

given item is the desired target. In the face of some variable amount of internal noise, each subject sets a termination threshold based on a desire to respond quickly and yet miss few target items.

None of these problems could be said to represent a "fatal flaw" in Tsotsos's work. Rather, each represents an area for possible revision of his current model. We find it encouraging that overall, Tsotsos's model, driven by complexity theory, already bears considerable resemblance to our data-driven guided search model.

Adaptation and attention

Steven W. Zucker

Research Center for Intelligent Machines, McGill University, Montreal, Quebec, Canada.

Electronic mail: zucker@larry.mrcim.mcgill.edu

Tsotsos argues that "an attentional scheme has as its main goal the selection of certain aspects of the input stimulus while causing the effects of other aspects of the stimulus to be minimized." Moran and Desimone (1985) empirically discovered that "the very structure of the receptive field, recently considered to be a fixed property of the neuron, can change from moment to moment in the behaving monkey depending on the immediate task and state of attention." The process of focusing attention is thus connected to dynamic variation in receptive field properties, a seemingly novel connection. But is this truly a novel phenomenon, and, if so, how might the mechanism be understood? We submit that analogous phenomena exist in a more primitive form as adaptation, and that the roots of attention can be illuminated by exploring the analogy with adaptation.

Adaptation exists in two forms: (1) intensity adaptation, by which the central excitatory region of a circular-surround (retinal ganglion) receptive field expands or contracts at the expense of the inhibitory surround as a function of photon intensity (Barlow et al. 1957); and (2) the effective operating range of cells in the visual cortex varies as a function of contrast (Sclar et al. 1989). That is, both (1) receptive field structure and (2) activity levels can vary as a function of stimulus properties. In this case the stimulus properties are physically based, and functionally they extend the sensitivity of the visual system to a broader range of operating environments.

The analogy between adaptation and attention arises as follows. Visual cortical neurons respond to contrast-encoded stimulus features, and exhibit a sigmoidal operating characteristic (plot of firing rate versus contrast). If all contrasts were in the saturated range then no structure would be visible. Adaptation is a primitive mechanism for preventing this, that is, for adjusting the operating range so that orientation (say) structure is detectable (pops out?) from the background. Analogously, attention is a mechanism for adjusting the feature context so that more complex structures are detectable from the background feature clutter. The increased responses to attended stimuli (Desimone et al., in press) would imply that the visual system has "adapted to" the unattended stimuli.

Although the analogy between adaptation and attention provides perspective, much remains before biologically plausible mechanisms can be specified. Adaptation is largely a *bottom-up* process, whereas attention may well be a myriad of processes many of which are *top-down*. Nevertheless, Tsotsos argues for a model of attention as an inhibitory process within a pyramidal beam, extending from large abstract receptive fields to tiny, low-level ones. A key reason for this is the positional accuracy to which attentional effects can be measured. It is almost as if the attended stimuli are described within a finer coordinate system

than the unattended ones. This accuracy may well reside in the visual descriptions being selected, however, and not in the details of the attentional beam. There is no evidence that abstract descriptions or their features are continuously distributed over the retinotopic array, with the attentional "beam" highlighting a well-delimited retinotopic subfield. In contrast, one might speculate that visual inferences are carried out by multistage processes, with feedforward and feedback loops between them. The initial stages could be coarse, local ones, and the latter stages precise, global ones. Attention could act as a gate between the early and later stages, effectively adapting away the unattended stimuli. There would then be no need to postulate a "beam" running through a pyramid of receptive fields.

Author's Response

A little complexity analysis goes a long way

John K. Tsotsos

Department of Computer Science, University of Toronto, Toronto, Ontario Canada M5S 1A4.

Electronic mail: tsotsos@wai.toronto.edu

1. Introduction

Commentators misunderstood several points in the target article. I will deal with these before addressing the many important and substantive issues raised in the commentaries.

I did not claim that computational complexity is "the key" to vision, as Dickinson states, only that it is an important and heretofore neglected dimension of study. This is explicit throughout the introductory section of the paper. Complexity analysis can reveal insights that no other method of analysis can, but it cannot even begin to address certain other equally important issues. Dickinson's commentary was dedicated to countering a view that was not expressed in the target article.

Siegel mistakenly believes that complexity theory is a top-down approach to vision. Marr (1982) describes the use of Laplacian operators; would Siegel consider Laplace's equation a bottom-up approach to vision? I hope not. Both complexity theory and Laplacian functions are tools. Complexity theory led me to develop a theory that has a significant top-down component; complexity provides its mathematical foundation. Strong likewise makes this unusual connection between tool and model, claiming that complexity theory as currently conceived is not adequate for modeling biological information processes even though in his own work he develops computational models of biological information processing. Complexity theory is one of the theoretical underpinnings of computation. Complexity theory does not model; it is a tool that provides a source of constraints for a model.

Heathcote & Mewhort provide an algorithm for unbounded visual search that they claim solves my only example in polynomial time. I did not give such an

example in the paper, however. The example in section 2.2 was provided as part of the general discussion of visual match to illustrate the definition; it preceded the formal definition of unbounded visual search. That definition does not include the target pattern and although Heathcote & Mewhort are not specific, I assume that their algorithm requires the target. **Kube** also makes this mistake (his third point). The target is not specified in the unbounded problem; the values of the functions are given only as mappings. Heathcote & Mewhort distinguish between matching and search incorrectly; as described in section 2.2, the former is a subproblem of the latter, and thus its difficulty must be included within the difficulty of the latter. If unbounded visual matching is NP-Complete, as Heathcote & Mewhort seem to agree, then unbounded visual search is necessarily NP-Complete, too.

A major component of the representations I use is the hierarchy – the simple variety that everyone understands. I did not claim that the simple hierarchy is the “best” mechanism for beating complexity (**Strong**). The argument was for sufficiency only, as clearly stated in section 3.1. Perhaps intrinsic parallelism (Holland 1975) is indeed more efficient. Although the simple hierarchy is logarithmically time-bounded, Holland’s scheme is exponential in the worst case. The worst case would be when the search creates the subset that contains the answer only after all other possible subsets have been examined. As **Uhr** points out, the logarithmic convergence in my model may be the best attack on the complexity problem; it is also biologically plausible.

Krueger & Tsay claim that I have misplaced the complexity and that the truly complex processes may be at the decisional level, yet they do not indicate whether or not they believe that the lower-level processes I have considered are tractable. I therefore assume that although they agree that perceptual processes are complicated, they believe that decisional ones are even more so. This may be so, but it is not obvious that perceptual processes, such as those in visual search, which require specific choices between response actions, require no decisional process.

Krueger & Tsay also note that I have not considered how “smart nonoccupational perceptual mechanisms” could eliminate complexity altogether and thus obviate my analysis. I cannot seriously entertain this suggestion. I have great difficulty in determining what “noncomputational” means in this context. The polar planimeter is not “noncomputational,” and Runeson (1977) does not label it so in his paper. Moreover, he notes that he is only proposing an analogy, lest anyone mistakenly infer any closer ties between the planimeter and the brain. The polar planimeter is indeed computational because there is a precise mechanistic algorithm for using it with well-defined input and output. It would be easy to simulate its operation with a computer program. Digital computers are only one manifestation of a device that computes; one must not equate computation with computers alone.

Krueger & Tsay argue that a really smart process would use only one measure for matching, either sameness or difference, not both, as I propose. They seem to have misunderstood the correlation measure, which is not just a measure of sameness but ensures the maximality of the match. A single correctly matching pixel

would pass a sameness or difference test, but not the two tests together as I have set them up. In addition, lightness and darkness have nothing to do with the validity of the computation. The formalism is capable of dealing with any type of physical measurement of a visual stimulus.

Krueger & Tsay go on to point out that vision may operate in “all kinds of undreamed of ways.” **Siegel** also makes the “what if things aren’t like this?” argument. This kind of criticism is easily leveled at any theory when the critics have no empirical counter evidence and offer no viable alternative theory. Is it really that complex? **Siegel** asks. After all, there aren’t any green elephants. Have you never watched cartoons nor enjoyed abstract art? **Siegel** describes his view as contradicting mine, yet I agree completely that “the beauty of the brain is that . . . it can solve really tough problems.” I am simply proposing a way of determining how tough the problems are and how they may be solved. More to the point, it is the very toughness of the problem that may force the brain to use the kinds of solutions I propose.

Both **Cave** and **Strong** claim that I have defined feature maps that operate independently. In section 2.4, I state that “types are not necessarily independent.” A map represents one type of visual parameter; maps are physically independent, but the types of parameters they represent are not necessarily so. Many physical visual maps have now been documented, and within each a variety of visual parameters seem to be represented, not all independent of each other (Maunsell & Newsome 1987).

Cave concludes his commentary by claiming that my model is not a serious one unless more detail of operation is provided. I had stated explicitly, however, that I would not address the operational level in this paper. All I intended was to provide a source of constraints and hypotheses.

Uhr points out what he calls a minor quibble – that the complexity is really $O(V^4)$. This is the correct order for the number of distinct images, not the number of data groupings. My analysis does not include the number of possible values of each type of visual information. This is of course an important issue, but the $O(2^{11})$ stands, for the purpose of my analysis.

Strong claims that my account requires one to assume that the entire image is in the head during processing and that this is a bad assumption. “Don’t carry anything you can readily find later,” he says. First of all, in the typical visual search experiment, there is no time to wait until later – the trial is over in a few hundred milliseconds. Second, how do you know what to discard and what to carry if you have not analyzed, at least to some degree, in the first place? **Strong** goes on to argue that this bad assumption leads to performance that does not agree with human data. He provides a figure as an example and claims that any human would see a perfectly good match to the target I define in Figure 1 of the target article. This is highly unusual because I certainly cannot see my target, an open black rectangle, in his figure!

Heathcote & Mewhort write that I cannot use Treisman’s data for comparison with my results because I use pixels and Treisman uses display items. In the early parts of the paper, I may have been unclear with this comparison; however, the relationship becomes explicit and clear in the description of the variables for Equation 23.

2. On computational modeling and visual science

Several commentators point out the impossibility of explaining biological phenomena with computational models. This argument has been made since the early days of artificial intelligence. Many have claimed that there is something special about implementations that are brain-like. This objection comes from at least two sources: those who follow Searle (1990) and those who work on neural networks. In the former case, the argument is rather nonspecific; in the latter case, it seems misguided. As **Uhr** points out, massive parallelism leads to greater speed and the ability to conceive radically different architectures than if one considers only von Neumann architectures. Most neural network research, however, is implemented on serial machines! Does this cause a problem? No, neural networks are Turing-equivalent, again as **Uhr** points out, and they are subject to the same results about computational complexity and computational theory as any other implementation. (See section 1.3 of the target article about the Church-Turing thesis.) It is important to note that relaxation processes are specific solutions to search problems in large parameter spaces and nothing more. Neural networks use variations of general search procedures called optimization techniques. If relaxation (or other optimization processes) is indeed the process by which real neurons perform some of their computation as **Siegel** suggests, it is subject to precisely the same considerations of computational complexity as any other search scheme.

Uttal points out that my particular theory cannot be an explanation of biological behavior and that it would at best be an analogy. Is there any other type of explanation? In physics, cosmology, or chemistry explanations and theories are put forward and the only requirement for their validity is that they account for the experimental observations. Would a cosmologist be required to create a universe in order for his theories to be taken seriously? Or a biologist, life? A theory that accounts for more observations than another is a better theory. Theories whose predictions are falsified are modified or rejected. In addition, computation itself plays a large role in modern theory construction even in the above disciplines. Computer simulation in particular has been a very powerful tool in the physical sciences. Yet, no cosmologist would claim that he is creating a universe and no one would criticize him for not doing so.

Is simulation of information processing particularly menacing for some reason? Or is it that in AI we have concentrated too much on toy examples and have not developed falsifiable theories and a solid experimental tradition, as in other disciplines? It is hard to say. The work presented in the target article, however, is intended to be one dimension of a framework for developing a general theory of biological and artificial perception. I have considered the dimension of computational complexity only, but other dimensions must also play a role, as many commentators have correctly noted. I have developed constraints that apply to all theories of perception and have tried to show one possible path of development that would satisfy those constraints.

Eklundh wonders whether the sort of analysis I propose can yield precise predictions or only provides constraints on the model space. He is right to ask. Complex-

ity level analysis only yields constraints, as I point out in the target article. My analysis is followed by one possible model conforming to those constraints. Competing models are encouraged; such models do lead to precise predictions.

The target article was guided by observations in psychology, neurophysiology, and neuroanatomy. The lifetime of my results will be determined by experimental confirmation or refutation from those disciplines and by their usefulness in designing machine-based perceptual systems. Many predictions were made in the target article, most of which no commentator criticized. Most of the predictions conform very closely to known findings in biology. I am pleased to see that such investigators as **Desimone**, **Cave**, **Wolfe**, and **Treisman** find such a close resemblance between my suggestions and their own. That was the whole point! A large interdisciplinary set of observations was tied together using the thread of complexity analysis.

3. Visual search within vision

It is suggested by **Lowe** that I have not shown the importance of visual search for vision in general. Indeed, I only state that visual search may be a very basic problem that is found in most other types of visual information processing. Elaboration is in order. Basic bounded visual search task seems to be precisely what any model-based computer vision system has as its goal: Given a target or set of targets (models), is there an instance of a target in the test display? **Lowe's** own work certainly falls into this category (**Lowe** 1987). Even basic visual operations, such as edge-finding, are also in this category: Given an edge-detection model (e.g. **Ballard & Brown** 1982), is there an instance of this edge in the test image? It is difficult to imagine any vision system that does not involve similar operations. My remark about the ubiquity of search in vision therefore seems to have merit. The point has not been rigorously proved, of course, but it is clear that these types of operations appear from the earliest levels of vision systems to the highest.

4. Complexity is even more complicated

Strong wonders about the relationship between the two Knapsack problems I present, one as an example in section 1.3 and one with a formal definition in section 2.3. The complexity literature indicates that the same problem can be formulated in various ways. Different instances will share certain basic features. So it is with the Knapsack problem. Many different statements of it are given in **Garey & Johnson** (1976). The example in section 1.3 was found in **Rosenkrantz & Stearns** (1983) as an easily understood example for a noncomputational readership. The intractability claim that **Kube** disputes came from that article. As defined, the statement is true; more on this later.

Kube proposes that the theorems I present concerning the intractability of unbounded visual search do not hold; he provides conditions under which Theorem 1 does not hold, noting that the Knapsack problem is not NP-Complete in the "strong sense." He is right; however, he goes on to say that unbounded visual search is consequently

not NP-Complete either. This is simply wrong. The problem is still NP-Complete and has exponential time complexity as defined, that is, with no a priori assumptions or bounds. My proof for unbounded visual search has been duplicated twice so far, each proof with slightly differing problem formulations (by Bart Selman, 1989, in our own department and by Ron Rensink at the University of British Columbia, personal communication, 1989).

Let us examine this a bit further. First, some definitions must be presented. Define two functions over the nonzero integers, *Length* and *Max*. The former is a function that maps any instance *I* of a problem to an integer corresponding to the number of symbols used to describe the instance under some reasonable encoding scheme for all instances. The latter maps an instance to an integer corresponding to the magnitude of the largest number in the instance. An NP-Completeness result does not necessarily rule out the possibility of solving a problem with a "pseudopolynomial" time algorithm. This is true only for "number problems," such as Knapsack. A problem is a number problem if there exists no polynomial *p* such that $\text{Max}[I] \leq p(\text{Length}[I])$ for all *I*. According to Kube, I assume that the magnitude of image values must increase exponentially with retinal size. I make no such assumption. Moreover, by definition, the relationship cannot be polynomial. Kube's comment does not fit the definitions. An algorithm that solves a problem is a pseudopolynomial-time algorithm if its time complexity function is bounded above by a polynomial function of the two variables *Length*[*I*] and *Max*[*I*]. Kube points out that Knapsack has a known polynomial-time algorithm if an assumption can be made about the magnitude of the numbers; but this is not the same as the problem being inherently polynomial. If it were, it would have proved that all NP-Complete problems have polynomial solutions, disproving the conjecture on which the entire theory of NP-Completeness depends. It turns out that this is a common mistake, but to show why one must determine the complexity function for the proposed solution and the length of an instance of unbounded visual search.

The polynomial-time solution to which Kube refers is presented by Dantzig (1957) based on a method first proposed by Bellman (1954). Lawler (1976) provides a different algorithm for Knapsack also based on Bellman's equations. Bellman motivates his solution by pointing out that practical experience with the problems is put to use. I wished to conduct an analysis that did not depend on such experience. After all, there was no experience to draw on before our visual systems had evolved. Dantzig carefully notes that although algorithms for approximate solutions also exist using techniques of linear programming, the solution by Bellman is intended for the derivation of exact solutions. As such, it is recommended when there are only a few items in the knapsack and only one kind of limitation. Moreover, Bellman says that because of the nonlinear functional relationships inherent in his equations, only special cases of them can be solved and, even then, solutions will not necessarily be unique. The algorithm, which relies on dynamic programming, seems to require $O(\theta \cdot |I|)$ operations where θ corresponds to one of the thresholds of the unbounded visual search problem defined in section 2.3, and $|I|$ is the number of elements in the test image set. To encode an instance of unbounded

visual search $O(|I| \cdot \log_2 \text{Max}[I])$ bits are needed. The number of operations, $O(\theta + |I|)$ is not bounded by any polynomial function of $|I| \cdot \log_2 \text{Max}[I]$ and thus the general problem does not have a polynomial-time algorithm.¹ It is still NP-Complete. The NP-Completeness depends on large inputs.

What sizes of numbers are present in unbounded visual search? This problem has three kinds of numbers: the values of the test image, and the values of the *diff* and *corr* functions. The human eye can discriminate over a luminance span of about 10 billion to 1 (Dowling 1987). Thus, image values should have this as a range; similarly, the *diff* function has this range while the *corr* function has a range of 1 to 10^2 billion because it is a product of two image values. Thus $\text{Max}[I]$ is at least 10^{20} . The retina has about 130 million photoreceptors. To binary encode one instance of unbounded visual search for humans would require $O(|\log_2 10^{20}| \cdot 1.3 \cdot 10^8 \cdot 3)$ bits or more than 20 million bits! This is certainly too large to be biologically plausible. According to the definition given earlier, an algorithm is pseudopolynomial if it has a time-complexity function bounded from above by a polynomial function of *Length*[*I*] and *Max*[*I*]. Using the estimates for *Max* and *Length* derived here, such a time-complexity function is of little help. This in fact exhibits a property of number problems that are NP-Complete yet have a pseudopolynomial-time solution: They display exponential behavior with large input numbers.

There is an additional problem with the pseudopolynomial time algorithm for Knapsack. That solution, together with all solutions based on Bellman's initial formulation² use the following clever observation: If we wish to solve a problem of size *N*, first determine the solution to same problem but of size *N*-1; the cost of determining the solution to the original problem then becomes easy because the decisions that must be made are only for the additional element. This line of reasoning can be extended from problems of size *N* all the way down to size 1. With this technique the number of operations becomes very small. Such solutions are known as recursive; each decision depends on decisions made for the problem of the next size down. This recursiveness poses a serious problem for biological plausibility. Bellman's functions are nonlinear; the algorithm that uses them involves two nested if-then-else conditions to decide which functions are used for each step based on the magnitude of the values determined in the previous step. Even though the solution may require polynomial rather than exponential time, it does not appear to be parallelizable because of the strong dependence of each step in the solution on the previous step.³ In a retina size problem this solution may necessarily require 130 million sequential steps.

Lowe, Krueger & Tsay, Uttal, and Wolfe all describe the importance of noise and probabilities in vision. I agree that research must pursue these considerations. Probabilistic complexity is not quite so well understood, however. Lowe and Uhr question the use of worst-case complexity. I first point out that worst case does indeed occur in practice. In any problem of fixed length, not necessarily large, it is quite possible that a search method will find a solution only after all other possible solutions have been tried. That is just as much a worst-case scenario as is the largest possible problem. Perceptual algorithms must be *time-bounded* to be useful to a perceiving sys-

tem. Worst-case complexity can provide this bound. Worst-case analysis can tell us about all instances of the problem; average-case analysis can only tell us about the average case; it is unclear what the average case could be for vision. Average-case and probabilistic analyses should also be attempted once the techniques are developed and we get a good enough idea of what the average visual input could be.

5. Complexity equations and the data

Several comments were made about the algorithm and explanation for visual search. Four experimental scenarios are addressed by the algorithm in section 5:

Type I: The target is the only item in the display to exhibit a specific feature; the target is known in advance.

Type II: The target is the only item in the display to exhibit a specific feature; the target is not known in advance.

Type III: The target is the only item in the display to exhibit a specific feature combination (two or more features); the target is known in advance.

Type IV: The target is the only item in the display to exhibit a specific feature combination (two or more features); the target is not known in advance.

Type I is the usual version of disjunctive search found in the literature; similarly, Type III is the usual version of conjunctive search. The target article is a bit vague about odd-man-out searches (Wolfe is justified in his criticism). In my defense, I have not seen too many experiments with Type II or IV conditions, Treisman and Sato (1990) being the only example. To help clarify the conclusions of the algorithm for visual search, I will give the time-complexity function for each of these conditions and comment on the relationship to the experimental data, where possible.

- Type I: Response Time varies as $|T| \cdot \Phi(\hat{M})/2$
- Type II: Response Time varies as $|R_a| \cdot |T| \cdot (2^{\Phi(\hat{M})} - 1)$
- Type III: Response Time varies as $|R_a| \cdot |T| \cdot \Phi(\hat{M})/2$
- Type IV: Response Time varies as $|R_a| \cdot |T| \cdot (2^{\Phi(\hat{M})} - 1)$

where $|R_a|$ stands for the total number of items in the display and $|R_s|$ represents the number of candidates left for matching after inhibitory tuning is applied. The other variables are as defined in the target article. In each case, the target may be present or absent in the test displays. Two targets rather than one would lead to a doubling of time to compute the visual response. Quinlan & Humphreys (1987) report similar effects. The story is not quite so neat, however. In section 5.2.6 I point out that the selection of candidates for matching may depend on their relative response strength. In other words, the ordering of candidates may be in descending order of response. Section 4.5 points out that inhibitory tuning based on the characteristics of the target leads to computational savings as well as larger responses and that the inhibition should be applied using a Gaussian weighting function over the feature dimension of interest, applying this weighting function multiplicatively. This mechanism manipulates the relative ranking of candidates for a search task. Consider the example in Figure 1. In the top half of the figure, the possible elements of a simple conjunction task are shown. For the given target, inhibition would rank the possible distractors depending on which features

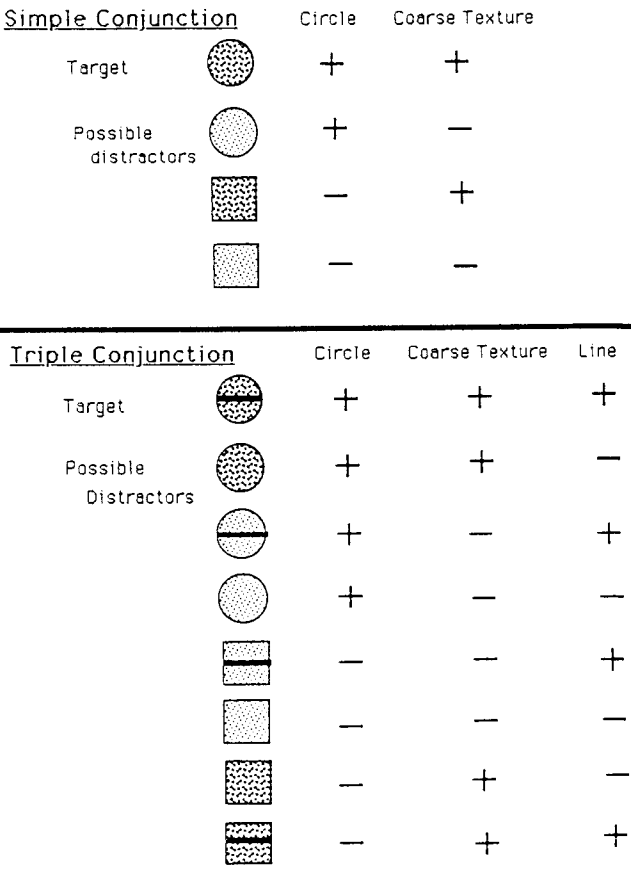


Figure 1. A comparison of possible relative effects of inhibitory tunings with known targets for simple and triple conjunction experiments. The “-” implies inhibition, the “+” denotes no change, both with respect to the relevant feature dimension. The magnitude of inhibition is not considered here; it would have an important effect in the actual ranking of candidates.

they possess (and to what degree). I have specified only a “+/-” scheme here; this is not to say that the ranks are equally spaced. The response depends on the relative strength of the item and the amount of inhibition applied. This, in turn, depends on the distance of the distractor feature from the comparable target feature along the same dimension. The weaker the distractor relative to the target, the smaller its final response; the farther away a feature from a target’s feature along the same dimension, the weaker its final response. The fact that features may not be independently computed (coarse-coding, or neurons that are selective for both color and orientation, for example) complicates the determination of “same” dimension.

In a typical conjunction display, some combination of target and distractors is presented. Each display poses a potentially different distribution of relative rankings of candidate elements. It cannot be assumed that each display is of precisely the same difficulty. This is even more evident in a triple conjunction where the possible distributions of candidate rankings are even more varied, as shown in the triple conjunction example of Figure 1. If search does proceed by selecting candidates in order according to response strength, then it is easy to see how triple conjunctions may be faster than simple conjunctions. All that is required is to ensure that the ranking

always leaves the target on top and that the distractors, even if ranked second, be distant seconds.

The proposal described in the preceding paragraph would lead to the observations of Egeth et al. (1984), who found that subjects can eliminate a feature dimension from consideration if instructed to do so. Treisman & Sato (1990) found that triple conjunctions can be fast if the target differs from distractors in two dimensions (representing two sources of inhibition for distractors) but that this is harder than a simple conjunction if the difference is only on one dimension. It also predicts the observations of Wolfe et al. (1989). Wolfe et al. always use size in their triple conjunctions and the target is always larger by at least double. It is easy to see how inhibition selective for scale can strongly favor the large element over the small. If Wolfe et al. repeated their experiments with small targets, my proposal predicts slower searches. It is odd that size plays such a large role in their experiments because Cave & Wolfe (in press) say that stereo and size are very effective for top-down guidance. Treisman & Sato (1990) report that conjunctions involving large size are faster than with small size. Burbeck & Yap (1990) recently reported that scale seems selectable based on context, with the largest response dominating. Further support for the proposal comes from Quinlan & Humphreys (1987), who observed that target-distractor discriminability influences the rate of conjunctive search. Another way of influencing the selection of candidates is to precue for location. Treisman (1985) reports a large advantage to precueing for location in conjunctive search whereas it is irrelevant in disjunctive searches. This, too, is consistent with the proposal.

Wolfe describes an odd-man-out problem that is surely "easy," i.e., parallel, with practice. If a unique item is created by one or more differences over the defining distractors, but all distractors are the same, then I must agree that the search appears easy, especially with practice.

It seems that my predictions for Type I and III agree well with observations. I have only one set of experimental data with which to compare with equations for Type II and IV; data supplied graciously by Anne Treisman (Treisman & Sato 1990). In that experiment, targets were unknown to subjects and displays were created with 4, 9 or 16 items. Targets consisted of (a) large items; (b) small items; (c) large-colored items; (d) small-colored items; (e) large-oriented items; and (f) small-oriented items. Using the standard method, the response times for each of these six conditions are plotted against display size and lead to linear relationships of varying slopes. These data are really three-dimensional, however, with the third dimension the feature dimension. My predictions for Type II and IV call for an exponential relationship in this dimension and a linear one in the display size dimension as observed. But how should we plot this feature dimension? It will not do to simply enumerate the features; it cannot be assumed that they are computed with equal ease. I fit exponential curves of the type predicted leaving the y-intercept and constraining the $\Phi(M)$ to have the same value across all display sizes for the same feature combination. The result is shown in Figure 2. The fit is very good.⁴ The values of the exponents for each condition are: (a) 4.54; (b) 4.66; (c) 4.73; (d) 4.82; (e) 5.77; and (f) 6.12. The y-intercepts are: for 4 items, 564; for 9 items,

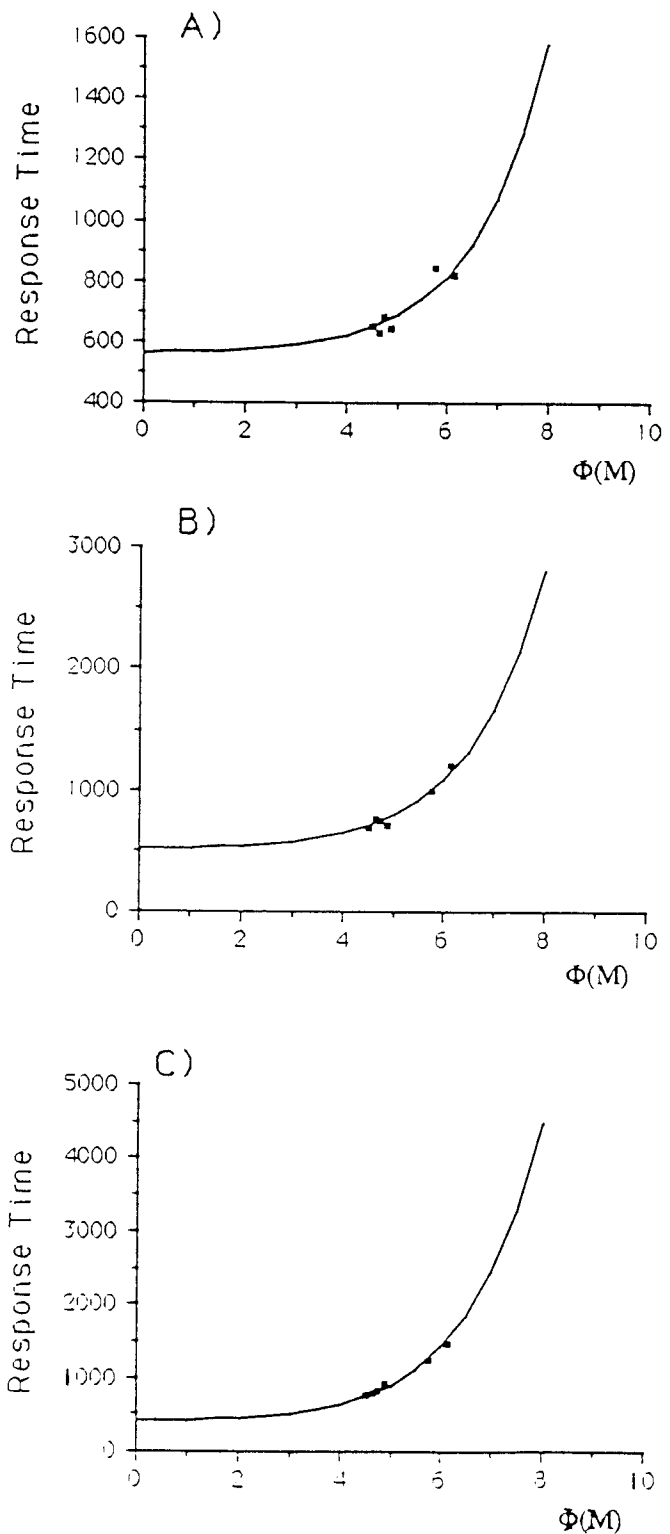


Figure 2. Plots of response time versus feature dimension $\Phi(M)$ for unknown target experiments of Treisman & Sato (1990). (A) 4 items in display; smooth curve is $RT = 564 + 4 \cdot (2^{\Phi(M)} - 1)$. (B) 9 items in display; smooth curve is $RT = 516 + 9 \cdot (2^{\Phi(M)} - 1)$. (C) 16 items in display; smooth curve is $RT = 411 + 16 \cdot (2^{\Phi(M)} - 1)$. The values of $\Phi(M)$ found for each of the display types are: large size: 4.54; small size: 4.66; large-color: 4.73; small-color: 4.82; large-orientation: 5.77; small-orientation: 6.12. These are common across display size.

the initial motivation for the beam arose out of the contradiction created by the determination of lower bounds on map size, the need is identical for the feature dimension. In the algorithm given in section 5.0, attention is applied in two key places, step 2 (section 5.2.2) and step 6 (section 5.2.6). In the first instance, it is used to "tune" the entire input hierarchy to expect features that are specified by the task and does not involve selection of spatial elements. One of the major effects of this tuning is to manipulate the distribution of competitors in the winner-take-all processes that are responsible for decision making. This manipulation changes the response characteristics of those processes, leading to enhanced response values achieved in shorter time. Haenny, Maunsell and Schiller (1988), as well as Desimone (1990), have observed a change in time course of response for attended units. Enhanced response is therefore a side-effect of an inhibitory mechanism. It is not necessarily the case that enhanced response implies an enhancement mechanism, as **Krueger & Tsay** suggest. Krueger & Tsay go on to point out that there is little evidence for any attentional effects on what features are extracted or compared. Their view is out of date and incompatible with the observations of Moran & Desimone, Haenny, Maunsell & Schiller and many others as cited in the target article.

The second instance of attention is the application of the beam as it was originally motivated, namely, to select spatial candidates for matching. The mechanism for this is currently being investigated. **Treisman** correctly criticizes the lack of detail for the implementation of the beam. I had intended it as an analogy only for the purpose of the target article, however, and a future paper will provide detail (Tsotsos, in preparation). Treisman points out that it is difficult to attend selectively to one feature at a particular location and to ignore others. How can one then account for the results of Moran & Desimone (1985) and other similar findings? In those experiments, exactly this type of selection occurs within individual receptive fields. I would suggest that some other mechanism must play a confounding role for the interferences Treisman cites. For example, coarse-coding of features would cause this effect to fall out naturally. Suppose a single unit codes both color and shape to some degree in a coarse-coded fashion. If only spatial selectivity is applied to that unit, it may indeed be difficult to select color over shape. This would be confirmed by my explanation too, as long as the only dimension of attention was location. Moran & Desimone and others showed that features selected can also be within units. My beam idea thus includes both of these aspects and the specification of the task determines which is used, if not both.

I still maintain that detectors in the early abstraction hierarchy are almost never in their "untuned" state; vision is almost always purposeful and if it is, the visual system will attempt to tune its resources in the direction most suited for that purpose. Optimizing the tuning of detectors for expected features will lead to faster responses for at least two reasons: competing items in the display are attenuated, and, winner-take-all results are speeded up as shown in section 4.5. **Treisman** points out that my alternative explanation for search asymmetries cannot be correct because I misinterpreted the "instructions" given to subjects. I did not misunderstand, but

certainly misstated my understanding in the article. I realize that subjects only see the stimuli and nothing else and are not told to search for a logical negation of features. Attentional tuning for search asymmetry cases would inhibit any detectors that would respond to the features that were not part of the target. In an attempt to put the requisite decisional process into computation terms, I used the term "logical negation" and thus created some confusion about my meaning. Treisman asks why "could the standard stimuli not be found simply by leaving the detectors in their untuned state, so that only the target is effective?" Does that mean during the course of an experiment? How is it proposed to turn off attentional or top-down influences in a conscious human subject? I agree, however, with Treisman's clarification of the difference between "discovering existence" and "specifying identity."

7. Representation

Desimone and **Zucker** question the need for certain representations of features within the framework that I have presented. They (along with several other commentators but on different issues) must be reminded of the caveat I made early in the target articles. The derived constraints and framework resulted from complexity considerations **alone**. This is the view of vision that complexity alone can yield – a considerable one, all things considered, but certainly not complete. Moreover, in my definition of features, I deliberately left them unspecified because the goal was simply to count how many were possible. This led to a lower bound for the number of physical feature maps. I fully recognize that feature representation is a "murky" area; I do not think I have contributed much to it other than to place constraints on numbers of features.

Zucker points out that the beam idea requires continuous representation of features across space. In the idealized framework I present this is true.⁴ I might point out that Zucker's own work, on curvature for example, also has this requirement and does not reflect the spatially fragmented nature of representations in the cortex. Do those breaks and gaps in representation have functional value, or are they artifacts of evolution or some other mechanism? We do not know at this point. I know of no model that has intentionally included the seemingly random gaps and anomalies of representation one finds in biology. How could it? We do not yet understand what a complete representation could be doing let alone one that seemingly cannot cover visual space adequately.

Treisman points out the need for object files, temporary ad hoc representations that are not hard-wired in a prelearned visual dictionary; she claims that my visual search algorithm has a problem because it does not include this. In step 1 of the algorithm (section 5.2.1), I describe the need to store a representation of the target. In step 7 (section 5.2.7), I argue for the need for a buffer representation, precisely because the wiring requirements would be too great. These temporary representations serve the same purposes (roughly) as Treisman's object files. **Mohnhaupt & Neumann** also point out the need for such an intermediate representation in vision, citing much relevant research.

8. Concluding remarks

The research described in the target article was first published as a technical report at the University of Toronto dated September 1987 (Tsotsos 1987a), shortly before that paper was submitted to *Behavioral and Brain Sciences*. The arguments of section 3 first appeared at the International Conference on Computer Vision, London, June 1987 (Tsotsos 1987b). Cave's comment about the newness of the results is not correct; at the time of submission, none of the researchers currently espousing visual search explanations involving inhibitory guidance (Wolfe, Cave, Treisman) were doing so. The idea was indeed new back then, as was the explanation for visual search.

Throughout my development, I attempted to include only minimal assumptions and very simple optimizations within the framework. Heathcote & Mewhort believe that I confuse simplicity with tractability. If tractability can be achieved simply, then the result is all the more powerful. If you need to hang a picture frame, do you use a jackhammer for that nail or a simple tack hammer? In addition, I certainly do not propose that further optimizations are not possible. Some commentators objected to my pointing out that the best use must be made of the tools provided or that I could have chosen more powerful or extensive optimizations (Strong, Cave,⁹ Wolfe¹⁰). One should remember that not all dimensions of a problem can be optimized simultaneously. How to choose which dimensions should be optimized and by how much is a judgment call – my intuition versus yours. I opted for a principle of least commitment. Who is right? Time will tell, of course. Science has always favored simple explanations for complex phenomena and it is our challenge to find them. If a solution is indeed too simple, then it should be easy to demonstrate this because it will not account for the experimental observations as well as another more sophisticated theory. This is how science progresses. I cannot claim at this point that I have found the complete and correct explanation; I can only hope that I have provided some useful constraints that delimit the future search for the solution and some hypotheses for one possible model. I was actually quite surprised to see how much can be explained with simple mechanisms and the single dimension of study on which I embarked.

Finally, I wish to emphasize strongly that complexity theory is as appropriate for the analysis of visual search specifically and of perception in general as any other analytic tool currently used by biological experimentalists. Experimental scientists attempt to explain their data and not just to describe it; it is not surprise that their explanations are typically well thought out and logically motivated, involving procedural steps or events. In this way, a proposed course of events is hypothesized to be responsible for the data observed. There is no appeal to nondeterminism or to oracles that guess the right answer or to undefined, unjustified, or "undreamed-of" mechanisms that solve difficult components. Can you imagine theories that do have these characteristics passing a peer-review procedure? They wouldn't pass such a procedure (at least not in our current view of science!). In proposing an explanation, experimental scientists attempt to provide an *algorithm* (using the definition of algorithm provided in my section 1.1) whose behavior leads to the

observed data. Because biological scientists provide algorithmic explanations, computational plausibility is not only an appropriate but a necessary consideration. One dimension of plausibility is satisfaction of the constraints imposed by the computational complexity of the problem, the resources available for the solution of the problem, and the specific algorithm proposed.

ACKNOWLEDGMENTS

I thank Bart Selman, Greg Dudek, and Joachim von zur Gathen for useful discussions.

NOTES

1. This line of reasoning is borrowed from Garey & Johnson (1976, pp. 90–91), who demonstrated that even though the Partition problem has a pseudopolynomial-time algorithm, it is still NP-Complete. The proof for the NP-Completeness of Knapsack involves a reduction from Partition.

2. It seems that the great majority of pseudopolynomial-time algorithms for NP-Complete number problems are derived using the methods outlined by Bellman (1954) and Dantzig (1957).

3. For example, no algorithms are known for linear programming that are parallelizable (Dobkin et al. 1979). Linear programming is used for approximate solutions to Knapsack.

4. Error data, etc. were unavailable for proper statistical analysis of the fit.

5. Van Essen and Anderson (1990) note that 24 visual areas are currently known.

6. The heavily myelinated zone of the superior temporal sulcus area that is direction-selective receiving input from striate cortex.

7. $MT^* (sq. mm.) = 14 \cdot body\ weight\ (kg.)$

8. Bob Desimone points out that recent experimentation has found a rather continuous representation of feature values along a given dimension, say color, at a given spatial location in V4 (personal communication).

9. Cave's suggestion for encoding the relevant maps with each prototype leaves open the problem of how to recognize colored objects in a black and white image, normally stationary objects that are moving, and other such exceptions from default settings.

10. Wolfe's "optimal" use of top-down guidance leads him and his colleagues to appeal to undefined noise effects to "fix" their model because it "works too well."

References

Letters "a" and "r" appearing before authors' initials refer to target article and response respectively.

- Albano, J. E., Mishkin, M., Westbrook, L. E. & Wurtz, R. H. (1982) Visuomotor deficits following ablation of monkey superior colliculus. *Journal of Neurophysiology* 48:338–51. [RD]
- Allman, J., Miezin, F. & McGuinnis, E. (1985) Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparisons in visual neurons. *Annual Review of Neuroscience* 8:407–30. [aJKT]
- Anderson, C. & van Essen, D. (1987) Shifter circuits: A computational strategy for dynamic aspects of visual processing. *Proceedings of the National Academy of Science USA* 84:6297–6301. [aJKT]
- Ballard, D. (1986) Cortical connections and parallel processing: Structure and function. *Behavioral and Brain Sciences* 9(1):67–90. [aJKT]
- (1989) Animate vision. *Proceedings of the International Joint Conferences on Artificial Intelligence*. Detroit. [aJKT]
- Ballard, D. & Brown, C. (1982) *Computer vision*. Prentice-Hall. [rJKT]
- Ballard, D., Hinton, G. & Sejnowski, T. (1983) Parallel visual computation. *Nature* 306(5935):21–26. [aJKT]

- Barlow, H. (1986) Why have multiple cortical areas? *Vision Research* 26(1):81-90. [aJKT]
- Barlow, H., Fitzhugh, R. & Kuffler, S. (1957) Change of organization in the receptive fields of the cat's retina during dark adaptation. *Journal of Physiology* 137:338-54. [SWZ]
- Barrow, H. & Tenenbaum, J. M. (1978) Recovering intrinsic scene characteristics from images. In: *Computer vision systems*, ed. A. Hanson & E. Riseman. Academic Press. [aJKT]
- Bellman, R. (1954) Some applications of the theory of dynamic programming: A review. *Operations Research* 2:275-88. [PRK, rJKT]
- Bennett, B., Hoffman, D. & Prakash, C. (1989) *Observer mechanics*. Academic Press. [RE]
- Biederman, I. (1988) Aspects and extensions of a theory of human image understanding. In: *Computational processes in human vision*, ed. Z. Pylyshyn. Ablex Publishing Corp. [aJKT]
- Broadbent, D. & Broadbent, M. (1978) From detection to identification: Response to multiple targets in rapid serial visual presentation. *Perception and Psychophysics* 42(2):105-13. [aJKT]
- Bureck, C. & Yap, Y. (1990) Spatial-filter selection in large-scale spatial-interval discrimination. *Vision Research* 30(2):263-72. [rJKT]
- Burt, P. & Adelson, E. (1983) The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications* 31:4:532-40. [aJKT]
- Cavanagh, P. (1987) Reconstructing the third dimension: Interactions between color, texture, motion, binocular disparity, and shape. *Computer Vision, Graphics, and Image Processing* 37:171-95. [PRK]
- Cave, K. R. & Wolfe, J. M. (in press) Modeling the role of parallel processing in visual search. *Cognitive Psychology*, [KRC, AH, JMW, rJKT]
- Chen, L. (1982) Topological structure in visual perception. *Science* 218:699. [MM]
- (1989) Topological perception: A challenge to computational approaches to vision. In: *Perspective in connectionism*, ed. R. Pfeiffer. Elsevier Science Publisher. [MM]
- Chignell, M. H. & Krueger, L. E. (1984) Further evidence for priming in perceptual matching: Temporal, not spatial, separation enhances the fast-same effect. *Perception & Psychophysics* 36:257-65. [LEK]
- Church, A. (1936) An unsolvable problem of elementary number theory. *American Journal of Mathematics* 58:345-63. [aJKT]
- Cook, S. (1971) The complexity of theorem-proving procedures. *Proceedings of the 3d Annual ACM Symposium on the Theory of Computing*. New York. [aJKT]
- Corbeil, J.-C. (1986) *The Stoddart visual dictionary*. Stoddart Publishing Co. [aJKT]
- Cowey, A. (1979) Cortical maps and visual perception. *Quarterly Journal of Experimental Psychology* 31:1-17. [aJKT]
- Crick, F. & Asunuma, C. (1986) Certain aspects of the anatomy and physiology of the cerebral cortex. In: *Parallel distributed processing*, ed. D. Rumelhart & J. McClelland. MIT Press. [aJKT]
- Daniel, P. & Whitteridge, D. (1961) The representation of the visual field on the cerebral cortex in monkeys. *Journal of Physiology* 159:203-21. [aJKT]
- Dantzig, G. B. (1957) Discrete-variable extremum problems. *Operations Research* 5:266-77. [rJKT]
- Davis, L. (1989) Mapping classifier systems into neural networks. In: *Advances in Neural Information Processing Systems 1*, ed. D. S. Touretzky. Morgan Kaufmann. [GWS]
- Desimone, R. (1990) Untitled abstract presented at Visual Search: Segmentation, Attention and Identification, January 19-21, 1990, Irvine, CA. [rJKT]
- Desimone, R., Chein, S., Moran, J. & Ungerleider, L. (1985) Contour, color, and shape analysis beyond the striate cortex. *Vision Research* 25(3):441-52. [aJKT]
- Desimone, R., Moran, J. & Spitzer, H. (in press) Neural mechanisms of attention in extrastriate cortex of monkeys. In: *Competition and cooperation in neural nets 2*, ed. M. A. Arbib. [SWZ]
- Desimone, R. & Ungerleider, L. G. (1989) Neural mechanisms of visual processing monkeys. In: *Handbook of Neuropsychology, vol. II.*, ed. E. Boller & J. Grafman. Elsevier Press. [RD]
- Desimone, R., Wessinger, M., Thomas, L. & Schneider, W. (1989) Effects of deactivation of lateral pulvinar or superior colliculus on the ability to selectively attend to a visual stimulus. *Society for Neuroscience Abstracts* 15:162. [RD]
- Dobkin, D., Lipton, R. & Reiss, S. (1979) Linear programming is log space hard for P. *Information Processing Letters* 5:96-97. [rJKT]
- Dobson, V. & Rose, D. (1985) Models and metaphysics: The nature of explanation revisited. In: *Models of the visual cortex*, ed. D. Rose & V. Dobson. John Wiley & Sons. [aJKT]
- Dowling, J. (1987) *The retina*. Belknap Press. [rJKT]
- Downing, C. & Pinker, S. (1985) The spatial structure of visual attention. In: *Attention and Performance XI*, ed. M. Posner & O. Marin. Lawrence Erlbaum. [aJKT]
- Duncan, J. (1980) The locus of interference in the perception of simultaneous stimuli. *Psychological Review* 87(3):272-300. [aJKT]
- Duncan, J. & Humphreys, G. W. (1989) Visual search and stimulus similarity. *Psychological Review* 96:433-58. [KRC]
- Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M. & Reitboeck, H. J. (1988) Coherent oscillations: A mechanism of feature detection in the visual cortex? Multiple electrode and correlation analysis in the cat. *Biological Cybernetics* 60:121-30. [RMS]
- Egeth, H., Virzi, R. & Garbart, H. (1984) Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance* 10(1):32-39. [rJKT]
- Eriksen, C. W., O'Hara, W. P. & Eriksen, B. A. (1982) Response competition effects in same-different judgments. *Perception & Psychophysics* 32:261-70. [LEK]
- Farah, M. J. (1985) Psychophysical evidence for a shared representational medium for mental images and percepts. *Journal on Experimental Psychology: General* 114:91-103. [MM]
- Feldman, J. (1985) Connectionist models and their applications. *Cognitive Science* 9(1):1-169. [aJKT]
- Feldman, J. & Ballard, D. (1982) Connectionist models and their properties. *Cognitive Science* 6:205-54. [aJKT]
- Finke, R. A. (1985) Theories relating mental imagery to perception. *Psychological Bulletin* 98:236-59. [MM]
- Fleet, D., Hallett, P. & Jepson, A. (1985) Spatio-temporal inseparability in early visual processing. *Biological Cybernetics* 52:153-64. [aJKT]
- Fleet, D. & Jepson, A. (1989) Hierarchical construction of orientation and velocity selective filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. II, 3:315-25. [aJKT]
- Fodor, J. A. (1985) Précis of *The modularity of mind*. *Behavioral and Brain Sciences* 8:1-42. [LEK]
- Funt, B. V. (1980) Problem solving with diagrammatic representations. *Artificial Intelligence* 13:201-30. [MM]
- Fuster, J. (1988) Attentional modulation of inferotemporal neuron responses to visual features. *Proceedings of the Society of Neuroscience*. Toronto. [aJKT]
- Gamble, E. & Koch, C. (1987) The dynamics of free calcium in dendritic spines in response to repetitive synaptic input. *Science* 236:1311-15. [RMS]
- Gardin, F. & Meltzer, B. (1989) The analogical representation of naive physics. *Artificial Intelligence* 38:139-59. [MM]
- Garey, M. & Johnson, D. (1979) *Computers and intractability: A guide to the theory of NP-completeness*. W. H. Freeman & Co. [arJKT, PRK]
- Gibson, J. J. (1966) *The senses considered as perceptual systems*. Houghton Mifflin. [LEK]
- (1979) *The ecological approach to visual perception*. Houghton Mifflin. [LEK]
- Gleitman, H. & Jonides, J. (1976) The cost of categorization in visual search: Incomplete processing of targets and field items. *Perception & Psychophysics* 20(4):281-88. [aJKT]
- Goldberg, D. E. (1989) *Genetic algorithms in search, optimization and machine learning*. Addison-Wesley. [GWS]
- Goldman-Rakic, P. S. (1988) Changing concepts of cortical connectivity: Parallel distributed cortical networks. In: *Neurobiology of neocortex*, ed. P. Rakic & W. Singer. John Wiley & Sons. [GWS]
- Gray, C. M. & Singer, W. (1989) Stimulus specific neuronal oscillations in orientation columns. *Proceedings of the National Academy of Science* 86:1698-1702. [RMS]
- Grimson, W. E. L. (1986) The combinatorics of local constraints in model-based recognition and localization from sparse data. *Journal of the Association for Computing Machinery* 33(4):658-86. [aJKT]
- Haenny, P., Maunsell, J. & Schiller, P. (1988) State dependent activity in monkey visual cortex II. Retinal and extraretinal factors in V4. *Experimental Brain Research* 69:245-59. [arJKT]
- Haenny, P. & Schiller, P. (1988) State dependent activity in monkey visual cortex I. Single cell activity in V1 and V4 on visual tasks. *Experimental Brain Research* 69:225-44. [aJKT]
- Hartline, H. (1940) The receptive fields of optic nerve fibers. *American Journal of Physiology* 130:690-99. [aJKT]
- Hebb, D. (1949) *The organization of behavior*. John Wiley & Sons. [aJKT]
- Hebb, D. O. (1958) Alice in Wonderland or psychology among the biological sciences. In: *Biological and biochemical bases of behavior*, ed. H. F. Harlow & C. N. Woolsey. University of Wisconsin Press. [AH]
- Hinton, G. (1981) Shape representation in parallel systems. *Proceedings of the International Joint Conference on Artificial Intelligence*, Vancouver. [aJKT]
- Hinton, G. E. (1989) Connectionist learning procedures. *Artificial Intelligence* 40:185-234. [DGL]

- Hoffman, J., Nelson, B. & Houck, M. (1983) The role of attentional resources in automatic detection. *Cognitive Psychology* 51:379–410. [aJKT]
- Hoffman, J. E. (1979) A two-stage model of visual search. *Perception & Psychophysics* 25:319–27. [KRC]
- Holland, J. H. (1975) *Adaptation in natural and artificial systems*. University of Michigan Press. [GWS, rJKT]
- Hopfield, J. J. & Tank, D. W. (1986) Computing with neural circuits: A model. *Science* 233:625–33. [RMS]
- Hubel, D. & Wiesel, T. (1977) Functional architecture of macaque visual cortex. *Proceedings of the Royal Society of London B* 198:1–59. [aJKT]
- Hummel, R. & Zucker, S. (1983) On the foundations of relaxation labeling processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 5:267–87. [aJKT]
- Jonides, J. & Gleitman, H. (1976) The benefit of categorization in visual search: Target location without identification. *Perception & Psychophysics* 20(4):289–98. [aJKT]
- Kaas, J. H. (1989) Why does the brain have so many visual areas? *Journal of Cognitive Neuroscience* 1(2):121–35. [aJKT]
- Kahneman, D. & Treisman, A. (1984) Changing views of attention and automaticity. In: *Varieties of attention*, ed. R. Parasuraman & J. Beatty. Academic Press. [AT]
- Kaufman, L., Okada, Y., Tripp, J. & Weinberg, H. (1984) Evoked neuromagnetic fields. In: *Brain and information: Event-related potentials*, ed. R. Karrer, J. Cohen, & P. Tuetting. *Annals of the New York Academy of Sciences* 425:722–42. [LEK]
- Kertzman, C. & Robinson, D. L. (1988) Contributions of the superior colliculus of the monkey to visual spatial attention. *Society for Neuroscience Abstracts* 14:831. [RD]
- Kosslyn, S. M. (1980) *Image and mind*. Harvard University Press. [MM]
- Krueger, L. E. (1978) A theory of perceptual matching. *Psychological Review* 85:278–304. [LEK]
- (1989) Cognitive impenetrability of perception. *Behavioral and Brain Sciences* 12(4):769–70. [LEK]
- Krueger, L. E. & Chignell, M. H. (1985) Same-different judgments under high speed stress: Missing-feature principle predominates in early processing. *Perception & Psychophysics* 38:183–93. [LEK]
- Kuffler, S. (1953) Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology* 16:37–68. [aJKT]
- Larkin, J. H. & Simon, H. A. (1987) Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science* 11:65–99. [MM]
- Lawler, E. (1976) *Combinatorial optimization: Networks and matroids*. Holt, Rinehart & Winston. [rJKT]
- Liu, L., Zhao, N. & Bian, Z. (1989) Can early stage vision detect topology? *Proceedings of the International Joint Conference on Artificial Intelligence* 11:1591–95. [MM]
- Llinas, R. & Yarom, Y. (1986) Oscillatory properties of guinea-pig inferior olivary neurones and their pharmacological modulation: An *in vitro* study. *Journal of Physiology* 376:163–82. [RMS]
- Lowe, D. G. (1985) *Perceptual organization and visual recognition*. Kluwer Academic Publishers. [DGL]
- (1987) Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence* 31:355–95. [DGL, rJKT]
- (1990) Visual recognition as probabilistic inference from spatial relations. In: *AI and the eye*, ed. A. Blake & T. Troscianko. Wiley. [DGL]
- Mackworth, A. & Freuder, E. (1985) The complexity of some polynomial network consistency algorithms for constraint satisfaction problems. *Artificial Intelligence* 25:65–74. [aJKT]
- Marr, D. (1982) *Vision: A computational investigation into the human representation and processing of visual information*. W. H. Freeman. [arJKT, LEK, RMS]
- Maunsell, J. (1989) Personal communication, June. [rJKT]
- Maunsell, J. & Newsome, W. (1987) Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience* 10:363–401. [arJKT]
- Maunsell, J., Sclar, G. & Nealey, T. (1988) Task-specific signals in area V4 of monkey visual cortex. *Proceedings of the Society of Neuroscience*, Toronto. [aJKT]
- Maunsell, J. & Van Essen, D. (1987) Topographic organization of the middle temporal visual area in the macaque monkey: Representational biases and the relationship to callosal connections and myeloarchitectonic boundaries. *Journal of Comparative Neurology* 266:535–55. [rJKT]
- Maxwell, N. (1985) Methodological problems of neuroscience. In: *Models of the visual cortex*, ed. D. Rose & V. Dobson. John Wiley & Sons. [aJKT]
- McLeod, P., Driver, J. & Crisp, J. (1988) Visual search for conjunctions of movements and form is parallel. *Nature* 332:154–55. [JMW]
- Miller, J. P., Rall, W. & Rinzel, J. (1985) Synaptic amplification by active membrane in dendritic spines. *Brain Research* 325:325–30. [RMS]
- Minsky, M. & Papert, S. (1969) *Perceptrons*. MIT Press. [MM]
- Mohnhaupt, M. & Neumann, B. (in press) Understanding object motion: Recognition, learning and spatio-temporal reasoning. *Journal of Robotics and Autonomous Systems*. North Holland. [MM]
- Moore, E. F. (1956) Gedanken-experiments on sequential machines. In: *Automata studies*, ed. C. E. Shannon & J. McCarthy. Princeton University Press. [WRU]
- Moran, J. & Desimone, R. (1985) Selective attention gates visual processing in the extrastriate cortex. *Science* 229:782–84. [arJKT, RD, SWZ]
- Motter, B. (1988) Responses of visual cortical neurons during a focal attentive task. *Proceedings of the Society of Neuroscience*, Toronto. [aJKT]
- Mountcastle, V. (1957) Modality and topographic properties of single neurons of cat's somatic sensory cortex. *Journal of Neurophysiology* 20:408–34. [aJKT]
- Mountcastle, V., Motter, B., Steinmetz, M. & Sestokas, A. (1987) Common and differential effects of attentive fixation on the excitability of parietal and prestriate (V4) cortical visual neurons in the Macaque monkey. *Journal of Neuroscience* 7(7):2239–55. [aJKT]
- Nakayama, K. & Silverman, G. H. (1986) Serial and parallel processing of visual feature conjunctions. *Nature* 320:264–65. [JMW]
- Neisser, U. (1967) *Cognitive psychology*. Appleton-Century-Crofts. [aJKT, KRC, J-OE, RMS]
- Okada, Y. C., Tanenbaum, R., Williamson, S. J. & Kaufman, L. (1984) Somatotopic organization of the human somatosensory cortex revealed by neuromagnetic measurements. *Experimental Brain Research* 56:197–205. [LEK]
- Parasuraman, R. & Davies, D., eds. (1984) *Varieties of attention*. Academic Press. [aJKT]
- Pashler, H. (1987) Detecting conjunctions of color and form: Reassessing the serial search hypothesis. *Perception and Psychophysics* 41:191–201. [AH]
- Pippenger, N. (1978) Complexity theory. *Scientific American* 238(6):114–24. [aJKT]
- Poggio, T. (1982) Visual algorithms. *AI Memo 683*. MIT Press. [aJKT]
- Posner, J. I., Choate, L. S., Rafal, R. K. & Vaughn, J. (1985) Inhibition of return: Neural mechanisms and function. *Cognitive Neuropsychology* 2:211–28. [RD]
- Pour-El, M. B. & Richards, I. (1981) The wave equation with computable initial data such that its unique solution is not computable. *Advances in Mathematics* 39:215–39. [BWD]
- (1982) Noncomputability in models of physical phenomena. *International Journal of Theoretical Physics* 21:553–55. [BWD]
- Proctor, R. W. (1981) A unified theory for matching-task phenomena. *Psychological Review* 88:291–326. [LEK]
- Proctor, R. W. & Rao, K. V. (1983) Evidence that the same-different disparity in letter matching is not attributable to response bias. *Perception & Psychophysics* 34:72–76. [LEK]
- Pylyshyn, Z. (1984) *Computation and cognition*. MIT Press/Bradford Books. [aJKT]
- (1987) *The robot's dilemma: The frame problem in AI*. Ablex. [RE]
- (1989) The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition* 32:65–97. [RE]
- Pylyshyn, Z. & Biederman, I. (1988) *Computational processes in human vision: An interdisciplinary perspective*. Ablex. [RE]
- Pylyshyn, Z. & Storm, R. (1988) Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision* 3:179–97. [RE]
- Quinlan, P. & Humphreys, G. (1987) Visual search for targets defined by combinations of color, shape and size: An examination of the task constraints on feature and conjunction searches. *Perception and Psychophysics* 41(5):455–72. [rJKT]
- Rabbitt, P. (1978) Sorting, categorization and visual search. In: *Handbook of perception: Perceptual processing*, vol. IX, ed. E. Carterette & M. Friedman. Academic Press. [aJKT]
- Ramachandran, V. S. (1985) Guest editorial: The neurobiology of perception. *Perception* 14:1–14. [RMS]
- Rensinck, R. (1989) Personal communication, University of British Columbia, September. [rJKT]
- Rosch, E. (1978) Principles of categorization. In: *Cognition and categorization*, ed. E. Rosch & B. B. Lloyd. Erlbaum. [MM]
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M. & Boyes-Bream, P. (1976) Basic objects in natural categories. *Cognitive Psychology* 8:382–439. [MM]
- Rosenfeld, A. (1962) Automatic recognition of basic terrain types from aerial photographs. *Photogrammetric Engineering* 28:115–32. [aJKT]
- Rosenkrantz, D. & Stearns, R. (1983) NP-complete problems. In: *The encyclopedia of computer science and engineering*, 2nd ed., ed. A. Ralston & E. Reilly. Van Nostrand Reinhold Co. [rJKT]

- Rubin, J. & Kanwisher, N. (1985) Topological perception: Holes in an experiment. *Perception and Psychophysics* 37. [MM]
- Rumelhart, D. & McClelland, J. (1986a) PDP models and general issues in cognitive science. In: *Parallel distributed processing*, ed. D. Rumelhart & J. McClelland. MIT Press. [aJKT]
- (1986b) *Parallel distributed processing*. MIT Press. [aJKT]
- Runeson, S. (1977) On the possibility of "smart" perceptual mechanisms. *Scandinavian Journal of Psychology* 18:172-79. [LEK, rJKT]
- Sagi, D. & Julesz, B. (1986) "Where" and "What" in vision. *Science* 228:1217-19. [aJKT]
- Schwartz, E. (1977) Spatial mapping in the primate sensory projection: Analytic structure and relevance to perception. *Biological Cybernetics* 25:181-94. [aJKT]
- Sclar, G., Lennie, P. & DePriest, D. (1989) Contrast adaptation in striate cortex of macaque. *Vision Research* 29:747-55. [SWZ]
- Searle, J. (1990) Is the brain's mind a computer program? *Scientific American* 262(1):26-31. [rJKT]
- Selman, B. (1989) Personal communication. Department of Computer Science, University of Toronto, January. [rJKT]
- Shepard, G. M. & Brayton, R. K. (1987) Logic operations are properties of computer simulated interactions between excitable dendritic spines. *Neuroscience* 23:151-66. [RMS]
- Siegel, R. M. (in press) Non-linear dynamical system theory and primary visual cortical processing. *Physica D*. [RMS]
- Simon, H. (1962) The architecture of complexity. *Proceedings of the American Philosophical Society* 106:467-82. [aJKT, GWS]
- Skarda, C. A. & Freeman, W. J. (1987) How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences* 10:161-96. [RMS]
- Slovan, A. (1975). Afterthoughts on analogical representations. In: *Proceedings on Theoretical Issues in Natural Language Processing* 164-68. [MM]
- Spitzer, H., Desimone, R. & Moran, T. (1988) Both behavioral and neuronal performance are improved by increased attention. *Proceedings of the Society for Neuroscience*, Toronto. [aJKT]
- (1988) Increased attention enhances both behavioral and neuronal performance. *Science* 240:338-40. [RD]
- Sporns, O., Gally, J. A., Recke, G. N. & Edelman, G. M. (1989) Recurrent signaling among simulated neuronal groups lead to coherency in their oscillatory activity. *Proceedings of the National Academy of Science* 86:7265-69. [RMS]
- Steels, L. (1988) Steps towards common sense. *Proceedings ECAI* 88:49-54. [MM]
- Stensaas, S., Eddington, D. & Dobelle, W. (1974) The topography and variability of the primary visual cortex in man. *Journal of Neurosurgery* 40:747-55. [aJKT]
- Stockmeyer, L. & Chandra, A. (1979) Intrinsically difficult problems. *Scientific American*, May. [aJKT]
- Stone, J., Dreher, B. & Leventhal, A. (1979) Hierarchical and parallel mechanisms in the organization of the visual cortex. *Brain Research Reviews* 1:345-94. [aJKT]
- Strong, G. W. & Whitehead, B. A. (1989) A solution to the tag assignment problem for neural networks. *Behavioral and Brain Sciences* 12:381-433. [GWS]
- Treisman, A. (1982) Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance* 8:194-214. [AH]
- (1985) Preattentive processing in vision. *Computer Vision, Graphics and Image Processing* 31:156-77. [arJKT, AH, AT]
- (1986) Features and objects in visual processing. *Scientific American* 255:144B-125. [KRC, AT]
- (1988) Features and objects: The fourteenth Bartlett memorial lecture. *The Quarterly Journal of Experimental Psychology* 40A(2):201-37. [aJKT, KRC, AH, AT]
- Treisman, A. & Gelade, G. (1980) A feature-integration theory of attention. *Cognitive Science* 12:99-136. [aJKT, KRC]
- Treisman, A. & Gormican, S. (1988) Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review* 95(1):15-48. [aJKT, AT]
- Treisman, A. & Sato, S. (1990) Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance*, 16. [arJKT, AT, JMW]
- Treisman, A. & Schmidt, H. (1982) Illusory conjunctions in the perception of objects. *Cognitive Psychology* 14:107-41. [aJKT]
- Treisman, A. & Souther, J. (1985) Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General* 114:285-310. [aJKT]
- Tsotsos, J. (1987a) Representational axes and temporal cooperative processes. In: *Vision, brain and cooperative computation*, ed. M. Arbib & A. Hansen. MIT Press/Bradford Books. [aJKT]
- (1987b) Image understanding. In: *The encyclopedia of artificial intelligence*, ed. S. Shapiro. John Wiley & Sons. [aJKT]
- (1987c) *Analyzing vision at the complexity level: Constraints on an architecture, an explanation for visual search performance, and computational justification for attentive processes*, RBCV-TR-87-20, Department of Computer Science, University of Toronto, September. [arJKT]
- (1988) A "complexity level" analysis of immediate vision. *International Journal of Computer Vision* 1(4):303-20. [arJKT]
- (1989) The complexity of perceptual search tasks. *Proceedings of the International Joint Conference on Artificial Intelligence*, Detroit. [aJKT]
- (in preparation) Attentional influences in vision: Applying an inhibitory attentional beam to intermediate level vision. [aJKT]
- Turing, A. (1937) On computable numbers with an application to the Entscheidungs problem. *Proceedings of the London Mathematical Society* 2(43):230-65. [aJKT]
- Uhr, L. (1972) Layered "recognition cone" networks that preprocess, classify and describe. *IEEE Transactions on Computers* C-21:758-68. [aJKT]
- (1980) Psychological motivation and underlying concepts. In: *Structured computer vision*, ed. S. Tanimoto & A. Klinger. Academic Press. [aJKT]
- Ullman, S. (1984) Visual routines. *Cognition* 18:97-159. [RE, J-OE]
- (1989) Aligning pictorial descriptions: An approach to object recognition. *Cognition* 32:193-254. [MM]
- Ungerleider, L. & Mishkin, M. (1982) Two cortical visual systems. In: *Analysis of visual behavior*, ed. D. Ingle, M. Goodale & R. Mansfield. MIT Press. [aJKT, RD]
- Van Doorn, A., van de Grind, W. & Koenderink, J., ed. (1984) *Limits in perception*. VNU Science Press. [aJKT]
- Van Essen, D. & Anderson, C. (in press) Information processing strategies and pathways in the primate retina and visual cortex. In: *Introduction to Neural and Electronic Networks*, ed. S. Zornetzer, J. Davis & C. Lau. Academic Press. [arJKT]
- Van Essen, D. & Maunsell, J. (1983) Hierarchical organization and functional streams in the visual cortex. *Trends in Neuroscience* 6:370-75. [aJKT]
- Vergis, A., Steiglitz, K. & Dickinson, B. (1986) The complexity of analog computation. *Mathematics and Computers in Simulation* 28:91-113. [BWD]
- Von Bekesy, G. (1956) Current status of theories of hearing. *Science* 123:779-83. [LEK]
- Watson, A. & Ahumada, A. (1987) An orthogonal oriented quadrature hexagonal image pyramid. *NASA Technical Memorandum* 100054. [aJKT]
- Wise, S. & Desimone, R. (1988) Behavioral neurophysiology: Insights into seeing and grasping. *Science* 242:736-41. [RD]
- Wolfe, J., Cave, K. & Franzel, S. (1989) Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance* 15:419-33. [arJKT, KRC, AH, AT, JMW]
- Wolfe, J., Cave, K. & Yu, K. (1988) Direction attention to complex objects. *Proceedings of the Society for Neuroscience*, Toronto. [aJKT]
- Zeki, S. (1978) Uniformity and diversity of structure and function in rhesus monkey prestriate visual cortex. *Journal of Physiology* 277:273-90. [aJKT]
- Zucker, S. (1985) Does connectionism suffice? *Behavioral and Brain Sciences* 8(2):301-2. [aJKT]
- Zucker, S. W. (1983) Cooperative grouping and early orientation selection. In: *Physical and biological processing of images*, ed. A. Sleigh. Springer. [RMS]