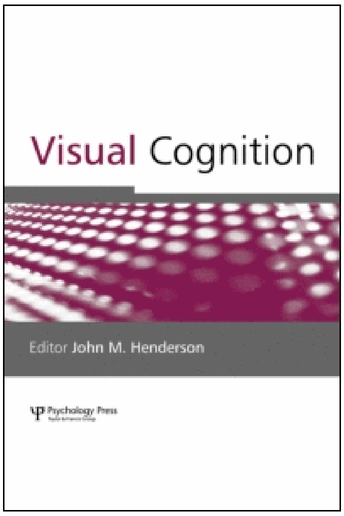

Visual Cognition
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713683696

Selection for Cognition: Cognitive Constraints on Visual Spatial Attention

Gordon D. Logan

To cite this Article Logan, Gordon D.(1999) 'Selection for Cognition: Cognitive Constraints on Visual Spatial Attention', Visual Cognition, 6: 1, 55 — 81
To link to this Article: DOI: 10.1080/713756797
URL: http://dx.doi.org/10.1080/713756797

VISUAL COGNITION, 1999, *6* (1), 55–81

Selection for Cognition: Cognitive Constraints on Visual Spatial Attention

Gordon D. Logan and N. Jane Zbrodoff
Department of Psychology, University of Illinois, Champaign, IL, USA

This article presents an argument that a primary function of attention in humans is to support cognition. Abundant evidence from cognitive science suggests that human cognition depends on "propositional representations". Propositions consist of compositional representations called predicates, and truth values. A primary function of attention, from a cognitive perspective, is to create propositions by assigning truth values to predicates. The attentional processes necessary to support propositional representations are identified and the implications for current and future theoretical and empirical approaches to attention are discussed.

INTRODUCTION

What is attention for? What functions does it serve? What biological advantage does it confer on the organisms in which it evolved? These are important questions because they place research on attention in a broader context of psychological research and because they motivated important theoretical developments in the recent past. This article presents a new perspective on what attention is for, arguing that, in humans at least, attention serves cognition. We sketch several implications of this new perspective, based on the properties of (propositional) cognitive representations, and suggest new directions for theoretical and empirical work in the study of attention.

Requests for reprints and correspondence concerning this article should be addressed to Gordon D. Logan, or N. Jane Zbrodoff, or both, care of Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign, IL 61820, USA. (E-mail: glogan@s.psych.uiuc.edu or jzbrodof@s.psych.uiuc.edu).

This research was supported by grant SBR 94-10406 to Gordon Logan from the National Science Foundation. The ideas expressed in this paper grew out of "Amsterdam seminars", a year-long discussion about the nature of attention and its relation to cognition and language that the authors had while on sabbatical leave in Amsterdam in 1992–93. We are grateful to Douwe Egberts and Hank DeVries for their contributions to these seminars.

Selection for Perception

Early work on attention, from Broadbent (1958) onward, focused on the role of attention in perception. The long-standing debate on the locus of selection (early vs. late) concerned the role of attention in perceptual identification. Advocates of early selection argued that attention was the mechanism that allowed objects to be identified, classified or categorized. Unattended objects were not identified (Broadbent, 1958). Advocates of late selection argued that objects could be identified without attention, that objects were selected on the basis of their identities rather than more primitive perceptual features (Deutsch & Deutsch, 1963; Norman, 1968).

The debate remains unsettled. Evidence appears to support both positions (for a review, see Kahneman & Treisman, 1984). In some cases, the same evidence is interpreted as supporting both positions. For example, Deutsch and Deutsch (1963) interpreted Moray's (1959) demonstration that people sometimes recognized their own name in an unattended channel as evidence against early selection theory, whereas Kahneman and Treisman (1984) interpreted it as evidence against late selection theory. The observed intrusion rate—about 35%—was too high for early selection theories, which predicted 0% intrusions, and too low for late selection theories, which predicted 100% intrusions. More recently, some theorists have argued that the question of early versus late selection is ill-formed (Logan, 1995) or confused (Van der Heijden, 1992). Nevertheless, many current theories of attention are concerned with the role of attention in perception and object identification (e.g. Bundesen, 1990; Cowan, 1995; LaBerge, 1995; Mozer, 1991), as if the major function of attention was to support perception.

Selection for Action

A decade ago, Allport (1987) and Neumann (1987) presented a new perspective on the function of attention, arguing that its main purpose was to support coherent action. Even if the capacity for identifying objects was unlimited, they argued, there would still be a need to focus action selectively on a single object or a small set of objects. Action is constrained by incontrovertible physical principles: One hand cannot be in two places at the same time; the two eyes cannot look in opposite directions at once (in humans); the body cannot go forwards and backwards at the same time. Selective attention is required to choose between alternative courses of action. Otherwise, the person would be frozen, motionless, lost in thought as the world passes by.

Selection for action should not be confused with the response selection process that appears in many stage models of reaction time (e.g. Sternberg, 1969). Response selection involves retrieving the appropriate motor command after the stimulus has been identified. Selection for action goes beyond simple choice of motor commands. It involves choice among stimuli and choice among ways of analysing stimuli as well as choice among alternative responses (cf. Treisman, 1969). The requirement to act on the world in a coherent fashion implies choice among stimuli, analyses and motor actions.

Selection for Perception and Action

Selection for perception and selection for action should be viewed as complementary rather than adversarial theoretical positions. There is room for both in the mechanisms of the mind. Perception can be selective even if there is no intention to act, as when we regard different aspects of a painting, watch a movie, or enjoy the view from an outdoor cafe. Action can be selective even if it is not directed towards perceptual objects, as in dance or paralinguistic gestures. The difference between the two kinds of selection is largely a matter of emphasis. The underlying issue is, "who is in charge?". If the goal is primarily to perceive, then selection for perception is the appropriate perspective; if the goal is primarily to act, then perception for action is more appropriate.

Some theorists, such as Van der Heijden (1992) and Logan (1996), try to accommodate both views, specifying the mechanisms that underlie each type of selection (see also Phaf, Van der Heijden, & Hudson, 1990). A popular tactic is to refer to the neurological distinction between "what" and "where" pathways (Ungerleider & Mishkin, 1982), assigning selection for perception to the "what" pathway and selection for action to the "where" pathway (e.g. Schneider, 1995).

Selection for Cognition

Theorists who integrate selection for perception and selection for action provide what appears to be a complete theory of attention, describing the flow of information from stimulus to response. We argue that this appearance is illusory, at least for humans. There is more to human attention than perception and action. Humans think, and there is more to thought than the simple stimulus-to-response mappings that are investigated in experiments on attention. Theories of attention may provide a complete account of behaviour in such experiments, but the behaviour in those experiments is not representative of the cognitive capacities manifest in everyday thought and language. In our view, a complete theory of attention must also explain significant aspects of thought and language.

The capacity for thought and the capacity for language must have imposed tremendous evolutionary pressure on our ancestors. The complex representations underlying thought would have made apparent patterns in the environment that could not be captured in simpler representations (for a discussion of the power of various representations, see Chomsky, 1963; Wasow, 1989). Those with the ability to apprehend such patterns and respond to them would have a strong advantage over those without the ability. Those better able to apprehend such patterns would have a strong advantage over those less able. Moreover, the capacity for abstract thought, not grounded in the current environment, would have introduced further advantages. Evolutionary pressures from imagined worlds may well have been as important as evolutionary pressures from the immediate environment in shaping hominid development. The advent of language must have strengthened these influences immensely. Early humans had to adapt to each other's imagined worlds as well as their own.

We argue that the capacity for thought and language influenced the evolution of human attention. Thought and language constrain attention, just as perception and action do. Attention must support thought and language, just as it supports perception and action. Our purpose, in the remainder of this article, is to identify some of the constraints that thought and language impose on attention and to suggest how attention might respond to those constraints. Through this process, we hope to identify mechanisms of attention that support thought and language.

WHAT IS ATTENTION?

William James (1980) said, "everyone knows what attention is". What he didn't say was that everyone has a different idea. Consequently, it is important for us to be clear about what we do and do not mean by attention. In our view, attention is an emergent property of the mind in action. Attention is a capacity or a capability to focus thought and action on some goal in some environment (see Logan, 1995). The capacity depends on several different processes that are recruited to fulfil the goal. The goal gives coherence to behaviour, causing perception, thought and action to be directed towards achievement of the goal. The goal implies selectivity of perception, thought and action. The goal biases perceptual organization of the environment and recruits perceptual processes required to fulfil it. The goal biases thought, suggesting ways to interpret the environment and inferences to generate from it. And the goal biases action, giving goal-directed actions priority over others. In our cognitive view, to be explicated throughout this article, the goal is generally to construct a proposition by assigning a truth value to a predicate. Attention is the selective processing involved in assigning the truth value.

We do not think of attention as a special mechanism, like a spotlight, a filter or a limited-capacity channel. Rather, attention is a kind of behaviour that emerges from the interaction of components like spotlights, filters and limited-capacity channels. In a sense, attention is a social concept, a description that people give to behaviour that is addressed in the fulfilment of a goal. When we talk about mechanisms of attention, we do not identify attention with those specific mechanisms to the exclusion of other mechanisms. In our view, any mechanism that participates in an act of attention is a mechanism of attention.

No single mechanism is involved in every act of attention, and different mechanisms may be involved in different acts.

We do not think of attention in terms of resources or capacity limitations (cf. Kahneman, 1973; Navon & Gopher, 1979). The energy metaphor that underlies those ideas appears to be theoretically bankrupt (Allport, 1987; Logan, 1987; Navon, 1984; Neumann, 1987). We are concerned more with characterizing the computations involved in the act of attention in terms of an interaction between representations and processes than with the describing the energy they consume. By analogy, the behaviour of a computer can be understood better by examining the program running on it than by examining the power it consumes in executing the program.

THOUGHT AND LANGUAGE

Language as a Window to Thought

An important feature of thought, for our purposes, is that it can be expressed in language. This is important, because language—linguistic behaviour—can be observed objectively. It can be recorded reliably and transcribed precisely. Language is thought made visible. There may be some thoughts that cannot be expressed in language. Certainly there are many that cannot be expressed easily. But a significant portion of thought can be expressed in language, and we can take advantage of that to gain theoretical leverage on our understanding of cognitive constraints on attention.

In the past 30 years, there have been major advances in the study of language in linguistics, artificial intelligence, mathematics, philosophy and psycholinguistics. These advances allow us to describe the structure of language objectively with a precision that rivals chemistry. Linguistic structures are important, from our perspective, because they reflect the structure of thought. More accurately, structures in language reflect the structure of the representations that underlie thought. Language is a product of thought, so structures could not exist in language if they did not exist first in thought (Fodor, 1975; Landau & Jackendoff, 1993).

Our strategy is to use language to tell us what is in the mind. If a structure exists in language, there must be mechanisms in the mind that are capable of supporting that structure. Thus, if language appears to require a certain attentional ability, then there must be mechanisms of attention with that kind of ability.

This strategy may provide us with “hits” but it does not protect us from “misses”. Language does not exhaust all of thought, so our strategy may miss mechanisms that are required to support thoughts that cannot be expressed in language, if those mechanisms are unique to such thoughts. However, our strategy will allow us to "hit" some important processes that are missed in current approaches to attention. We believe this to be an important advance.

Our ideas are related to Fodor's (1975) "language of thought" hypothesis, but they are not identical to it. Fodor argued that the medium of thought was a language with a structure at least as rich as one's native language. The language of thought was not the same as a natural language, like English or Dutch, but it had the same representational power. Fodor's argument for this claim was based on the idea that the observable complexity of natural languages could not arise from simpler representations with less representational power, or else the complexity must arise magically from thin air (see Chomsky, 1963; Wasow, 1989). Syntactic and semantic distinctions in natural languages must be supported by corresponding distinctions in the language of thought. Fodor argued from observations of the complexity of adult language and he argued from observations about language learning. He claimed that young children could not learn their native natural language unless they first had a language of thought available to them that would support the distinctions made in their native language (also see Clark, 1973).

We agree with these claims, and we base our theory on similar arguments (see Landau & Jackendoff, 1993). However, Fodor went further to argue that the language of thought pervaded all of cognition. While we are sympathetic to that view, we need not go that far to make our point. There may be non-linguistic thoughts and non-linguistic ways to direct attention voluntarily. Our point is that language directs attention some of the time, and we can learn a lot about attention by examining those occasions. To ignore the influence of language on attention is to ignore an important part of human cognition.

Feature Lists, Templates, Structured Descriptions and Propositional Representations

A cognitive theory explains behaviour in terms of the interaction of representations and processes. The representations are data structures and the processes are operations performed on the data structures. Representations by themselves do not do anything, and neither do processes. To affect behaviour, representations need processes to operate on them, and processes need representations to operate on. One cannot be studied without the other (Anderson, 1978).

Theories of attention—even those based on neuroscience—are cognitive theories, and, consequently, they make assumptions about representation and process. Current theories of attention rely on two kinds of representations: feature lists (e.g. Treisman & Gormican, 1988) and templates (Bundesen, 1990). Feature lists represent knowledge as a set of elementary properties, such as colour, shape and orientation. Object identity is based on the features of the objects, either singly (as in "the red thing") or in combination (as in "the red

X"). Templates represent knowledge in terms of single, prototypic representations. Object identity is based on similarity to the template or the prototype. Modern connectionist theories of attention make essentially the same representational assumptions. Localist connectionist models that represent objects as single nodes (e.g. Cohen, Dunbar, & McClelland, 1990) assume template representation. Distributed connectionist models that represent objects as "vectors" (e.g. Phaf et al., 1990) assume feature list representations.

These representational assumptions of theories of attention are problematic, as theories of cognition, because feature lists and templates are not adequate representations of objects. Curiously, this insufficiency has been known since the 1960s (e.g. Neisser, 1967), yet modern theorists continue to make the same representational assumptions (see Pinker, 1984).[1] The alternative is the idea of a "structured description", in which knowledge is represented in terms of parts or features and the relations between them. The parts or features represent the elements of the description; the relations represent the structure (see Biederman, 1987; Marr & Nishihara, 1978; Pinker, 1984). In language, the structured descriptions that represent knowledge are known as "propositions".

An important property of structured descriptions in general and propositions in particular is their "compositionality". Compositional representations have internal structure and their meaning depends jointly on the parts of that structure and on the relations between the parts (Barsalou, 1993; Fodor & Pylyshyn, 1988). Thus, the representation of "the cat is on the mat" has three parts, one representing *cat*, one representing *mat*, and one representing the relation between them—*on*. Fodor and Pylyshyn (1988) argue convincingly that compositionality is an essential property of linguistic representations. Feature lists and templates fail because they are not compositional representations.

There is broad consensus among linguists, philosophers, cognitive scientists and students of artificial intelligence that the representation of meaning is based on propositions. A proposition is a predicate with a truth value. A predicate is a structured description—are representation that specifies the relation among a set of elements. The truth value specifies the extent to which the elements and the relation are present in a particular domain. So, for example, a proposition

[1]Template matching appears to be making a comeback in the object-recognition literature. Several researchers have argued that the problems with template matching can be overcome by transforming the input to bring it into alignment with the template or transforming the template to bring it into alignment with the input (e.g. Bülthoff, Edelman, & Tarr, 1995; Edelman, 1995; Ullman, 1996). Empirical support for these ideas comes from effects of viewing angle on object recognition. However, theories based on structural descriptions can also account for the effects of viewing angle (e.g. Hummel & Biederman, 1992), so the jury is still out. Regardless of the outcome, conceptual representations of objects may still be based on structured descriptions. The representations used to recognize an object may be different from the representation used to think about it or talk about it (Smith & Medin, 1981). In many cases, parts of objects and relations between parts are important in our interactions with them (e.g. cars).

representing "the cat is on the mat" could be written as the predicate, *on*(*cat, mat*). The truth value of this predicate depends on (a) whether there is a cat, (b) whether there is a mat, and (c) whether the cat is on the mat. A rat on a mat, a cat on a rug, or a mat on a cat would result in lower truth values than a cat on a mat. Classically, truth values are binary—true or false—but more recent theorizing suggests it is useful to characterize truth values as matters of degree, varying, say, from 0 (surely false) to 1 (surely true; Zadeh, 1965).

The crux of our argument rests on the idea that humans have propositional representations. We know they have propositional representations because propositions are expressed in language. We can observe (and record) what people say, and recover the propositions underlying what they say through linguistic analysis. The capacity to have propositional representations has important implications for theories of attention, which we draw out in the remainder of this article.

COGNITIVE CONSTRAINTS ON ATTENTION

What is Attention for?

From our cognitive perspective, the function of attention is to create propositions. The act of attention assigns a truth value to a predicate, creating a proposition (Logan, 1990; Logan & Etherton, 1994). Our view is analogous in some ways to Neisser's (1967) view of attention as a creative process that synthesized interpretations of perceptual analyses. We differ from Neisser primarily in being more explicit about the thing that attention creates—a proposition. This may be largely an accident of history: Propositional representations were not well known in psychology when Neisser wrote his book.

The acts of attention that create propositions are numerous and varied. The following are a few examples.

Propositions in Identification. Consider the case in which a person's goal is to identify an object that appears in front of him or her. The person might be getting off an aeroplane looking for someone in his family to give him a ride home. The person might be performing an identification or naming experiment, in which she must name the letter that appears on a computer screen. In our analysis, the person seeing the letter must choose among predicates like *is*(*letter, A*), *is*(*letter, B*), and so on. From our cognitive perspective, this choice involves assigning a truth value to one of a set of candidate predicates representing letter identities, creating a proposition. At the moment the truth value is assigned, the predicate becomes a proposition (Logan, 1990; 1995; Logan & Etherton, 1994). Moreover, we believe that the person's response—keypress or vocalization—should be interpreted as a "speech act" (Searle, 1969) that communicates a proposition to the experimenter about what she believes to be the identity of the letter, for example, *and*(*contains*(*display, letter*), *is*(*letter, T*)).

Propositions in Search. Consider a case in which a person has a hypothesis in mind about a possible state of the world. The person might be trying to decide whether it is raining outside, evaluating the truth of the predicate *is*(*raining*). Alternatively, the person might be performing a visual search experiment, trying to decide whether an upcoming display contains the letter "T". According to our analysis, he or she would begin the trial with the predicate *contains*(*display, T*) in mind, without knowing whether that predicate was true or false. When the display appears, the person selects perceptual operations to perform on the display to decide whether or not it contains a "T". The result of that selective perceptual activity (selection-for-perception) would be a decision about the truth of the predicate, creating a proposition. After making the decision, the person usually presses a key to communicate it to the experimenter. We argue that the key press is a speech act that communicates a proposition to the experimenter. Pressing one key says "I believe the display contains a T" and pressing the other says "I believe the display does not contain a T".

Propositions in Noticing. What about cases in which perception and cognition are less goal-directed? A tourist in an outdoor cafe may be actively exploring the environment with no specific purpose other than entertainment—"Look at that!" In cases like these, the environment itself may suggest propositions. The environment activates knowledge structures (predicates) in the person's mind and some of those may lead the person to form propositions about states of the environment, such as *and*(*has*(*person, hair*), *colour-of*(*hair, green*)).

The important point here is a distinction between looking and seeing, between receiving information and noticing it. "Look" means to point one's eyes in some direction, towards some object; "see" means to identify or categorize the thing one is looking at. Receiving information is a passive, bottom-up process; noticing is an active, top-down choice of an interpretation of the received information. In our view, seeing involves creating propositions. Noticing entails taking note of something, and the "note" that is "taken" is a proposition.

Propositions in Thinking. We interpret thought as internally directed attention, where the objects of attention are not the things provided by current perceptual activity, but rather, propositions themselves or the arguments of propositions. Thinking involves creating propositions, just as searching, identifying and noticing do, but the truth values are often hypothetical rather than "empirical". We can use thought to generate inferences from things we know to be true (e.g. “it is true that it's raining; that implies we'll get wet if we go outside”) and we can also use it to generate inferences about what might be true (e.g. “if we borrowed an umbrella, we might not get wet”).

Attentional Routines

Our cognitive analysis assumes (at least) two representations—a perceptual representation that contains modality-specific information about objects that impinge on sensory surfaces, and a conceptual representation (i.e. a predicate) that contains symbols that represent perceptual objects and the relations among them. To establish the truth of the predicate, the perceptual representation must be mapped onto the conceptual one. This mapping process involves attentional selection in several ways. Perceptual objects corresponding to the arguments of the predicate have to be selected; interpretations of, or identities for, the selected objects have to be chosen; and the identified objects have to be connected or bound to symbols that represent the arguments. Once this is done (or as it is being done), the relation between the arguments has to be evaluated. This involves carrying out operations on the perceptual representation that are specified in some way by the symbol in the predicate that represents the relation. The end-product of all of this activity is the creation of a proposition (i.e. by assigning a truth value to the predicate). We call the set processes involved in creating a proposition an “attentional routine”, extending Ullman's (1984) idea of a visual routine to conceptual processing as well as perceptual processing.

Attentional Routines in Identification. Many researchers believe that simple basic-level concepts (Rosch, 1978) can be identified without attention (e.g. Van der Heijden, 1992). We argue, to the contrary, that identification requires attention even at the basic level. Our disagreement may stem from a different interpretation of identification. Other researchers may interpret identification as the activation of knowledge structures in memory (e.g. Van der Heijden, 1992); we interpret identification as the formation of an explicit propositional representation that asserts “objects *x* is a member of category *i*”. In our view, knowledge structures (e.g. predicates) may become active without attention, but they do not become propositions unless some attentional routine is executed that assigns them an explicit truth value. Basic-level categorization may require only a simple, single-step attentional routine to assign a truth value, but the result is an explicit proposition.

Simple perceptual categorizations about elementary stimulus features may involve simple, single-step attentional routines, but more complex categorizations about combinations or conjunctions of features require more extensive attentional routines (e.g. Treisman & Gelade, 1980). For example, Logan (1996) argues that conjunctions of elementary properties required separate decisions about each of the properties. Complex concepts that depend on relations between parts of objects should also require complex attentional routines (e.g. Logan, 1994).

Attentional Routines in Visual Search. Visual search requires attentional routines, because the same elementary relation must be computed iteratively for each item in the display until the target is found. Few theories address the process that controls search, which chooses which item to inspect next and decides when to stop. An important exception is "guided search theory" (Cave & Wolfe, 1990, Wolfe, 1994; Wolfe, Cave, & Franzel, 1989), which claims that display items are prioritized pre-attentively in terms of the likelihood that they are targets, and search proceeds serially through the display in descending order of priority until a target is found or the set of likely targets is exhausted. Chun and Wolfe (1996) proposed a theory to explain how search is terminated, arguing that subjects adopt some criterion for resemblance between perceptual objects and targets, and stop searching after they examine all of the objects that meet that criterion.

What remains to be specified in theories of search is how subjects manage to restrict their search to previously unexamined items, or, conversely, how they manage to avoid examining items they have already examined. Spatial indexing may be involved in choosing the next item to examine (Pylyshyn, 1984, 1994), but some kind of marking process is required to distinguish examined items from unexamined ones (Ullman, 1984; Watson & Humphreys, 1997).

Attentional Routines in Cueing Tasks. Many current theories of attention address cueing tasks, in which subjects must report an item designated by a cue (e.g. Averbach & Coriell, 1961; Eriksen & Hoffman, 1972). Such theories typically do not address the process by which the cue is found or the process by which attention is directed from the cue to the target. Instead, they focus primarily on the elementary relations computed when the target is found. The work of Posner and colleagues (e.g. Posner & Cohen, 1984) is an important exception. They distinguished three processes in moving attention: "disengagement", which releases attention from its current focus; "movement", which changes the focus of attention from its current position to the intended position; and "re-engagement", which enables attentional processing at the new focus.

Posner's theory is general, applying to "exogenous" cues that draw attention to themselves, as well as "endogenous" cues that instruct subjects to direct attention to a different location. Logan (1995) presented a theory of endogenous cueing, arguing that endogenous cues require subjects to compute a spatial relation between the cue and the target. The target is often *next-to*, *above* or *below* the cue, and these relations are complex, involving at least three parts (i.e. two arguments—the cue and the target—and a relation between them).

Logan (1995) argued that finding the target in a cueing experiment involves the same attentional routines required to compute spatial relations, which we discuss below.

Exogenous cueing is much simpler computationally than endogenous cueing. Orienting to the cue is sufficient, and orienting is a relatively simple operation. A bilaterally symmetric organism, like a flatworm, can orient by turning itself so that both sides of its body are stimulated equally, and then "climbing the gradient" to increase stimulation. Even single-celled organisms, such as amoebas, exhibit tropisms like these. The same simple computation (carried out internally, of course) may be sufficient to explain exogenous cueing, but it is not sufficient for endogenous cueing. The person may orient to the cue, but then he or she must move from the cue to the target in a specific direction, often orthogonal to the line of regard, and this requires computing a spatial relation.

Theories of attention that have failed to explain how a cue directs attention may have done so because it does not matter much. Most studies use only one cueing relation, so the cueing relation may not contribute much systematic variance to the experiment. We argue that the issue is vitally important and that cueing relations contribute large amounts of systematic variance.

Logan (1995) presented subjects with a bar marker cue outside an array drawn on an imaginary circle and compared four cueing relations: *next-to, opposite, clockwise* and *counterclockwise. Next-to* required them to report the item closest to the cue; *opposite* required them to report the item across the array from the one closest to the cue; *clockwise* required them to report the item adjacent to the closest one to the cue, going in a clockwise direction; and *counterclockwise* required them to report the adjacent item in the counterclockwise direction. There were large differences in reaction time between cueing conditions. In three replications, the difference between *next-to* on the one hand and *clockwise* and *counterclockwise* on the other was greater than 200msec. In the same experiments, varying the distance between the cue and the target over a range of 4° of visual angle produced effects smaller than 50msec, largely attributable to acuity.

Logan (1995) presented an asterisk or a picture of a head as a cue in the centre of a circular array and had subjects report items *above, below, in front of, behind, left of* and *right of* the cue. *Above* and *below* were faster than *in front of* and *behind*, which were faster than *left of* and *right of.* In eight experiments, the difference between *above* and *below* on the one hand and *left of* and *right of* on the other hand ranged from 100 to 400msec. Interestingly, Tversky and colleagues have found similar differences between the same relations when using them to cue imagined objects in imagined scenes (Bryant, Tversky, & Franklin, 1992; Franklin & Tversky, 1990).

Differences of this magnitude are too large to ignore. Such results suggest that cueing relations contribute very substantial amounts of systematic variance to attention experiments, and theories that fail to account for this variance are incomplete.

Attentional Routines in Apprehending Spatial Relations. Spatial concepts like *above, below, left of* and *right of* are complex, compositional structures that assert a relation between two or more arguments (Clark, 1973; Jackendoff, 1983; Landau & Jackendoff, 1993; Levelt, 1984; Miller & Johnson-Laird, 1976; Talmy, 1983). They require several kinds of selection, and the activities underlying the various acts of selection must be coordinated. Each of the arguments must be selected (spatially indexed), identified and set into correspondence with the symbols representing the arguments in the predicate that represents the concept. For example, in apprehending "the dog is beside the fireplace", an object corresponding to the dog and an object corresponding to the fireplace must be selected and identified. The object corresponding to the dog must be bound to the argument, *dog*, in the predicate *beside*(*dog, fireplace*), and the object corresponding to the fireplace must be bound to the argument, *fireplace*, in the same predicate. Selection of objects for these bindings may involve a search process, in which candidate objects are selected and compared iteratively with conceptual representations of *dog* and *fireplace* until appropriate matches are found.

Once the candidate objects are found and bound to the arguments of the predicate, the relation between the objects must be computed. The process by which relations are computed involves two further steps: First, a reference frame must be projected onto one of the objects. This is essential because spatial relations are defined in terms of reference frames applied to one of the objects (Clark, 1973; Garnham, 1989; Herskovits, 1986; Levelt, 1984; Talmy, 1983). This step involves a further choice between objects. Spatial relations describe the location of one object, the "located object", relative to the location of another object, the "reference object" (Landau & Jackendoff, 1993; Talmy, 1983), and the reference frame must be applied to the reference object. In "the dog is beside the fireplace", the dog is the located object and the fireplace is the reference object, because the dog's location is being described relative to the fireplace.

The reference frame that is applied to the reference object is a three-dimensional coordinate system that defines three axes: up–down, front–back and left–right. A reference frame has four parameters: origin, orientation, direction and scale. The difference between orientation and direction can be seen in the relative difficulty of deciding how a person's sides are aligned with something (orientation) versus deciding which side is left and which is right (direction).

Several different kinds of reference frames are important in apprehending spatial relations. The most important ones, from a linguistic perspective, are a person's "egocentric" reference frame, a "deictic" reference frame and an "intrinsic" reference frame. The egocentric reference frame is centred on the person, defining the person's top and bottom, front and back, and left and right sides. The egocentric reference frame is used in spatial indexing. The location of a single object is defined with respect to the person (Garnham, 1989). Deictic and intrinsic reference frames are used in apprehending relations between two objects external to the viewer. The deictic reference frame is the viewer's egocentric reference frame projected onto a reference object. To say "the ball is left of the tree" means that, if the viewer were to walk up to the tree, the ball would be on his or her left. Intrinsic reference frames are extracted from the reference objects themselves. However, not all objects have intrinsic reference frames. People and animals, cars and ships have intrinsic tops, bottoms, fronts, backs, left sides and right sides, so they have intrinsic reference frames. Balls do not have intrinsic tops, bottoms, etc., so they do not have intrinsic reference frames. Trees have intrinsic tops and bottoms but not intrinsic fronts, backs, left sides and right sides. Bullets have intrinsic fronts and backs but no intrinsic tops, bottoms, left sides and right sides. The choice between deictic and intrinsic reference frames depends on the goals of the viewer and the nature of the reference object.

Once a reference frame is chosen and projected on, or extracted from, the reference object, the second step begins: deciding the appropriateness or goodness-of-fit of the spatial relation. Logan and Sadler (1996) argued that this step was accomplished by aligning a "spatial template" with the reference frame of the reference object. A spatial template is a representation that parses the space surrounding the reference object into regions that correspond to good, acceptable and bad examples of the relation in question (see Carlson-Radvansky & Logan, 1997). For *above*, for example, the good region extends directly upward from the reference object, the acceptable region includes parts of space higher than the top of the reference object (relative to the reference object's reference frame) but not directly above it, and the bad region includes parts of space beside and below the reference object. Logan and Sadler (1996) argued that each lexicalized spatial relation had its own spatial template. They distinguished spatial templates for 12 different relations.

Once the spatial template is centred on the reference object and aligned with its reference frame, the goodness-of-fit of the located object is determined. This involves determining whether the located object falls in a good, acceptable or bad region of the spatial template. Once goodness-of-fit is determined, the truth value of the relation can be assigned. The truth value is higher if the located objects fall in a good region than if it falls in an acceptable region, and higher if it falls in an acceptable region than if it falls in a bad region.

This analysis underscores our point that apprehension of spatial relations involves complex attentional routines. Logan (1994) found behavioural evidence supporting this claim, showing that search for targets defined in terms of spatial relations between their parts was very difficult. This analysis should also underscore our point that endogenous cueing involves attentional routines (see Logan, 1995). To find the target, subjects must first find the cue and then compute the spatial relationship between it and several candidate targets, choosing the one that best exemplifies the relation.

Attentional Mechanisms

We suggest that any mechanisms involved in an attentional routine—involved in creating a proposition—is an attentional mechanism. Nevertheless, two deserve special mention because of the central role they play in our theorizing: spatial indexing and reference frame adjustments.

Spatial Indexing. Spatial indexing involves choosing a perceptual objects and establishing correspondence between it and a symbol—a spatial index—in the conceptual representation. This mapping process is a solution to the cognitive version of the binding problem that pervades research on attention: How does the system connect one kind of representation to another? In attention research, the binding problem often concerns connecting different attributes of the same object, such as colour and form (e.g. Treisman & Gelade, 1980; Treisman & Schmidt, 1982). We suggest broadening the conception to include connecting percepts to symbols and symbols to symbols (see Pylyshyn, 1984, 1994; Ullman, 1984).

A spatial index acts as an address for the perceptual object, providing the cognitive system with a means of accessing perceptual information about the object. The spatial index does not specify the identify of the object or any of the properties of the object. It is simply as assertion that the object exists. It can serve as a "storing house" to which information about the object can be attached, like Kahneman and Treisman's (1984) "object files" (see also Kahneman, Treisman, & Gibbs, 1992). If the spatial index is *X*, then propositions that include *X* as an argument can be viewed as attaching information to the index for *X*. Thus, *above*(*X*, *Y*) and *colour-of*(*X*, *red*) attach information to the index *X*.

The process of spatial indexing can be interpreted as an act of attention. It involves the same kind of selection among alternatives as traditional approaches to attention. Spatial indexing provides selective access to perceptual information, just like attentional spotlights do (Eriksen & Eriksen, 1974; Eriksen & Hoffman, 1972; Posner & Cohen, 1984; Treisman & Gormican, 1988). Spatial indexing allows the cognitive system to assess perceptual objects, just like object-based attention does (Duncan, 1984; Kahneman & Henik, 1981). Spatial indexing goes beyond traditional approaches, however, in distinguishing between perceptual and conceptual representations and in construing attention as mapping between these representations.

The problem of mapping between representations occurs in other contexts besides connecting conceptual representations to visual representations. It exists in other modalities and it exists within the conceptual domain itself.

Comprehending an analogy, such as “an atom is like the solar system”, involves establishing correspondence between parts of the domains that are arguments of the analogy (e.g. between electrons and planets, between the nucleus and the sun; Gentner, 1983). The processes that solve the mapping problem in these other domains should also be construed as attentional mechanisms.

Reference Frame Adjustment. We suggest that deictic and intrinsic reference frames are mechanisms of attention, like spotlights and spatial indices (also see Logan, 1995). Four arguments lead us to this conclusion. First, reference frames serve computational functions that are similar to those served by these other mechanisms; reference frames orient attention to space, whereas spotlights and spatial indices orient attention to objects. In principle, an infinite number of coordinate axes could be used to describe any two- or three-dimensional perceptual space, differing in origin, orientation, direction and scale. The space itself confers no privilege on any one of them. The choice of a particular reference frame—a particular setting for each parameter—provides a unique perspective on space.

Second, the use of a reference frame involves acts of selection that parallel the acts of selection involved in using other kinds of attentional mechanisms. Spatial relations take their meaning from reference frames centred on the reference object; choice of a reference object is much the same as input selection in traditional analyses of attention (Treisman, 1969). The contrast between “Maggie is left of Felix” and “Felix is right of Maggie” is largely a choice between reference objects—whether to use a reference frame centred on Felix or Maggie, respectively. Moreover, the choice between deictic and intrinsic reference frames is much the same as analyser selection in traditional analyses of attention (Treisman, 1969). The choice may be constrained somewhat by the properties of the reference object—we cannot chose an intrinsic reference frame for an object that cannot support it—but, in many cases, we have a choice and we exploit it.

Third, reference frames have the kind of flexibility associated with attentional mechanisms, like spotlights and spatial indices. The origin can be moved around space at will. Orientation and direction can be set at will, as can scale. This is not to say that all parameter settings are equally easy to attain; deictic and intrinsic reference frames are harder to deal with when they conflict with one's egocentric reference frame (Carlson-Radvansky & Irwin, 1994; Logan, 1995). The same is true of spotlights and spatial indices; it is harder to attend to locations contralateral to eye movements than to locations to ipsilateral to eye movements (e.g. Irwin & Gordon, 1998). The important point is that reference frames can be adjusted flexibly and voluntarily.

Finally, Logan's (1995) experiments suggest that reference frame adjustment may be an important step in directing attention from cues to targets. We suggest that the mechanisms used to direct attention should be considered to be mechanisms of attention.

SELECTION FOR COGNITION: IMPLICATIONS

The idea that attention is the set of processes involved in creating propositional representations has several implications for research on attention and cognition. The implications stem from the fact that propositions take several arguments, and these arguments must be instantiated explicitly in apprehending, comprehending and working with propositions.

Multiple Spatial Indices: Working Memory

The predicates underlying propositional representations take several arguments, usually between one and three. Each argument requires a spatial index to connect it to the perceptual object it correspond to. Predicates that take multiple arguments require multiple spatial indices,one for each argument. This implies that the attention system must be able to provide more than one spatial index at a time (see Pylyshyn, 1989, 1994; Trick & Pylyshyn, 1994). The multiplicity of spatial indices implies a kind of working memory—some process that keeps track of the different indices and their mappings to different perceptual objects. Without such a memory, propositional thought would not be possible.

The connection between multiple spatial indices and working memory provides an explanation of the connection between attention and primary (or short-term) memory that was first articulated by James (1890) and later endorsed by several theorists (e.g. Broadbent, 1958; Cowan, 1995; Norman, 1968; Shiffrin & Schneider, 1977). Working memory "holds" the different things that are currently being attended. The connection between multiple spatial indices and multi-argument propositions provides a new interpretation of the idea that working memory contains the contents of consciousness (James, 1890; Posner & Klein, 1973). Usually, people can talk about the propositions they entertain, and the ability to report the contents of cognition is a hallmark of conscious awareness (e.g. Dennett, 1991; Dulany, 1996; Hollender, 1986).

Ordering of Arguments: Conceptual Direction

The distinction between located objects and reference objects (Jackendoff, 1983; Talmy, 1983) imparts a conceptual direction to spatial relations. The relation points from the reference object to the located object. Many spatial relations, like *above*(*X, Y*) or *on*(*X, Y*), are not commutative; their meaning depends on the order of the arguments. "The lamp is above the table" does not have the same meaning as "the table is above the lamp". Other relations, such as *beside*(*X, Y*) or *near*(*X, Y*), are commutative, so the order of the arguments does not appear to affect their meaning. This may be true with respect to the semantics of these relations, but it is not true in terms of the pragmatics of their use in language. “Sam is beside the ball” means something different from “the ball is beside Sam”; the former focuses on Sam's location; whereas the latter focuses on the ball's location.

Speakers typically choose reference objects that the listener knows already or that the listener can find easily because of their prominence or perceptual salience (Talmy, 1983). Thus, we say “the bicycle is beside the church” because we expect the listener to know where the church is or to be able to find it easily. Implicit in this constraint on how we speak is the idea that attention goes first to the reference object and then to the located object. Thus, the contrast between located object and reference object provides direction to movements of attention (Logan, 1995).

Focus, Background and Supportive Attention

The contrast between located object and reference object provides a conceptual focus to spatial relations. The located object is focal and the reference object is backgrounded. The point of saying “the lamp is above the desk” is to focus on the location of the lamp. The speaker mentions the desk only because it serves to specify the location of the lamp. Focus is an important feature of non-spatial propositions as well. Usually, one argument is the main point behind a speaker's utterance and the other arguments are mentioned to help highlight it (Clark & Clark, 1977).

The contrast between focal and background information in propositional representations has important implications for human attention. It suggests that people must pay attention to both kinds of information in apprehending, comprehending and working with propositions. Indeed, we just argued that people must keep all of the arguments of a proposition active (i.e. represented explicitly) in working memory. The focal argument may be the most important to represent explicitly, but the background arguments must be represented explicitly as well, to support perception of the focal argument.

The necessity of representing both focal and background information explicitly is especially apparent in “relational concepts”, whose meaning is defined in terms of other concepts. A “hypotenuse”, for example, is defined relationally as the side of a right-angle triangle that is opposite the right angle. One cannot draw a picture of a hypotenuse outside the context of a right-angle triangle. To depict a hypotenuse, one must draw the other two sides of the triangle and indicate somehow that the angle opposite the hypotenuse is a right angle. A natural way to represent the meaning of hypotenuse would be to draw a right-angle triangle and use a bold line for the hypotenuse. The important point, for our purposes, is that to focus on the hypotenuse as a hypotenuse, one must also instantiate the other parts of the right-angle triangle that give the hypotenuse its meaning. The hypotenuse may be focal and the other lines and the right angle may be backgrounded, but the backgrounded features must nevertheless be instantiated to "see" the hypotenuse. The backgrounded elements serve a "supportive" function, enabling perception of the focal element.

Many other concepts have this property. Spatial relations clearly do. To see the lamp *above* the table, we must also see the table. Kinship relations have this property as well; to see someone as a father, we must at some level instantiate children and a mother. Many social roles are defined relationally; a professor is someone who teaches students; a performer is not a performer without an audience; and so on.

This analysis suggests there may be three levels of attention (focal, supportive and ignored) rather than the two levels distinguished in traditional theories (focal and ignored). Traditional theories distinguish between attended and unattended objects, arguing that attended objects receive attention and unattended objects receive none. Our cognitive analyses distinguishes between two kinds of attended objects: those receiving focal attention and those in the background receiving supportive attention, which enables perception of the focal object. Objects that receive focal and supportive attention are represented explicitly in working memory; unattended or ignored objects may activate knowledge representations implicitly, but they are not represented explicitly in working memory.

The contrast between focal and supportive attention has to do with the conceptual focus of attention, not with the amount of attention allocated at any particular moment in time. It does not imply any difference in the amount of attention (i.e. resources) allocated to focal and backgrounded objects. Nor does it imply a gradation of attention in the way that spotlight and gradient theories of attention do.

Spotlight and gradient theories of attention distinguish between different amounts of attention given to regions of space. They argue that attention is concentrated most heavily in the centre of the spotlight or gradient, tapering off towards the edges (e.g. Eriksen & Eriksen, 1974; LaBerge, 1995). One could distinguish three levels of attention in such theories—the high concentration in the centre of the beam, the low concentration near the edge of the beam, and the null attention outside the beam—but these three levels are not the same as the three levels we wish to distinguish. The distinction between levels of attention in the centre and the fringe of the beam is determined by the shape of the spotlight or gradient and depends primarily on proximity (i.e. the fringe must be adjacent to the centre). By contrast, our distinction between focal, backgrounded and ignored objects is determined by the conceptual representation that drives attention and depends on semantics rather than proximity. For example, the truth of spatial relations like *above* does not depend on the proximity of the arguments (Logan & Compton, 1996). The ceiling, the clouds and the stars are all above the desk, even though none of them are close to it. Nevertheless, to talk about them focally in relation to the desk requires some instantiation of the desk as a background.

An important implication of our distinction between three levels of attention is that subjects must pay some supportive attention to cues in cueing tasks. To see the target in relation to the cue, they must attend to both cue and target. We argue that they focus on the target and background the cue. Attention to the cue should have important effects on target processing.

The Flanker Task. Consider, for example, the flanker task of Eriksen and Eriksen (1974), in which subjects classify the central item in a display while attempting to ignore the distractors on either side. Performance is affected by the compatibility of the flankers and the target. If the flanking letters lead to the same response as the central target letter, reaction time and accuracy are facilitated; if the flanking letters lead to the opposite response, reaction time and accuracy are inhibited. These effects diminish when the distance between the target and the flankers increases (Eriksen & Eriksen, 1974) and they diminish when the target is cued by a bar marker that appears before the array (Eriksen & Hoffman, 1972). Eriksen and Eriksen interpreted these results in terms of a spotlight model. When the flankers are close to the target, they intrude on the spotlight beam and are processed along with the target. When the target is cued in advance, the spotlight can be focused sharply on the target location and mitigate the effects of the flankers.

We would interpret these effects differently, arguing that subjects sometimes locate the target in relation to the flankers. When they do, they pay focal attention to the target and supportive attention to the flankers. Supportive attention to the flankers activates their response mappings and causes facilitation when the flankers are compatible and inhibition when the flankers are incompatible. Subjects may have alternative ways of locating the target. When it appears in a constant position with respect to the fixation point, as in Eriksen and Eriksen's (1974) experiment, they may locate the target with respect to the flankers or with respect to the fixation point. Increasing the distance between the flankers and the target may bias subjects towards locating the target with respect to the fixation point. The fixation point may receive supportive attention rather than the flankers, and because it is neither compatible nor incompatible, facilitation and interference are reduced. Decreasing the distance between the flankers and the target may bias subjects towards locating the target with respect to the flankers, increasing supportive attention to the flankers, and therefore increasing facilitation and interference.

A similar argument can be made in experiments in which target position varies but is cued by a bar marker. The earlier the bar marker appears, the more likely subjects are to define target location with respect to the bar marker and the less likely they are to define it with respect to the flankers. Less supportive attention may be paid to the flankers when the bar marker appears earlier, so there may be less facilitation and less interference.

Of course, these arguments are speculative. Nevertheless, it should be possible to test our interpretation empirically, pitting it against spotlight theory to see how much of the effect can be accounted for by the competing interpretations.[2] Our approach has the virtue of explaining how subjects know which item is the target, which is something that is not addressed in current spotlight theories.

Early Versus Late Selection. Part of the controversy over early versus late selection may have stemmed from theorists distinguishing only two levels of attention (i.e. focal and ignored) when they should have distinguished three (i.e. focal, supportive and ignored). Classically, the distinction between early and late selection concerns the level of processing reached by ignored stimuli. Advocates of early selection argue that ignored stimuli should receive only cursory processing of elementary physical features (Broadbent, 1958), whereas advocates of late selection argue that ignored stimuli should be processed to the level of identity or meaning (Deutsch & Deutsch, 1963; Norman, 1968).

We suggest that some of the controversy may have arisen from a confusion between ignored stimuli and those that receive supportive attention. Evidence that appears to support late selection (i.e. evidence that non-focal stimuli are identified) could actually support early selection if it turns out that the non-focal stimuli were not ignored, but instead received supportive attention.

As our discussion of the flanker task suggests, different experiments using the same paradigm may place greater or lesser emphasis on supportive attention to the so-called ignored stimuli. What tends to be ignored in one version of an experiment may tend to receive supportive attention in another. Thus, evidence of early and late selection that appears to come and go capriciously may actually be related systematically to tendencies to give supportive attention to non-focal stimuli.

Our analysis of the problems with early and late selection is also speculative. Nevertheless, it should be possible to address it empirically, if one is clever enough to design the right experiments.

[2]One possibility would be to present subjects with a large array of white characters and have them report the identity of one that lies between two red characters. The red characters would be cues, and so should receive supportive attention, while the cued character should receive focal attention. The red characters could be response-compatible or response-incompatible with the target, following the procedure of Eriksen and Eriksen (1974). Our analysis would predict strong compatibility effects from the red cue letters that receive supportive attention than from ignored white letters an equal distance away from the target. We have some preliminary data supporting this prediction.

DISCUSSION

We began this article by arguing that selection for cognition is an important aspect of attention in humans, possibly just as important as selection for perception and selection for action. We identified several characteristics of selection for cognition, starting with the idea that humans must have mental capacities (representations and processes) to support the kinds of structures that we can observe objectively in human language. Our analysis of language suggested the proposition as an important linguistic structure that imposes several strong constraints on human attention.

We argued that the purpose of selection for cognition was to create propositional representations. Acts of attention attach truth values to predicates, which turns them into propositions. Propositions are compositional representations, consisting of interrelated parts. Propositions take their meaning jointly from the meaning of the parts and from the relations among the parts. This compositionality has important implications for attention. First, it raises the binding problem, which is already well known in the attention literature, but it provides a new perspective. Compositional representations require binding between the symbols that comprise them and parts of perceptual representations to which the symbols refer. This kind of binding is also known as "spatial indexing" (Pylyshyn, 1984, 1994; Ullman, 1984). Second, compositional representations require several spatial indices, one for each argument (at least), and this implies the necessity of a working memory capable of holding two or three spatial indices simultaneously.

Propositional representations consist of relations as well as arguments, and the relations have important implications for attention. There must be some way to compute them. Simple atomic relations may be computed in one step, but more complex (molecular?) relations require multi-step attentional routines (analogous to visual routines; Ullman, 1984).

The arguments of propositional representation are ordered, and this provides a conceptual direction, from one argument to another. In the context of spatial relations, the conceptual direction is accompanied by a reference frame, which serves as a map between the conceptual representation of the relation and the perceptual representation of the objects that instantiate it. We argued that the reference frame is a mechanism of attention because it orients attention to space, because it implies two kinds of selectivity that are classically associated with attention (input selection and analyser selection), and because reference frames are used with the kind of flexibility and voluntary control associated with other mechanisms of attention, such as spotlights and spatial indices.

The ordering of spatial relations also led us to distinguish between the conceptual focus and the conceptual background, and that led us to distinguish between three levels of attention (focal, supportive and ignored), whereas traditional analyses only distinguished between two (focal and ignored). The idea of supportive attention gave us new interpretations of phenomena in attention cueing and of problematic issues in the debate over the locus of selection (early *vs* late).

Selection for Perception, Cognition and Action

In promoting selection for cognition, our attention is not to replace previous views of the purpose of attention. Ideally, theories that address selection for cognition should be integrated with theories that address selection for perception and selection for action. The three positions should be viewed as complementary rather than adversarial; human attention is constrained by the need to perceive and act as well as to talk and think. Nevertheless, we feel it is important to focus on cognitive constraints because they have been largely neglected in the recent history of research on attention.

This was not always so. In the beginning of the cognitive revolution, research on attention and cognition was one and the same. Broadbent's (1958) first book presented an integrated view of attention that addressed short-term and long-term memory as well as perception. Neisser's (1967) classic book presented a theory of cognition that ranged from perception to language and thought, with attention intimately involved in every stage.

Things began to change around 1970. Disgruntled by the difficulty of determining the locus of selection, theories of attention began to focus on the "energy" requirements of cognitive processing rather than the computations that consumed the energy (e.g. Kahneman, 1973; Moray, 1967; Posner & Boies, 1971). At the same time, cognitive psychologists because seriously interested in representation. Psycholinguistics became a dominant paradigm (e.g. Dixon & Horton, 1968; Fodor, Bever, & Garrett, 1974) and computational approaches to cognition flourished (Anderson & Bower, 1973; Newell & Simon, 1972). Attention and cognitive science parted company. Current theories of attention, based on selection for perception and selection for action, present a picture of cognitive science that was current around 1970 (see Pinker, 1984).

Our goal in promoting selection for cognition is to bring attention and cognitive science together again. Remarkable advances have been made in the study of language, memory, reasoning and problem solving that are directly relevant to attention. Attention researchers should be aware of them and attention theorists should account for them. On the other side, theories of attention expressed (implicitly) in much of cognitive science date to the 1970s. Attention researchers have much to offer their cognitive-science colleagues. In our view, the way to begin to bridge the gap between the fields is to focus on cognitive constraints on attention. After 50 years of studying the constraints imposed by perception and action and 10 years of studying constraints imposed by neuroscience, it is time to examine the constraints imposed by thought and language and give them serious attention.

REFERENCES

Allport, D.A. (1987). Selection for action: Some behavioural and neurophysiological considerations of attention and action. In H. Heuer & A.F. Sanders (Eds.), *Perspectives on perception and action* (pp. 395–419). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Anderson, J.R. (1978). Arguments concerning representations for mental imagery. *Psychological Review, 85*, 249–277.

Anderson, J.R., & Bower, G.H. (1973). *Human associative memory*. New York: John Wiley.

Averbach, E., & Coriell, A.S. (1961). Short-term memory in vision. *Bell System Technical Journal, 40*, 309–328.

Barsalou, L.W. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In A.C. Collins, S.E. Gathercole, & M.A. Conway (Eds), *Theories of memories* (pp. 29–101). Hove: Lawrence Erlbaum Associates Ltd.

Biederman, I. (1987). Recognition-by-computers: A theory of human image processing. *Psychological Review, 94*, 65–96.

Broadbent, D.E. (1958). *Perception and communication*. London: Pergamon Press.

Bryant, D.J., Tversky, B., & Franklin, N. (1992). Internal and external spatial frameworks for representing described scenes. *Journal of Memory and Language, 31*, 74–98.

Bülthoff, H.H., Edelman, S.Y., & Tarr, M.J. (1995). How are three-dimensional objects represented in the brain? *Cerebral Cortex, 5*, 247–260.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*, 523–547.

Carlson-Radvansky, L.A., & Irwin, D.E. (1994). Reference frame activation during spatial term assignment. *Journal of Memory and Language, 33*, 646–671.

Carlson-Radvansky, L.A., & Logan, G.D. (1997). The influence of reference frame selection on spatial template construction. *Journal of Memory and Language, 37*, 411–437.

Cave, K.R., & Wolfe, J.M. (1990). Modelling the role of parallel processing in visual search. *Cognitive Psychology, 22*, 225–271.

Chomsky, N. (1963). Formal properties of grammars. In R.D. Luce, R.R. Bush, & E. Galanter (Eds), *Handbook of mathematical psychology*, Vol. 2 (pp. 323–418). New York: John Wiley.

Chun, M.M., & Wolfe, J.M. (1996). Just say no: How are visual searches terminated when there is no target present? *Cognitive Psychology, 30*, 39–78.

Clark, H.H. (1973). Space, time, semantics, and the child. In T.E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 27–63). New York: Academic Press.

Clark, H.H., & Clark, E.V. (1977). *Psychology and language*. New York: Harcourt Brace Jovanovich.

Cohen, J.D., Dunbar, K., & McClelland, J.L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97*, 323–361.

Cowan, N. (1995). *Attention and memory: An integrated framework*. New York: Oxford University Press.

Dennett, D. (1991). *Consciousness explained*. Boston, MA: Little Brown.

Deutsch, J.A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review, 70*, 80–90.

Dixon, T.R. & Horton, D.L. (Eds) (1968). *Verbal behavior and general behavior theory.* Englewood Cliffs, NJ: Prentice Hall.

Dulany, D.E. (1996). Consciousness in the explicit (deliberative) and implicit (evocative). In J. Cohen & J. Schooler (Eds), *Scientific approaches to consciousness* (pp. 179–212). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113*, 501–517.

Edelman, S. (1995). Class similarity and viewpoint invariance in the recognition of 3D objects. *Biological Cybernetics, 72*, 207–220.

Eriksen, B.A. & Eriksen, C.W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics, 16*, 143–149.

Eriksen, C.W., & Hoffman, J.E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. *Perception and Psychophysics, 12*, 201–204.

Fodor, J.A. (1975). *The language of thought.* Cambridge, MA: Harvard University Press.

Fodor, J.A., Bever, J.T., & Garret, M.R. (1974). *The psychology of language: An introduction to psycholinguistics and generative grammar.* New York: McGraw-Hill.

Fodor, J.A., & Pylyshyn, Z.W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*, 3–71.

Franklin, N., & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General, 119*, 63–76.

Garnham, A. (1989). A unified theory of the meaning of some spatial relational terms. *Cognition, 31*, 45–60.

Gentner, D. (1983). Structure mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155–170.

Herskovits, A. (1986). *Language and spatial cognition: An interdisciplinary study of the prepositions in English.* Cambridge: Cambridge University Press.

Hollender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behavioural and Brain Science, 9*, 1–66.

Hummel, J.E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review, 99*, 480–517.

Irwin, D.E., & Gordon, R.D. (1998). Eye movements, attention, and transsaccadic memory. *Visual Cognition, 5,* 127–155.

Jackendoff, R. (1983). *Semantics and cognition.* Cambridge, MA: MIT Press.

James, W. (1890). *Principles of psychology.* New York: Holt.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy & J.R. Pomerantz (Eds), *Perceptual organization* (pp. 181–211). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D.R. Davies (Eds), *Varieties of attention* (pp. 29–61). New York: Academic Press.

Kahneman, D., Treisman, A., & Gibbs, B. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*, 175–219.

LaBerge, D. (1995). *Attentional processing: The brain's art of mindfulness.* Cambridge, MA: Harvard University Press.

Landau, B., & Jackendoff, R. (1993). "What" and "Where" in spatial cognition and spatial language. *Behavioural and Brain Sciences, 16*, 217–238.

Levelt, W.J.M. (1984). Some perceptual limitations on talking about space. In A.J. van Doorn, W.A. de Grind, & J.J. Koenderink (Eds), *Limits on perception* (pp. 323–358). Utrecht, VNU Press.

Logan, G.D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology, 22*, 1–35.

Logan, G.D. (1994). Spatial attention and the apprehension of spatial relations. *Journal of Experimental Psychology: Perception and Performance, 10*, 1015–1036.

Logan, G.D. (1995). Linguistic and conceptual control of visual spatial attention. *Cognitive Psychology, 28*, 103–174.

Logan, G.D. (1996). The CODE theory of visual attention: A theoretical integration of space-based and object-based attention. *Psychological Review, 103*, 603–649.

Logan, G.D. (1997). The automaticity of academic life: Unconscious applications of an implicit theory. In R.S. Wyer (Ed.), *Advances in social cognition*, Vol. 10 (pp. 157–179). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Logan, G.D., & Compton, B.J. (1996). Distance and distraction effects in the apprehension of spatial relations. *Journal of Experimental Psychology: Human Perception and Performaance, 22*, 159–172.

Logan, G.D., & Etherton, J.L. (1994). What is learned during automatization? The role of attention in constructing an instance. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20*, 1022–1050.

Logan, G.D., & Sadler, D. (1996). A computational analysis of the apprehension of spatial relations. In P. Bloom, M. Peterson, M. Garrett, & L. Nadel (Eds), *Language and space* (pp. 493–529). Cambridge, MA: MIT Press.

Marr, D., & Nishihara, H.K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London, 200*, 269–294.

Miller, G.A., & Johnson-Laird, P.N. (1976). *Language and perception*. Cambridge, MA: Harvard University Press.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology, 11*, 56–60.

Moray, N. (1967). Where is capacity limited? A survey and a model. *Acta Psychologia, 27*, 84–92.

Mozer, M.C. (1991). *The perception of multiple objects: A connectionist approach*. Cambridge, MA: MIT Press.

Navon, D. (1984). Resources—a theoretical soup stone? *Psychological Review, 91*, 216–234.

Navon, D., & Gopher, D. (1979). On the economy of the human processing system. *Psychological Review, 86*, 214–255.

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.

Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer & A.F. Sanders, (Eds), *Perspectives on perception and action* (pp. 361–394). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Newell, A., & Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Norman, D.A. (1968). Toward a theory of memory and attention. *Psychological Review, 75*, 522–536.

Phaf, R.H., Van der Heijden, A.H.C., & Hudson, P.T.W. (1990). SLAM: A connectionist model for attention in visual selection tasks. *Cognitive Psychology, 22*, 273–341.

Pinker, S. (1984). Visual cognition: An introduction. *Cognition, 18*, 1–63.

Posner, M.I., & Boies, S.J. (1971). Components of attention. *Psychological Review, 78*, 391–408.

Posner, M.I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. Bouwhuis (Eds), *Attention and performance X* (pp. 531–556). Hove, UK: Lawrence Erlbaum Associates Ltd.

Posner, M.I., & Klein, R. (1973). On the functions of consciousness. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 21–35). New York: Academic Press.

Pylyshyn, Z.W. (1984). *Computation and cognition*. Cambridge, MA: MIT Press.

Pylyshyn, Z.W. (1989). The role of location indices in spatial perception: A sketch of the FINST spatial-index model. *Spatial Vision, 3*, 179–197.

Pylyshyn, Z.W. (1994). Some primitive mechanisms of spatial attention. *Cognition, 50*, 363–384.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B.B. Lloyd (Eds), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Lawrence Erlbaum Associates Ltd.

Schneider, W.X. (1995). VAM: A neuro-cognitive model for visual attention, control of segmentation, object recognition, and space-based motor action. *Visual Cognition, 2*, 331–375.

Searle, J.R. (1969). *Speech acts.* Cambridge: Cambridge University Press.

Shiffrin, R.M. & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127–190.

Smith, E.E., & Medin, D.L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donder's method. In W.G. Koster (Ed.), *Attention and performance II* (pp. 276–315). Amsterdam: North-Holland.

Talmy, L. (1983). How language structures space. In H.L. Pick & L.P. Acredolo (Eds), *Spatial orientation: Theory, research and application* (pp. 225–282). New York: Plenum Press.

Treisman, A. (1969). Strategies and models of selective attention. *Psychological Review, 76*, 282–299.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97–136.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review, 95*, 14–48.

Treisman, A., & Schmidt, H. (1982). Illusory conjuctions in the perception of objects. *Cognitive Psychology, 14*, 107–141.

Trick, L.M., & Pylyshyn, Z.W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101*, 80–102.

Ullman, S. (1984). Visual routines. *Cognition, 18*, 97–159.

Ullman, S. (1996). *High-level vision.* Cambridge, MA: MIT Press.

Ungerleider, L.G., & Mishkin, M. (1982). Two cortical visual systems. In D.J. Ingle, M.A. Goodale, & R.W.J. Mansfield (Eds), *Analysis of visual behaviour* (pp. 549–586). Cambridge, MA: MIT Press.

Van der Heijden, A.H.C. (1992). *Selective attention in vision.* New York: Routledge.

Wasow, T. (1989). Grammatical theory. In M.I. Posner (Ed.), *Foundations of cognitive science* (pp. 161–205). Cambridge, MA: MIT Press.

Watson, D.G., & Humphreys, G.W. (1997). Visual marking: Prioritizing selection for new objects by top-down attentional inhibition of old objects? *Psychological Review, 104*, 90–122.

Wolfe, J.M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review, 1*, 202–238.

Wolfe, J.M., Cave, K.R., & Franzel, S.L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 419–433.

Zadeh, L.A. (1965). Fuzzy sets. *Information and Control, 8*, 338–353.

Manuscript received February 1997
Revised manuscript received February 1998