does not create new physical entities connected with the phenomena of attention, but it does direct researchers as to what the phenomena are, thereby creating entities in their understanding and conceptualization of attention. Scientists' reasoning about the workings of attention is shaped by their knowledge of the source domain structure. As has been shown, given the cross-domain mapping, we expect to find certain elements in the target domain that follow from what we know about the corresponding entities in the source domain. Within the spotlight metaphor, we understand, conceptualize, and reason about attention through what we know about the way light illuminates objects in visual fields.

A typical skeptical response to such strong constitutive claims is that scientists are aware of the limitations of their metaphors, and so they do not, and need not, take their metaphors seriously as being constitutive of their knowledge. According to this view, researchers explicitly acknowledge that the use of a metaphor does not entail that the object of study and its coun-

terpart share all of their features. LaBerge (1995) reminded us that “one problem. . . [of the spotlight metaphor]. . . is that properties of the device may be inappropriately included in the model, along with features that do usefully describe how attention operates” (p. 38).

Obviously, no researcher believes that attention is literally a spotlight. However, as we have just argued, scientists cannot conduct their research independent of such metaphors. Scientists choose research topics, identify phenomena that need to be explained, frame their research hypotheses, understand the phenomena, and interpret their findings under the guidance of the spotlight metaphor, the limited resource metaphor, the biased competition metaphor, or some other conceptual metaphor. The logic of our understanding of attention is the logic of such metaphors.

The crucial point is the following: It is not as though scientists simply know what attention is, independent of their metaphorical understanding of it. There may be some structure in the target domain before the metaphorical conceptualization of it, and there are observed behaviors to be explained. However, we do not have any idea about how to circumscribe and to think about those phenomena—how to carve them up and reason about them—without guidance from metaphor mapping.

This point is overlooked by those who see no essential role for metaphor in scientific reasoning. They suppose that, without metaphors, we know exactly what the phenomena are that we want to study, how to individuate the relevant entities, and what needs to be explained about these entities. But this is simply not the case. Try, for example, to think in any serious way, beyond the most thin and elementary notion, about attention without the spotlight metaphor, the limited resource metaphor, or some other set of metaphors. It cannot be done, at least not in any way that really furthers our knowledge of the subject. Any literal concept one might try to frame about attention will be too underspecified to generate people's actual understanding of and reasoning about attention. Thus, although metaphors do not bring attention into existence, they are constitutive of our very concept of attention and of what we can know about it. This is the sense in which they are constitutive, indispensable, and irreducible.

Metaphors do not act in isolation; they are not the whole story of scientific understanding. Instead, they are immersed within an ecology of scientific and social practices. Technical advances, cultural influences, and empirical data all work together to constrain how metaphors are used. For example, empirical data have influenced the evolution of scientific metaphors, in some cases confirming and in others challenging the predictions entailed by a given metaphor. Mathematical and computational simulations, in their quest for internal consistency, have imposed their own constraints on how metaphors are interpreted. Science is influenced by many forces other than metaphor, and scientific theories are not an exact reiteration of their underlying metaphors.

Sometimes the existence of other forces is mistakenly taken as evidence against the constitutive role of metaphors in science. Critics argue that although metaphors are important for the process of discovery, mature theories will use only mathematics and formal logic, thereby transcending metaphor. Someone who thinks of psychology as an immature science might argue that, should psychology finally come of age, all of this metaphorical thinking would be replaced with rigorous neurocomputational and mathematical models. On this view, the use of metaphor in science is thought to be merely an intermediate stage on the way to literal scientific truth about the mind. But this is false for several reasons.

First, such a literalist view ignores the fact that conceptual metaphors and other structures of imagination are characteristic of abstract conceptualization and reasoning in general (Gibbs, 1994). Over the past two decades, there has emerged in the cognitive sciences a substantial body of research showing the central and indispensable role of conceptual metaphor in many aspects of abstract conceptualization and reasoning (Lakoff & Johnson, 1999). Thus, it should be no surprise that the structures of conceptualization and reasoning in the sciences will use the same cognitive mechanisms used in ordinary, everyday thinking. And those cognitive mechanisms include metaphors that structure our abstract concepts.

Second, the idea that mature computational and mathematical theories of mind will not involve metaphor is mistaken, because mathematics is itself a massive interlocking system of

metaphors (Lakoff & Núñez, 2000). And, finally, even if it were the case (which it is not) that mathematics consisted only of arbitrary symbols manipulated according to strictly algorithmic operations, any mathematics would be meaningless unless it could be interpreted by means of mathematical ideas. Mathematical models would need to be interpreted as applying to models of mind. And it is here that metaphor would once again be relevant, as is evident in any actual model of the mind that uses mathematics.

The question of whether a scientific theory is structured by metaphors and other imaginative devices is an empirical one. It cannot be settled adequately by arguing from a priori philosophical assumptions about meaning, concepts, or rationality. It requires looking to see whether the concepts themselves, along with the inferences they support, are explainable through conceptual metaphors. In this article, we have given examples of one kind of analysis that is relevant to such an inquiry. Metaphor in scientific reasoning is a fact. It is the very means by which scientists pursue their hypotheses and make sense of things. It defines the relevant phenomena, generates inferences and directions for research, and forms an important basis for scientific knowledge.

References

- Baars, B. J. (1997). *In the theater of consciousness: The workspace of the mind*. New York: Oxford University Press.
- Behrmann, M., & Haimson, C. (1999). The cognitive neuroscience of visual attention. *Current Opinion in Neurobiology*, 9, 158–163.
- Brefczynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the 'spotlight' of visual attention. *Nature Neuroscience*, 2, 370–374.
- Broadbent, D. (1958). *Perception and communication*. New York: Pergamon Press.
- Carpenter, P. A., Just, M. A., Keller, T. A., Eddy, W., & Thulborn, K. (1999). Graded functional activation in the visuospatial system with the amount of task demand. *Journal of Cognitive Neuroscience*, 11, 9–24.
- Case, R. (1985). *Intellectual development: Birth to adulthood*. New York: Academic Press.
- Cave, K. R., & Bichot, N. P. (1999). Visuospatial attention: Beyond a spotlight model. *Psychonomic Bulletin & Review*, 6, 204–223.
- Chawla, D., Rees, G., & Friston, K. J. (1999). The biological bases of attentional modulation in the

- extrastriate visual cortex. *Nature Neuroscience*, 2, 671–676.
- Christie, J., & Klein, R. M. (1996). Assessing the evidence for novel pop-out. *Journal of Experimental Psychology: General*, 125, 201–207.
- Cohen, J. D., Romero, R. D., Servan-Schreiber, D., & Farah, M. J. (1994). Mechanisms of spatial attention: The relation of macrostructure to microstructure in parietal neglect. *Journal of Cognitive Neuroscience*, 6, 377–383.
- Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary attention is dissociated from target detection in the human posterior parietal cortex. *Nature Neuroscience*, 3, 292–297.
- Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik & S. Trehub (Eds.), *Aging and cognitive processes* (pp. 191–211). New York: Plenum.
- Dennet, D. C., & Kinsbourne, M. (1992). Time and the observer: The where and when of consciousness in the brain. *Behavioral and Brain Sciences*, 15, 183–247.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193–222.
- DeWeerd, P., Peralta, M. R., Desimone, R., & Ungerleider, L. G. (1999). Loss of attentional stimulus selection after extrastriate cortical lesions in macaques. *Nature Neuroscience*, 2, 753–758.
- Farah, M. J. (1994). Neuropsychological inference with an interactive brain: A critique of the “locality” assumption. *Behavioral and Brain Sciences*, 17, 43–104.
- Fernandez-Duque, D., & Johnson, M. L. (1999). Attention metaphors: How metaphors guide the cognitive psychology of attention. *Cognitive Science*, 23, 83–116.
- Gibbs, R. W. (1994). *The poetics of mind: Figurative thought, language, and understanding*. New York: Cambridge University Press.
- Hanson, N. R. (1958). *The patterns of discovery*. Cambridge, England: Cambridge University Press.
- Harnishfeger, K. K. (1995). The development of cognitive inhibition: Theories, definitions, and research evidence. In F. N. Dempster & C. J. Brainerd (Eds.), *New perspectives on interference and inhibition in cognition* (pp. 175–204). San Diego, CA: Academic Press.
- Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 108, 356–388.
- Hesse, M. B. (1966). *The explanatory function of metaphor*. Notre Dame, IN: University of Notre Dame Press.
- Hopfinger, J. B., Buonocore, M. H., & Mangun, G. R. (2000). The neural mechanisms of top-down attentional control. *Nature Neuroscience*, 3, 284–291.
- James, W. (1950). *The principles of psychology*. New York: Dover. (Original work published 1890)
- Johnson, M. L. (1987). *The body in the mind*. Chicago: University of Chicago Press.
- Johnston, W. A., & Dark, V. J. (1986). Selective attention. *Annual Review of Psychology*, 37, 43–75.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.
- Kastner, S., DeWeerd, P., Desimone, R., & Ungerleider, L. G. (1998). Mechanisms of directed attention in the extrastriate cortex as revealed by functional MRI. *Science*, 282, 108–111.
- Kinsbourne, M. (1977). Hemineglect and hemisphere rivalry. *Advances in Neurology*, 18, 41–49.
- Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- LaBerge, D. (1995). *Attentional processing: The brain's art of mindfulness*. Cambridge, MA: Harvard University Press.
- Làdavas, E., Pretronicio, A., & Umiltà, C. (1990). The deployment of visual attention in the intact field of hemineglect patients. *Cortex*, 26, 307–317.
- Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- Lakoff, G., & Johnson, M. L. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Lakoff, G., & Johnson, M. L. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. New York: Basic Books.
- Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York: Basic Books.
- Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 451–468.
- Logothetis, N. K., Leopold, D. A., & Sheinberg, D. L. (1996). What is rivaling during binocular rivalry? *Nature*, 380, 621–624.
- Martinez, A., Anllo-Vento, L., Sereno, M. I., Frank, L. R., Buxton, R. B., Dubowitz, D. J., et al. (1999). Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nature Neuroscience*, 2, 364–369.
- Moray, N. (1967). Where is capacity limited? A survey and a model. *Acta Psychologica*, 27, 84–92.
- Müller, M. M., Teder-Sälejärvi, W., & Hillyard, S. A. (1998). The time course of cortical facilitation during cued shifts of spatial attention. *Nature Neuroscience*, 1, 631–634.
- Navon, D. (1984). Attention division or attention sharing. In M. I. Posner & O. Marin (Eds.), *Atten-*

- tion and performance XI (pp. 133–146). Hillsdale, NJ: Erlbaum.
- Navon, D., & Gopher, D. (1979). On the economy of the human processing system. *Psychological Review*, 86, 214–255.
- Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, 7, 44–64.
- O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature*, 401, 584–587.
- Pashler, H. E. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.
- Posner, M. I. (1994). Attention: The mechanisms of consciousness. *Proceedings of the National Academy of Sciences*, 91, 7398–7402.
- Posner, M. I., & Rothbart, M. K. (1986). The concept of energy in psychological theory. In G. R. Hockey, A. Gaillard, & M. G. Coles (Eds.), *Energetics and human information processing* (pp. 23–40). Dordrecht, the Netherlands: Martinus Nijhoff.
- Posner, M. I., & Tudela, P. (1997). Imaging resources. *Biological Psychology*, 45, 95–107.
- Rees, G., Frackowiak, R., & Frith, C. (1997). Two modulatory effects of attention that mediate object categorization in human cortex. *Science*, 275, 835–838.
- Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *Journal of Neuroscience*, 19, 1736–1753.
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84, 1–66.
- Schwartz, M. F., Buxbaum, L. J., Montgomery, M. W., Fitzpatrick-DeSalme, E., Hart, T., Ferraro, M., et al. (1999). Naturalistic action production following right hemisphere stroke. *Neuropsychologia*, 37, 51–66.
- Seyal, M., Ro, T., & Rafal, R. (1995). Increased sensitivity to ipsilateral cutaneous stimuli following transcranial magnetic stimulation of the parietal lobe. *Annals of Neurology*, 38, 264–267.
- Shiffrin, R. M. (1988). Attention. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology: Learning and cognition* (pp. 739–811). New York: Wiley.
- Treue, S., & Martinez, J. C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, 399, 575–579.
- Umiltà, C. (1988). Orienting of attention. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (pp. 175–192). Amsterdam: Elsevier.
- Watchel, P. L. (1967). Conceptions of broad and narrow attention. *Psychological Bulletin*, 68, 417–429.
- Wickens, R. C. (1984). Processing resources in attention. In R. D. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 120–142). Orlando, FL: Academic Press.
- Wolfe, J. M. (1996). Resolving perceptual ambiguity. *Nature*, 380, 587–588.
- Woodman, G. F., & Luck, S. J. (1999). Electrophysiological measurement of rapid shifts of attention during visual search. *Nature*, 400, 867–869.

Received March 20, 2001

Revision received August 17, 2001

Accepted August 27, 2001 ■